Calculus on Manifolds

A Solution Manual for Spivak (1965)

Jianfei Shen

School of Economics, The University of New South Wales

Sydney, Australia 2010

Contents

1

FUNCTIONS ON EUCLIDEAN SPACE

1.1 Norm and Inner Product

► Exercise 1 (1-1*). *Prove that* $\|\boldsymbol{x}\| \leqslant \sum_{i=1}^{n} \left| x^i \right|$.

Proof. Let $\boldsymbol{x} = \left(x^1, \ldots, x^n\right)$. Then

$$\left(\sum_{i=1}^{n} \left| x^i \right| \right)^2 = \sum_{i=1}^{n} \left(x^i \right)^2 + \sum_{i \neq j} \left| x^i x^j \right| \geqslant \sum_{i=1}^{n} \left(x^i \right)^2 = \|\boldsymbol{x}\|^2 .$$

Taking the square root of both sides gives the result. □

► Exercise 2 (1-2). *When does equality hold in Theorem 1-1 (3)* $\left[\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \right]$*?*

Proof. We reprove that $\left| \langle \boldsymbol{x}, \boldsymbol{y} \rangle \right| \leqslant \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$ for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. Obviously, if $\boldsymbol{x} = \boldsymbol{0}$ or $\boldsymbol{y} = \boldsymbol{0}$, then $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\| = 0$. So we assume that $\boldsymbol{x} \neq \boldsymbol{0}$ and $\boldsymbol{y} \neq \boldsymbol{0}$. We first find some $\boldsymbol{w} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ such that $\langle \boldsymbol{w}, \alpha \boldsymbol{y} \rangle = 0$. Write $\boldsymbol{w} = \boldsymbol{x} - \alpha \boldsymbol{y}$. Then

$$0 = \langle \boldsymbol{w}, \alpha \boldsymbol{y} \rangle = \langle \boldsymbol{x} - \alpha \boldsymbol{y}, \alpha \boldsymbol{y} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle - \alpha^2 \|\boldsymbol{y}\|^2$$

implies that

$$\alpha = \langle \boldsymbol{x}, \boldsymbol{y} \rangle \big/ \|\boldsymbol{y}\|^2 .$$

Then

$$\|\boldsymbol{x}\|^2 = \|\boldsymbol{w}\|^2 + \|\alpha \boldsymbol{y}\|^2 \geqslant \|\alpha \boldsymbol{y}\|^2 = \left(\frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{y}\|} \right)^2 .$$

Hence, $\left| \langle \boldsymbol{x}, \boldsymbol{y} \rangle \right| \leqslant \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|$. Particularly, the above display holds with equality if and only if $\|\boldsymbol{w}\| = 0$, if and only if $\boldsymbol{w} = \boldsymbol{0}$, if and only if $\boldsymbol{x} - \alpha \boldsymbol{y} = \boldsymbol{0}$, if and only if $\boldsymbol{x} = \alpha \boldsymbol{y}$.

Since

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{y}\|^2 = \langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y} \rangle = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2 \langle \boldsymbol{x}, \boldsymbol{y} \rangle &\leqslant \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2 \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\| \\ &= \left(\|\boldsymbol{x}\| + \|\boldsymbol{y}\| \right)^2 , \end{aligned}$$

equality holds precisely when $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \|\boldsymbol{x}\| \cdot ||\boldsymbol{y}||$, i.e., when one is a nonnegative multiple of the other. □

► EXERCISE 3 (1-3). *Prove that* $\|\boldsymbol{x} - \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$. *When does equality hold?*

PROOF. By Theorem 1-1 (3) we have $\|\boldsymbol{x} - \boldsymbol{y}\| = \|\boldsymbol{x} + (-\boldsymbol{y})\| \leqslant \|\boldsymbol{x}\| + \|-\boldsymbol{y}\| = \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$. The equality holds precisely when one vector is a non-positive multiple of the other. □

► EXERCISE 4 (1-4). *Prove that* $\big| \|\boldsymbol{x}\| - \|\boldsymbol{y}\| \big| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|$.

PROOF. We have $\|\boldsymbol{x} - \boldsymbol{y}\|^2 = \sum_{i=1}^n (x_i - y_i)^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\sum_{i=1}^n x_i y_i \geqslant \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\|\boldsymbol{x}\| \|\boldsymbol{y}\| = (\|\boldsymbol{x}\| - \|\boldsymbol{y}\|)^2$. Taking the square root of both sides gives the result. □

► EXERCISE 5 (1-5). *The quantity* $\|\boldsymbol{y} - \boldsymbol{x}\|$ *is called the distance between* $\boldsymbol{x}$ *and* $\boldsymbol{y}$. *Prove and interpret geometrically the "triangle inequality":* $\|\boldsymbol{z} - \boldsymbol{x}\| \leqslant \|\boldsymbol{z} - \boldsymbol{y}\| + \|\boldsymbol{y} - \boldsymbol{x}\|$.

PROOF. The inequality follows from Theorem 1-1 (3):

$$\|\boldsymbol{z} - \boldsymbol{x}\| = \|(\boldsymbol{z} - \boldsymbol{y}) + (\boldsymbol{y} - \boldsymbol{x})\| \leqslant \|\boldsymbol{z} - \boldsymbol{y}\| + \|\boldsymbol{y} - \boldsymbol{x}\|.$$

Geometrically, if $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ are the vertices of a triangle, then the inequality says that the length of a side is no larger than the sum of the lengths of the other two sides. □

► EXERCISE 6 (1-6). *If* f *and* g *be integrable on* $[a, b]$.

a. *Prove that* $\left|\int_a^b f \cdot g\right| \leqslant \left(\int_a^b f^2\right)^{\frac{1}{2}} \cdot \left(\int_a^b g^2\right)^{\frac{1}{2}}$.

b. *If equality holds, must* $f = \lambda g$ *for some* $\lambda \in \mathbb{R}$? *What if* f *and* g *are continuous?*

c. *Show that Theorem 1-1 (2) is a special case of (a).*

PROOF.

a. Theorem 1-1 (2) implies the inequality of Riemann sums:

$$\left|\sum_i f(x_i)\, g(x_i)\, \Delta x_i\right| \leqslant \left(\sum_i f(x_i)^2\, \Delta x_i\right)^{1/2} \left(\sum_i g(x_i)^2\, \Delta x_i\right)^{1/2}.$$

Taking the limit as the mesh approaches 0, one gets the desired inequality.

b. No. We could, for example, vary f at discrete points without changing the values of the integrals. If f and g are continuous, then the assertion is true. In fact, suppose that for each $\lambda \in \mathbb{R}$, there is an $x \in [a, b]$ with

$[f(x) - \lambda g(x)]^2 > 0$. Then the inequality holds true in an open neighborhood of x since f and g are continuous. So $\int_a^b (f - \lambda g)^2 > 0$ since the integrand is always non-negative and is positive on some subinterval of $[a, b]$. Expanding out gives $\int_a^b f^2 - 2\lambda \int_a^b f \cdot g + \lambda^2 \int_a^b g^2 > 0$ for all λ. Since the quadratic has no solutions, it must be that its discriminant is negative.

c. Let $a = 0$, $b = n$, $f(x) = x_i$ and $g(x) = y_i$ for all $x \in [i-1, i)$ for $i = 1, \ldots, n$. Then part (a) gives the inequality of Theorem 1-1 (2). Note, however, that the equality condition does not follow from (a). □

► Exercise 7 (1-7). *A linear transformation* $\mathrm{M} : \mathbb{R}^n \to \mathbb{R}^n$ *is called norm preserving if* $\|\mathrm{M}\boldsymbol{x}\| = \|\boldsymbol{x}\|$*, and inner product preserving if* $\langle \mathrm{M}\boldsymbol{x}, \mathrm{M}\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$*.*

a. *Prove that* M *is norm preserving if and only if* M *is inner product preserving.*

b. *Prove that such a linear transformation* M *is 1-1 and* M^{-1} *is of the same sort.*

Proof.

(a) If M is norm preserving, then the polarization identity together with the linearity of M give:

$$\begin{aligned}\langle \mathrm{M}\boldsymbol{x}, \mathrm{M}\boldsymbol{y}\rangle &= \frac{\|\mathrm{M}\boldsymbol{x} + \mathrm{M}\boldsymbol{y}\|^2 - \|\mathrm{M}\boldsymbol{x} - \mathrm{M}\boldsymbol{y}\|^2}{4} \\ &= \frac{\|\mathrm{M}(\boldsymbol{x} + \boldsymbol{y})\|^2 - \|\mathrm{M}(\boldsymbol{x} - \boldsymbol{y})\|^2}{4} \\ &= \frac{\|\boldsymbol{x} + \boldsymbol{y}\|^2 - \|\boldsymbol{x} - \boldsymbol{y}\|^2}{4} \\ &= \langle \boldsymbol{x}, \boldsymbol{y}\rangle .\end{aligned}$$

If M is inner product preserving, then one has by Theorem 1-1 (4):

$$\|\mathrm{M}\boldsymbol{x}\| = \sqrt{\langle \mathrm{M}\boldsymbol{x}, \mathrm{M}\boldsymbol{x}\rangle} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle} = \|\boldsymbol{x}\| .$$

(b) Take any $\mathrm{M}\boldsymbol{x}, \mathrm{M}\boldsymbol{y} \in \mathbb{R}^n$ with $\mathrm{M}\boldsymbol{x} = \mathrm{M}\boldsymbol{y}$. Then $\mathrm{M}\boldsymbol{x} - \mathrm{M}\boldsymbol{y} = \mathbf{0}$ and so

$$0 = \langle \mathrm{M}\boldsymbol{x} - \mathrm{M}\boldsymbol{y}, \mathrm{M}\boldsymbol{x} - \mathrm{M}\boldsymbol{y}\rangle = \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}\rangle ;$$

but the above equality forces $\boldsymbol{x} = \boldsymbol{y}$; that is, M is 1-1.

Since $\mathrm{M} \in \mathfrak{L}(\mathbb{R}^n)$ and M is injective, it is invertible; see Axler (1997, Theorem 3.21). Hence, $\mathrm{M}^{-1} \in \mathfrak{L}(\mathbb{R}^n)$ exists. For every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, we have

$$\|\mathrm{M}^{-1}\boldsymbol{x}\| = \left\|\mathrm{M}\left(\mathrm{M}^{-1}\boldsymbol{x}\right)\right\| = \|\boldsymbol{x}\| ,$$

and

$$\left\langle \mathrm{M}^{-1}\boldsymbol{x}, \mathrm{M}^{-1}\boldsymbol{y}\right\rangle = \left\langle \mathrm{M}\left(\mathrm{M}^{-1}\boldsymbol{x}\right), \mathrm{M}\left(\mathrm{M}^{-1}\boldsymbol{y}\right)\right\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle .$$

Therefore, M^{-1} is also norm preserving and inner product preserving. □

► EXERCISE 8 (1-8). *If $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ are non-zero, the* angle *between $\boldsymbol{x}$ and $\boldsymbol{y}$, denoted $\angle(\boldsymbol{x}, \boldsymbol{y})$, is defined as* $\arccos\left(\langle \boldsymbol{x}, \boldsymbol{y}\rangle \big/ \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|\right)$, *which makes sense by Theorem 1-1 (2). The linear transformation* T *is* angle preserving *if* T *is 1-1, and for $\boldsymbol{x}, \boldsymbol{y} \neq \mathbf{0}$ we have $\angle(\mathrm{T}\boldsymbol{x}, \mathrm{T}\boldsymbol{y}) = \angle(\boldsymbol{x}, \boldsymbol{y})$.*

a. *Prove that if* T *is norm preserving, then* T *is angle preserving.*

b. *If there is a basis $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ of $\mathbb{R}^n$ and numbers $\lambda_1, \ldots, \lambda_n$ such that $\mathrm{T}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$, prove that* T *is angle preserving if and only if all $|\lambda_i|$ are equal.*

c. *What are all angle preserving $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^n$?*

PROOF.

(a) If T is norm preserving, then T is inner product preserving by the previous exercise. Hence, for $\boldsymbol{x}, \boldsymbol{y} \neq \mathbf{0}$,

$$\angle(\mathrm{T}\boldsymbol{x}, \mathrm{T}\boldsymbol{y}) = \arccos\left(\frac{\langle \mathrm{T}\boldsymbol{x}, \mathrm{T}\boldsymbol{y}\rangle}{\|\mathrm{T}\boldsymbol{x}\| \cdot \|\mathrm{T}\boldsymbol{y}\|}\right) = \arccos\left(\frac{\langle \boldsymbol{x}, \boldsymbol{y}\rangle}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|}\right) = \angle(\boldsymbol{x}, \boldsymbol{y}).$$

(b) We first suppose that T is angle preserving. Since $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ is a basis of $\mathbb{R}^n$, all $\boldsymbol{x}_i$'s are nonzero. Since

$$\begin{aligned}\angle\left(\mathrm{T}\boldsymbol{x}_i, \mathrm{T}\boldsymbol{x}_j\right) = \arccos\left(\frac{\langle \mathrm{T}\boldsymbol{x}_i, \mathrm{T}\boldsymbol{x}_j\rangle}{\|\mathrm{T}\boldsymbol{x}_i\| \cdot \left\|\mathrm{T}\boldsymbol{x}_j\right\|}\right) &= \arccos\left(\frac{\langle \lambda_i\boldsymbol{x}_i, \lambda_j\boldsymbol{x}_j\rangle}{\|\lambda_i\boldsymbol{x}_i\| \cdot \left\|\lambda_j\boldsymbol{x}_j\right\|}\right) \\ &= \arccos\left(\frac{\lambda_i\lambda_j\left\langle \boldsymbol{x}_i, \boldsymbol{x}_j\right\rangle}{|\lambda_i| \cdot \left|\lambda_j\right| \cdot \|\boldsymbol{x}_i\| \cdot \|\boldsymbol{x}_j\|}\right) \\ &= \angle\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right),\end{aligned}$$

it must be the case that

$$\lambda_i\lambda_j = |\lambda_i| \cdot \left|\lambda_j\right|.$$

Then λ_i and λ_j have the same signs. □

► EXERCISE 9 (1-9). *If $0 \leqslant \theta < \pi$, let $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^2$ have the matrix*

$$\mathbf{A} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

Show that T *is angle preserving and if $\boldsymbol{x} \neq \mathbf{0}$, then $\angle(\boldsymbol{x}, \mathrm{T}\boldsymbol{x}) = \theta$.*

PROOF. For every $(x, y) \in \mathbb{R}^2$, we have

$$\mathrm{T}(x, y) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta + y\sin\theta \\ -x\sin\theta + y\cos\theta \end{pmatrix}.$$

Therefore,

$$\left\|\mathrm{T}(x, y)\right\|^2 = x^2 + y^2 = \left\|(x, y)\right\|^2;$$

that is, T is norm preserving. Then it is angle preserving by Exercise 8 (a).

Let $\boldsymbol{x} = (a, b) \neq \mathbf{0}$. We first have

$$\langle \boldsymbol{x}, \mathrm{T}\boldsymbol{x} \rangle = a\,(a\cos\theta + b\sin\theta) + b\,(-a\sin\theta + b\cos\theta) = \left(a^2 + b^2\right)\cos\theta.$$

Hence,

$$\angle\,(\boldsymbol{x}, \mathrm{T}\boldsymbol{x}) = \arccos\left(\frac{\langle \boldsymbol{x}, \mathrm{T}\boldsymbol{x} \rangle}{\|\boldsymbol{x}\| \cdot \|\mathrm{T}\boldsymbol{x}\|}\right) = \arccos\left(\frac{\left(a^2 + b^2\right)\cos\theta}{a^2 + b^2}\right) = \theta. \qquad \square$$

► EXERCISE 10 (1-10*). *If* $\mathrm{M} : \mathbb{R}^m \to \mathbb{R}^n$ *is a linear transformation, show that there is a number* M *such that* $\|\mathrm{M}\,\boldsymbol{h}\| \leqslant M\,\|\boldsymbol{h}\|$ *for* $\boldsymbol{h} \in \mathbb{R}^m$.

PROOF. Let M's matrix be

$$\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} := \begin{pmatrix} \boldsymbol{a}^1 \\ \vdots \\ \boldsymbol{a}^n \end{pmatrix}.$$

Then

$$\mathrm{M}\,\boldsymbol{h} = \mathbf{A}\boldsymbol{h} = \begin{pmatrix} \langle \boldsymbol{a}^1, \boldsymbol{h} \rangle \\ \vdots \\ \langle \boldsymbol{a}^n, \boldsymbol{h} \rangle \end{pmatrix},$$

and so

$$\|\mathrm{M}\,\boldsymbol{h}\|^2 = \sum_{i=1}^{n} \left\langle \boldsymbol{a}^i, \boldsymbol{h} \right\rangle^2 \leqslant \sum_{i=1}^{n} \left(\|\boldsymbol{a}^i\| \cdot \|\boldsymbol{h}\|\right)^2 = \left(\sum_{i=1}^{n} \|\boldsymbol{a}^i\|^2\right) \cdot \|\boldsymbol{h}\|^2,$$

that is,

$$\|\mathrm{M}\,\boldsymbol{h}\| \leqslant \left(\sqrt{\sum_{i=1}^{n} \|\boldsymbol{a}^i\|}\right) \cdot \|\boldsymbol{h}\|.$$

Let $M = \sqrt{\sum_{i=1}^{n} \|\boldsymbol{a}^i\|}$ and we get the result. $\square$

► EXERCISE 11 (1-11). *If* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ *and* $\boldsymbol{z}, \boldsymbol{w} \in \mathbb{R}^m$, *show that* $\langle (\boldsymbol{x}, \boldsymbol{z}), (\boldsymbol{y}, \boldsymbol{w}) \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{z}, \boldsymbol{w} \rangle$ *and* $\|(\boldsymbol{x}, \boldsymbol{z})\| = \sqrt{\|\boldsymbol{x}\|^2 + \|\boldsymbol{z}\|^2}$.

PROOF. We have $(\boldsymbol{x}, \boldsymbol{z}), (\boldsymbol{y}, \boldsymbol{w}) \in \mathbb{R}^{n+m}$. Then

$$\langle (\boldsymbol{x}, \boldsymbol{z}), (\boldsymbol{y}, \boldsymbol{w}) \rangle = \sum_{i=1}^{n} x_i y_i + \sum_{j=1}^{m} z_j w_j = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{z}, \boldsymbol{w} \rangle,$$

and

$$\|(\boldsymbol{x}, \boldsymbol{z})\|^2 = \langle (\boldsymbol{x}, \boldsymbol{z}), (\boldsymbol{x}, \boldsymbol{z}) \rangle = \langle \boldsymbol{x}, \boldsymbol{x} \rangle + \langle \boldsymbol{z}, \boldsymbol{z} \rangle = \|\boldsymbol{x}\|^2 + \|\boldsymbol{z}\|^2. \qquad \square$$

► EXERCISE 12 (1-12*). *Let* $(\mathbb{R}^n)^*$ *denote the dual space of the vector space* $\mathbb{R}^n$. *If* $\boldsymbol{x} \in \mathbb{R}^n$, *define* $\varphi_{\boldsymbol{x}} \in (\mathbb{R}^n)^*$ *by* $\varphi_{\boldsymbol{x}}(\boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$. *Define* $\mathrm{M} : \mathbb{R}^n \to (\mathbb{R}^n)^*$ *by* $\mathrm{M}\,\boldsymbol{x} = \varphi_{\boldsymbol{x}}$. *Show that* M *is a 1-1 linear transformation and conclude that every* $\varphi \in (\mathbb{R}^n)^*$ *is* $\varphi_{\boldsymbol{x}}$ *for a unique* $\boldsymbol{x} \in \mathbb{R}^n$.

PROOF. We first show M is linear. Take any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$. Then

$$\mathrm{M}(a\boldsymbol{x} + b\boldsymbol{y}) = \varphi_{a\boldsymbol{x}+b\boldsymbol{y}} = a\varphi_{\boldsymbol{x}} + b\varphi_{\boldsymbol{y}} = a\,\mathrm{M}\,\boldsymbol{x} + b\,\mathrm{M}\,\boldsymbol{y},$$

where the second equality holds since for every $\boldsymbol{z} \in \mathbb{R}^n$,

$$\varphi_{a\boldsymbol{x}+b\boldsymbol{y}}(\boldsymbol{z}) = \langle a\boldsymbol{x} + b\boldsymbol{y}, \boldsymbol{z} \rangle = a\langle \boldsymbol{x}, \boldsymbol{z} \rangle + b\langle \boldsymbol{y}, \boldsymbol{z} \rangle = a\varphi_{\boldsymbol{x}}(\boldsymbol{z}) + b\varphi_{\boldsymbol{y}}(\boldsymbol{z}).$$

To see M is 1-1, we need only to show that $\mathcal{N}_{\mathrm{M}} = \{\boldsymbol{0}\}$, where $\mathcal{N}_{\mathrm{M}}$ is the null set of M. But this is clear and so M is 1-1. Since $\dim(\mathbb{R}^n)^* = \dim \mathbb{R}^n$, M is also onto. This proves the last claim. □

► EXERCISE 13 (1-13*). *If* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, *then* $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are called perpendicular (or orthogonal) if* $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$. *If* $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *are perpendicular, prove that* $\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2$.

PROOF. If $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, we have

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 = \langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y} \rangle = \|\boldsymbol{x}\|^2 + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \|\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2. \qquad \square$$

1.2 SUBSETS OF EUCLIDEAN SPACE

► EXERCISE 14 (1-14*). *Simple. Omitted.*

► EXERCISE 15 (1-15). *Prove that* $\{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{a}\| < r\}$ *is open.*

PROOF. For any $\boldsymbol{y} \in \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{a}\| < r\} =: \mathbb{B}(\boldsymbol{a}; r)$, let $\varepsilon = r - \|\boldsymbol{a}, \boldsymbol{y}\|$. We show that $\mathbb{B}(\boldsymbol{y}; \varepsilon) \subseteq \mathbb{B}(\boldsymbol{a}; r)$. Take any $\boldsymbol{z} \in \mathbb{B}(\boldsymbol{y}; \varepsilon)$. Then

$$\|\boldsymbol{a}, \boldsymbol{z}\| \leqslant \|\boldsymbol{a}, \boldsymbol{y}\| + \|\boldsymbol{y}, \boldsymbol{z}\| < \|\boldsymbol{a}, \boldsymbol{y}\| + \varepsilon = r. \qquad \square$$

► EXERCISE 16 (1-16). *Simple. Omitted.*

► EXERCISE 17 (1-17). *Omitted.*

► EXERCISE 18 (1-18). *If* $A \subset [0,1]$ *is the union of open intervals* (a_i, b_i) *such that each rational number in* $(0,1)$ *is contained in some* (a_i, b_i)*, show that* $\partial A = [0,1] \smallsetminus A$*.*

PROOF. Let $X := [0,1]$. Obviously, A is open since $A = \bigcup_i (a_i, b_i)$. Then $X \smallsetminus A$ is closed in X and so $\overline{X \smallsetminus A} = X \smallsetminus A$. Since $\partial A = \overline{A} \cap \overline{X \smallsetminus A} = \overline{A} \cap (X \smallsetminus A)$, it suffices to show that

$$X \smallsetminus A \subseteq \overline{A}. \tag{1.1}$$

But (1.1) holds if and only if $\overline{A} = X$. Now take any $x \in X$ and any open nhood U of x in X. Since $\mathbb{Q}$ is dense, there exists $y \in U$. Since there exists some i such that $y \in (a_i, b_i)$, we know that $U \cap (a_i, b_i) \neq \varnothing$, which means that $U \cap A \neq \varnothing$, which means that $x \in \overline{A}$. Hence, $X = \overline{A}$, i.e., A is dense in X. □

► EXERCISE 19 (1-19*). *If* A *is a closed set that contains every rational number* $r \in [0,1]$*, show that* $[0,1] \subset A$*.*

PROOF. Take any $r \in (0,1)$ and any open interval $r \in I \subset (0,1)$. Then there exists $q \in \mathbb{Q} \cap (0,1)$ such that $q \in I$. Since $q \in A$, we know that $r \in \overline{A} = A$. Since $0, 1 \in A$, the claim holds. □

► EXERCISE 20 (1-20). *Prove the converse of Corollary 1-7: A compact subset of* $\mathbb{R}^n$ *is closed and bounded.*

PROOF. To show A is closed, we prove that A^c is open. Assume that $\boldsymbol{x} \notin A$, and let $G_m = \{\boldsymbol{y} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{y}\| > 1/m\}$, $m = 1, 2, \ldots$. If $\boldsymbol{y} \in A$, then $\boldsymbol{x} \neq \boldsymbol{y}$; hence, $\|\boldsymbol{x} - \boldsymbol{y}\| > 1/m$ for some m; therefore $\boldsymbol{y} \in G_m$ (see Figure 1.1). Thus, $A \subseteq \bigcup_{m=1}^{\infty} G_m$, and by compactness we have a finite subcovering. Now observe that the G_m for an increasing sequence of sets: $G_1 \subseteq G_2 \subseteq \cdots$; therefore, a finite union of some of the G_m is equal to the set with the highest index. Thus, $K \subseteq G_s$ for some s, and it follows that $\mathbb{B}(\boldsymbol{x}; 1/s) \subseteq A^c$. Therefore, A^c is open.

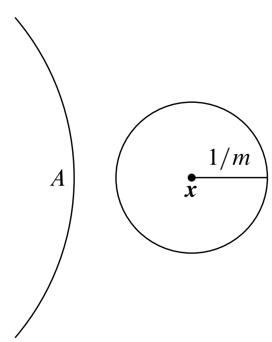

FIGURE 1.1. A compact set is closed

Let A be compact. We first show that A is bounded. Let

$$\mathcal{O} = \left\{(-i,i)^n : i \in \mathbb{N}\right\}$$

be an open cover of A. Then there is a finite subcover $\left\{(-i_1,i_1)^n, \ldots, (-i_m,i_m)^n\right\}$ of A. Let $i' = \max\{i_1,\ldots,i_m\}$. Hence, $A \subset \left(-i', i'\right)$; that is, A is bounded. □

► EXERCISE 21 (1-21*).

a. *If A is closed and $x \notin A$, prove that there is a number $d > 0$ such that $\|y - x\| \geqslant d$ for all $y \in A$.*

b. *If A is closed, B is compact, and $A \cap B = \varnothing$, prove that there is $d > 0$ such that $\|y - x\| \geqslant d$ for all $y \in A$ and $x \in B$.*

c. *Give a counterexample in $\mathbb{R}^2$ if A and B are closed but neither is compact.*

PROOF.

(a) A is closed implies that A^c is open. Since $x \in A^c$, there exists an open ball $\mathbb{B}(x;d)$ with $d > 0$ such that $x \in \mathbb{B}(x;d) \subset A^c$. Then $\|y - x\| \geqslant d$ for all $y \in A$.

(b) For every $x \in B$, there exists $d_x > 0$ such that $x \in \mathbb{B}(x; d_x/2) \subset A^c$ and $\|y - x\| \geqslant d_x$ for all $y \in A$. Then the family $\{\mathbb{B}(x; d_x/2) : x \in B\}$ is an open cover of B. Since B is compact, there is a finite set $\{x_1, \ldots, x_n\}$ such that $\left\{\mathbb{B}(x_1; d_{x_1}/2), \ldots, \mathbb{B}(x_n; d_{x_n}/2)\right\}$ covers B as well. Now let

$$d = \min\left\{d_{x_1}/2, \ldots, d_{x_n}/2\right\}\Big/2.$$

Then for any $x \in B$, there is an open ball $\mathbb{B}(x_i; x_i/2)$ containing x and $\|y - x_i\| \geqslant d_i$. Hence,

$$\|y - x\| \geqslant \|y - x_i\| - \|x_i - x\| \geqslant d_i - d_i/2 = d_i/2 \geqslant d.$$

(c) See Figure 1.2.

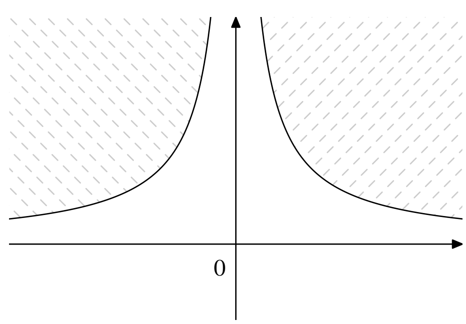

FIGURE 1.2.

□

► EXERCISE 22 (1-22*). *If U is open and $C \subset U$ is compact, show that there is a compact set D such that $C \subset D^{\circ}$ and $D \subset U$.*

PROOF. □

1.3 Functions and Continuity

► EXERCISE 23 (1-23). *If $f : A \to \mathbb{R}^m$ and $a \in A$, show that* $\lim_{x \to a} f(x) = b$ *if and only if* $\lim_{x \to a} f^i(x) = b^i$ *for $i = 1, \ldots, m$.*

PROOF. Let $f : A \to \mathbb{R}^m$ and $a \in A$.

If: Assume that $\lim_{x \to a} f^i(x) = b^i$ for $i = 1, \ldots, m$. Then for every $\varepsilon / \sqrt{m} > 0$, there is a number $\delta_i > 0$ such that $\left\| f^i(x) - b^i \right\| < \varepsilon / \sqrt{m}$ for all $x \in A$ which satisfy $0 < \|x - a\| < \delta_i$, for every $i = 1, \ldots, m$. Put

$$\delta = \min \{\delta_1, \ldots, \delta_m\}.$$

Then for all $x \in A$ satisfying $0 < \|x - a\| < \delta$,

$$\left\| f^i(x) - b^i \right\| < \frac{\varepsilon}{\sqrt{m}}, \quad i = 1, \ldots, m.$$

Therefore, for every $x \in A$ which satisfy $0 < \|x - a\| < \delta$,

$$\|f(x) - b\| = \sqrt{\sum_{i=1}^{m} \left(f^i(x) - b_i \right)^2} < \sqrt{\sum_{i=1}^{m} \left(\varepsilon^2 / m \right)} = \varepsilon;$$

that is, $\lim_{x \to a} f(x) = b$.

Only if: Now suppose that $\lim_{x \to a} f(x) = b$. Then for every number $\varepsilon > 0$ there is a number $\delta > 0$ such that $\|f(x) - b\| < \varepsilon$ for all $x \in A$ which satisfy $0 < \|x - a\| < \delta$. But then for every $i = 1, \ldots, m$,

$$\left\| f^i(x) - b^i \right\| \leqslant \|f(x) - b\| < \varepsilon,$$

i.e. $\lim_{x \to a} f^i(x) = b^i$. □

► EXERCISE 24 (1-24). *Prove that $f : A \to \mathbb{R}^m$ is continuous at a if and only if each f^i is.*

PROOF. By definition, f is continuous at a if and only if $\lim_{x \to a} f(x) = f(a)$; it follows from Exercise 23 that $\lim_{x \to a} f(x) = f(a)$ if and only if $\lim_{x \to a} f^i(x) = f^i(a)$ for every $i = 1, \ldots, m$; that is, if and only if f^i is continuous at a for each $i = 1, \ldots, m$. □

► EXERCISE 25 (1-25). *Prove that a linear transformation* $\mathrm{T} : \mathbb{R}^n \to \mathbb{R}^m$ *is continuous.*

PROOF. Take any $\boldsymbol{a} \in \mathbb{R}^n$. Then, by Exercise 10 (1-10), there exists $M > 0$ such that

$$\mathrm{T}\boldsymbol{x} - \mathrm{T}\boldsymbol{a} = \mathrm{T}(\boldsymbol{x} - \boldsymbol{a}) \leqslant M \|\boldsymbol{x} - \boldsymbol{a}\|.$$

Hence, for every $\varepsilon > 0$, let $\delta = \varepsilon/M$. Then $\mathrm{T}\boldsymbol{x} - \mathrm{T}\boldsymbol{a} < \varepsilon$ when $\boldsymbol{x} \in \mathbb{R}^n$ and $0 < \|\boldsymbol{x} - \boldsymbol{a}\| < \delta = \varepsilon/M$; that is, $\lim_{\boldsymbol{x}\to\boldsymbol{a}} \mathrm{T}\boldsymbol{x} = \mathrm{T}\boldsymbol{a}$, and so T is continuous. □

► EXERCISE 26 (1-26). *Let* $A = \left\{(x, y) \in \mathbb{R}^2 : x > 0 \text{ and } 0 < y < x^2\right\}$.

a. *Show that every straight line through* $(0, 0)$ *contains an interval around* $(0, 0)$ *which is in* $\mathbb{R}^2 \smallsetminus A$.

b. *Define* $f : \mathbb{R}^2 \to \mathbb{R}$ *by* $f(\boldsymbol{x}) = 0$ *if* $\boldsymbol{x} \notin A$ *and* $f(\boldsymbol{x}) = 1$ *if* $x \in A$. *For* $\boldsymbol{h} \in \mathbb{R}^2$ *define* $g_{\boldsymbol{h}} : \mathbb{R} \to \mathbb{R}$ *by* $g_{\boldsymbol{h}}(t) = f(t\boldsymbol{h})$. *Show that each* $g_{\boldsymbol{h}}$ *is continuous at* 0, *but* f *is not continuous at* $(0, 0)$.

PROOF.

(a) Let the line through $(0, 0)$ be $y = ax$. If $a \leqslant 0$, then the whole line is in $\mathbb{R}^2 \smallsetminus A$. If $a > 0$, then ax intersects x^2 at (a, a^2) and $(0, 0)$ and nowhere else; see Figure 1.3.

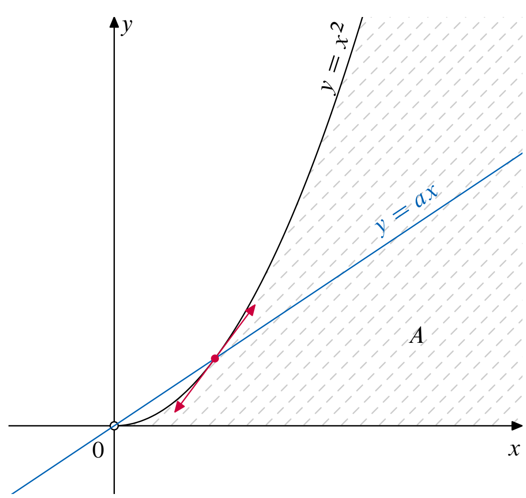

FIGURE 1.3.

(b) We first show that f is not continuous at $\mathbf{0}$. Clearly, $f(\mathbf{0}) = 0$ since $\mathbf{0} \notin A$. For every $\delta > 0$, there exists $\boldsymbol{x} \in A$ satisfying $0 < \|\boldsymbol{x}\| < \delta$, but $|f(\boldsymbol{x}) - f(\mathbf{0})| = 1$.

We next show $g_{\boldsymbol{h}}(t) = f(t\boldsymbol{h})$ is continuous at 0 for every $\boldsymbol{h} \in \mathbb{R}^2$. If $\boldsymbol{h} = \mathbf{0}$, then $g_{\mathbf{0}}(t) = f(\mathbf{0}) = 0$ and so is continuous. So we now assume that $\boldsymbol{h} \neq \mathbf{0}$. It is clear that

$$g_{\boldsymbol{h}}(0) = f(\mathbf{0}) = 0.$$

The result is now from (a) immediately. □

▶ EXERCISE 27 (1-27). *Prove that* $\left\{ \boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{a}\| < r \right\}$ *is open by considering the function* $f : \mathbb{R}^n \to \mathbb{R}$ *with* $f(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}\|$.

PROOF. We first show that f is continuous. Take a point $\boldsymbol{b} \in \mathbb{R}^n$. For any $\varepsilon > 0$, let $\delta = \varepsilon$. Then for every $\boldsymbol{x}$ satisfying $\|\boldsymbol{x} - \boldsymbol{b}\| < \delta$, we have

$$|f(\boldsymbol{x}) - f(\boldsymbol{b})| = \big| \|\boldsymbol{x} - \boldsymbol{a}\| - \|\boldsymbol{b} - \boldsymbol{a}\| \big| \leqslant \|\boldsymbol{x} - \boldsymbol{a}\| - \|\boldsymbol{b} - \boldsymbol{a}\| \leqslant \|\boldsymbol{x} - \boldsymbol{b}\| < \delta = \varepsilon.$$

Hence, $\left\{ \boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{a}\| < r \right\} = f^{-1}(-\infty, r)$ is open in $\mathbb{R}^n$. □

▶ EXERCISE 28 (1-28). *If* $A \subset \mathbb{R}^n$ *is not closed, show that there is a continuous function* $f : A \to \mathbb{R}$ *which is unbounded.*

PROOF. Take any $\boldsymbol{x} \in \partial A$. Let $f(\boldsymbol{y}) = 1/\|\boldsymbol{y} - \boldsymbol{x}\|$ for all $\boldsymbol{y} \in A$. □

▶ EXERCISE 29 (1-29). *Simple. Omitted.*

▶ EXERCISE 30 (1-30). *Let* $f : [a, b] \to \mathbb{R}$ *be an increasing function. If* $x_1, \ldots, x_n \in [a, b]$ *are distinct, show that* $\sum_{i=1}^{n} \mathfrak{o}\left(f, x_i\right) < f(b) - f(a)$.

PROOF. □

2

DIFFERENTIATION

2.1 Basic Definitions

► Exercise 31 (2-1*). *Prove that if $f: \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $\boldsymbol{a} \in \mathbb{R}^n$, then it is continuous at $\boldsymbol{a}$.*

Proof. Let f be differentiable at $\boldsymbol{a} \in \mathbb{R}^n$; then there exists a linear map $\lambda: \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{\boldsymbol{h} \to \boldsymbol{0}} \frac{\|f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - \lambda(\boldsymbol{h})\|}{\|\boldsymbol{h}\|} = 0,$$

or equivalently,

$$f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) = \lambda(\boldsymbol{h}) + r(\boldsymbol{h}), \tag{2.1}$$

where the remainder $r(\boldsymbol{h})$ satisfies

$$\lim_{\boldsymbol{h} \to \boldsymbol{0}} \|r(\boldsymbol{h})\| \big/ \|\boldsymbol{h}\| = 0. \tag{2.2}$$

Let $\boldsymbol{h} \to \boldsymbol{0}$ in (2.1). The error term $r(\boldsymbol{h}) \to \boldsymbol{0}$ by (2.2); the linear term $\lambda(\boldsymbol{h})$ aslo tends to $\boldsymbol{0}$ because if $\boldsymbol{h} = \sum_{i=1}^n h_i \boldsymbol{e}_i$, where $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ is the standard basis of $\mathbb{R}^n$, then by linearity we have $\lambda(\boldsymbol{h}) = \sum_{i=1}^n h_i \lambda(\boldsymbol{e}_i)$, and each term on the right tends to $\boldsymbol{0}$ as $\boldsymbol{h} \to \boldsymbol{0}$. Hence,

$$\lim_{\boldsymbol{h} \to \boldsymbol{0}} \left[f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) \right] = \boldsymbol{0};$$

that is, $\lim_{\boldsymbol{h} \to \boldsymbol{0}} f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a})$. Thus, f is continuous at $\boldsymbol{a}$. □

► Exercise 32 (2-2). *A function $f: \mathbb{R}^2 \to \mathbb{R}$ is independent of the second variable if for each $x \in \mathbb{R}$ we have $f(x, y_1) = f(x, y_2)$ for all $y_1, y_2 \in \mathbb{R}$. Show that f is independent of the second variable if and only if there is a function $g: \mathbb{R} \to \mathbb{R}$ such that $f(x, y) = g(x)$. What is $f'(a, b)$ in terms of g'?*

Proof. The first assertion is trivial: if f is independent of the second variable, we can let g be defined by $g(x) = f(x, 0)$. Conversely, if $f(x, y) = g(x)$, then $f(x, y_1) = g(x) = f(x, y_2)$.

If f is independent of the second variable, then

$$\begin{aligned}\lim_{(h,k)\to\mathbf{0}} \frac{\left|f(a+h,b+k)-f(a,b)-g'(a)h\right|}{\|(h,k)\|} &= \lim_{(h,k)\to\mathbf{0}} \frac{\left|g(a+h)-g(a)-g'(a)h\right|}{\|(h,k)\|} \\ &\leqslant \lim_{h\to 0} \frac{\left|g(a+h)-g(a)-g'(a)h\right|}{|h|} \\ &= 0;\end{aligned}$$

hence, $f'(a,b) = (g'(a), 0)$. □

► EXERCISE 33 (2-3). *Define when a function $f: \mathbb{R}^2 \to \mathbb{R}$ is independent of the first variable and find $f'(a,b)$ for such f. Which functions are independent of the first variable and also of the second variable?*

PROOF. We have $f'(a,b) = (0, g'(b))$ with a similar argument as in Exercise 32. If f is independent of the first and second variable, then for any (x_1, y_1), $(x_2, y_2) \in \mathbb{R}^2$, we have $f(x_1, y_1) = f(x_2, y_1) = f(x_2, y_2)$; that is, f is constant. □

► EXERCISE 34 (2-4). *Let g be a continuous real-valued function on the unit circle $\left\{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\| = 1\right\}$ such that $g(0,1) = g(1,0) = 0$ and $g(-\boldsymbol{x}) = -g(\boldsymbol{x})$. Define $f: \mathbb{R}^2 \to \mathbb{R}$ by*

$$f(\boldsymbol{x}) = \begin{cases} \|\boldsymbol{x}\| \cdot g\left(\boldsymbol{x}/\|\boldsymbol{x}\|\right) & \text{if } \boldsymbol{x} \neq \mathbf{0}, \\ 0 & \text{if } \boldsymbol{x} = \mathbf{0}. \end{cases}$$

a. *If $\boldsymbol{x} \in \mathbb{R}^2$ and $h: \mathbb{R} \to \mathbb{R}$ is defined by $h(t) = f(t\boldsymbol{x})$, show that h is differentiable.*

b. *Show that f is not differentiable at $(0,0)$ unless $g = 0$.*

PROOF. (a) If $\boldsymbol{x} = \mathbf{0}$ or $t = 0$, then $h(t) = f(\mathbf{0}) = 0$; if $\boldsymbol{x} \neq \mathbf{0}$ and $t > 0$,

$$h(t) = f(t\boldsymbol{x}) = t\,\|\boldsymbol{x}\| \cdot g\left(\frac{t\boldsymbol{x}}{t\,\|\boldsymbol{x}\|}\right) = \left[\|\boldsymbol{x}\| \cdot g\left(\boldsymbol{x}/\|\boldsymbol{x}\|\right)\right] \cdot t = f(\boldsymbol{x})t;$$

finally, if $\boldsymbol{x} \neq \mathbf{0}$ and $t < 0$,

$$\begin{aligned} h(t) = f(t\boldsymbol{x}) = -t\,\|\boldsymbol{x}\| \cdot g\left(\frac{t\boldsymbol{x}}{-t\,\|\boldsymbol{x}\|}\right) &= -t\,\|\boldsymbol{x}\| \cdot g\left(-\boldsymbol{x}/\|\boldsymbol{x}\|\right) \\ &= \left[\|\boldsymbol{x}\| \cdot g\left(\boldsymbol{x}/\|\boldsymbol{x}\|\right)\right] \cdot t \\ &= f(\boldsymbol{x})t. \end{aligned}$$

Therefore, $h(t) = f(\boldsymbol{x})t$ for every given $\boldsymbol{x} \in \mathbb{R}^2$, and so is differentiable: $\mathbb{D}h = h$.

(b) Since $g(1,0) = 0$ and $g(-\boldsymbol{x}) = -g(\boldsymbol{x})$, we have $g(-1,0) = g(-(1,0)) = -g(1,0) = 0$. If f is differentiable at $(0,0)$, there exists a matrix (a,b) such that $\mathbb{D}f(0,0)(h,k) = ah + bk$. First consider any sequence $(h,0) \to (0,0)$. Then

$$
\begin{aligned}
0 = \lim_{h \to 0} \frac{|f(h,0) - f(0,0) - ah|}{|h|} &= \lim_{h \to 0} \frac{\left| |h| \cdot g\left(h/|h|,0\right) - ah \right|}{|h|} \\
&= \lim_{h \to 0} \frac{\left| |h| \cdot g(\pm 1, 0) - ah \right|}{|h|} \\
&= |a|
\end{aligned}
$$

implies that $a = 0$. Next let us consider $(0,k) \to (0,0)$. Then

$$
0 = \lim_{k \to 0} \frac{|f(0,k) - f(0,0) - bk|}{|k|} = \lim_{k \to 0} \frac{\left| |k| \cdot g\left(0, k/|k|\right) - bk \right|}{|k|} = |b|
$$

forces that $b = 0$. Therefore, $f'(0,0) = (0,0)$ and $\mathbb{D}f(0,0)(x,y) = 0$. If $g(\boldsymbol{x}) \neq 0$, then

$$
\lim_{\boldsymbol{x} \to \boldsymbol{0}} \frac{|f(\boldsymbol{x}) - f(\boldsymbol{0}) - 0|}{\|\boldsymbol{x}\|} = \lim_{\boldsymbol{x} \to \boldsymbol{0}} \frac{\left| \|\boldsymbol{x}\| \cdot g\left(\boldsymbol{x}/\|\boldsymbol{x}\|\right) \right|}{\|\boldsymbol{x}\|} = \lim_{\boldsymbol{x} \to \boldsymbol{0}} \left| g\left(\boldsymbol{x}/\|\boldsymbol{x}\|\right) \right| \neq 0,
$$

and so f is not differentiable.

Of course, if $g(\boldsymbol{x}) = 0$, then $f(\boldsymbol{x}) = 0$ and is differentiable. □

► EXERCISE 35 (2-5). *Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by*

$$
f(x,y) = \begin{cases} x\,|y| \Big/ \sqrt{x^2 + y^2} & \text{if } (x,y) \neq \boldsymbol{0}, \\ 0 & \text{if } (x,y) = \boldsymbol{0}. \end{cases}
$$

Show that f is a function of the kind considered in Exercise 34, so that f is not differentiable at $(0,0)$.

PROOF. If $(x,y) \neq \boldsymbol{0}$, we can rewrite $f(x,y)$ as

$$
f(x,y) = \frac{x \cdot |y|}{\sqrt{x^2 + y^2}} = \frac{x \cdot |y|}{\|(x,y)\|} = \|(x,y)\| \cdot \left(\frac{x}{\|(x,y)\|} \cdot \frac{|y|}{\|(x,y)\|} \right). \tag{2.3}
$$

If we let $g\colon \left\{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\| = 1\right\} \to \mathbb{R}$ be defined as $g(x,y) = x \cdot |y|$, then (2.3) can be rewritten as

$$
f(x,y) = \|(x,y)\| \cdot g((x,y)/\|(x,y)\|).
$$

It is easy to see that

$$
g(0,1) = g(1,0) = 0, \quad \text{and} \quad g(-x,-y) = -x|-y| = -x|y| = -f(x,y);
$$

that is, g satisfies all of the properties listed in Exercise 34. Since $g(\boldsymbol{x}) \neq 0$ unless $x = 0$ or $y = 0$, we know that f is not differentiable at $\boldsymbol{0}$. A direct proof can be found in Berkovitz (2002, Section 1.11). □

► EXERCISE 36 (2-6). *Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = \sqrt{|xy|}$. Show that f is not differentiable at* $(0,0)$.

PROOF. It is clear that

$$\lim_{h\to 0}\frac{|f(h,0)|}{|h|} = 0 = \lim_{k\to 0}\frac{|f(0,k)|}{|k|};$$

hence, if f is differentiable at $(0,0)$, it must be that $\mathbb{D}f(0,0)(x,y) = 0$ since derivative is unique if it exists. However, if we let $h = k > 0$, and take a sequence $\{(h,h)\} \to (0,0)$, we have

$$\lim_{(h,h)\to(0,0)}\frac{|f(h,h)-f(0,0)-0|}{\|(h,h)\|} = \lim_{(h,h)\to(0,0)}\frac{\sqrt{h^2}}{\|(h,h)\|} = \frac{1}{\sqrt{2}} \neq 0.$$

Therefore, f is not differentiable. □

► EXERCISE 37 (2-7). *Let $f:\mathbb{R}^2\to\mathbb{R}$ be a function such that $|f(\boldsymbol{x})| \leqslant \|\boldsymbol{x}\|^2$. Show that f is differentiable at $\mathbf{0}$.*

PROOF. $|f(\mathbf{0})| \leqslant \|\mathbf{0}\|^2 = 0$ implies that $f(\mathbf{0}) = 0$. Since

$$\lim_{\boldsymbol{x}\to\mathbf{0}}\frac{|f(\boldsymbol{x})-f(\mathbf{0})|}{\|\boldsymbol{x}\|} = \lim_{\boldsymbol{x}\to\mathbf{0}}\frac{|f(\boldsymbol{x})|}{\|\boldsymbol{x}\|} \leqslant \lim_{\boldsymbol{x}\to\mathbf{0}}\|\boldsymbol{x}\| = 0,$$

$\mathbb{D}f(\mathbf{0})(x,y) = 0$. □

► EXERCISE 38 (2-8). *Let $f:\mathbb{R}\to\mathbb{R}^2$. Prove that f is differentiable at $a\in\mathbb{R}$ if and only if f^1 and f^2 are, and that in this case*

$$f'(a) = \begin{pmatrix}(f^1)'(a)\\(f^2)'(a)\end{pmatrix}.$$

PROOF. Suppose that f is differentiable at a with $f'(a) = \begin{pmatrix}c^1\\c^2\end{pmatrix}$. Then for $i = 1,2$,

$$0 \leqslant \lim_{h\to 0}\frac{\left|f^i(a+h)-f^i(a)-c^i\cdot h\right|}{|h|} \leqslant \lim_{h\to 0}\frac{\|f(a+h)-f(a)-\mathbb{D}f(a)(h)\|}{|h|} = 0$$

implies that f^i is differentiable at a with $(f^i)'(a) = c^i$.

Now suppose that both f^1 and f^2 are differentiable at a, then by Exercise 1,

$$0 \leqslant \frac{\|f(a+h)-f(a)-\mathbb{D}f(a)(h)\|}{|h|} \leqslant \sum_{i=1}^{2}\frac{\left|f^i(a+h)-f^i(a)-(f^i)'(a)\cdot h\right|}{|h|}$$

implies that f is differentiable at a with $f'(a) = \begin{pmatrix}(f^1)'(a)\\(f^2)'(a)\end{pmatrix}$. □

► EXERCISE 39 (2-9). *Two functions $f,g:\mathbb{R}\to\mathbb{R}$ are equal up to n-th order at a if*

$$\lim_{h\to 0} \frac{f(a+h)-g(a+h)}{h^n} = 0.$$

a. *Show that f is differentiable at a if and only if there is a function g of the form $g(x) = a_0 + a_1(x-a)$ such that f and g are equal up to first order at a.*

b. *If $f'(a), \dots, f^{(n)}(a)$ exist, show that f and the function g defined by*

$$g(x) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i$$

are equal up to n-th order at a.

PROOF. (a) If f is differentiable at a, then by definition,

$$\lim_{h\to 0} \frac{f(a+h) - \left[f(a) + f'(a)\cdot h\right]}{h} = 0,$$

so we can let $g(x) = f(a) + f'(a)\cdot(x-a)$.

On the other hand, if there exists a function $g(x) = a_0 + a_1(x-a)$ such that

$$\lim_{h\to 0} \frac{f(a+h)-g(a+h)}{h} = \lim_{h\to 0} \frac{f(a+h)-a_0-a_1 h}{h} = 0,$$

then $a_0 = f(a)$, and so f is differentiable at a with $f'(a) = a_1$.

(b) By Taylor's Theorem[1] we rewrite f as

$$f(x) = \sum_{i=0}^{n-1} \frac{f^{(i)}(a)}{i!}(x-a)^i + \frac{f^{(n)}(y)}{n!}(x-a)^n,$$

where y is between a and x. Thus,

$$\begin{aligned}
\lim_{x\to a} \frac{f(x)-g(x)}{(x-a)^n} &= \lim_{x\to a} \frac{\frac{f^{(n)}(y)}{n!}(x-a)^n - \frac{f^{(n)}(a)}{n!}(x-a)^n}{(x-a)^n} \\
&= \lim_{x\to a} \frac{f^{(n)}(y) - f^{(n)}(x)}{n!} \\
&= 0.
\end{aligned}$$

□

[1] (Rudin, 1976, Theorem 5.15) Suppose f is a real function on $[a,b]$, n is a positive integer, $f^{(n-1)}$ is continuous on $[a,b]$, $f^{(n)}$ exists for every $t \in (a,b)$. Let α, β be distinct points of $[a,b]$, and define

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t-\alpha)^k.$$

Then there exists a point x between α and β such that

$$f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta-\alpha)^n.$$

2.2 Basic Theorems

▶ Exercise 40 (2-10). *Use the theorems of this section to find f' for the following:*

a. $f(x,y,z) = x^y$.

b. $f(x,y,z) = (x^y, z)$.

c. $f(x,y) = \sin(x \sin y)$.

d. $f(x,y,z) = \sin\big(x \sin(y \sin z)\big)$.

e. $f(x,y,z) = x^{y^z}$.

f. $f(x,y,z) = x^{y+z}$.

g. $f(x,y,z) = (x+y)^z$.

h. $f(x,y) = \sin(xy)$.

i. $f(x,y) = \big[\sin(xy)\big]^{\cos 3}$.

j. $f(x,y) = \Big(\sin(xy), \sin\big(x \sin y\big), x^y\Big)$.

Solution. Compare this with Exercise 47.

(a) We have $f(x,y,z) = x^y = e^{\ln x^y} = e^{y \ln x} = \exp\circ(\pi^2 \cdot \ln \pi^1)(x,y,z)$. It follows from the Chain Rule that

$$\begin{aligned} f'(a,b,c) &= \exp'\Big[(\pi^2 \ln \pi^1)(a,b,c)\Big] \cdot \Big(\pi^2 \ln \pi^1\Big)'(a,b,c) \\ &= \exp(b \ln a) \cdot \Big[(\ln \pi^1)(\pi^2)' + \pi^2 (\ln \pi^1)'\Big](a,b,c) \\ &= a^b \cdot \Big[\big(0, \ln a, 0\big) + \big(b/a, 0, 0\big)\Big] \\ &= \Big(a^{b-1}b \quad a^b \ln a \quad 0\Big). \end{aligned}$$

(b) By (a) and Theorem 2-3(3), we have

$$f'(a,b,c) = \begin{pmatrix} a^{b-1}b & a^b \ln a & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(c) We have $f(x,y) = \sin\circ(\pi^1 \sin(\pi^2))$. Then, by the chain rule,

$$\begin{aligned} f'(a,b) &= \sin'\Big[(\pi^1 \sin(\pi^2))(a,b)\Big] \cdot \Big[\pi^1 \sin(\pi^2)\Big]'(a,b) \\ &= \cos(a \sin b) \cdot \Big[(\sin \pi^2)(\pi^1)' + \pi^1 (\sin \pi^2)'\Big](a,b) \\ &= \cos(a \sin b) \cdot \big[\sin b\,(1,0) + a\,(0, \cos b)\big] \\ &= \Big(\cos(a \sin b) \cdot \sin b \quad a \cdot \cos(a \sin b) \cdot \cos b\Big). \end{aligned}$$

(d) Let $g(y,z) = \sin(y \sin z)$. Then

$$f(x, y, z) = \sin\left(x \cdot g(y, z)\right) = \sin(\pi^1 \cdot g(\pi^2, \pi^3)).$$

Hence,

$$\begin{aligned} f'(a,b,c) &= \sin'\left(ag\,(b,c)\right) \cdot (\pi^1 g(\pi^2,\pi^3))'(a,b,c) \\ &= \cos\left(ag\,(b,c)\right) \cdot \left[g\,(b,c)\,(\pi^1)' + ag'(\pi^2,\pi^3)\right](a,b,c) \\ &= \cos\left(ag\,(b,c)\right) \cdot \left[\left(g\,(b,c)\,,0,0\right) + ag'(\pi^2,\pi^3)(a,b,c)\right]. \end{aligned}$$

It follows from (c) that

$$g'(\pi^2,\pi^3)(a,b,c) = \begin{pmatrix} 0 & \cos\left(b\sin c\right)\cdot\sin c & b\cdot\cos\left(b\sin c\right)\cdot\cos c \end{pmatrix}.$$

Therefore,

$$\begin{aligned} &f'(a,b,c) \\ &= \cos\left(a\sin\left(b\sin c\right)\right)\begin{pmatrix} \sin\left(b\sin c\right) & a\cos\left(b\sin c\right)\sin c & ab\cos\left(b\sin c\right)\cos c \end{pmatrix}. \end{aligned}$$

(e) Let $g(x, y) = x^y$. Then

$$f(x, y, z) = x^{g(y,z)} = g\left(x, g(y, z)\right) = g\left(\pi^1, g(\pi^2, \pi^3)\right).$$

Then

$$\mathbb{D}f(a,b,c) = \mathbb{D}g\left(a, g\,(b,c)\right) \circ \left[\mathbb{D}\pi^1, \mathbb{D}g(\pi^2,\pi^3)\right](a,b,c).$$

By (a),

$$\begin{aligned} \mathbb{D}g\left(a, g\,(b,c)\right)(x,y,z) &= \begin{pmatrix} a^{g(b,c)}g\,(b,c)\,/a & a^{g(b,c)}\ln a & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \frac{a^{b^c}b^c}{a}x + \left(a^{b^c}\ln a\right)y, \\ \mathbb{D}\pi^1(a,b,c)(x,y,z) &= x, \end{aligned}$$

and

$$\begin{aligned} \mathbb{D}g(\pi^2,\pi^3)(a,b,c)(x,y,z) &= \mathbb{D}g\,(b,c) \circ \left(\mathbb{D}\pi^2, \mathbb{D}\pi^3\right)(a,b,c)(x,y,z) \\ &= \frac{b^c c}{b}y + \left(b^c\ln b\right)z. \end{aligned}$$

Hence,

$$\mathbb{D}f(a,b,c)(x,y,z) = \frac{a^{b^c}b^c}{a}x + \left(a^{b^c}\ln a\right)\left[\frac{b^c c}{b}y + \left(b^c\ln b\right)z\right],$$

and

$$f'(a,b,c) = \left(a^{b^c} b^c / a \quad a^{b^c} b^c c \ln a / b \quad a^{b^c} b^c \ln a \ln b \right).$$

(f) Let $g(x,y) = x^y$. Then $f(x,y,z) = x^{y+z} = g\left(x, y+z\right) = g\left(\pi^1, \pi^2 + \pi^3\right)$. Hence,

$$\begin{aligned}
\mathbb{D} f(a,b,c)(x,y,z) &= \mathbb{D} g\,(a, b+c) \circ \left(\mathbb{D}\pi^1, \mathbb{D}\pi^2 + \mathbb{D}\pi^3 \right)(a,b,c)(x,y,z) \\
&= \mathbb{D} g\,(a, b+c) \circ \left(x, y+z \right) \\
&= \frac{a^{b+c}\,(b+c)}{a} x + \left(a^{b+c} \ln a \right)\left(y+z \right),
\end{aligned}$$

and

$$f'(a,b,c) = \left(\tfrac{a^{b+c}(b+c)}{a} \quad a^{b+c} \ln a \quad a^{b+c} \ln a \right).$$

(g) Let $g(x,y) = x^y$. Then

$$f(x,y,z) = (x+y)^z = g\left(x+y, z\right) = g\left(\pi^1 + \pi^2, \pi^3\right).$$

Hence,

$$\begin{aligned}
\mathbb{D} f(a,b,c)(x,y,z) &= \mathbb{D} g\,(a+b, c) \circ \left[\mathbb{D}\pi^1 + \mathbb{D}\pi^2, \mathbb{D}\pi^3 \right](a,b,c)(x,y,z) \\
&= \mathbb{D} g\,(a+b, c) \circ \left(x+y, z \right) \\
&= \frac{(a+b)^c\, c}{(a+b)} (x+y) + \left((a+b)^c \ln(a+b) \right) z,
\end{aligned}$$

and

$$f'(a,b,c) = \left(\tfrac{(a+b)^c c}{(a+b)} \quad \tfrac{(a+b)^c c}{(a+b)} \quad (a+b)^c \ln(a+b) \right).$$

(h) We have $f(x,y) = \sin(xy) = \sin \circ \left(\pi^1 \pi^2\right)$. Hence,

$$\begin{aligned}
f'(a,b) &= (\sin)'(ab) \cdot \left[b(\pi^1)'(a,b) + a(\pi^2)'(a,b) \right] \\
&= \cos(ab) \cdot \left[b\,(1,0) + a(0,1) \right] \\
&= \cos(ab) \cdot (b,a) \\
&= \left(b \cdot \cos(ab) \quad a \cdot \cos(ab) \right).
\end{aligned}$$

(i) Straightforward.

(j) By Theorem 2-3 (3), we have

$$\begin{aligned}
f'(a,b,c) &= \begin{pmatrix} \left[\sin(xy)\right]'(a,b,c) \\ \left[\sin\left(x \sin y\right)\right]'(a,b,c) \\ \left[x^y\right]'(a,b,c) \end{pmatrix} \\
&= \begin{pmatrix} b \cdot \cos(ab) & a \cdot \cos(ab) \\ \cos(a \sin b) \cdot \sin b & a \cdot \cos(a \sin b) \cdot \cos b \\ a^{b-1} b & a^b \ln a \end{pmatrix}.
\end{aligned}$$

□

► EXERCISE 41 (2-11). *Find f' for the following (where $g\colon \mathbb{R} \to \mathbb{R}$ is continuous):*

a. $f(x,y)=\int_a^{x+y} g$.

b. $f(x,y)=\int_a^{xy} g$.

c. $f(x,y,z)=\int_{xy}^{\sin\left(x\sin(y\sin z)\right)} g$.

SOLUTION. (a) Let $h(t)=\int_a^t g$. Then $f(x,y)=\left[h\circ(\pi_1+\pi_2)\right](x,y)$, and so

$$
\begin{aligned}
f'(a,b) &= h'(a+b)\cdot\left[(\pi^1+\pi^2)'(a,b)\right] \\
&= g(a+b)\cdot(1,1) \\
&= \begin{pmatrix} g(a+b) & g(a+b)\end{pmatrix}.
\end{aligned}
$$

(b) Let $h(t)=\int_a^t g$. Then $f(x,y)=\int_a^{xy} g=h(xy)=\left[h\circ\left(\pi^1\cdot\pi^2\right)\right](x,y)$. Hence,

$$
\begin{aligned}
f'(a,b) &= h'(ab)\cdot\left[b\cdot(\pi^1)'(a,b)+a\cdot(\pi^2)'(a,b)\right] \\
&= g(ab)\cdot(b,a) \\
&= \begin{pmatrix} b\cdot g(ab) & a\cdot g(ab)\end{pmatrix}.
\end{aligned}
$$

(c) We can rewrite $f(x,y,z)$ as

$$
f(x,y,z)=\int_{xy}^{a} g+\int_a^{\sin\left(x\sin(y\sin z)\right)} g=\int_a^{\sin\left(x\sin(y\sin z)\right)} g-\int_a^{xy} g.
$$

Let $\gamma(x,y,z)=\sin\left(x\sin(y\sin z)\right)$, $k(x,y,z)=\int_a^{\gamma(x,y,z)} g$, and $h(x,y,z)=\int_a^{xy} g$. Then $f(x,y,z)=k(x,y,z)-h(x,y,z)$, and so

$$
f'(a,b,c)=k'(a,b,c)-h'(a,b,c).
$$

It follows from Exercise 40 (d) that

$$
k'(a,b,c)=k'\left(\gamma(a,b,c)\right)\cdot\gamma'(a,b,c).
$$

The other parts are easy. □

► EXERCISE 42 (2-12). *A function $f:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^p$ is* bilinear *if for $\boldsymbol{x},\boldsymbol{x}_1,\boldsymbol{x}_2\in\mathbb{R}^n$, $\boldsymbol{y},\boldsymbol{y}_1,\boldsymbol{y}_2\in\mathbb{R}^m$, and $a\in\mathbb{R}$ we have*

$$
\begin{aligned}
f(a\boldsymbol{x},\boldsymbol{y}) &= af(\boldsymbol{x},\boldsymbol{y})=f(\boldsymbol{x},a\boldsymbol{y}), \\
f(\boldsymbol{x}_1+\boldsymbol{x}_2,\boldsymbol{y}) &= f(\boldsymbol{x}_1,\boldsymbol{y})+f(\boldsymbol{x}_2,\boldsymbol{y}), \\
f(\boldsymbol{x},\boldsymbol{y}_1+\boldsymbol{y}_2) &= f(\boldsymbol{x},\boldsymbol{y}_1)+f(\boldsymbol{x},\boldsymbol{y}_2).
\end{aligned}
$$

a. *Prove that if f is bilinear, then*

$$
\lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}}\frac{\|f(\boldsymbol{h},\boldsymbol{k})\|}{\|(\boldsymbol{h},\boldsymbol{k})\|}=\boldsymbol{0}.
$$

b. *Prove that* $\mathbb{D}f(\boldsymbol{a},\boldsymbol{b})(\boldsymbol{x},\boldsymbol{y}) = f(\boldsymbol{a},\boldsymbol{y}) + f(\boldsymbol{x},\boldsymbol{b})$.

c. *Show that the formula for* $\mathbb{D}p(\boldsymbol{a},\boldsymbol{b})$ *in Theorem 2-3 is a special case of (b).*

PROOF. (a) Let $\left(\boldsymbol{e}_1^n,\dots,\boldsymbol{e}_n^n\right)$ and $\left(\boldsymbol{e}_1^m,\dots,\boldsymbol{e}_m^m\right)$ be the stand bases for $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Then for any $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{y} \in \mathbb{R}^m$, we have

$$\boldsymbol{x} = \sum_{i=1}^{n} x^i \boldsymbol{e}_n^i, \quad \text{and} \quad \boldsymbol{y} = \sum_{j=1}^{m} y^j \boldsymbol{e}_m^j.$$

Therefore,

$$\begin{aligned}
f(\boldsymbol{x},\boldsymbol{y}) = f\left(\sum_{i=1}^{n} x^i \boldsymbol{e}_n^i, \sum_{j=1}^{m} y^j \boldsymbol{e}_m^j\right) &= \sum_{i=1}^{n} f\left(x^i \boldsymbol{e}_n^i, \sum_{j=1}^{m} y^j \boldsymbol{e}_m^j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} f\left(x^i \boldsymbol{e}_n^i, y^j \boldsymbol{e}_m^j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} x^i y^j f\left(\boldsymbol{e}_n^i, \boldsymbol{e}_m^j\right).
\end{aligned}$$

Then, by letting $M = \sum_{i,j} \left\| f\left(\boldsymbol{e}_n^i, \boldsymbol{e}_m^j\right)\right\|$, we have

$$\begin{aligned}
\|f(\boldsymbol{x},\boldsymbol{y})\| = \left\| \sum_{i,j} x^i y^j f\left(\boldsymbol{e}_n^i, \boldsymbol{e}_m^j\right)\right\| &\leqslant \sum_{i,j} \left| x^i y^j \right| \left\| f\left(\boldsymbol{e}_n^i, \boldsymbol{e}_m^j\right)\right\| \\
&\leqslant M\left[\max_i\left\{\left|x^i\right|\right\} \max_j\left\{\left|y^j\right|\right\}\right] \\
&\leqslant M\,\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|.
\end{aligned}$$

Hence,

$$\begin{aligned}
\lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}} \frac{\|f(\boldsymbol{h},\boldsymbol{k})\|}{\|(\boldsymbol{h},\boldsymbol{k})\|} &\leqslant \lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}} \frac{M\,\|\boldsymbol{h}\|\,\|\boldsymbol{k}\|}{\|(\boldsymbol{h},\boldsymbol{k})\|} \\
&= \lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}} \frac{M\,\|\boldsymbol{h}\|\,\|\boldsymbol{k}\|}{\sqrt{\sum_{i,j}\left[\left(h^i\right)^2 + \left(k^j\right)^2\right]}} \\
&= \lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}} \frac{M\,\|\boldsymbol{h}\|\,\|\boldsymbol{k}\|}{\sqrt{\|\boldsymbol{h}\|^2 + \|\boldsymbol{k}\|^2}}.
\end{aligned}$$

Now

$$\|\boldsymbol{h}\|\,\|\boldsymbol{k}\| \leqslant \begin{cases} \|\boldsymbol{h}\|^2 & \text{if } \|\boldsymbol{k}\| \leqslant \|\boldsymbol{h}\| \\ \|\boldsymbol{k}\|^2 & \text{if } \|\boldsymbol{h}\| \leqslant \|\boldsymbol{k}\|. \end{cases}$$

Hence $\|\boldsymbol{h}\|\,\|\boldsymbol{k}\| \leqslant \|\boldsymbol{h}\|^2 + \|\boldsymbol{k}\|^2$, and so

$$\lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}} \frac{M\,\|\boldsymbol{h}\|\,\|\boldsymbol{k}\|}{\sqrt{\|\boldsymbol{h}\|^2+\|\boldsymbol{k}\|^2}} \leqslant \lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}} M\sqrt{\|\boldsymbol{h}\|^2+\|\boldsymbol{k}\|^2} = 0.$$

(b) We have

$$\begin{aligned}
&\lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}} \frac{\|f(\boldsymbol{a}+\boldsymbol{h},\boldsymbol{b}+\boldsymbol{k})-f(\boldsymbol{a},\boldsymbol{b})-f(\boldsymbol{a},\boldsymbol{k})-f(\boldsymbol{h},\boldsymbol{b})\|}{\|(\boldsymbol{h},\boldsymbol{k})\|}\\
&=\lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}} \frac{\|f(\boldsymbol{a},\boldsymbol{b})+f(\boldsymbol{a},\boldsymbol{k})+f(\boldsymbol{h},\boldsymbol{b})+f(\boldsymbol{h},\boldsymbol{k})-f(\boldsymbol{a},\boldsymbol{b})-f(\boldsymbol{a},\boldsymbol{k})-f(\boldsymbol{h},\boldsymbol{b})\|}{\|(\boldsymbol{h},\boldsymbol{k})\|}\\
&=\lim_{(\boldsymbol{h},\boldsymbol{k})\to\boldsymbol{0}} \frac{\|f(\boldsymbol{h},\boldsymbol{k})\|}{\|(\boldsymbol{h},\boldsymbol{k})\|}\\
&=0
\end{aligned}$$

by (a); hence, $\mathbb{D}f(\boldsymbol{a},\boldsymbol{b})(\boldsymbol{x},\boldsymbol{y}) = f(\boldsymbol{a},\boldsymbol{y}) + f(\boldsymbol{x},\boldsymbol{b})$.

(c) It is easy to check that $p\colon \mathbb{R}^2 \to \mathbb{R}$ defined by $p(x,y) = xy$ is bilinear. Hence, by (b), we have

$$\mathbb{D}p(a,b)(x,y) = p\left(a,y\right) + p\left(x,b\right) = ay + xb. \qquad \square$$

► EXERCISE 43 (2-13). *Define* $\mathrm{IP}\colon \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ *by* $\mathrm{IP}(\boldsymbol{x},\boldsymbol{y}) = \langle \boldsymbol{x},\boldsymbol{y}\rangle$.

a. *Find* $\mathbb{D}\,(\mathrm{IP})\,(\boldsymbol{a},\boldsymbol{b})$ *and* $(\mathrm{IP})'\,(\boldsymbol{a},\boldsymbol{b})$.

b. *If* $f,g\colon \mathbb{R}\to\mathbb{R}^n$ *are differentiable and* $h\colon\mathbb{R}\to\mathbb{R}$ *is defined by* $h(t) = \langle f(t), g(t)\rangle$, *show that*

$$h'(a) = \left\langle f'(a)^{\mathrm{T}}, g(a)\right\rangle + \left\langle f(a), g'(a)^{\mathrm{T}}\right\rangle.$$

c. *If* $f\colon\mathbb{R}\to\mathbb{R}^n$ *is differentiable and* $\|f(t)\| = 1$ *for all* t, *show that* $\left\langle f'(t)^{\mathrm{T}}, f(t)\right\rangle = 0$.

d. *Exhibit a differentiable function* $f\colon\mathbb{R}\to\mathbb{R}$ *such that the function* $|f|$ *defined by* $|f|\,(t) = |f(t)|$ *is not differentiable.*

PROOF. (a) It is evident that IP is bilinear; hence, by Exercise 42 (b), we have

$$\begin{aligned}
\mathbb{D}\,(\mathrm{IP})\,(\boldsymbol{a},\boldsymbol{b})(\boldsymbol{x},\boldsymbol{y}) &= \mathrm{IP}\,(\boldsymbol{a},\boldsymbol{y}) + \mathrm{IP}\,(\boldsymbol{x},\boldsymbol{b})\\
&= \langle \boldsymbol{a},\boldsymbol{y}\rangle + \langle \boldsymbol{x},\boldsymbol{b}\rangle\\
&= \langle \boldsymbol{b},\boldsymbol{x}\rangle + \langle \boldsymbol{a},\boldsymbol{y}\rangle,
\end{aligned}$$

and so $(\mathrm{IP})'\,(\boldsymbol{a},\boldsymbol{b}) = (\boldsymbol{b},\boldsymbol{a})$.

(b) Since $h(t) = \mathrm{IP}\circ\left(f,g\right)(t)$, by the chain rule, we have

$$\begin{aligned}
\mathbb{D}h(a)\,(x) &= \mathbb{D}\,(\mathrm{IP})\left(f(a),g(a)\right)\circ\left(\mathbb{D}f(a)\,(x)\,,\mathbb{D}g(a)\,(x)\right)\\
&= \langle g(a), \mathbb{D}f(a)\,(x)\rangle + \langle f(a), \mathbb{D}g(a)\,(x)\rangle\\
&= \left\langle g(a), f'(a)\right\rangle x + \left\langle f(a), g'(a)\right\rangle x.
\end{aligned}$$

(c) Let $h(t) = \langle f(t), f(t)\rangle$ with $\|f(t)\| = 1$ for all $t\in\mathbb{R}$. Then

$$h(t) = \|f(t)\|^2 = 1$$

is constant, and so $h'(a) = 0$; that is,

$$0 = \left\langle f'(a)^{\mathrm{T}}, f(a) \right\rangle + \left\langle f(a), f'(a)^{\mathrm{T}} \right\rangle = 2\left\langle f'(a)^{\mathrm{T}}, f(a) \right\rangle,$$

and so $\left\langle f'(a)^{\mathrm{T}}, f(a) \right\rangle = 0$.

(d) Let $f(t) = t$. Then f is linear and so is differentiable: $\mathbb{D}f = t$. However,

$$\lim_{t \to 0^+} \frac{|t|}{t} = 1, \quad \lim_{t \to 0^-} \frac{|t|}{t} = -1;$$

that is, $|f|$ is not differentiable at 0. □

► Exercise 44 (2-14). *Let $\mathbb{E}_i$, $i = 1, \dots, k$ be Euclidean spaces of various dimensions. A function $f \colon \mathbb{E}_1 \times \cdots \times \mathbb{E}_k \to \mathbb{R}^p$ is called* multilinear *if for each choice of $\boldsymbol{x}_j \in \mathbb{E}_j$, $j \neq i$ the function $g \colon \mathbb{E}_i \to \mathbb{R}^p$ defined by $g(\boldsymbol{x}) = f(\boldsymbol{x}_1, \dots, \boldsymbol{x}_{i-1}, \boldsymbol{x}, \boldsymbol{x}_{i+1}, \dots, \boldsymbol{x}_k)$ is a linear transformation.*

a. *If f is multilinear and $i \neq j$, show that for $\boldsymbol{h} = (\boldsymbol{h}_1, \dots, \boldsymbol{h}_k)$, with $\boldsymbol{h}_\ell \in \mathbb{E}_\ell$, we have*

$$\lim_{\boldsymbol{h} \to \boldsymbol{0}} \frac{\left\| f\left(\boldsymbol{a}_1, \dots, \boldsymbol{h}_i, \dots, \boldsymbol{h}_j, \dots, \boldsymbol{a}_k\right) \right\|}{\|\boldsymbol{h}\|} = 0.$$

b. *Prove that*

$$\mathbb{D}f\left(\boldsymbol{a}_1, \dots, \boldsymbol{a}_k\right)\left(\boldsymbol{x}_1, \dots, \boldsymbol{x}_k\right) = \sum_{i=1}^{k} f\left(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{i-1}, \boldsymbol{x}_i, \boldsymbol{a}_{i+1}, \dots, \boldsymbol{a}_k\right).$$

Proof.

(a) To light notation, define

$$\boldsymbol{a}_{-i-j} := \left(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{i-1}, \boldsymbol{a}_{i+1}, \dots, \boldsymbol{a}_{j-1}, \boldsymbol{a}_{j+1}, \dots, \boldsymbol{a}_k\right).$$

Let $g \colon \mathbb{E}_i \times \mathbb{E}_j \to \mathbb{R}^p$ be defined as $g\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = f\left(\boldsymbol{a}_{-i-j}, \boldsymbol{x}_i, \boldsymbol{x}_j\right)$. Then g is bilinear and so

$$\lim_{\boldsymbol{h} \to \boldsymbol{0}} \frac{\left\| g\left(\boldsymbol{a}_{-i-j}, \boldsymbol{h}_i, \boldsymbol{h}_j\right) \right\|}{\|\boldsymbol{h}\|} \leqslant \lim_{\boldsymbol{h} \to \boldsymbol{0}} \frac{\left\| g\left(\boldsymbol{a}_{-i-j}, \boldsymbol{h}_i, \boldsymbol{h}_j\right) \right\|}{\left\| \left(\boldsymbol{h}_i, \boldsymbol{h}_j\right) \right\|} = 0$$

by Exercise 42 (a).

(b) It follows from Exercise 42 (b) immediately. □

► Exercise 45 (2-15). *Regard an $n \times n$ matrix as a point in the n-fold product $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$ by considering each row as a member of $\mathbb{R}^n$.*

a. *Prove that* $\det \colon \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ *is differentiable and*

$$\mathbb{D}\left(\det\right)\left(\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n\right)\left(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\right)=\sum_{i=1}^{n}\det\begin{pmatrix}\boldsymbol{a}_1\\ \vdots\\ \boldsymbol{x}_i\\ \vdots\\ \boldsymbol{a}_n\end{pmatrix}.$$

b. *If $a_{ij}\colon\mathbb{R}\to\mathbb{R}$ are differentiable and $f(t)=\det\left(a_{ij}(t)\right)$, show that*

$$f'(t)=\sum_{j=1}^{n}\det\begin{pmatrix}a_{11}(t)&\cdots&a_{1n}(t)\\ \vdots&\ddots&\vdots\\ a'_{j1}(t)&\cdots&a'_{jn}(t)\\ \vdots&\ddots&\vdots\\ a_{n1}(t)&\cdots&a_{nn}(t)\end{pmatrix}.$$

c. *If $\det\left(a_{ij}(t)\right)\neq 0$ for all t and $b_1,\ldots,b_n\colon\mathbb{R}\to\mathbb{R}$ are differentiable, let $s_1,\ldots,s_n\colon\mathbb{R}\to\mathbb{R}$ be the functions such that $s_1(t),\ldots,s_n(t)$ are the solutions of the equations*

$$\sum_{j=1}^{n}a_{ji}(t)s_j(t)=b_i(t),\quad i=1,\ldots,n.$$

Show that s_i is differentiable and find $s_i'(t)$.

Proof.

(a) It is easy to see that $\det\colon\mathbb{R}^n\times\cdots\times\mathbb{R}^n\to\mathbb{R}$ is multilinear; hence, the conclusion follows from Exercise 44.

(b) By (a) and the chain rule,

$$\begin{aligned}f'(t)&=\left(\det\right)'\left(a_{ij}(t)\right)\circ\left[\boldsymbol{a}_1'(t),\ldots,\boldsymbol{a}_n'(t)\right]\\ &=\sum_{j=1}^{n}\det\begin{pmatrix}a_{11}(t)&\cdots&a_{1n}(t)\\ \vdots&\ddots&\vdots\\ a'_{j1}(t)&\cdots&a'_{jn}(t)\\ \vdots&\ddots&\vdots\\ a_{n1}(t)&\cdots&a_{nn}(t)\end{pmatrix}.\end{aligned}$$

(c) Let

$$\mathbf{A}=\begin{pmatrix}a_{11}(t)&\cdots&a_{n1}(t)\\ \vdots&\ddots&\vdots\\ a_{1n}(t)&\cdots&a_{nn}(t)\end{pmatrix},\quad\mathbf{s}=\begin{pmatrix}s_1(t)\\ \vdots\\ s_n(t)\end{pmatrix},\quad\text{and}\quad\mathbf{b}=\begin{pmatrix}b_1(t)\\ \vdots\\ b_n(t)\end{pmatrix}.$$

Then

$$\mathbf{A}\mathbf{s}=\mathbf{b},$$

and so

$$s_i(t) = \frac{\det(\mathbf{B}_i)}{\det(\mathbf{A})},$$

where $\mathbf{B}_i$ is obtained from $\mathbf{A}$ by replacing the i-th column with the $\mathbf{b}$. It follows from (b) that $s_i(t)$ is differentiable. Define $f(t) = \det(\mathbf{A})$ and $g_i(t) = \det(\mathbf{B}_i)$. Then

$$f'(t) = \sum_{j=1}^{n} \det \begin{pmatrix} a_{11}(t) & \cdots & a_{n1}(t) \\ \vdots & \ddots & \vdots \\ a'_{1j}(t) & \cdots & a'_{nj}(t) \\ \vdots & \ddots & \vdots \\ a_{1n}(t) & \cdots & a_{nn}(t) \end{pmatrix},$$

and

$$g'_i(t) = \sum_{j=1}^{n} \begin{pmatrix} a_{11}(t) & \cdots & a_{i-1,1}(t) & b_1(t) & a_{i+1,1}(t) & \cdots & a_{n1}(t) \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a'_{1j}(t) & \cdots & a'_{i-1,j}(t) & b'_j(t) & a'_{i+1,j}(t) & \cdots & a'_{nj}(t) \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n}(t) & \cdots & a_{i-1,n}(t) & b_n(t) & a_{i+1,n}(t) & \cdots & a_{nn}(t) \end{pmatrix}.$$

Therefore,

$$s'_i(t) = \frac{f'(t)g'_i(t) - f(t)g'_i(t)}{f^2(t)}. \qquad \square$$

► EXERCISE 46 (2-16). *Suppose $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is differentiable and has a differentiable inverse $f^{-1}\colon \mathbb{R}^n \to \mathbb{R}^n$. Show that $\left(f^{-1}\right)'(a) = \left[f'\left(f^{-1}(a)\right)\right]^{-1}$.*

PROOF. We have $f \circ f^{-1}(\boldsymbol{x}) = \boldsymbol{x}$. On the one hand $\mathbb{D}\left(f \circ f^{-1}\right)(\boldsymbol{a})(\boldsymbol{x}) = \boldsymbol{x}$ since $f \circ f^{-1}$ is linear; on the other hand,

$$\mathbb{D}\left(f \circ f^{-1}\right)(\boldsymbol{a})(\boldsymbol{x}) = \left[\mathbb{D}f\left(f^{-1}(\boldsymbol{a})\right) \circ \mathbb{D}f^{-1}(\boldsymbol{a})\right](\boldsymbol{x}).$$

Therefore, $\mathbb{D}f^{-1}(\boldsymbol{a}) = \left[\mathbb{D}f\left(f^{-1}(\boldsymbol{a})\right)\right]^{-1}$. $\square$

2.3 PARTIAL DERIVATIVES

► EXERCISE 47 (2-17). *Find the partial derivatives of the following functions:*

a. $f(x, y, z) = x^y$.

b. $f(x, y, z) = z$.

c. $f(x, y) = \sin(x \sin y)$.

d. $f(x, y, z) = \sin\left(x \sin(y \sin z)\right)$.

e. $f(x, y, z) = x^{y^z}$.

f. $f(x, y, z) = x^{y+z}$.

g. $f(x, y, z) = (x + y)^z$.

h. $f(x, y) = \sin(xy)$.

i. $f(x, y) = \left[\sin(xy)\right]^{\cos 3}$.

SOLUTION. Compare this with Exercise 40.

(a) $\mathbb{D}_1 f(x, y, z) = yx^{y-1}$, $\mathbb{D}_2 f(x, y, z) = x^y \ln x$, and $\mathbb{D}_3 f(x, y, z) = 0$.

(b) $\mathbb{D}_1 f(x, y, z) = \mathbb{D}_2 f(x, y, z) = 0$, and $\mathbb{D}_3 f(x, y, z) = 1$.

(c) $\mathbb{D}_1 f(x, y) = (\sin y) \cos(x \sin y)$, and $\mathbb{D}_2 f(x, y) = x \cos y \cos(x \sin y)$.

(d) $\mathbb{D}_1 f(x, y, z) = \sin(y \sin z) \cos(x \sin(y \sin z))$,
$\mathbb{D}_2 f(x, y, z) = \cos(x \sin(y \sin z))\, x \cos(y \sin z) \sin z$, and
$\mathbb{D}_3 f(x, y, z) = \cos(x \sin(y \sin z))\, x \cos(y \sin z) y \cos z$.

(e) $\mathbb{D}_1 f(x, y, z) = y^z x^{y^z-1}$, $\mathbb{D}_2 f(x, y, z) = x^{y^z} z y^{z-1} \ln x$, and $\mathbb{D}_3 f(x, y, z) = y^z \ln y \left(x^{y^z} \ln x\right)$.

(f) $\mathbb{D}_1 f(x, y, z) = (y + z)\, x^{y+z-1}$, and $\mathbb{D}_2 f(x, y, z) \mathbb{D}_3 f(x, y, z) = x^{y+z} \ln x$.

(g) $\mathbb{D}_1 f(x, y, z) = \mathbb{D}_2 f(x, y, z) = z(x + y)^{z-1}$, and
$\mathbb{D}_3 f(x, y, z) = (x + y)^z \ln(x + y)$.

(h) $\mathbb{D}_1 f(x, y) = y \cos(xy)$, and $\mathbb{D}_2 f(x, y) = x \cos(xy)$.

(i) $\mathbb{D}_1 f(x, y) = \cos 3 \left[\sin(xy)\right]^{\cos 3-1} y \cos(xy)$, and
$\mathbb{D}_2 f(x, y) = \cos 3 \left[\sin(xy)\right]^{\cos 3-1} x \cos(xy)$. □

► EXERCISE 48 (2-18). *Find the partial derivatives of the following functions (where $g\colon \mathbb{R} \to \mathbb{R}$ is continuous):*

a. $f(x, y) = \int_a^{x+y} g$.

b. $f(x, y) = \int_y^x g$.

c. $f(x, y) = \int_a^{xy} g$.

d. $f(x, y) = \int_a^{\left(\int_b^y g\right)} g$.

SOLUTION.

(a) $\mathbb{D}_1 f(x, y) = \mathbb{D}_2 f(x, y) = g(x + y)$.

(b) $\mathbb{D}_1 f(x, y) = g(x)$, and $\mathbb{D}_2 f(x, y) = -g(y)$.

(c) $\mathbb{D}_1 f(x, y) = yg(xy)$, and $\mathbb{D}_2 f(x, y) = xg(xy)$.

(d) $\mathbb{D}_1 f(x,y) = 0$, and $\mathbb{D}_2 f(x,y) = g(y) \cdot g\left(\int_b^y g\right)$. □

▶ EXERCISE 49 (2-19). *If*

$$f(x,y) = x^{x^{x^{x^y}}} + (\ln x)\left(\arctan\left(\arctan\left(\arctan\left(\sin(\cos xy) - \ln(x+y)\right)\right)\right)\right)$$

find $\mathbb{D}_2 f(1,y)$.

SOLUTION. Putting $x = 1$ into $f(x,y)$, we get $f(1,y) = 1$. Then $\mathbb{D}_2 f(1,y) = 0$. □

▶ EXERCISE 50 (2-20). *Find the partial derivatives of f in terms of the derivatives of g and h if*

a. $f(x,y) = g(x)h(y)$.

b. $f(x,y) = g(x)^{h(y)}$.

c. $f(x,y) = g(x)$.

d. $f(x,y) = g(y)$.

e. $f(x,y) = g(x+y)$.

SOLUTION.

(a) $\mathbb{D}_1 f(x,y) = g'(x)\, h(y)$, and $\mathbb{D}_2 f(x,y) = g(x)h'(y)$.

(b) $\mathbb{D}_1 f(x,y) = h(y)\, g(x)^{h(y)-1} g'(x)$, and $\mathbb{D}_2 f(x,y) = h'(y)\, g(x)^{h(y)} \ln g(x)$.

(c) $\mathbb{D}_1 f(x,y) = g'(x)$, and $\mathbb{D}_2 f(x,y) = 0$.

(d) $\mathbb{D}_1 f(x,y) = 0$, and $\mathbb{D}_2 f(x,y) = g'(y)$.

(e) $\mathbb{D}_1 f(x,y) = \mathbb{D}_2 f(x,y) = g'(x+y)$. □

▶ EXERCISE 51 (2-21*). *Let $g_1, g_2 \colon \mathbb{R}^2 \to \mathbb{R}$ be continuous. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by*

$$f(x,y) = \int_0^x g_1(t,0)\,\mathrm{d}t + \int_0^y g_2(x,t)\,\mathrm{d}t.$$

a. *Show that* $\mathbb{D}_2 f(x,y) = g_2(x,y)$.

b. *How should f be defined so that* $\mathbb{D}_1 f(x,y) = g_1(x,y)$?

c. *Find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\mathbb{D}_1 f(x,y) = x$ and $\mathbb{D}_2 f(x,y) = y$. Find one such that $\mathbb{D}_1 f(x,y) = y$ and $\mathbb{D}_2 f(x,y) = x$.*

PROOF.

(a) $\mathbb{D}_2 f(x, y) = 0 + g_2(x, y) = g_2(x, y)$.

(b) We should let

$$f(x, y) = \int_0^x g_1\left(t, y\right) \mathrm{d}t + \int_0^y g_2\left(a, t\right) \mathrm{d}t,$$

where $t \in \mathbb{R}$ is a constant.

(c) Let

- $f(x, y) = \left(x^2 + y^2\right)/2$.
- $f(x, y) = xy$. □

► EXERCISE 52 (2-22*). *If $f\colon \mathbb{R}^2 \to \mathbb{R}$ and $\mathbb{D}_2 f = 0$, show that f is independent of the second variable. If $\mathbb{D}_1 f = \mathbb{D}_2 f = 0$, show that f is constant.*

PROOF. Fix any $x \in \mathbb{R}$. By the *mean-value theorem*, for any $y_1, y_2 \in \mathbb{R}$, there exists a point $y^* \in \left(y_1, y_2\right)$ such that

$$f\left(x, y_2\right) - f\left(x, y_1\right) = \mathbb{D}_2 f\left(x, y^*\right)\left(y_2 - y_1\right) = 0.$$

Hence, $f\left(x, y_1\right) = f\left(x, y_2\right)$; that is, f is independent of y.

Similarly, if $\mathbb{D}_1 f = 0$, then f is independent of x. The second claim is then proved immediately. □

► EXERCISE 53 (2-23*). *Let $A = \left\{(x, y) \in \mathbb{R}^2 : x < 0, \text{ or } x \geqslant 0 \text{ and } y \neq 0\right\}$.*

a. *If $f\colon A \to \mathbb{R}$ and $\mathbb{D}_1 f = \mathbb{D}_2 f = 0$, show that f is constant.*

b. *Find a function $f\colon A \to \mathbb{R}$ such that $\mathbb{D}_2 f = 0$ but f is not independent of the second variable.*

PROOF.

(a) As in Figure 2.1, for any $(a, b), (c, d) \in \mathbb{R}^2$, we have

$$f(a, b) = f\left(-1, b\right) = f\left(-1, d\right) = f\left(c, d\right).$$

(b) For example, we can let

$$f(x, y) = \begin{cases} 0 & \text{if } x < 0 \text{ or } y < 0 \\ x & \text{otherwise.} \end{cases}$$

□

► EXERCISE 54 (2-24). *Define $f\colon \mathbb{R}^2 \to \mathbb{R}$ by*

$$f(x, y) = \begin{cases} xy\frac{x^2 - y^2}{x^2 + y^2} & (x, y) \neq \mathbf{0}, \\ 0 & (x, y) = \mathbf{0}. \end{cases}$$

a. *Show that $\mathbb{D}_2 f\left(x, 0\right) = x$ for all x and $\mathbb{D}_1 f\left(0, y\right) = -y$ for all y.*

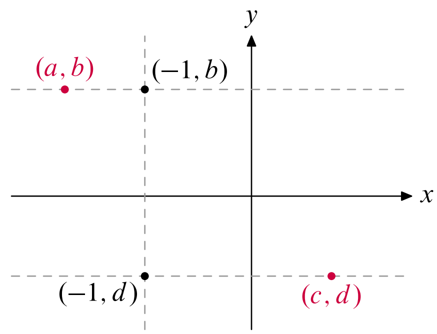

FIGURE 2.1. f is constant

b. *Show that* $\mathbb{D}_{1,2} f(0,0) \neq \mathbb{D}_{2,1} f(0,0)$.

PROOF.

(a) We have

$$\mathbb{D}_2 f(x,y) = \begin{cases} \dfrac{x\left(x^4-y^4-4x^2y^2\right)}{\left(x^2+y^2\right)^2} & (x,y) \neq \mathbf{0}, \\ 0 & (x,y) = \mathbf{0}, \end{cases}$$

and

$$\mathbb{D}_1 f(x,y) = \begin{cases} \dfrac{-y\left(y^4-x^4-4x^2y^2\right)}{\left(x^2+y^2\right)^2} & (x,y) \neq \mathbf{0}, \\ 0 & (x,y) = \mathbf{0}. \end{cases}$$

Hence, $\mathbb{D}_2 f\,(x,0) = x$ and $\mathbb{D}_1 f\,\left(0,y\right) = -y$.

(b) By (a), we have $\mathbb{D}_{1,2} f(0,0) = \mathbb{D}_2\left(\mathbb{D}_1 f\,\left(0,y\right)\right)(0) = -1$; but $\mathbb{D}_{2,1} f(0,0) = \mathbb{D}_1\left(\mathbb{D}_2\,(x,0)\right)(0) = 1$. □

► EXERCISE 55 (2-25*). *Define* $f : \mathbb{R} \to \mathbb{R}$ *by*

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

Show that f *is a* C^∞ *function, and* $f^{(i)}\,(0) = 0$ *for all* i.

PROOF. Figure 2.2 depicts $f\,(x)$. We first show that $f \in C^\infty$.

Let $p_n\left(y\right)$ be a polynomial with degree n with respect to y. For $x \neq 0$ and $k \in \mathbb{N}$, we show that $f^{(k)}\,(x) = p_{3k}\left(x^{-1}\right) e^{-x^{-2}}$. We do this by induction.

Step 1 Clearly, $f'\,(x) = 2x^{-3}e^{-x^{-2}}$.

Step 2 Suppose that $f^{(k)}\,(x) = p_{3k}\left(x^{-1}\right) e^{-x^{-2}}$.

Step 3 Then by the chain rule,

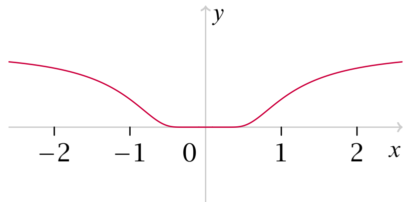

FIGURE 2.2.

$$\begin{aligned}
f^{(k+1)}(x) &= \left[f^{(k)}(x)\right]' \\
&= p_{3k}'\left(x^{-1}\right)\cdot\left(-x^{-2}\right)\cdot e^{-x^{-2}} + p_{3k}\left(x^{-1}\right)\cdot 2x^{-3}\cdot e^{-x^{-2}} \\
&= \left[p_{3k}'\left(x^{-1}\right)\cdot\left(-x^{-2}\right) + p_{3k}\left(x^{-1}\right)\cdot 2x^{-3}\right]\cdot e^{-x^{-2}} \\
&= \left[q_{3k+1}\left(x^{-1}\right) + q_{3k+3}\left(x^{-1}\right)\right]\cdot e^{-x^{-2}} \\
&= p_{3(k+1)}\left(x^{-1}\right)\cdot e^{-x^{-2}},
\end{aligned}$$

where q_{3k+1} and q_{3k+3} are polynomials.

Therefore, $f(x) \in C^\infty$ for all $x \neq 0$. It remains to show that $f^{(k)}(x)$ is defined and continuous at $x = 0$ for all k.

Step 1 Obviously,

$$f'(0) = \lim_{x\to 0}\frac{f(x)-f(0)}{x} = \lim_{x\to 0}\frac{e^{-x^{-2}}}{x} = \lim_{x\to 0} 2x^{-3}e^{-x^{-2}} = 0$$

by L'Hôpital's rule.

Step 2 Suppose that $f^{(k)}(0) = 0$.

Step 3 Then,

$$\begin{aligned}
f^{(k+1)}(0) &= \lim_{x\to 0}\frac{f^{(k)}(x) - f^{(k)}(0)}{x} \\
&= \lim_{x\to 0} p_{3k+1}\left(x^{-1}\right)e^{-x^{-2}} \\
&= \lim_{x\to 0}\frac{p_{3k+1}\left(x^{-1}\right)}{e^{x^{-2}}}.
\end{aligned}$$

Hence, if we use L'Hôpital's rule $3k+1$ times, we get $f^{(k+1)}(0) = 0$.

A similar computation shows that $f^{(k)}(x)$ is continuous at $x = 0$. □

► EXERCISE 56 (2-26*). *Let*

$$f(x) = \begin{cases} e^{-(x-1)^{-2}}\cdot e^{-(x+1)^{-2}} & x \in (-1,1), \\ 0 & x \notin (-1,1). \end{cases}$$

a. *Show that* $f : \mathbb{R} \to \mathbb{R}$ *is a* C^∞ *function which is positive on* $(-1, 1)$ *and* 0 *elsewhere.*

b. *Show that there is a* C^∞ *function* $g : \mathbb{R} \to [0, 1]$ *such that* $g(x) = 0$ *for* $x \leqslant 0$ *and* $g(x) = 1$ *for* $x \geqslant \varepsilon$.

c. *If* $\boldsymbol{a} \in \mathbb{R}^n$, *define* $g : \mathbb{R}^n \to \mathbb{R}$ *by*

$$g(\boldsymbol{x}) = f\left(\frac{x^1 - a^1}{\varepsilon}\right) \cdots f\left(\frac{x^n - a^n}{\varepsilon}\right).$$

Show that g *is a* C^∞ *function which is positive on*

$$\left(a^1 - \varepsilon, a^1 + \varepsilon\right) \times \cdots \times \left(a^n - \varepsilon, a^n + \varepsilon\right)$$

and zero elsewhere.

d. *If* $A \subset \mathbb{R}^n$ *is open and* $C \subset A$ *is compact, show that there is a non-negative* C^∞ *function* $f : A \to \mathbb{R}$ *such that* $f(x) > 0$ *for* $x \in C$ *and* $f = 0$ *outside of some closed set contained in* A.

e. *Show that we can choose such an* f *so that* $f : A \to [0, 1]$ *and* $f(x) = 1$ *for* $x \in C$.

PROOF.

(a) If $x \in (-1, 1)$, then $x - 1 \neq 0$ and $x + 1 \neq 0$. It follows from Exercise 55 that $e^{-(x-1)^{-2}} \in C^\infty$ and $e^{-(x+1)^{-2}} \in C^\infty$. Then it is straightforward to check that $f \in C^\infty$. See Figure 2.3

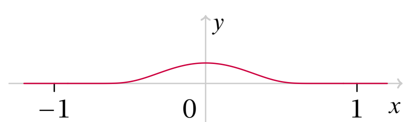

FIGURE 2.3.

(b) By letting $z = x + 1$, we derive a new function $j : \mathbb{R} \to R$ from f as follows:

$$j(z) = \begin{cases} e^{-(z-2)^{-2}} \cdot e^{-z^{-2}} & z \in (0, 2), \\ 0 & z \notin (0, 2). \end{cases}$$

By letting $w = \varepsilon z/2$, we derive a function $k : \mathbb{R} \to \mathbb{R}$ from j as follows:

$$k(w) = \begin{cases} e^{-(2w/\varepsilon-2)^{-2}} \cdot e^{-(2w/\varepsilon)^{-2}} & w \in (0, \varepsilon), \\ 0 & w \notin (0, \varepsilon). \end{cases}$$

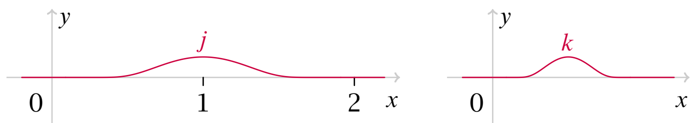

FIGURE 2.4.

It is easy to see that $k \in C^\infty$, which is positive on $(0, \varepsilon)$ and 0 elsewhere. Now let

$$g(x) = \left(\int_0^x k(x)\right) \Big/ \left(\int_0^\varepsilon k(x)\right).$$

Then $g \in C^\infty$; it is 0 for $x \leqslant 0$, increasing on $(0, \varepsilon)$, and 1 for $x \geqslant \varepsilon$.

(c) It follows from (a) immediately.

(d) For every $x \in C$, let $R_x := (-\varepsilon, \varepsilon)^n$ be a rectangle containing x, and $\overline{R_x}$ is contained in A (we can pick such a rectangle since A is open and $C \subset A$). Then $\{R_x : x \in C\}$ is an open cover of C. Since C is compact, there exists $\{x_1, \dots, x_m\} \subset C$ such that $\{R_{x_1}, \dots, R_{x_m}\}$ covers C. For every x_i, $i = 1, \dots, n$, we define a function $g_i : R_{x_i} \to \mathbb{R}$ as

$$g_i(\boldsymbol{x}) = f\left(\frac{x_i^1 - a_i^1}{\varepsilon}\right) \cdots f\left(\frac{x_i^n - a_i^n}{\varepsilon}\right),$$

where $\left(a_i^1, \dots, a_i^n\right) \in \mathbb{R}^n$ is the middle point of R_{x_i}.

Finally, we define $g : R_{x_1} \cup \cdots \cup R_{x_m} \to \mathbb{R}$ as follows:

$$g(x) = \sum_{i=1}^m g_i(x).$$

Then $g \in C^\infty$; it is positive on C, and 0 outside $\overline{R_{x_1}} \cup \cdots \cup \overline{R_{x_m}}$.

(e) Follows the hints. □

► EXERCISE 57 (2-27). *Define* $g, h : \left\{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\| \leqslant 1\right\} \to \mathbb{R}^3$ *by*

$$g(x, y) = \left(x, y, \sqrt{1 - x^2 - y^2}\right),$$
$$h(x, y) = \left(x, y, -\sqrt{1 - x^2 - y^2}\right).$$

Show that the maximum of f *on* $\left\{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\| = 1\right\}$ *is either the maximum of* $f \circ g$ *or the maximum of* $f \circ h$ *on* $\left\{x \in \mathbb{R}^2 : \|x\| \leqslant 1\right\}$.

PROOF. Let $A := \left\{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\| \leqslant 1\right\}$ and $B := \left\{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\| = 1\right\}$. Then $B = g(A) \cup h(A)$. □

2.4 Derivatives

▶ Exercise 58 (2-28). *Find expressions for the partial derivatives of the following functions:*

a. $F(x, y) = f\left(g(x)k\left(y\right), g(x) + h\left(y\right)\right)$.

b. $F(x, y, z) = f\left(g(x + y), h\left(y + z\right)\right)$.

c. $F(x, y, z) = f\left(x^y, y^z, z^x\right)$.

d. $F(x, y) = f\left(x, g(x), h(x, y)\right)$.

Proof.

(a) Letting $\boldsymbol{a} := g(x)k\left(y\right), g(x) + h\left(y\right)$, we have

$$\begin{aligned}
\mathbb{D}_1 F(x, y) &= \mathbb{D}_1 f(\boldsymbol{a}) \cdot g'(x) \cdot k\left(y\right) + \mathbb{D}_2 f(\boldsymbol{a}) \cdot g'(x), \\
\mathbb{D}_2 F(x, y) &= \mathbb{D}_1 f(\boldsymbol{a}) \cdot g(x) \cdot k'\left(y\right) + \mathbb{D}_1 f(\boldsymbol{a}) \cdot h'\left(y\right).
\end{aligned}$$

(b) Letting $\boldsymbol{a} := g(x + y), h\left(y + z\right)$, we have

$$\begin{aligned}
\mathbb{D}_1 F(x, y, z) &= \mathbb{D}_1 f(\boldsymbol{a}) \cdot g'(x + y), \\
\mathbb{D}_2 F(x, y, z) &= \mathbb{D}_1 f(\boldsymbol{a}) \cdot g'(x + y) + \mathbb{D}_2 f(\boldsymbol{a}) \cdot h'\left(y + z\right), \\
\mathbb{D}_3 F(x, y, z) &= \mathbb{D}_2 f(\boldsymbol{a}) \cdot h'\left(y + z\right).
\end{aligned}$$

(c) Letting $\boldsymbol{a} := x^y, y^z, z^x$, we have

$$\begin{aligned}
\mathbb{D}_1 F(x, y, z) &= \mathbb{D}_1 f(\boldsymbol{a}) \cdot y x^{y-1} + \mathbb{D}_3 f(\boldsymbol{a}) \cdot z^x \ln z, \\
\mathbb{D}_2 F(x, y, z) &= \mathbb{D}_1 f(\boldsymbol{a}) \cdot x^y \ln x + \mathbb{D}_2 f(\boldsymbol{a}) \cdot z y^{z-1}, \\
\mathbb{D}_3 F(x, y, z) &= \mathbb{D}_2 f(\boldsymbol{a}) \cdot y^z \ln y + \mathbb{D}_3 f(\boldsymbol{a}) \cdot x z^{x-1}.
\end{aligned}$$

(d) Letting $\boldsymbol{a} := x, g(x), h(x, y)$, we have

$$\begin{aligned}
\mathbb{D}_1 F(x, y) &= \mathbb{D}_1 f(\boldsymbol{a}) + \mathbb{D}_2 f(\boldsymbol{a}) \cdot g'(x) + \mathbb{D}_3 f(\boldsymbol{a}) \cdot \mathbb{D}_1 h(x, y) \\
\mathbb{D}_2 F(x, y) &= \mathbb{D}_3 f(\boldsymbol{a}) \cdot \mathbb{D}_2 h(x, y).
\end{aligned}$$

□

▶ Exercise 59 (2-29). *Let $f : \mathbb{R}^n \to \mathbb{R}$. For $\boldsymbol{x} \in \mathbb{R}^n$, the limit*

$$\lim_{t \to 0} \frac{f\left(\boldsymbol{a} + t\boldsymbol{x}\right) - f(\boldsymbol{a})}{t},$$

if it exists, is denoted $\mathbb{D}_{\boldsymbol{x}} f(\boldsymbol{a})$, and called the directional derivative *of f at $\boldsymbol{a}$, in the direction $\boldsymbol{x}$.*

a. *Show that $\mathbb{D}_{\boldsymbol{e}_i} f(\boldsymbol{a}) = \mathbb{D}_i f(\boldsymbol{a})$.*

b. *Show that $\mathbb{D}_{t\boldsymbol{x}} f(\boldsymbol{a}) = t\mathbb{D}_{\boldsymbol{x}} f(\boldsymbol{a})$.*

c. *If f is differentiable at $\boldsymbol{a}$, show that $\mathbb{D}_{\boldsymbol{x}} f(\boldsymbol{a}) = \mathbb{D} f(\boldsymbol{a})(\boldsymbol{x})$ and therefore $\mathbb{D}_{\boldsymbol{x}+\boldsymbol{y}} f(\boldsymbol{a}) = \mathbb{D}_{\boldsymbol{x}} f(\alpha) + \mathbb{D}_{\boldsymbol{y}} f(\boldsymbol{a})$.*

PROOF.

(a) For $\boldsymbol{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$, we have

$$\begin{aligned}\mathbb{D}_{\boldsymbol{e}_i} f(\boldsymbol{a}) &= \lim_{t\to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{e}_i) - f(\boldsymbol{a})}{t}\\ &= \lim_{t\to 0} \frac{f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) - f(\boldsymbol{a})}{t}\\ &= \mathbb{D}_i f(\boldsymbol{a})\end{aligned}$$

by definition.

(b) We have

$$\mathbb{D}_{t\boldsymbol{x}} f(\boldsymbol{a}) = \lim_{s\to 0} \frac{f(\boldsymbol{a} + st\boldsymbol{x}) - f(\boldsymbol{a})}{s} = \lim_{st\to 0} t\frac{f(\boldsymbol{a} + st\boldsymbol{x}) - f(\boldsymbol{a})}{st} = t\mathbb{D}_{\boldsymbol{x}} f(\boldsymbol{a}).$$

(c) If f is differentiable at $\boldsymbol{a}$, then for any $\boldsymbol{x} \neq \boldsymbol{0}$ we have

$$\begin{aligned}0 &= \lim_{t\to\boldsymbol{0}} \frac{|f(\boldsymbol{a} + t\boldsymbol{x}) - f(\boldsymbol{a}) - \mathbb{D} f(\boldsymbol{a})(t\boldsymbol{x})|}{\|t\boldsymbol{x}\|}\\ &= \lim_{t\to 0} \frac{|f(\boldsymbol{a} + t\boldsymbol{x}) - f(\boldsymbol{a}) - t\cdot\mathbb{D} f(\boldsymbol{a})(\boldsymbol{x})|}{|t|}\cdot\frac{1}{\|\boldsymbol{x}\|}\\ &= \lim_{t\to 0}\left|\frac{f(\boldsymbol{a} + t\boldsymbol{x}) - f(\boldsymbol{a})}{t} - \mathbb{D} f(\boldsymbol{a})(\boldsymbol{x})\right|\cdot\frac{1}{\|\boldsymbol{x}\|},\end{aligned}$$

and so

$$\mathbb{D}_{\boldsymbol{x}} f(\boldsymbol{a}) = \lim_{t\to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{x}) - f(\boldsymbol{a})}{t} = \mathbb{D} f(\boldsymbol{a})(\boldsymbol{x}).$$

The case of $\boldsymbol{x} = \boldsymbol{0}$ is trivial. Therefore,

$$\begin{aligned}\mathbb{D}_{\boldsymbol{x}+\boldsymbol{y}} f(\boldsymbol{a}) &= \mathbb{D} f(\boldsymbol{a})(\boldsymbol{x} + \boldsymbol{y})\\ &= \mathbb{D} f(\boldsymbol{a})(\boldsymbol{x}) + \mathbb{D} f(\boldsymbol{a})(\boldsymbol{y})\\ &= \mathbb{D}_{\boldsymbol{x}} f(\boldsymbol{a}) + \mathbb{D}_{\boldsymbol{y}} f(\boldsymbol{a}).\end{aligned}$$

□

► EXERCISE 60 (2-30). *Let f be defined as in Exercise 34. Show that $\mathbb{D}_{\boldsymbol{x}} f(0,0)$ exists for all $\boldsymbol{x}$, but if $g \neq 0$, then $\mathbb{D}_{\boldsymbol{x}+\boldsymbol{y}} f(0,0) \neq \mathbb{D}_{\boldsymbol{x}} f(0,0) + \mathbb{D}_{\boldsymbol{y}} f(0,0)$ for all $\boldsymbol{x}, \boldsymbol{y}$.*

PROOF. Take any $\boldsymbol{x} \in \mathbb{R}^2$.

$$\lim_{t\to 0} \frac{f(t\boldsymbol{x}) - f(0,0)}{t} = \lim_{t\to 0} \frac{|t|\cdot\|\boldsymbol{x}\|\cdot g\left(t\boldsymbol{x}\Big/\big(|t|\cdot\|\boldsymbol{x}\|\big)\right)}{t}.$$

Therefore, $\mathbb{D}_{\boldsymbol{x}} f(0,0)$ exists for any $\boldsymbol{x}$.

Now let $g \neq 0$; then, $\mathbb{D}_{(0,1)} f(0,0) = \mathbb{D}_{(1,0)} f(0,0) = 0$, but $\mathbb{D}_{(1,0)+(0,1)} f(0,0) = \mathbb{D}_{(1,1)} f(0,0) \neq 0$. □

▶ EXERCISE 61 (2-31). *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined as in Exercise 26. Show that $\mathbb{D}_{\boldsymbol{x}} f(0,0)$ exists for all $\boldsymbol{x}$, although f is not even continuous at $(0,0)$.*

PROOF. For any $\boldsymbol{x} \in \mathbb{R}^2$, we have

$$\lim_{t\to 0} \frac{f(t\boldsymbol{x}) - f(\mathbf{0})}{t} = \lim_{t\to 0} \frac{f(t\boldsymbol{x})}{t} = 0$$

by Exercise 26 (a). □

▶ EXERCISE 62 (2-32).

a. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by*

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0. \end{cases}$$

Show that f is differentiable at 0 but f' is not continuous at 0.

b. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by*

$$f(x,y) = \begin{cases} \left(x^2+y^2\right) \sin \dfrac{1}{\sqrt{x^2+y^2}} & (x,y) \neq \mathbf{0} \\ 0 & (x,y) = \mathbf{0}. \end{cases}$$

Show that f is differentiable at $(0,0)$ but $\mathbb{D}_i f$ is not continuous at $(0,0)$.

PROOF.

(a) We have

$$\lim_{x\to 0} \frac{f(x) - f(0)}{x} = \lim_{x\to 0} \frac{x^2 \sin \frac{1}{x}}{x} = \lim_{x\to 0} x \sin \frac{1}{x} = 0.$$

Hence, $f'(0) = 0$. Further, for any $x \neq 0$, we have

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}.$$

It is clear that $\lim_{x\to 0} f'(x)$ does not exist. Therefore, f' is not continuous at 0.

(b) Since

$$\lim_{(x,y)\to(0,0)} \frac{\left(x^2+y^2\right) \sin \dfrac{1}{\sqrt{x^2+y^2}}}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \sqrt{x^2+y^2} \sin \frac{1}{\sqrt{x^2+y^2}} = 0,$$

we know that $f'(0,0) = (0,0)$. Now take any $(x,y) \neq (0,0)$. Then

$$\mathbb{D}_1 f(x,y) = 2x \sin \frac{1}{\sqrt{x^2+y^2}} - 2x \cos \frac{1}{\sqrt{x^2+y^2}}.$$

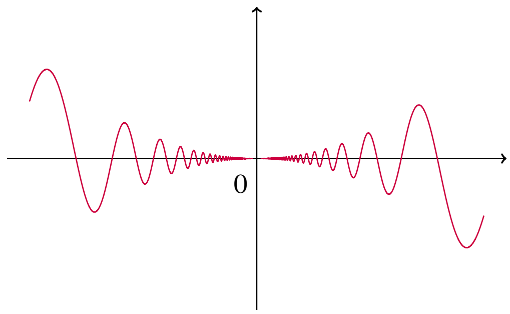

FIGURE 2.5.

As in (a), $\lim_{x\to 0} \mathbb{D}_1 f(x,0)$ does not exist. Similarly for $\mathbb{D}_2 f$. □

► EXERCISE 63 (2-33). *Show that the continuity of $\mathbb{D}_1 f^j$ at $\boldsymbol{a}$ may be eliminated from the hypothesis of Theorem 2-8.*

PROOF. It suffices to see that for the first term in the sum, we have, by letting $\left(a^2, \dots, a^n\right) =: \boldsymbol{a}_{-1}$,

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{\left|f\left(a^1+h^1, \boldsymbol{a}_{-1}\right) - f(\boldsymbol{a}) - \mathbb{D}_1 f(\boldsymbol{a})\cdot h^1\right|}{\|\boldsymbol{h}\|}$$
$$\leqslant \lim_{h^1\to 0} \frac{\left|f\left(a^1+h^1, \boldsymbol{a}_{-1}\right) - f(\boldsymbol{a}) - \mathbb{D}_1 f(\boldsymbol{a})\cdot h^1\right|}{\left|h^1\right|} = 0.$$

See aslo Apostol (1974, Theorem 12.11). □

► EXERCISE 64 (2-34). *A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is* homogeneous *of degree m if $f(t\boldsymbol{x}) = t^m f(\boldsymbol{x})$ for all $\boldsymbol{x}$. If f is also differentiable, show that*

$$\sum_{i=1}^n x^i \mathbb{D}_i f(\boldsymbol{x}) = mf(\boldsymbol{x}).$$

PROOF. Let $g(t) = f(t\boldsymbol{x})$. Then, by Theorem 2-9,

$$g'(t) = \sum_{i=1}^n \mathbb{D}_i f(t\boldsymbol{x})\cdot x^i. \tag{2.4}$$

On the other hand, $g(t) = f(t\boldsymbol{x}) = t^m f(\boldsymbol{x})$; then

$$g'(t) = mt^{m-1} f(\boldsymbol{x}). \tag{2.5}$$

Combining (2.4) and (2.5), and letting $t = 1$, we then get the result. □

► EXERCISE 65 (2-35). *If $f\colon \mathbb{R}^n \to \mathbb{R}$ is differentiable and $f(\boldsymbol{0}) = 0$, prove that there exist $g_i\colon \mathbb{R}^n \to \mathbb{R}$ such that*

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} x^i g_i(\boldsymbol{x}).$$

Proof. Let $h_{\boldsymbol{x}}(t) = f(t\boldsymbol{x})$. Then

$$\int_0^1 h'_{\boldsymbol{x}}(t)\,\mathrm{d}t = h_{\boldsymbol{x}}(1) - h_{\boldsymbol{x}}(0) = f(\boldsymbol{x}) - f(\boldsymbol{0}) = f(\boldsymbol{x}).$$

Hence,

$$\begin{aligned} f(\boldsymbol{x}) = \int_0^1 h'_{\boldsymbol{x}}(t)\,\mathrm{d}t = \int_0^1 f'(t\boldsymbol{x})\,\mathrm{d}t = \int_0^1 \left[\sum_{i=1}^{n} x_i \mathbb{D}_i f(t\boldsymbol{x})\right] \mathrm{d}t \\ = \sum_{i=1}^{n} x^i \int_0^1 \mathbb{D}_i f(t\boldsymbol{x})\,\mathrm{d}t \\ = \sum_{i=1}^{n} x^i g_i(\boldsymbol{x}), \end{aligned}$$

where $g_i(\boldsymbol{x}) = \int_0^1 \mathbb{D}_i f(t\boldsymbol{x})\,\mathrm{d}t$. □

2.5 Inverse Functions

For this section, Rudin (1976, Section 9.3 and 9.4) is a good reference.

► Exercise 66 (2-36*). *Let $A \subset \mathbb{R}^n$ be an open set and $f: A \to \mathbb{R}^n$ a continuously differentiable 1-1 function such that* $\det\left(f'(\boldsymbol{x})\right) \neq 0$ *for all $\boldsymbol{x}$. Show that $f(A)$ is an open set and $f^{-1}: f(A) \to A$ is differentiable. Show also that $f(B)$ is open for any open set $B \subset A$.*

Proof. For every $\boldsymbol{y} \in f(A)$, there exists $\boldsymbol{x} \in A$ such that $f(\boldsymbol{x}) = \boldsymbol{y}$. Since $f \in \mathscr{C}'(A)$ and $\det\left(f'(\boldsymbol{x})\right) \neq 0$, it follows from the Inverse Function Theorem that there is an open set $V \subset A$ containing $\boldsymbol{x}$ and an open set $W \subset \mathbb{R}^n$ containing $\boldsymbol{y}$ such that $W = f(V)$. This proves that $f(A)$ is open.

Since $f: V \to W$ has a continuous inverse $f^{-1}: W \to V$ which is differentiable, it follows that f^{-1} is differentiable at $\boldsymbol{y}$; since $\boldsymbol{y}$ is chosen arbitrary, it follows that $f^{-1}: f(A) \to A$ is differentiable.

Take any open set $B \subset A$. Since $f \upharpoonright B \in \mathscr{C}'(B)$ and $\det\left(\left(f \upharpoonright B\right)'(\boldsymbol{x})\right) \neq 0$ for all $\boldsymbol{x} \in B \subset A$, it follows that $f(B)$ is open. □

► Exercise 67 (2-37).

a. *Let $f: \mathbb{R}^2 \to \mathbb{R}$ be a continuously differentiable function. Show that f is* not *1-1.*

b. *Generalize this result to the case of a continuously differentiable function $f: \mathbb{R}^n \to \mathbb{R}^m$ with $m < n$.*

PROOF.

(a) Let $f \in \mathcal{C}'$. Then both $\mathbb{D}_1 f$ and $\mathbb{D}_2 f$ are continuous. Assume that f is 1-1; then both $\mathbb{D}_1 f$ and $\mathbb{D}_2 f$ cannot not be constant and equal to 0. So suppose that there is $(x_0, y_0) \in \mathbb{R}^2$ such that $\mathbb{D}_1 f\left(x_0, f_0\right) \neq 0$. The continuity of $\mathbb{D}_1 f$ implies that there is an open set $A \subset \mathbb{R}^2$ containing (x_0, y_0) such that $\mathbb{D}_1 f(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in A$.

Define a function $g: A \to \mathbb{R}^2$ with

$$g(x, y) = \left(f(x, y), y\right).$$

Then for all $(x, y) \in A$,

$$g'(x, y) = \begin{pmatrix} \mathbb{D}_1 f(x, y) & \mathbb{D}_2 f(x, y) \\ 0 & 1 \end{pmatrix},$$

and so $\det\left(g'(x, y)\right) = \mathbb{D}_1 f(x, y) \neq 0$; furthermore, $g \in \mathcal{C}'(A)$ and g is 1-1. Then by Exercise 66, we know that $g(A)$ is open. We now show that $g(A)$ cannot be open actually.

Take a point $\left(f\left(x_0, y_0\right), \tilde{y}\right) \in g(A)$ with $y \neq y_0$. Then for any $(x, y) \in A$, we must have

$$g(x, y) = \left(f(x, y), y\right) = \left(f\left(x_0, y_0\right), \tilde{y}\right) \Longrightarrow (x, y) = \left(x_0, y_0\right);$$

that is, there is no $(x, y) \in A$ such that $g(x, y) = \left(f\left(x_0, y_0\right), \tilde{y}\right)$. This proves that f cannot be 1-1.

(b) We can write $f: \mathbb{R}^n \to \mathbb{R}^m$ as $f = \left(f^1, \dots, f^m\right)$, where $f^i: \mathbb{R}^n \to \mathbb{R}$ for every $i = 1, \dots, m$. As in (a), there is a mapping, say, f^1, a point $\boldsymbol{a} \in \mathbb{R}^n$, and an open set A containing $\boldsymbol{a}$ such that $\mathbb{D}_1 f^1(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in A$. Define $g: A \to \mathbb{R}^m$ as

$$g\left(x^1, \boldsymbol{x}^{-1}\right) = \left(f(\boldsymbol{x}), \boldsymbol{x}^{-1}\right),$$

where $\boldsymbol{x}^{-1} := \left(x^2, \dots, x^n\right)$. Then as in (a), it follows that f cannot be 1-1. □

► EXERCISE 68 (2-38).

a. *If $f: \mathbb{R} \to \mathbb{R}$ satisfies $f'(a) \neq 0$ for all $a \in \mathbb{R}$, show that f is 1-1 (on all of $\mathbb{R}$).*

b. *Define $f: \mathbb{R}^2 \to \mathbb{R}^2$ by $f(x, y) = \left(e^x \cos y, e^x \sin y\right)$. Show that $\det\left(f'(x, y)\right) \neq 0$ for all (x, y) but f is not 1-1.*

PROOF.

(a) Suppose that f is not 1-1. Then there exist $a, b \in \mathbb{R}$ with $a < b$ such that $f(a) = f(b)$. It follows from the mean-value theorem that there exists $c \in (a, b)$ such that

$$0 = f(b) - f(a) = f'(c)(b - a),$$

which implies that $f'(c) = 0$. A contradiction.

(b) We have

$$\begin{aligned} f'(x, y) &= \begin{pmatrix} \mathbb{D}_x e^x \cos y & \mathbb{D}_y e^x \cos y \\ \mathbb{D}_x e^x \sin y & \mathbb{D}_y e^x \sin y \end{pmatrix} \\ &= \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}. \end{aligned}$$

Then

$$\det\left(f'(x, y)\right) = e^{2x}\left(\cos^2 y + \sin^2 y\right) = e^{2x} \neq 0.$$

However, $f(x, y)$ is not 1-1 since $f(x, y) = f\left(x, y + 2k\pi\right)$ for all $(x, y) \in \mathbb{R}^2$ and $k \in \mathbb{N}$.

This exercise shows that the non-singularity of $\mathbb{D}f$ on A implies that f is locally 1-1 at each point of A, but it does not imply that f is 1-1 on all of A. See Munkres (1991, p. 69). □

► Exercise 69 (2-39). *Use the function* $f : \mathbb{R} \to \mathbb{R}$ *defined by*

$$f(x) = \begin{cases} \frac{x}{2} + x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

to show that continuity of the derivative cannot be eliminated from the hypothesis of Theorem 2-11.

Proof. If $x \neq 0$, then

$$f'(x) = \frac{1}{2} + 2x \sin \frac{1}{x} - \cos \frac{1}{x};$$

if $x = 0$, then

$$f'(0) = \lim_{h \to 0} \frac{h/2 + h^2 \sin\left(1/h\right)}{h} = \frac{1}{2}.$$

Hence, $f'(x)$ is not continuous at 0. It is easy to see that f is not injective for any neighborhood of 0 (see Figure 2.6).

2.6 Implicit Functions

► Exercise 70 (2-40). *Use the implicit function theorem to re-do Exercise 45 (c).*

Proof. Define $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ by

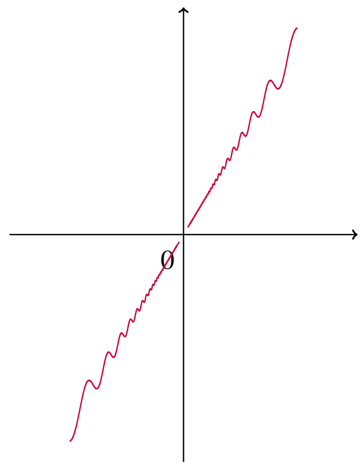

FIGURE 2.6.

□

$$f^i(t,\mathbf{s}) = \sum_{j=1}^{n} a_{ji}(t)s^j - b_i(t),$$

for $i = 1,\ldots,n$. Then

$$\mathbf{M} := \begin{pmatrix} \mathbb{D}_2 f^1(t,\mathbf{s}) & \cdots & \mathbb{D}_{1+n} f^1(t,\mathbf{s}) \\ \vdots & \ddots & \vdots \\ \mathbb{D}_2 f^n(t,\mathbf{s}) & \cdots & \mathbb{D}_{1+n} f^n(t,\mathbf{s}) \end{pmatrix} = \begin{pmatrix} a_{11}(t) & \cdots & a_{n1}(t) \\ \vdots & \ddots & \vdots \\ a_{1n}(t) & \cdots & a_{nn}(t) \end{pmatrix},$$

and so $\det(\mathbf{M}) \neq 0$.

It follows from the Implicit Function Theorem that for each $t \in \mathbb{R}$, there is a unique $\mathbf{s}(t) \in \mathbb{R}^n$ such that $f(t,\mathbf{s}(t)) = 0$, and $\mathbf{s}$ is differentiable. □

► EXERCISE 71 (2-41). *Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be differentiable. For each $x \in \mathbb{R}$ define $g_x : \mathbb{R} \to \mathbb{R}$ by $g_x(y) = f(x,y)$. Suppose that for each x there is a unique y with $g_x'(y) = 0$; let $c(x)$ be this y.*

a. *If $\mathbb{D}_{2,2} f(x,y) \neq 0$ for all (x,y), show that c is differentiable and*

$$c'(x) = -\frac{\mathbb{D}_{2,1} f(x,c(x))}{\mathbb{D}_{2,2} f(x,c(x))}.$$

b. *Show that if $c'(x) = 0$, then for some y we have*

$$\mathbb{D}_{2,1} f(x,y) = 0,$$
$$\mathbb{D}_2 f(x,y) = 0.$$

c. *Let $f(x,y) = x(y\log y - y) - y\log x$. Find*

$$\max_{1/2\leqslant x\leqslant 2}\left[\min_{1/3\leqslant y\leqslant 1} f(x,y)\right].$$

PROOF.

(a) For every x, we have $g'_x(y) = \mathbb{D}_2 f(x,y)$. Since for every x there is a unique $y = c(x)$ such that $\mathbb{D}_2 f(x, c(x)) = 0$, the solution $c(x)$ is the same as obtained from the Implicit Function Theorem; hence, $c(x)$ is differentiable, and by differentiating $\mathbb{D}_2 f(x, c(x)) = 0$ with respect to x, we have

$$\mathbb{D}_{2,1} f(x, c(x)) + \mathbb{D}_{2,2} f(x, c(x)) \cdot c'(x) = 0;$$

that is,

$$c'(x) = -\frac{\mathbb{D}_{2,1} f(x, c(x))}{\mathbb{D}_{2,2} f(x, c(x))}.$$

(b) It follows from (a) that if $c'(x) = 0$, then $\mathbb{D}_{2,1} f(x, c(x)) = 0$. Hence, there exists some $y = c(x)$ such that $\mathbb{D}_{2,1} f(x,y) = 0$. Furthermore, by definition, $\mathbb{D}_2(x, c(x)) = \mathbb{D}_2 f(x,y) = 0$.

(c) We have

$$\mathbb{D}_2 f(x,y) = x\ln y - \ln x.$$

Let $\mathbb{D}_2 f(x,y) = 0$ we have $y = c(x) = x^{1/x}$. Also, $\mathbb{D}_{2,2} f(x,y) = x/y > 0$ since $x, y > 0$. Hence, for every fixed $x \in [1/2, 2]$,

$$\min_y f(x,y) = f(x, c(x)).$$

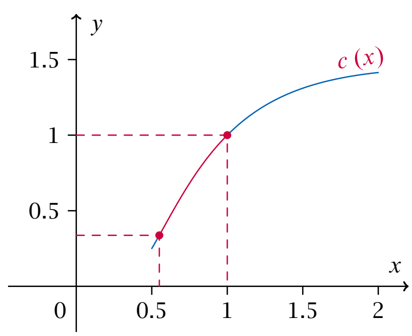

FIGURE 2.7.

It is easy to see that $c'(x) > 0$ on $[1/2, 2]$, $c(1) = 1$, and $c(a) = 1/3$ for some $a > 1/2$ (see Figure 2.7). Therefore,

$$\min_{1/3\leqslant y\leqslant 1} f(x,y) = f(x, y^*(x)),$$

where (see Figure 2.8)

$$y^*(x) = \begin{cases} 1/3 & \text{if } 1/2 \leqslant x \leqslant a \\ c(x) = x^{1/x} & \text{if } a < x \leqslant 1 \\ 1 & \text{if } 1 < x \leqslant 2. \end{cases}$$

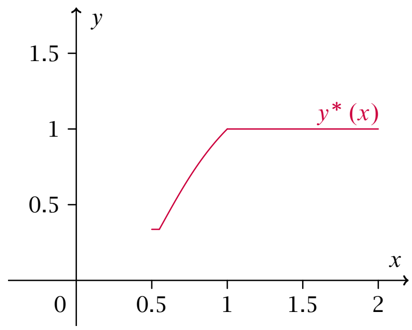

FIGURE 2.8.

$1/2 \leqslant x \leqslant a$ In this case, our problem is

$$\max_{1/2 \leqslant x \leqslant a} f(x, 1/3) = -\left(\frac{1+\ln 3}{3}\right) x - \frac{1}{3} \ln x.$$

It is easy to see that $x^* = 1/2$, and so $f(x^*, 1/3) = \ln(4/3e)/6$.

$a < x \leqslant 1$ In this case, our problem is

$$\max_{a < x \leqslant 1} f\left(x, x^{1/x}\right) = -x^{1+1/x}.$$

It is easy to see that the maximum of f occurs at $x^* = a$ and $y^*(x^*) = 1/3$.

$1 < x \leqslant 2$ In this case, our problem is

$$\max_{1 < x \leqslant 2} f(x, 1) = -x - \ln x.$$

The maximum of f occurs at $x^* = 1$.

Now, as depicted in Figure 2.9, we have $x^* = 1/2$, $y^* = 1/3$, and $f(x^*, y^*) = \ln(4/3e)/6$. □

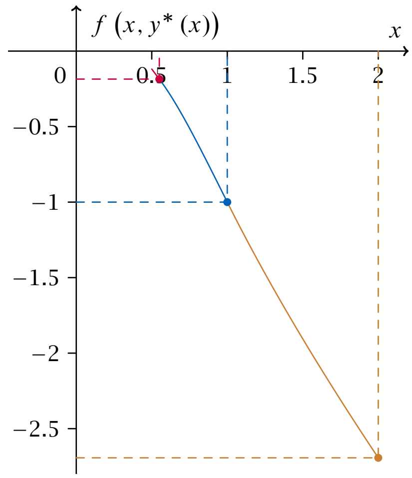

FIGURE 2.9.

3

INTEGRATION

3.1 Basic Definitions

► Exercise 72 (3-1). *Let $f\colon [0,1]\times[0,1]\to\mathbb{R}$ be defined by*

$$f(x,y)=\begin{cases}0 & \text{if } 0\leqslant x<1/2\\ 1 & \text{if } 1/2\leqslant x\leqslant 1.\end{cases}$$

Show that f is integrable and $\int_{[0,1]\times[0,1]} f=1/2$.

Proof. Consider a partition $P=(P_1,P_2)$ with $P_1=P_2=(0,1/2,1)$. Then $L(f,P)=U(f,P)=1/2$. It follows from Theorem 3-3 (the Riemann condition) that f is integrable and $\int_{[0,1]\times[0,1]} f=1/2$. □

► Exercise 73 (3-2). *Let $f\colon A\to\mathbb{R}$ be integrable and let $g=f$ except at finitely many points. Show that g is integrable and $\int_A f=\int_A g$.*

Proof. Fix an $\varepsilon>0$. It follows from the Riemann condition that there is a partition P of A such that

$$U(f,P)-L(f,P)<\frac{\varepsilon}{2}.$$

Let P' be a refinement of P such that:

- for every $\boldsymbol{x}\in A$ with $g(\boldsymbol{x})\neq f(\boldsymbol{x})$, it belongs to 2^n subrectangles of P', i.e., $\boldsymbol{x}$ is a corner of each subrectangle.
- for every subrectangle S of P',

  $$v(S)<\frac{\varepsilon}{2^{n+1}d\,(u-\ell)},$$

  where

$$\begin{aligned}
d &= \left|\left\{\boldsymbol{x} : f(\boldsymbol{x}) \neq g(\boldsymbol{x})\right\}\right|, \\
u &= \sup_{\boldsymbol{x}\in A}\{g(\boldsymbol{x})\} - \inf_{\boldsymbol{x}\in A}\{f(\boldsymbol{x})\}, \\
\ell &= \inf_{\boldsymbol{x}\in A}\{g(\boldsymbol{x})\} - \sup_{\boldsymbol{x}\in A}\{f(\boldsymbol{x})\}.
\end{aligned}$$

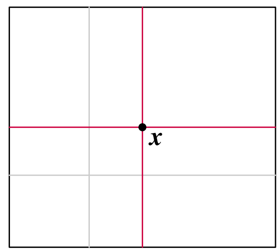

FIGURE 3.1.

With such a choice of partition of A, we have

$$\begin{aligned}
U(g, P') - U(f, P') &= \sum_{i=1}^{d}\left[\sum_{j=1}^{2^n}\left[M_{S_{ij}}(g) - M_{S_{ij}}(f)\right] v(S_{ij})\right] \\
&\leqslant d2^n uv,
\end{aligned}$$

where $v := \sup_{S\in P'}\{v(S)\}$ is the least upper bound of the volumes of the subrectangles of P'. Similarly,

$$\begin{aligned}
L(g, P') - L(f, P') &= \sum_{i=1}^{d}\left[\sum_{j=1}^{2^n}\left[m_{S_{ij}}(g) - m_{S_{ij}}(f)\right] v(S_{ij})\right] \\
&\geqslant d2^n \ell v.
\end{aligned}$$

Therefore,

$$\begin{aligned}
U(g, P') - L(g, P') &\leqslant \left[U(f, P') + d2^n uv\right] - \left[L(f, P') + d2^n \ell v\right] \\
&\leqslant \frac{\varepsilon}{2} + d2^n (u-\ell) v \\
&= \frac{\varepsilon}{2} + d2^n (u-\ell) \frac{\varepsilon}{2^{n+1} d (u-\ell)} \\
&= \varepsilon;
\end{aligned}$$

that is, g is integrable. It is easy to see now that $\int_A g = \int_A f$. □

► EXERCISE 74 (3-3). *Let $f, g\colon A \to \mathbb{R}$ be integrable.*

a. *For any partition P of A and subrectangle S, show that $m_S(f) + m_S(g) \leqslant m_S(f+g)$ and $M_S(f+g) \leqslant M_S(f) + M_S(g)$ and therefore $L(f, P) + L(g, P) \leqslant L(f+g, P)$ and $U(f+g, P) \leqslant U(f, P) + U(g, P)$.*

b. *Show that* $f+g$ *is integrable and* $\int_A (f+g) = \int_A f + \int_A g$.

c. *For any constant* c, *show that* $\int_A cf = c\int_A f$.

PROOF.

(a) We show that $m_S(f) + m_S(g)$ is a lower bound of $\left\{(f+g)(x) : x \in S\right\}$. It is clear that $m_S(f) \leqslant f(x)$ and $m_S(g) \leqslant g(x)$ for any $x \in S$. Then for every $x \in S$ we have

$$m_S(f) + m_S(g) \leqslant f(x) + g(x) = (f+g)(x).$$

Hence, $m_S(f) + m_S(g) \leqslant m_S(f+g)$.

Similarly, for every $x \in S$ we have $M_S(f) \geqslant f(x)$ and $M_S(g) \geqslant g(x)$; hence, $(f+g)(x) = f(x) + g(x) \leqslant M_S(f) + M_S(g)$ and so $M_S(f+g) \leqslant M_S(f) + M_S(g)$.

Now for any partition P of A we have

$$\begin{aligned} L(f,P) + L(g,P) &= \sum_{S\in P} m_S(f)\, v(S) + \sum_{S\in P} m_S(g) \\ &= \sum_{S\in P} \left[m_S(f) + m_S(g)\right] v(S) \\ &\leqslant \sum_{S\in P} m_S(f+g)\, v(S) \\ &= L(f+g,P), \end{aligned} \tag{3.1}$$

and

$$\begin{aligned} U(f,P) + U(g,P) &= \sum_{S\in P} M_S(f)\, v(S) + \sum_{S\in P} M_S(g)\, v(S) \\ &= \sum_{S\in P} \left[M_S(f) + M_S(g)\right] v(S) \\ &\geqslant \sum_{S\in P} M_S(f+g)\, v(S) \\ &= U(f+g,P). \end{aligned} \tag{3.2}$$

(b) It follows from (3.1) and (3.2) that for any partition P,

$$\begin{aligned} U(f+g,P) - L(f+g,P) &\leqslant \left[U(f,P) + U(g,P)\right] - \left[L(f,P) + L(g,P)\right] \\ &= \left[U(f,P) - L(f,P)\right] + \left[U(g,P) - L(g,P)\right]. \end{aligned}$$

Since f and g are integrable, there exist P' and P'' such that for any $\varepsilon > 0$, we have $U(f,P') - L(f,P') < \varepsilon/2$ and $U(g,P'') - L(g,P'') < \varepsilon/2$. Let $\overline{P}$ refine both P' and P''. Then

$$U\left(f,\overline{P}\right) - L\left(f,\overline{P}\right) < \frac{\varepsilon}{2} \quad \text{and} \quad U\left(g,\overline{P}\right) - L\left(g,\overline{P}\right) < \frac{\varepsilon}{2}.$$

Hence,

$$U\left(f+g,\bar{P}\right)-L\left(f+g,\bar{P}\right)<\varepsilon,$$

and so $f+g$ is integrable.

Now, by definition, for any $\varepsilon>0$, there exists a partition P (by using a common refinement partition if necessary) such that $\int_A f < L(f,P)+\varepsilon/2$, $\int_A g < L(g,P)+\varepsilon/2$, $U(f,P)<\int_A f+\varepsilon/2$, and $U(g,P)<\int_A g+\varepsilon/2$. Therefore,

$$\begin{aligned}\int_A f+\int_A g-\varepsilon < L(f,P)+L(g,P)\leqslant L(f+g,P)&\leqslant \int_A (f+g)\\ &\leqslant U(f+g,P)\\ &\leqslant U(f,P)+U(g,P)\\ &<\int_A f+\int_A g+\varepsilon.\end{aligned}$$

Hence, $\int_A (f+g)=\int_A f+\int_A g$.

(c) First, suppose that $c>0$. Then for any partition P and any subrectangle S, we have $m_S(cf)=cm_S(f)$ and $M_S(cf)=cM_S(f)$. But then $L(cf,P)=cL(f,P)$ and $U(cf,P)=cU(f,P)$. Since f is integrable, for any $\varepsilon>0$ there exists a partition P such that $U(f,P)-L(f,P)<\varepsilon/c$. Therefore,

$$U(cf,P)-L(cf,P)=c\left[U(f,P)-L(f,P)\right]<\varepsilon;$$

that is, cf is integrable. Further,

$$\begin{aligned}c\int_A f-\frac{\varepsilon}{c}<cL(f,P)=L(cf,P)\leqslant\int_A cf\leqslant U(cf,P)&=cU(f,P)\\ &<c\int_A f+\frac{\varepsilon}{c},\end{aligned}$$

i.e., $\int_A cf=c\int_A f$.

Now let $c<0$. Then for any partition P of A, we have $m_S(cf)=cM_S(f)$ and $M_S(cf)=cm_S(f)$. Hence $L(cf,P)=cU(f,P)$ and $U(cf,P)=cL(f,P)$. Since f is integrable, for every $\varepsilon>0$, choose P such that $U(f,P)-L(f,P)<-\varepsilon/c$. Then

$$U(cf,P)-L(cf,P)=-c\left[U(f,P)-L(f,P)\right]<\varepsilon;$$

that is, cf is integrable. Furthermore,

$$\begin{aligned}-c\int_A f+\frac{\varepsilon}{c}<-cL(f,P)=-U(cf,P)\leqslant-\int_A cf\leqslant-L(cf,P)&=-cL(f,P)\\ &<-c\int_A f-\frac{\varepsilon}{c},\end{aligned}$$

i.e., $\int_A cf=c\int_A f$. □

► EXERCISE 75 (3-4). *Let $f : A \to \mathbb{R}$ and let P be a partition of A. Show that f is integrable if and only if for each subrectangle S the function $f \restriction S$ is integrable, and that in this case $\int_A f = \sum_S \int_S f \restriction S$.*

PROOF. Let P be a partition of A, and S be a subrectangle with respect to P.

Only if: Suppose that f is integrable. Then there exists a partition P_1 of A such that $U\left(f, P_1\right) - L\left(f, P_1\right) < \varepsilon$ for any given $\varepsilon > 0$. Let P_2 be a common refinement of P and P_1. Then

$$U\left(f, P_2\right) - L\left(f, P_2\right) \leqslant U\left(f, P_1\right) - L\left(f, P_1\right) < \varepsilon,$$

and there are rectangles $\left\{S_2^1, \ldots, S_2^n\right\} =: \mathcal{S}_2(S)$ with respect to P_2, such that $S = \bigcup_{i=1}^n S_2^i$. Therefore,

$$\begin{aligned} U\left(f, P_2\right) - L\left(f, P_2\right) &= \sum_{S_2} \left[M_{S_2}(f) - m_{S_2}(f)\right] v(S_2) \\ &\geqslant \sum_{S_2 \in \mathcal{S}_2(S)} \left[M_{S_2}(f) - m_{S_2}(f)\right] v(S_2) \\ &= U\left(f \restriction S, P_2\right) - L\left(f \restriction S, P_2\right); \end{aligned}$$

that is, $f \restriction S$ is integrable.

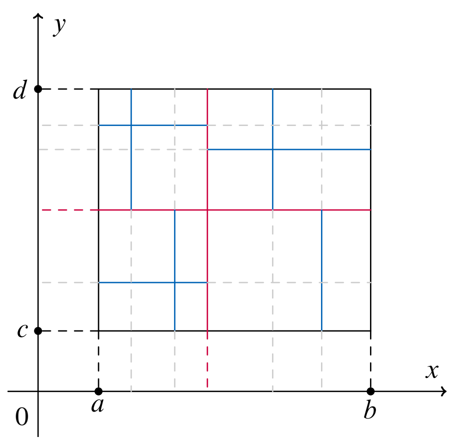

FIGURE 3.2.

If: Now suppose that $f \restriction S$ is integrable for each S. For each partition P', let $\left|P'\right|$ be the number of subrectangles induced by P'. Let P_S be a partition such that

$$U\left(f \restriction S, P_S\right) - L\left(f \restriction S, P_S\right) < \frac{\varepsilon}{2|P|}.$$

Let P' be the partition of A obtained by taking the union of all the subsequences defining the partitions of the P_S; see Figure 3.2. Then there are refinements P'_S of P_S whose rectangles are the set of all subrectangles of P' which are contained in S. Hence,

$$\begin{aligned}\sum_S \int_S f\upharpoonright S - \varepsilon < \sum_S L\left(f\upharpoonright S, P_S\right) \leqslant \sum_S L\left(f\upharpoonright S, P'_S\right) &= L\left(f, P'\right) \\ &\leqslant U\left(f, P'\right) \\ &= \sum_S U\left(f\upharpoonright S, P'_S\right) \\ &\leqslant \sum_S U\left(f\upharpoonright S, P_S\right) \\ &< \sum_S \int_S f\upharpoonright S + \varepsilon.\end{aligned}$$

Therefore, f is integrable, and $\int_A f = \sum_S \int_S f\upharpoonright S$. □

► EXERCISE 76 (3-5). *Let $f, g\colon A \to \mathbb{R}$ be integrable and suppose $f \leqslant g$. Show that $\int_A f \leqslant \int_A g$.*

PROOF. Since f is integrable, the function $-f$ is integrable by Exercise 74 (c); then $g - f$ is integrable by Exercise 74 (b). It is easy to see $\int_A (g - f) \geqslant 0$ since $g \geqslant f$. It follows from Exercise 74 that $\int_A (g - f) = \int_A \big(g + (-f)\big) = \int_A g + \int_A (-f) = \int_A g - \int_A f$; hence, $\int_A f \leqslant \int_A g$. □

► EXERCISE 77 (3-6). *If $f\colon A \to \mathbb{R}$ is integrable, show that $|f|$ is integrable and $\left|\int_A f\right| \leqslant \int_A |f|$.*

PROOF. Let $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$. Then

$$f = f^+ - f^- \quad \text{and} \quad |f| = f^+ + f^-.$$

It is evident that for any partition P of A, both $U\left(f^+, P\right) - L\left(f^+, P\right) \leqslant U\left(f, P\right) - L\left(f, P\right)$ and $U\left(f^-, P\right) - L\left(f^-, P\right) \leqslant U\left(f, P\right) - L\left(f, P\right)$; hence, both f^+ and f^- are integrable if f is. Further,

$$\begin{aligned}\left|\int_A f\right| = \left|\int_A \left(f^+ - f^-\right)\right| = \left|\int_A f^+ - \int_A f^-\right| &\leqslant \int_A f^+ + \int_A f^- \\ &= \int_A \left(f^+ + f^-\right) \\ &= \int_A |f|.\end{aligned}$$

□

► EXERCISE 78 (3-7). *Let $f\colon [0, 1] \times [0, 1] \to \mathbb{R}$ be defined by*

$$f(x,y) = \begin{cases} 0 & x \text{ irrational} \\ 0 & x \text{ rational, } y \text{ irrational} \\ 1/q & x \text{ rational, } y = p/q \text{ is lowest terms.} \end{cases}$$

Show that f is integrable and $\int_{[0,1]\times[0,1]} f = 0$.

PROOF. □

3.2 Measure Zero and Content Zero

► EXERCISE 79 (3-8). *Prove that $[a_1, b_1] \times \cdots \times [a_n, b_n]$ does not have content 0 if $a_i < b_i$ for each i.*

PROOF. Similar to the $[a,b]$ case. □

► EXERCISE 80 (3-9).

a. *Show that an unbounded set cannot have content 0.*

b. *Give an example of a closed set of measure 0 which does not have content 0.*

PROOF.

(a) Finite union of bounded sets is bounded.

(b) $\mathbb{Z}$ or $\mathbb{N}$. □

► EXERCISE 81 (3-10).

a. *If C is a set of content 0, show that the boundary of C has content 0.*

b. *Give an example of a bounded set C of measure 0 such that the boundary of C does not have measure 0.*

PROOF. □

3.3 Fubini's Theorem

► EXERCISE 82 (3-27). *If $f: [a,b] \times [a,b] \to \mathbb{R}$ is continuous, show that*

$$\int_a^b \int_a^y f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \int_x^b f(x,y)\,\mathrm{d}y\,\mathrm{d}x.$$

PROOF. As illustrated in Figure 3.3,

$$\begin{aligned} C &= \left\{ (x, y) \in [a, b]^2 : a \leqslant x \leqslant y \text{ and } a \leqslant y \leqslant b \right\} \\ &= \left\{ (x, y) \in [a, b]^2 : a \leqslant x \leqslant b \text{ and } x \leqslant y \leqslant b \right\}. \end{aligned}$$

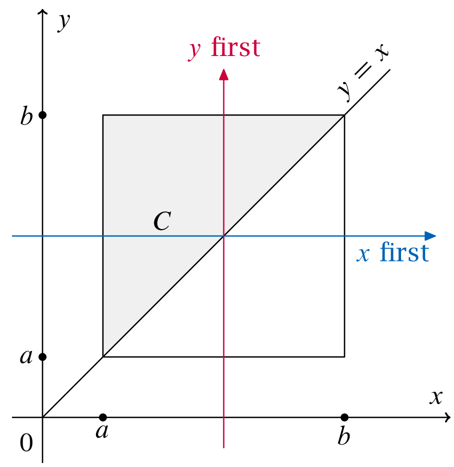

FIGURE 3.3. Fubini's Theorem

□

► EXERCISE 83 (3-30). *Let C be the set in Exercise 17. Show that*

$$\int_{[0,1]} \left(\int_{[0,1]} \mathbb{1}_C (x, y) \, \mathrm{d}x \right) \mathrm{d}y = \int_{[0,1]} \left(\int_{[0,1]} \mathbb{1}_C (x, y) \, \mathrm{d}y \right) \mathrm{d}x = 0.$$

PROOF. There must be typos. □

► EXERCISE 84 (3-31). *If $A = [a_1, b_1] \times \cdots \times [a_n, b_n]$ and $f : A \to \mathbb{R}$ is continuous, define $F : A \to \mathbb{R}$ by*

$$F(\boldsymbol{x}) = \int_{[a_1, x^1] \times \cdots \times [a_n, x^n]} f.$$

What is $\mathbb{D}_i F(\boldsymbol{x})$, for $x \in \operatorname{int}(A)$?

SOLUTION. Let $\boldsymbol{c} \in \operatorname{int}(A)$. Then

$$\begin{aligned}
\mathbb{D}_i F(\boldsymbol{c}) &= \lim_{h\to 0} \frac{F\left(\boldsymbol{c}^{-i}, c^i + h\right) - F(\boldsymbol{c})}{h} \\
&= \lim_{h\to 0} \frac{\int_{[a_1,c^1]\times\cdots\times[a_i,c^i+h]\times\cdots\times[a_n,c^n]} f - F(\boldsymbol{c})}{h} \\
&= \lim_{h\to 0} \frac{\int_{a_i}^{c^i+h} \left(\int_{[a_1,c^1]\times\cdots\times[a_{i-1},x^{i-1}]\times[a_{i+1},c^{i+1}]\times\cdots\times[a_n,c^n]} f\right) \mathrm{d}x_i - F(\boldsymbol{c})}{h} \\
&= \lim_{h\to 0} \frac{\int_{c^i}^{c^i+h} \left(\int_{[a_1,c^1]\times\cdots\times[a_{i-1},c^{i-1}]\times[a_{i+1},c^{i+1}]\times\cdots\times[a_n,c^n]} f\right) \mathrm{d}x_i}{h} \\
&= \int_{[a_1,c^1]\times\cdots\times[a_{i-1},c^{i-1}]\times[a_{i+1},c^{i+1}]\times\cdots\times[a_n,c^n]} f\left(\boldsymbol{x}^{-i}, c^i\right). \qquad \square
\end{aligned}$$

► EXERCISE 85 (3-32*). *Let* $f\colon [a,b]\times[c,d] \to \mathbb{R}$ *be continuous and suppose* $\mathbb{D}_2 f$ *is continuous. Define* $F(y) = \int_a^b f(x,y)\,\mathrm{d}x$. *Prove* Leibnitz's rule*:* $F'(y) = \int_a^b D_2 f(x,y)\,\mathrm{d}x$.

PROOF. We have

$$\begin{aligned}
F'(y) &= \lim_{h\to 0} \frac{F(y+h) - F(y)}{h} \\
&= \lim_{h\to 0} \frac{\int_a^b f(x,y+h)\,\mathrm{d}x - \int_a^b f(x,y)\,\mathrm{d}x}{h} \\
&= \lim_{h\to 0} \int_a^b \frac{f(x,y+h) - f(x,y)}{h}\,\mathrm{d}x.
\end{aligned}$$

By DCT, we have

$$\begin{aligned}
F'(y) &= \int_a^b \left[\lim_{h\to 0} \frac{f(x,y+h) - f(x,y)}{h}\right] \mathrm{d}x \\
&= \int_a^b \mathbb{D}_2 f(x,y)\,\mathrm{d}x. \qquad \square
\end{aligned}$$

► EXERCISE 86 (3-33). *If* $f\colon [a,b]\times[c,d] \to \mathbb{R}$ *is continuous and* $\mathbb{D}_2 f$ *is continuous, define* $F(x,y) = \int_a^x f(t,y)\,\mathrm{d}t$.

a. *Find* $\mathbb{D}_1 F$ *and* $\mathbb{D}_2 F$.

b. *If* $G(x) = \int_a^{g(x)} f(t,x)\,\mathrm{d}t$, *find* $G'(x)$.

SOLUTION.

(a) $\mathbb{D}_1 F(x,y) = f(x,y)$, and $\mathbb{D}_2 F = \int_a^x \mathbb{D}_2 f(t,y)\,\mathrm{d}t$.

(b) It follows that $G(x) = F(g(x), x)$. Then

$$\begin{aligned}
G'(x) &= g'(x)\,\mathbb{D}_1 F(g(x), x) + \mathbb{D}_2 F(g(x), x) \\
&= g'(x)\, f(g(x), x) + \int_a^{g(x)} \mathbb{D}_2 f(t,x)\,\mathrm{d}t. \qquad \square
\end{aligned}$$

4

INTEGRATION ON CHAINS

4.1 Algebraic Preliminaries

► Exercise 87 (4-1*). *Let $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ be the usual basis of $\mathbb{R}^n$ and let $\varphi_1, \dots, \varphi_n$ be the dual basis.*

a. *Show that $\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(\boldsymbol{e}_{i_1}, \dots, \boldsymbol{e}_{i_k}) = 1$. What would the right side be if the factor $(k+\ell)!/k!\ell!$ did not appear in the definition of $\wedge$?*

b. *Show that $\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(v_1, \dots, v_k)$ is the determinant of the $k \times k$ minor of $\begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}$ obtained by selecting columns $i_1, \dots, i_k$.*

Proof.

(a) Since $\varphi_{i_j} \in \mathcal{T}(\mathbb{R}^n)$, for every $j = 1, \dots, k$, we have

$$\begin{aligned}
\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}(\boldsymbol{e}_{i_1}, \dots, \boldsymbol{e}_{i_k}) &= \frac{k!}{1! \cdots 1!} \operatorname{Alt}\left(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k}\right)(\boldsymbol{e}_{i_1}, \dots, \boldsymbol{e}_{i_k}) \\
&= \sum_{\sigma \in S_k} (\operatorname{sgn}(\sigma)) \varphi_{i_1}(\boldsymbol{e}_{\sigma(i_1)}) \cdots \varphi_{i_k}(\boldsymbol{e}_{\sigma(i_k)}) \\
&= 1.
\end{aligned}$$

If the factor $(k+\ell)!/k!\ell!$ did not appear in the definition of $\wedge$, then the solution would be $1/k!$.

(b) □

► Exercise 88 (4-9*). *Deduce the following properties of the cross product in $\mathbb{R}^3$.*

a.
$$\begin{array}{lll}
e_1 \times e_1 = 0 & e_2 \times e_1 = -e_3 & e_3 \times e_1 = e_2 \\
e_1 \times e_2 = e_3 & e_2 \times e_2 = 0 & e_3 \times e_2 = -e_1 \\
e_1 \times e_3 = -e_2 & e_2 \times e_3 = e_1 & e_3 \times e_3 = 0
\end{array}$$

Proof.

(a) We just do the first line.

$$\langle w, z\rangle = \begin{Vmatrix} e_1 \\ e_1 \\ w \end{Vmatrix} = 0 \Longrightarrow z = e_1 \times e_1 = \mathbf{0},$$

$$\langle w, z\rangle = \begin{Vmatrix} e_2 \\ e_1 \\ w \end{Vmatrix} = -w_3 \Longrightarrow e_2 \times e_1 = -e_3,$$

$$\langle w, z\rangle = \begin{Vmatrix} e_3 \\ e_1 \\ w \end{Vmatrix} = w_2 \Longrightarrow e_3 \times e_1 = e_2.$$

□

References

[1] Apostol, Tom M. (1974) *Mathematical Analysis*: Pearson Education, 2nd edition. [37]

[2] Axler, Sheldon (1997) *Linear Algebra Done Right*, Undergraduate Texts in Mathematics, New York: Springer-Verlag, 2nd edition. []

[3] Berkovitz, Leonard D. (2002) *Convexity and Optimization in* $\mathbb{R}^n$, Pure and Applied Mathematics: A Wiley-Interscience Series of Texts, Monographs and Tracts, New York: Wiley-Interscience. [15]

[4] Munkres, James R. (1991) *Analysis on Manifolds*, Boulder, Colorado: Westview Press. [40]

[5] Rudin, Walter (1976) *Principles of Mathematical Analysis*, New York: McGraw-Hill Companies, Inc. 3rd edition. [17, 38]

[6] Spivak, Michael (1965) *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus*, Boulder, Colorado: Westview Press. [i]

Index