Spivak Solutions

Max Chien

February 2024

Contents

Introduction

This document contains exercises to selected solutions in the book *Calculus on Manifolds*, by Michael Spivak. These solutions were written as part of personal self-study during Summer 2024.

Chapter 1

Solutions to Selected Exercises

1.1 Chapter 1 Exercises

Exercise 1-1 Prove that $|x| \leq \sum_{i=1}^n |x_i|$ for any $x \in \mathbb{R}^n$.

Proof. Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ be arbitrary. For each $1 \leq i \leq n$, let us denote by (x_i) the vector $\begin{bmatrix} 0 & \ldots & x_i & \ldots & 0 \end{bmatrix}^T$, with the x_i term in the ith coordinate. Then for each i, we have the following:

$$|(x_i)| = \sqrt{(x_i)^2} = |x_i|$$

Moreover, by construction we have $x = (x_1) + \ldots + (x_n)$. By repeated application of the triangle inequality, we have

$$|x| = |\sum_{i=1}^n (x_i)| \leq \sum_{i=1}^n |(x_i)| = \sum_{i=1}^n |x_i| \qquad \square$$

Exercise 1-2 When does equality hold for the triangle inequality?

I claim that $|x + y| = |x| + |y|$ if and only if $y = \lambda x$ for some $\lambda \geq 0$, or $x = \mathbf{0}$.

Proof. $x = \mathbf{0}$ clearly satisfies the triangle inequality, so assume $x \neq \mathbf{0}$. Following the proof of the triangle inequality given by Spivak, we already see that x, y being linearly dependent is certainly a necessary condition. Thus, assume that $y = \lambda x$ for some $\lambda \in \mathbb{R}$. Then

$$\begin{aligned}
|x+y|^2 &= \sum_{i=1}^{n}(x_i+y_i)^2 \\
&= \sum_{i=1}^{n}(x_i+\lambda x_i)^2 \\
&= \sum_{i=1}^{n} x_i^2 + \lambda^2 x_i^2 + 2\lambda \sum_{i=1}^{n} x_i^2 \\
&= |x|^2 + |\lambda x|^2 + 2\lambda |x|^2
\end{aligned}$$

When $\lambda \geq 0$ we have

$$|x|^2 + |\lambda x|^2 + 2\lambda|x| = |x|^2 + 2|x||\lambda x| + |\lambda x|^2 = (|x| + |\lambda x|)^2 = (|x| + |y|)^2$$

where equality follows by taking the square root on both sides.

When $\lambda < 0$ this becomes

$$|x|^2 + |\lambda x|^2 + 2\lambda|x| = |x|^2 + -2|x||\lambda x| + |\lambda x|^2 = (|x| - |\lambda x|)^2 = (|x| - |y|)^2$$

By taking square roots on both sides, we have $|x+y| = |x| - |y| \neq |x| + |y|$ where the inequality holds since $x \neq \mathbf{0}$, $\lambda \neq 0$ means that $|y| \neq 0$. Thus $y = \lambda x$ for $\lambda \geq 0$, or $x = \mathbf{0}$ is a necessary and sufficient condition. □

Exercise 1-3 Prove that $|x - y| \leq |x| + |y|$ for any $x, y \in \mathbb{R}^n$.

Proof. Let $x, y \in \mathbb{R}^n$ be arbitrary. Then

$$|x - y| = |x + (-1 * y)| \leq |x| + |(-1) * y| = |x| + |-1||y| = |x| + |y|$$ □

Exercise 1-4 Prove that $||x| - |y|| \leq |x - y|$.

Proof. We expand:

$$\begin{aligned}
|x-y|^2 &= \sum_{i=1}^{n}(x_i-y_i)^2 \\
&= \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} y_i^2 - 2\sum_{i=1}^{n} x_i y_i \\
&\geq |x|^2 + |y|^2 - 2|x||y| \text{(by Cauchy-Schwarz)} \\
&= (|x| - |y|)^2
\end{aligned}$$

Taking square roots on both sides (using the fact that it is order-preserving), we get

$$|x - y| \geq ||x| - |y||$$ □

Exercise 1-5 The quantity $|y-x|$ is called the **distance** between x and y. Prove and interpret geometrically the inequality $|z-x| \le |z-y| + |y-x|$.

Proof. Noting that $|z-x| = |(z-y)+(y-x)|$, this is a simple application of the triangle inequality. This says that the sum of the lengths of any two sides of a triangle must be greater than the length of the third. □

Exercise 1-6 Let f, g be integrable on $[a, b]$.

(a) Prove that $|\int_a^b fg| \le (\int_a^b f^2)^{\frac{1}{2}} (\int_a^b g^2)^{\frac{1}{2}}$.

(b) If equality holds, must it be true that $f = \lambda g$ for some $\lambda \in \mathbb{R}$? What if f, g are required to be continuous?

(c) Show that the Cauchy-Schwarz inequality is a special case of (a).

(a) *Proof.* We consider the cases $0 = \int_a^b (f-\lambda g)^2$ for some $\lambda \in \mathbb{R}$, and $0 < \int_a^b (f-\lambda g)^2$ for all λ.

Case 1: Here, we have

$$0 = \int_a^b (f-\lambda g)^2 = \int_a^b f^2 - 2\lambda fg + \lambda^2 g^2 = \int_a^b f^2 - 2\lambda \int_a^b fg + \lambda^2 \int_a^b g^2$$

if $\lambda = 0$, then f (and thus fg) is zero on a set of measure 1, immediately making both sides of the inequality 0. Thus assume that $\lambda \neq 0$, which implies

$$\int_a^b fg = \frac{1}{2\lambda} \int_a^b f^2 + \frac{\lambda}{2} \int_a^b g^2$$

so

$$\begin{aligned}
\left(\int_a^b fg\right)^2 &= \left(\frac{1}{2\lambda}\int_a^b f^2\right)^2 + \left(\frac{\lambda}{2}\int_a^b g^2\right)^2 + \frac{1}{2}\left(\int_a^b f^2\right)\left(\int_a^b g^2\right) \\
&\le \frac{1}{2}\left(\int_a^b f^2\right)\left(\int_a^b g^2\right) \\
&\le \left(\int_a^b f^2\right)\left(\int_a^b g^2\right)
\end{aligned}$$

Taking the square root on both sides gives $|\int_a^b fg| \le (\int_a^b f^2)^{\frac{1}{2}} (\int_a^b g^2)^{\frac{1}{2}}$, as desired.

Case 2: Here, we have

$$\int_a^b (f-g)^2 > 0 \implies \int_a^b f^2 + \int_a^b g^2 > 2\int_a^b fg$$

Squaring both sides,

$$\left(\int_a^b fg\right)^2 < \left(\frac{1}{2}\int_a^b f^2\right) + \left(\frac{1}{2}\int_a^b g^2\right) + \frac{1}{2}\left(\int_a^b f^2\right)\left(\int_a^b g^2\right)$$

and the rest of the proof is identical to the first case. □

(b) *Proof.* Examining the proof of part (a), we must have $0 = \int_a^b (f - \lambda g)^2$ for equality to hold. This implies $f - \lambda g$ is 0 almost everywhere, so $f = \lambda g$ almost everywhere. However, it may not be the case that $f = \lambda g$ everywhere (consider $f = 0$ and $g = 0$ except at countably many points). When f, g are required to be continuous, then they cannot differ on a set of measure zero, so equality implies $f = \lambda g$ for some $\lambda \in \mathbb{R}$. □

(c) *Proof.* Let $x, y \in \mathbb{R}^n$ be arbitrary. Define $f : [0, n) \to \mathbb{R}$ such that $f = x_i$ on the interval $[i-1, i)$ and define g similarly for y. Then

$$\int_0^n f^2 = \sum_{i=1}^n x_i^2 = |x|^2, \int_0^n g^2 = \sum_{i=1}^n y_i^2 = |y|^2, \int_0^n fg = \sum_{i=1}^n x_i y_i$$

Then by part a,

$$\left|\sum_{i=1}^n x_i y_i\right| = \left|\int_0^n fg\right| \leq \left(\int_a^b f^2\right)^{\frac{1}{2}} \left(\int_a^b g^2\right)^{\frac{1}{2}} = |x||y|$$

□

Exercise 1-7 A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is **norm preserving** if $|T(x)| = |x|$ for all $x \in \mathbb{R}^n$, and **inner product preserving** if $\langle Tx, Ty\rangle = \langle x, y\rangle$ for all $x, y \in \mathbb{R}^n$.

(a) Prove that T is norm preserving if and only if T is inner product preserving.

(b) Prove that such a linear transformation T is one-to-one and T^{-1} is of the same sort.

(a) *Proof.* ($\Longrightarrow$) Suppose T is norm preserving. Then for any $x, y \in \mathbb{R}^n$, we use bilinearity of the inner product:

$$\begin{aligned}
\langle Tx, Ty\rangle &= \langle Tx - Ty + Ty, Ty\rangle \\
&= \langle Tx - Ty, Ty\rangle + \langle Ty, Ty\rangle \\
&= \langle Tx - Ty, Ty - Tx + Tx\rangle + |Ty| \\
&= \langle Tx - Ty, Ty - Tx\rangle + \langle Tx - Ty, Tx\rangle + |Ty| \\
&= |Tx| - \langle Ty, Tx\rangle + |Ty| - |Tx - Ty|
\end{aligned}$$

which gives

$$\begin{aligned}
\langle Tx, Ty\rangle &= \frac{1}{2}\left(|Tx| + |Ty| - |T(x-y)|\right) && \text{(by linearity of } T\text{)} \\
&= \frac{1}{2}(|x| + |y| - |x - y|) && \text{(by norm preserving)} \\
&= \langle x, y\rangle
\end{aligned}$$

where the last line follows through a similar calculation as the first part.

$(\Longleftarrow)$ Suppose T is inner product preserving. Then for any $x \in \mathbb{R}^n$,

$$|Tx| = \langle Tx, Tx\rangle = \langle x, x\rangle = |x|$$

where the second equality follows since T preserves inner products. □

(b) *Proof.* Suppose T is inner product/norm preserving. Suppose $Tx = Ty$. Since T is linear, we have $T(x-y) = 0$. So $|T(x-y)| = 0$. But T is norm preserving, so $|x-y| = 0$, which occurs only when $x - y = 0$, showing that $x = y$. So T is one-to-one.

Let T^{-1} denote the inverse of T (which exists since T is an injective endomorphism on finite dimensional vector spaces). Then let $x \in \mathbb{R}^n$ be arbitrary. Since T is norm preserving, we have

$$|T^{-1}x| = |TT^{-1}x| = |x|$$

so T^{-1} is norm preserving as well. □

Exercise 1-8 If $x, y \in \mathbb{R}^n$ are nonzero, then the angle between x and y is denoted $\angle(x, y)$, which is defined as $\arccos\left(\frac{\langle x,y\rangle}{|x|\cdot|y|}\right)$. This is well-defined since $\left|\frac{\langle x,y\rangle}{|x|\cdot|y|}\right| \leq 1$ by Cauchy-Schwarz. The linear transformation T is angle preserving if T is one-to-one and for any $x, y \neq 0$ we have $\angle(Tx, Ty) = \angle(x, y)$.

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

(b) If there is a basis $x_1, \ldots, x_n$ of $\mathbb{R}^n$ and numbers $\lambda_1, \ldots, \lambda_n$ such that $Tx_i = \lambda_i x_i$, prove that T is angle preserving only if all $|\lambda_i|$ are equal. (**Note:** Spivak's original exercise has an if and only if here, but this is false.)

(c) What are all angle preserving $T : \mathbb{R}^n \to \mathbb{R}^n$?

(a) *Proof.* Since T is both norm preserving and inner product preserving by Exercise 1-7, we have

$$\frac{\langle Tx, Ty\rangle}{|Tx|\cdot|Ty|} = \frac{\langle x, y\rangle}{|x|\cdot|y|}$$

so

$$\angle(Tx, Ty) = \arccos\left(\frac{\langle Tx, Ty\rangle}{|Tx|\cdot|Ty|}\right) = \arccos\left(\frac{\langle x, y\rangle}{|x|\cdot|y|}\right) = \angle(x, y)$$ □

(b) *Proof.* Proof by contrapositive. Suppose $|\lambda_i| \neq |\lambda_j|$ for some $i \neq j$. Then consider the vectors

$$v_1 = x_i + x_j, v_2 = x_i - \frac{|x_i|}{|x_j|}x_j$$

Since x_i, x_j are linearly independent, neither v_1 or v_2 is the zero vector. Then we have

$$\begin{aligned}
\cos\angle(v_1, v_2) &= \cos\arccos\left(\frac{\left\langle x_i + x_j, x_i - \frac{|x_i|}{|x_j|}x_j\right\rangle}{|x_i + x_j||x_i - \frac{|x_i|}{|x_j|}x_j|}\right) \\
&= \frac{|x_i|^2 - \frac{|x_i|^2}{|x_j|^2}|x_j|^2}{|x_i + x_j||x_i - \frac{|x_i|}{|x_j|}x_j|} \\
&= 0
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\cos\angle(T(v_1), T(v_2)) &= \cos\angle(\lambda_i x_i + \lambda_j x_j, \lambda_i x_i - \lambda_j \frac{|x_i|}{|x_j|}x_j) \\
&= \frac{\lambda_i^2|x_i|^2 - \lambda_j^2|x_i|^2}{|\lambda_i x_i + \lambda_j x_j||\lambda_i x_i - \lambda_j \frac{|x_i|}{|x_j|}x_j|} \neq 0
\end{aligned}$$

where the last inequality holds since $|\lambda_i| \neq |\lambda_j| \implies \lambda_i^2 \neq \lambda_j^2$. So if $|\lambda_i| \neq |\lambda_j|$, then T is not angle preserving. So T is angle preserving only if $|\lambda_i| = |\lambda_j|$ for all i, j. □

(c) Intuitively, the answer is that T must consist of only rotation and scaling by a constant factor. More rigorously, the singular values of T must all be $\sigma_1 = \ldots = \sigma_n = k$ for some $k > 0$. We do not provide a full proof here.

Exercise 1-9 If $0 \leq \theta < \pi$, then let $T : \mathbb{R}^2 \to \mathbb{R}^2$ have the matrix in the standard basis given by

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Show that T is angle preserving, and that for any $x \neq 0$, $\angle(x, Tx) = \theta$.

Proof. To show that T is one-to-one, we instead prove that T is invertible. Consider the matrix

$$T' = \begin{bmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{bmatrix}$$

Then

$$\begin{aligned}
TT' &= \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \cos\theta\sin\theta \\ -\cos\theta\sin\theta + \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}$$

Since T is square and $TT' = I$, we have $T'T = I$ so T is invertible and thus must be one-to-one.

Let $x, y \neq 0 \in \mathbb{R}^2$ be arbitrary. Suppose $x = (x_1, x_2)$, $y = (y_1, y_2)$. Then

$$\cos \angle(x, y) = \frac{x_1 y_1 + x_2 y_2}{\sqrt{x_1^2 + x_2^2}\sqrt{y_1^2 + y_2^2}}$$

Moreover, $Tx = (x_1 \cos\theta + x_2 \sin\theta, x_2 \cos\theta - x_1 \sin\theta)$ and $Ty = (y_1 \cos\theta + y_2 \sin\theta, y_2 \cos\theta - y_1 \sin\theta$. Then we have

$$\begin{aligned}
\langle Tx, Ty\rangle &= (x_1 \cos\theta + x_2 \sin\theta)(y_1 \cos\theta + y_2 \sin\theta) + (x_2 \cos\theta - x_1 \sin\theta)(y_2 \cos\theta - y_1 \sin\theta) \\
&= x_1 y_1 \underbrace{(\cos^2\theta + \sin^2\theta)}_{=1} + x_1 y_2 \underbrace{(\cos\theta \sin\theta - \sin\theta\cos\theta)}_{=0} \\
&\qquad + \underbrace{x_2 y_1 (\sin\theta\cos\theta - \sin\theta\cos\theta)}_{=0} + x_2 y_2 \underbrace{(\sin^2\theta + \cos^2\theta)}_{=1} \\
&= x_1 y_1 + x_2 y_2 = \langle x, y\rangle
\end{aligned}$$

and

$$\begin{aligned}
|Tx| &= \sqrt{(x_1 \cos\theta + x_2 \sin\theta)^2 + (x_2 \cos\theta - x_1 \sin\theta)^2} \\
&= \sqrt{x_1^2(\cos^2\theta + \sin^2\theta) + x_2^2(\sin^2\theta + \cos^2\theta)} \\
&= \sqrt{x_1^2 + x_2^2} \\
&= |x|
\end{aligned}$$

Similarly,

$$|Ty| = |y|$$

Then

$$\begin{aligned}
\angle(Tx, Ty) &= \arccos\left(\frac{\langle Tx, Ty\rangle}{|Tx||Ty|}\right) \\
&= \arccos\left(\frac{\langle x, y\rangle}{|x||y|}\right) \\
&= \angle(x, y)
\end{aligned}$$

Lastly, using the fact that $|x| = |Tx|$,

$$\begin{aligned}
\angle(x, Tx) &= \arccos\left(\frac{\langle x, Tx\rangle}{|x||Tx|}\right) \\
&= \arccos\left(\frac{x_1^2 \cos\theta + x_1 x_2 \sin\theta + x_2^2 \cos\theta - x_1 x_2 \sin\theta}{|x|^2}\right) \\
&= \arccos\left(\cos\theta \frac{x_1^2 + x_2^2}{|x|^2}\right) \\
&= \arccos\cos\theta \\
&= \theta
\end{aligned}$$

□

Exercise 1-10 If $T : \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation, show that there is a number M such that $|T(h)| \leq M|h|$ for $h \in \mathbb{R}^m$.

Proof. By singular value decomposition, there are orthonormal bases $\mathcal{B} = \{u_1, \ldots, u_m\} \subseteq \mathbb{R}^m$ and $\mathcal{C} = \{v_1, \ldots, v_n\} \subseteq \mathbb{R}^n$ as well as scalars $\sigma_1 \geq \ldots \geq \sigma_m \geq 0$ such that $Tu_i = \sigma_i v_i$ for all i (with $Tu_j = 0$ for any $j \geq n$). Then for any $h \in \mathbb{R}^m$, if we suppose that $h = a_1u_1 + \ldots + a_mu_m$, then we have

$$\begin{aligned}|Th| &= |T(a_1u_1 + \ldots + a_mu_m)|\\ &= |a_1Tu_1 + \ldots + a_mTu_m|\\ &= |a_1\sigma_1v_1 + \ldots + a_m\sigma_mv_m|\end{aligned}$$

(where the indices only run to n if $n < m$). Now since $\mathcal{C}$ is orthonormal, the Pythagorean identity gives

$$|a_1\sigma_1v_1 + \ldots + a_m\sigma_mv_m|^2 = a_1^2\sigma_1^2 + \ldots + a_m^2\sigma_m^2 \leq (\sigma_m)^2(a_1^2 + \ldots + a_m^2)$$

But since $\mathcal{B}$ is also orthonormal, we have $(a_1^2 + \ldots + a_m^2) = |h|^2$. So

$$|Th|^2 \leq \sigma_1^2|h|^2 \implies |Th| \leq \sigma_1|h|$$

so our choice of $M = \sigma_1$ works. $\square$

Exercise 1-11 If $x, y \in \mathbb{R}^n$ and $z, w \in \mathbb{R}^m$, show that $\langle(x, z),(y, w)\rangle = \langle x, y\rangle + \langle z, w\rangle$, and that $|(x, z)| = \sqrt{|x|^2 + |z|^2}$. Recall that $(x, z) \in \mathbb{R}^{n+m}$ is the concatenation of x and z.

Proof. For the first statement,

$$\begin{aligned}\langle(x, z)(y, w)\rangle &= \sum_{i=1}^{n+m}(x, z)_i(y, w)_i\\ &= \sum_{i=1}^{n}(x, z)_i(y, w)_i + \sum_{j=1}^{m}(x, z)_{n+j}(y, w)_{n+j}\\ &= \sum_{1=1}^{n}x_iy_i + \sum_{j=1}^{m}z_jw_j\\ &= \langle x, y\rangle + \langle z, w\rangle\end{aligned}$$

For the second statement,

$$|(x, z)|^2 = \langle(x, z),(x, z)\rangle = \langle x, x\rangle + \langle z, z\rangle = |x|^2 + |z|^2$$

where the second equality is by the first statement. Taking square roots on both sides recovers $|(x, z)| = \sqrt{|x|^2 + |z|^2}$. $\square$

Exercise 1-12 Let $(\mathbb{R}^n)^*$ denote the dual space of $\mathbb{R}^n$, which is the space of all linear functions $f : \mathbb{R}^n \to \mathbb{R}$. If $x \in \mathbb{R}^n$, then define $\phi_x \in (\mathbb{R}^n)^*$ such that $\phi_x(y) \coloneqq \langle x, y \rangle$. Define $T : \mathbb{R}^n \to (\mathbb{R}^n)^*$ such that $T(x) = \phi_x$. Show that T is one-to-one and conclude that each $\phi \in (\mathbb{R}^n)^*$ is ϕ_x for a unique $x \in \mathbb{R}^n$.

Proof. Suppose $\phi_x = \phi_y$. Then $\langle x, z \rangle = \langle y, z \rangle$ for all $z \in \mathbb{R}^n$. Choosing $z = x - y$, this gives

$$0 = \langle x, z \rangle - \langle y, z \rangle = \langle x - y, z \rangle = \langle x - y, x - y \rangle = |x - y|$$

which implies that $|x - y|$ is the zero vector. So $x = y$. The rest of the proof follows since $\dim \mathbb{R}^n = \dim(\mathbb{R}^n)^*$, so T is injective between vector spaces of the same dimension and is thus surjective and bijective. ☐

Exercise 1-13 (Pythagorean Identity) If $x, y \in \mathbb{R}^n$, then x and y are called orthogonal if $\langle x, y \rangle = 0$. If x and y are orthogonal, prove that $|x + y|^2 = |x|^2 + |y|^2$.

Proof. By the definition of the norm and bilinearity of the inner product,

$$\begin{aligned} |x + y|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle y, y \rangle + 2 \underbrace{\langle x, y \rangle}_{=0} \\ &= |x|^2 + |y|^2 \end{aligned}$$

☐

Exercise 1-14 Prove that the arbitrary union of open sets is open. Prove that the finite intersection of open sets is open. Show that an infinite union of open sets need not be open.

Proof. Let $U = \bigcup_{i \in I} U_i$ be the union of some open sets over an arbitrary indexing set I. Then for any $x \in U$, $x \in U_i$ for some i. Then $x \in B \subseteq U_i$ for some open rectangle B. Since $B \subseteq U_i$, $B \subseteq U$, so $x \in B \subseteq U$. So U is open.

Let $U = U_1 \cap U_2$ for some open sets U_1, U_2. Let $x \in U$ be arbitrary. Then $x \in B_{r_1}(x) \subseteq U_1$ and $x \in B_{r_2}(x) \subseteq U_2$ for some radii r_1, r_2. Taking $r = \min\{r_1, r_2\} > 0$, we have $x \in B_r(x) \subseteq B_{r_1} \subseteq U_1$ and $B_r(x) \subseteq B_{r_2} \subseteq U_2$, so $x \in B_r(x) \subseteq U$. By induction, this extends to any finite intersection.

The intersection of the sets $(-1/n, 1/n)$ for $n \in \mathbb{N}$ is the singleton $\{0\}$, which is not open. ☐

Exercise 1-15 Prove that the open ball $B_r(a) \coloneqq \{x \in \mathbb{R}^n : |x - a| < r\}$ is indeed open.

Proof. When $r = 0$, $B_r(a) = \varnothing$ which is vacuously open. If $r > 0$, then pick some $x \in B_r(a)$. Let $r' = r - |x - a|$. Then if $x = (x_1, \ldots, x_n)$, consider the box B with sides

$(x_1 - r'/n, x_1 + r'/n) \times \ldots \times (x_n - r'/n, x_n + r'/n)$. For any other $y \in B$, we have $|x_i - y_i| \leq r'/n$ by construction, so

$$|x - y| \leq |x_1 - y_1| + \ldots + |x_n - y_n| \leq r'$$

By the triangle inequality,

$$|y - a| = |y - x - (a - x)| \leq |y - x| + |a - x| \leq r' + |a - x| = r - |x - a| + |x - a| = r$$

So $y \in B_r(a)$, and thus $B \subseteq B_r(a)$. So $B_r(a)$ is open. □

> **Exercise 1-16** Find the interior, exterior, and boundary of the following sets:
>
> 1. $A := \{x \in \mathbb{R}^n : |x| \leq 1\}$
> 2. $B := \{x \in \mathbb{R}^n : |x| = 1\}$
> 3. $C := \{x \in \mathbb{R}^n : \text{ each coordinate } x_i \in \mathbb{Q}\}$

1. We proved in Exercise 1-15 that $B_1(\mathbf{0}) \subseteq A$ is open. So $B_1(\mathbf{0}) \subseteq \operatorname{int} A$.

   I claim that $\mathbb{R}^n \setminus A = \operatorname{ext} A$. Let $x \in \mathbb{R}^n \setminus A$. Then take the open ball $B_{|x|-1}(x)$. For any $y \in B_{|x|-1}(x)$, the reverse triangle inequality tells us

   $$|y| \geq ||y - x| - |x||$$

   Since $y \in B_{|x|-1}(x)$, $|y - x| \leq |x| - 1$. So $|y - x| - |x| \leq -1$, and thus

   $$|y| \geq ||y - x| - |x|| \geq 1$$

   so $y \in \mathbb{R}^n \setminus A$. Thus $\mathbb{R}^n \setminus A \subseteq \operatorname{ext} A$, but $\operatorname{ext} A \subseteq \mathbb{R}^n \setminus A$ (this is easy to see based on the definition of $\operatorname{ext} A$), so $\mathbb{R}^n \setminus A = \operatorname{ext} A$.

   Lastly, for any x with $|x| = 1$, pick any open ball $B_r(x)$. Then the point $y = x + \frac{r}{2}x$ has

   $$|y - x| = \left|\frac{r}{2}x\right| = \frac{r}{2}\underbrace{|x|}_{=1} < r$$

   So $y \in B_r(x)$. Moreover,

   $$|y| = \left(1 + \frac{r}{2}\right)\underbrace{|x|}_{=1} > 1$$

   so $y \in \mathbb{R}^n \setminus A$. On the other hand, a similar calculation shows that $z = x - \frac{r}{2}x \in B_r(x)$ is in A. So the set of points with $|x| = 1$ is a subset of ∂A. But $\operatorname{int} A \sqcup \partial A \sqcup \operatorname{ext} A = \mathbb{R}^n$, and we have already partitioned $\mathbb{R}^n$, so our subsets must be equalities and we must have $\operatorname{int} A = \{x : |x| < 1\}$, $\partial A = \{x : |x| = 1\}$, $\operatorname{ext} A = \{x : |x| > 1$.

2. By the same argument as before, the set of $|x| > 1$ is a subset of $\operatorname{ext} B$. By a similar argument, the set of $|x| < 1$ is also a subset of $\operatorname{ext} B$. Lastly, the same argument shows that B itself is not a subset of $\operatorname{int} B$. But B cannot be in $\operatorname{ext} B$, so we must have $\operatorname{int} B = \varnothing$, $\partial B = \{x : |x| = 1\}$, $\operatorname{ext} B = \{x : |x| \neq 1\}$.

3. Let $x \in \mathbb{R}^n$ be arbitrary. Then let $D = (y_1, z_1) \times \ldots \times (y_n, z_n)$ be an arbitrary open rectangle containing x. By the density of $\mathbb{Q}$ in $\mathbb{R}$, we can pick rational numbers $q_i \in (y_i, z_i)$. Then the point $q = (q_1, \ldots, q_n) \in C$ and $q \in D$, so D contains points of C. Similarly, we can construct a point with all irrational coordinates $p = (p_1, \ldots, p_n) \notin C$ and $p \in D$, so D contains points of $\mathbb{R}^n \setminus C$. Thus $x \in \partial C$. x was arbitrary, so $\partial C = \mathbb{R}^n$ and $\operatorname{int} C = \operatorname{ext} C = \varnothing$.

Exercise 1-17 Construct a set $A \subseteq [0,1] \times [0,1]$ such that A contains at most one point on each horizontal and each vertical line but has $\operatorname{ext} A = [0,1] \times [0,1]$.

We construct sets recursively as follows: for A_1, pick a point in each quadrant of $[0,1] \times [0,1]$, such that none lie on the same horizontal or vertical line. For A_2, pick a point in each sixteenth of $[0,1] \times [0,1]$ such that none lie on the same horizontal or vertical line, and none lie on the same horizontal or vertical line as the points in A_1. Continue doing this, picking 4^i points for A_i such that no point $x \in A_i$ shares a vertical or horizontal line with a point $y \in \bigcup_{k=1}^{i} A_k$. This is possible because each choice of point removes only a single vertical line and horizontal line from our possible choices, which is a set of measure zero, so we always have a set of measure one to choose from. Then take our set to be $A = \bigcup_{i=1}^{\infty} A_i$. By construction, this set satisfies the vertical/horizontal line property. This set has no interior, since a nontrivial open rectangle being a subset of A would violate the vertical/horizontal line condition. Moreover, for any point $x \in [0,1] \times [0,1]$ and any radius r, we simply look in a (4^i)-ant of length $r/2$ or less in order to find a point y that is close to x. So $\partial A = [0,1] \times [0,1]$.

Exercise 1-18 If $A \subseteq [0,1]$ is the union of open intervals (a_i, b_i) such that any rational number in $(0,1)$ is in (a_i, b_i) for some i, prove that $\partial A = [0,1] \setminus A$.

Proof. Since A is the union of open intervals, A is open and thus $\operatorname{int} A = A$. I claim that $\operatorname{ext} A = \mathbb{R}^n \setminus [0,1]$. Clearly $\mathbb{R}^n \setminus [0,1] \subseteq \operatorname{ext} A$. Then take some point $x \in [0,1]$. For any open interval (a,b) containing x, the density of $\mathbb{Q}$ tells us that there is a rational number in $(a,b) \cap [0,1]$, so $x \notin \operatorname{ext} A$. So $\operatorname{ext} = \mathbb{R}^n \setminus [0,1]$, $\operatorname{int} A = A$, and this forces $\partial A = [0,1] \setminus A$. $\square$

Exercise 1-19 If A is a closed set that contains every rational $r \in [0,1]$, show that $[0,1] \subseteq A$.

Proof. Suppose not. Then there is some $x \in [0,1]$ with $x \in \mathbb{R}^n \setminus A$. x must be in $(0,1)$, which is open. Moreover, $x \in \mathbb{R}^n \setminus A$, which is open since A is closed, so $x \in (\mathbb{R}^n \setminus A) \cap (0,1)$ which is open (since the finite intersection of open sets is open). Take some open interval $x \in (a,b) \subseteq (\mathbb{R}^n \setminus A) \cap (0,1)$. By the density of $\mathbb{Q}$, there is a rational r in (a,b). But $r \in A$ by definition, so $(a,b) \not\subseteq \mathbb{R}^n \setminus A$, so $\mathbb{R}^n \setminus A$ isn't open, which contradicts the assumption that A is closed. So we must have $[0,1] \subseteq A$. $\square$

Exercise 1-20 Prove that a compact subset of $\mathbb{R}^n$ is closed and bounded.

Proof. Suppose $K \subseteq \mathbb{R}^n$ is compact. The collection of open rectangles $(i-1, i+1) \times (j-1, j+1) \ldots \times (k-1, k+1)$ for $i, j, \ldots, k \in \mathbb{Z}$ covers $\mathbb{R}$, so it covers K. Then a finite number of these boxes covers K, so it is bounded.

We wish to show that $\mathbb{R}^n \setminus K$ is open. Suppose it is not. Then there is some $x \in \mathbb{R}^n \setminus K$ such that for all open balls $B_r(x)$, $B_r(x) \cap K \neq \varnothing$. We can construct a sequence of points $y_1, y_2, \ldots \in K$ as follows: Pick some r_1, say $r_1 = 1$. Then $B_{r_1}(x)$ contains some point $y_1 \in K$. Let $r_2 = |y_1 - x|$ (note this is strictly less than r_1 since $y_1 \in B_{r_1}(x) \implies |y_1 - x| < r_1$). Next, $B_{r_2}(x)$ contains some other point $y_2 \in K$, and $|y_2 - x| < r_2 = |y_1 - x|$. Continue this to construct a sequence of points $y_1, y_2, \ldots \in K$ such that $|y_1 - x| > |y_2 - x| > \ldots$.

We use this sequence to create an open cover of K. Let $r_i = |y_i - x|$. Let C be the closed ball with radius r_2 and center x. The set $\mathbb{R}^n \setminus C$ is open, since its complement C is closed. Now let $R_i \coloneqq \{y : r_{i+2} < |y - x| < r_i\}$ be the open ring with outer radius r_i and inner radius r_{i+2}. Then $\bigcup R_i = \{y : |y - x| < r_1\} = B_{r_1}(x)$ contains all points with distance $|y - x| < r_1$. $\mathbb{R}^n \setminus C$ contains all points with distance $|y - x| > r_2$. But $r_2 < r_1$, so $\mathbb{R}^n \setminus C \cup \bigcup R_i = \mathbb{R}^n$.

Thus the collection $\mathcal{O} = \{\mathbb{R}^n \setminus C, R_1, R_2, \ldots\}$ covers $\mathbb{R}^n$ and thus K. But if we pick only a finite number of these, then there is some R_i in the finite subcover such that i is maximal in the subcover, so the points $y_{i+2}, y_{i+3}, \ldots$ are not contained in the subcover, and thus K is not compact. So if K is compact, then it is closed. □

Exercise 1-21

1. If A is closed and $x \notin A$ prove that there is a number $d > 0$ such that $y - x \geq d$ for all $y \in A$.
2. If A is closed, B is compact, and $A \cap B = \varnothing$, prove that there is $d > 0$ such that $|y - x| \geq d$ for all $y \in A$ and $x \in B$.
3. Give a counterexample in $\mathbb{R}^2$ if A and B are closed but neither is compact.

1. *Proof.* Since A is closed, $\mathbb{R}^n \setminus A$ is open. Let $x \notin A$. Then $x \in \mathbb{R}^n \setminus A$, so there is an open ball $B_r(x) \subseteq \mathbb{R}^n \setminus A$. Then we have $|x - y| < r \implies y \in B_r(x) \subseteq \mathbb{R}^n \setminus A \implies y \notin A$, and thus for any $y \in A$ we must have $|x - y| \geq r$. □

2. *Proof.* For each point $b \in B$, part (a) tells us there is a distance r_b such that $|b - y| \geq r_b$ for any $y \in A$. Consider the collection of open balls $(B_{r_b/2}(b))_{b \in B}$. This collection covers B, so we pick a finite subcover $\{B_{r_{b_1}/2}(b_1), B_{r_{b_2}/2}(b_2), \ldots, B_{r_{b_n}/2}(b_n)\}$. For any $x \in B$, $x \in B_{r_{b_i}/2}(b_i)$ for some i. Then by the reverse triangle inequality, for any $y \in A$, we have

   $$|y - x| = |y - b_i - (x - b_i)| \geq ||y - b_i| - |x - b_i||$$

   Since $y \in A$, $|y - b_i| \geq r_{b_i}$. Since $x \in B_{r_{b_i}/2}(b_i)$, $|x - b_i| \leq r_{b_i}/2 \leq r_{b_i} \leq |y - b_i|$. So the quantity $|y - b_i| - |x - b_i|$ is positive, so

   $$|y - x| \geq |y - b_i| - |x - b_i| \geq r_{b_i} - \frac{r_{b_i}}{2} = \frac{r_{b_i}}{2} \geq \frac{\min_{1 \leq i \leq n} r_{b_i}}{2}$$

Since $r_{b_i} \geq 0$ for all i and there are finite i, $\min r_{b_i}$ is well defined and positive. Thus for arbitrary $y \in A$, $x \in B$, we have $|y - x| \geq \min r_{b_i}/2 = d > 0$. □

3. We define two sets as follows: first, pick $A = \mathbb{N}$. Next, pick $B = \{x_1, x_2, \ldots\}$, where $x_i = i + \frac{1}{i+1}$. Since x_i is never an integer, $A \cap B = \varnothing$. However, let $r > 0$ be arbitrary. Then pick i large enough that $\frac{1}{i+1} < r$. Choosing $x = x_i$, $y = i$, we have

$$|x - y| = |x_i - i| = \left|\frac{1}{i+1}\right| = \frac{1}{i+1} < r$$

Exercise 1-22 If U is open and $C \subseteq U$ is compact, show that there is a compact set D such that $C \subseteq \operatorname{int} D$ and $D \subseteq U$.

Proof. Since U is open, $\mathbb{R}^n \setminus U$ is closed. Thus by Exercise 1-21 part (b), there is a distance d such that $|y - x| < d$ for any $x \in C$ and $y \in \mathbb{R}^n \setminus U$. Let $B_x = B_d(x)$ be the open ball of radius d and center x. Let $\overline{B_x} = \overline{B_d(x)} = \{y : |y - x| \leq d\}$ be the closed ball of radius d and center x.

The collection $(B_x)_{x \in C}$ is an open cover of C compact, so we pick a finite subcollection $B_{x_1}, \ldots, B_{x_n}$. Then let $D = \overline{B_{x_1}} \cup \ldots \cup \overline{B_{x_n}}$. We have $\overline{B_{x_i}} \supseteq B_{x_i}$ for all i, so

$$D = \bigcup_{i=1}^{n} \overline{B_{x_i}} \supseteq \bigcup_{i=1}^{n} B_{x_i} \supseteq C$$

so $C \subseteq D$. Moreover, for any point $y \in \mathbb{R}^n \setminus U$ and $x \in D$, $x \in B_{x_i}$ for some i. Then $|x - x_i| \leq d/2$, and $|y - x_i| \geq d$, so

$$|y - x| \geq ||y - x_i| - |x - x_i|| \geq d - \frac{d}{2} = \frac{d}{2} > 0$$

so $D \cap \mathbb{R}^n \setminus U = \varnothing$ and thus $D \subseteq U$. □

Exercise 1-23 If $f : A \to \mathbb{R}^m$ and $a \in A$, show that $\lim_{x \to a} f(x) = b = (b_1, \ldots, b_m)$ if and only if $\lim_{x \to a} f^i(x) = b_i$ for each i (recall f^i is the ith component function).

Proof. ($\implies$) Suppose $\lim_{x \to a} f(x) = b$. Then for any $\varepsilon > 0$, there is $\delta > 0$ such that $|x - a| < \delta$ and $x \in A$ implies $|f(x) - b| < \varepsilon$. Then for any such x, we have $|f^i(x) - b_i|^2 \leq \sum_{j=1}^{m} |f^j(x) - b_j|^2 = |f(x) - b|^2 < \varepsilon^2$ so $|f^i(x) - b_i| < \varepsilon$. So $\lim_{x \to a} f^i(x) = b_i$.

Suppose $\lim_{x \to a} f^i(x) = b_i$ for each i. Then for any $\varepsilon > 0$, pick $\delta_i > 0$ for each i such that $|x - a| < \delta_i \implies |f^i(x) - b_i| < \varepsilon/\sqrt{m}$. Let $\delta = \min \delta_i$. Then for any x with $|x - a| < \delta$,

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

so $|f(x) - b| < \varepsilon$ and thus $\lim_{x \to a} f(x) = b$. □

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

Proof. Immediate from Exercise 1-23. □

Exercise 1-25 Prove that a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is continuous.

Proof. From Exercise 1-10, we know that there exists $M > 0$ such that $|T(h)| \leq M|h|$ for all h. Then at any point $a \in \mathbb{R}^n$, let $\varepsilon > 0$ be arbitrary. Set $\delta = \varepsilon/M$. Then for any $x \in \mathbb{R}^n$ with $|x - a| < \delta$, we have

$$|T(x) - T(a)| = |T(x - a)| \leq M|x - a| < M\frac{\varepsilon}{M} = \varepsilon$$

□

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

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

(b) Define $f : \mathbb{R}^2 \to \mathbb{R}$ by $f(x) = 0$ if $x \notin A$ and $f(x) = 1$ if $x \in A$. For $h \in \mathbb{R}^2$ define $g_h : \mathbb{R} \to \mathbb{R}$ by $g_h(t) = f(th)$. Show that each g_h is continuous at 0, but f is not continuous at $(0, 0)$. (This problem shows that f is continuous in any direction, but not continuous as a two-variable function).

(a) *Proof.* Suppose $y = mx$ defines a straight line through $(0, 0)$. When $m = 0$ one can verify that that the entire line is in $\mathbb{R}^n \setminus A$ since $y = 0$. (For a vertical line we similarly have $x = 0$ so the line is in $\mathbb{R}^n \setminus A$). Then consider the interval $[-|m|, |m|]$. The portion of the line with $x \leq 0$ is automatically in $\mathbb{R}^2 \setminus A$, but for any $x \in (0, |m|]$,

$$x^2 \leq |m|x = y$$

so the entire interval $[-|m|, |m|]$ is in $\mathbb{R}^2 \setminus A$. □

(b) *Proof.* Pick some g_h. By part (a), there is an interval about 0 such that $th \in \mathbb{R}^2 \setminus A$, so $g_h(t) = 0$. So $g_h(t) = 0$ on an interval about 0, so $\lim_{t \to 0} g_h(t) = 0 = g_h(0)$. Thus each g_h is continuous at 0.

To show f is not continuous at 0, pick $\varepsilon = 1/2$. Let $\delta > 0$ be arbitrary. Assume $\delta < 1$ since this will automatically prove larger δ. Then the point $x = (\delta/2, \delta^2/5)$ is in A, so $f(x) = 1$. Moreover,

$$|x - 0| = |x| \leq \frac{\delta}{2} + \frac{\delta^2}{5} \leq \frac{\delta}{2} + \frac{\delta}{5} < \delta$$

But $|f(x) - f(0)| = |1| = 1 > \varepsilon$, so f is not continuous at 0. □

Exercise 1-27 Prove that $\{x \in \mathbb{R}^n : |x - a| < r\}$ is open using the topological condition.

Proof. Consider the function $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) = |x - a|$. To prove f is continuous, pick some point y. Then let $\varepsilon > 0$ and set $\delta = \varepsilon$. Then we have

$$|x - y| < \delta \implies |f(x) - f(y)| = ||x - a| - |y - a|| \leq |x - a - (y - a)| = |x - y| < \delta = \varepsilon$$

so f is continuous. Thus the preimage of the open ball $B_r(0)$ under f is open, but this is precisely the set $\{x \in \mathbb{R}^n : f(x) = |x - a| < r\}$.t □

> **Exercise 1-28** Suppose $A \subseteq \mathbb{R}^n$ is not closed. Show that there exists an unbounded continuous function $f : A \to \mathbb{R}$.

Proof. Let $A \subseteq \mathbb{R}^n$ be not closed. Then $\mathbb{R}^n \setminus A$ is not open, so there exists a point $x \in \mathbb{R}^n \setminus A$ such that every $B_r(x)$ contains a point in A. Then define $f : A \to \mathbb{R}$ by

$$f(y) = \frac{1}{|y - x|}$$

To verify that this function is continuous, first consider the function $|y - x|$. Letting $a \in A$ be arbitrary, for any $\varepsilon > 0$ set $\delta = \varepsilon$. Then for any $b \in A$ with $|b - a| < \delta$,

$$|f(b) - f(a)| = ||b - x| - |a - x|| \leq |b - a| < \delta = \varepsilon$$

So $y \mapsto |y - x|$ is continuous. Then since $|y - x| \neq 0$ for $y \in A$, and f is the quotient of nonzero continuous functions, f is continuous.

To show that f is unbounded, pick $M > 0$. Then by our choice of x, the ball $B_{1/M}(x)$ contains a point $y \in A$. Then

$$f(y) = \frac{1}{|y - x|} \geq \frac{1}{\frac{1}{M}} = M$$ □

> **Exercise 1-29** Let $K \subseteq \mathbb{R}^n$ be compact, and let $f : K \to \mathbb{R}$ be continuous. Show that f attains a maximum and minimum value.

Proof. Since K is compact and f is continuous, $f(K)$ is compact. Specifically, it is bounded, so let $\alpha = \sup f(K)$. We want to show $\alpha \in f(K)$. By way of contradiction, suppose $\alpha \notin f(K)$. Then since $f(K)$ is closed, $\mathbb{R} \setminus f(K)$ is open, so there is an interval $(\alpha - \varepsilon, \alpha + \varepsilon)$ that doesn't intersect $f(K)$. But then $\alpha - \varepsilon$ is also an upper bound for $f(K)$, contradicting that fact that $\alpha = \sup f(K)$. So we must have $\sup f(K) = \max f(K) \in f(K)$, and thus there is a $y \in K$ such that $f(y) = \max f(K)$. The proof for the minimum is similar. □

> **Exercise 1-30** Let $f : [a, b] \to \mathbb{R}$ be increasing. Let $x_1, \ldots, x_n \in [a, b]$ be distinct. Show that
>
> $$\sum_{i=1}^{n} o(f, x_i) \leq f(b) - f(a)$$

Proof. Note that since f is increasing, for any $[c,d] \subseteq [a,b]$, we have

$$\max_{[c,d]} f(x) = f(d), \min_{[c,d]} f(x) = f(c)$$

In particular, $M(f,x,\delta) = f(x+\delta)$ and $m(f,x,\delta) = f(x-\delta)$, so $f(x+\delta) - f(x-\delta) \geq o(f,x)$.

We may suppose that the x_i are ordered, so that $x_1 < \ldots < x_n$. Pick δ small enough that $|x_{i+1} - x_i| < \delta$ for all δ. This gives us disjoint intervals $[x_1 - \delta, x_1 + \delta], \ldots, [x_n - \delta, x_n + \delta]$. Then we have

$$\begin{aligned}
\sum_{i=1}^{n} o(f,x_i) &\leq \sum_{i=1}^{n} f(x_i + \delta) - f(x_i - \delta) \\
&= f(x_n + \delta) - f(x_n - \delta) + \ldots + f(x_1 + \delta) - f(x_1 - \delta) \\
&\leq \underbrace{f(b) - f(x_n + \delta)}_{\geq 0} + f(x_n + \delta) - f(x_n - \delta) + \underbrace{f(x_n - \delta) - f(x_{n-1} + \delta)}_{\geq 0} \\
&\quad + f(x_{n-1} + \delta) - \ldots - f(x_1 - \delta) + \underbrace{f(x_1 - \delta) - f(a)}_{\geq 0} \\
&= f(b) - f(a)
\end{aligned}$$

The first and last intervals may be adjusted slightly for the case where $x_1 = a$ or $x_n = b$. □

1.2 Chapter 2 Exercises

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

Proof. Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$. Then $Df(a)$ is linear transformation. By Exercise 1-10, there exists a number $M > 0$ such that

$$\frac{|Df(a)(h)|}{|h|} \geq M, \quad \forall h \in \mathbb{R}^n$$

Then since f is differentiable at a, there exists $\delta > 0$ such that for any $|h| < \delta$,

$$\frac{|f(a+h) - f(a) - Df(a)(h)|}{|h|} < 1$$

Now let $\varepsilon > 0$ be arbitrary, and pick $\delta' = \min\left\{\delta, \frac{\varepsilon}{M+1}\right\}$. Then for any x with $|x - a| < \delta'$ we have

$$\begin{aligned}
|f(x) - f(a)| &= |f(x) - f(a) - Df(a)(x-a) + Df(a)(x-a)| \\
&\leq |f(x) - f(a) - Df(a)(x-a)| + |Df(a)(x-a)| \\
&< |x-a| + M|x-a| \\
&< (M+1)\frac{\varepsilon}{M+1} = \varepsilon
\end{aligned}$$

□

Exercise 2-2 A function $f : \mathbb{R}^2 \to \mathbb{R}$ is **independent of the second variable** if for any $x \in \mathbb{R}$ and $y_1, y_2 \in \mathbb{R}$ we have $f(x, y_1) = f(x, y_2)$. 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. ($\Longrightarrow$) Suppose f is independent of the second variable. Then define $g(x) = f(x, 0)$. For any x, y we have

$$f(x, y) = f(x, 0) = g(x)$$

($\Longleftarrow$) Suppose $g(x) = f(x, y)$. Then let $x, y_1, y_2 \in \mathbb{R}$ be arbitrary. We have

$$f(x, y_1) = g(x) = f(x, y_2)$$

□

Claim: $f'(a, b) = \begin{bmatrix} g'(a) & 0 \end{bmatrix}$.

Proof. Fix $(a, b) \in \mathbb{R}^2$. Then let $\varepsilon > 0$. Since g is differentiable at a, there exists $\delta > 0$ such that for any $|h| < \delta$,

$$\frac{|g(a + h) - g(a) - g'(a)(h)|}{|h|} < \varepsilon$$

Then if $h = (h_1, h_2)$ satisfies $|h| < \delta$, it must also be the case that $|h_1| \leq |(h_1, h_2)| < \delta$. Thus for any $|(h_1, h_2)| = |h| < \delta$, we have

$$\begin{aligned} \frac{\left| f(a + h_1, b + h_2) - f(a, b) - \begin{bmatrix} g'(a) & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \right|}{|h|} &= \frac{|g(a + h_1) - g(a) - g'(a)(h_1)|}{|h|} \\ &\leq \frac{|g(a + h_1) - g(a) - g'(a)(h_1)|}{|h_1|} < \varepsilon \end{aligned}$$

Thus we have $f'(a, b) = \begin{bmatrix} g'(a) & 0 \end{bmatrix}$. □

Exercise 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 both the first and second variables?

A function $f; \mathbb{R}^2 \to \mathbb{R}$ is independent of the first variable if for any $x_1, x_2, y \in \mathbb{R}$ we have $f(x_1, y) = f(x_2, y)$, or equivalently if there exists $h : \mathbb{R} \to \mathbb{R}$ such that $f(x, y) = h(y)$. In this case, $f'(a, b) = \begin{bmatrix} 0 & h'(b) \end{bmatrix}$. If a function is independent of both variables, then $f(a_1, b_1) = f(a_2, b_1) = f(a_2, b_2)$ for any $(a_1, b_1), (a_2, b_2) \in \mathbb{R}^2$ so f is constant.

Exercise 2-4 Let g be a continuous real-valued function on the unit circle such that $g(0,1) = g(1,0) = 0$ and $g(-x) = -g(x)$. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x) = \begin{cases} |x| g\left(\frac{x}{|x|}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

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

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

(a) *Proof.* If $x = 0$ then h is identically 0 and is differentiable. If $x \neq 0$, then for $t \neq 0$ we have

$$h(t) = |tx| g\left(\frac{tx}{|tx|}\right) = \underbrace{|t||x| g\left(\text{sign}(t) \frac{x}{|x|}\right) = |t| \text{sign}(t) |x| g\left(\frac{x}{|x|}\right)}_{g(-x) = -g(x)} = t\left[|x| g\left(\frac{x}{|x|}\right)\right]$$

We also have $h(0) = f(0) = 0 = 0|x| g\left(\frac{x}{|x|}\right)$ so h is a linear function of t. Thus it is differentiable from single-variable analysis. □

(b) *Proof.* Suppose that f can be differentiated. Then since $Df(0,0)$ is linear, it is uniquely determined by its behavior on the basis $\{e_1, e_2\}$. In particular, pick $\varepsilon > 0$. Then there exists a $\delta > 0$ such that whenever $0 < |h| < \delta$ we have

$$\frac{|f(h) - f(0,0) - Df(0,0)(h)|}{|h|} < \varepsilon$$

Then picking some $h_1 \in \mathbb{R}$ with $0 < |h_1| < \delta$,

$$|Df(0,0)(e_1)| = \frac{|h_1 Df(0,0)(e_1)|}{|h_1|} = \frac{|\underbrace{f(h_1 e_1) - f(0,0)}_{=0} - Df(0,0)(h_1 e_1)|}{|h_1 e_1|} < \varepsilon$$

This works for all epsilon, so $Df(0,0)(e_1) = 0$. Similarly, $Df(0,0)(e_2) = 0$, so $Df(0,0)$ is the zero transformation. Now suppose $g(x) \neq 0$ for some x. Then for $\varepsilon = g(x)$ and arbitrary, δ,

$$\frac{|f\left(\frac{\delta x}{2}\right) - \underbrace{f(0,0) - Df(0,0)\left(\frac{\delta x}{2}\right)}_{=0}|}{\left|\frac{\delta x}{2}\right|} = \frac{\frac{\delta}{2} g\left(\frac{\delta x/2}{\delta/2}\right)}{\frac{\delta}{2}} = g(x) \geq \varepsilon$$

so f is not differentiable. Thus f is only differentiable when $g(x) = 0$ everywhere. □

Exercise 2-5 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} \frac{x|y|}{\sqrt{x^2+y^2}}, & (x,y) \neq 0 \\ 0 & (x,y) = 0 \end{cases}$$

Show that f is a function of the kind considered in Exercise 2-4, so that f is not differentiable at $(0,0)$.

Proof. Let

$$g(x,y) = \begin{cases} \frac{x|y|}{x^2+y^2}, & (x,y) \neq 0 \\ 0, & (x,y) = 0 \end{cases}$$

Then for $(x,y) \neq 0$ we have

$$\begin{aligned} |(x,y)|g\left(\frac{(x,y)}{|(x,y)|}\right) &= \sqrt{x^2+y^2}g\left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right) \\ &= \sqrt{x^2+y^2}\frac{x|y|}{x^2+y^2} \\ &= \frac{x|y|}{\sqrt{x^2+y^2}} \\ &= f(x,y) \end{aligned}$$

Moreover,

$$g(1,0) = \frac{0}{\sqrt{1}} = 0 = \frac{|0|}{\sqrt{1}} = g(0,1)$$

and

$$g(-x,-y) = \frac{-x|-y|}{(-x)^2+(-y)^2} = -\frac{x|y|}{x^2+y^2} = -g(x,y)$$

so f is of the form in Exercise 2-4. However,

$$g\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2} \neq 0$$

so g is not 0 everywhere and thus f is not differentiable at $(0,0)$. $\square$

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

Proof. Following the proof of Exercise 2-4 part (a), first suppose f is differentiable at $(0,0)$. Then $Df(0,0)$ exists, and it is determined by its behavior on the basis $\{e_1, e_2\}$. Letting $\varepsilon > 0$ be arbitrary, there must exist $\delta > 0$ such that for any $0 < |h| < \delta$,

$$\frac{|f(h) - f(0,0) - Df(0,0)(h)|}{|h|} < \varepsilon$$

Pick some $h_1 \in \mathbb{R}$ with $0 < |h_1| < \delta$. Then

$$|Df(0,0)(e_1)| = \frac{|h_1 Df(0,0)(e_1)|}{|h_1|} = \frac{|f(h_1e_1) - f(0,0) - Df(0,0)(h_1e_1)|}{|h_1e_1|} < \varepsilon$$

So $|Df(0,0)(e_1)| < \varepsilon$ for all ε, and thus $Df(0,0)(e_1) = 0$. Similarly, $Df(0,0)(e_2) = 0$, so $Df(0,0)$ is the zero transformation. However, let $\varepsilon = \frac{1}{\sqrt{2}}$, and let $\delta > 0$ be arbitrary. Then the point $(x,y) = (\frac{\delta}{\sqrt{3}}, \frac{\delta}{\sqrt{3}})$ satisfies $0 < |(x,y)| < \delta$, but

$$\frac{|f(x,y) - f(0,0) - Df(0,0)(x,y)|}{|(x,y)|} = \frac{\sqrt{\frac{\delta^2}{3}}}{\sqrt{\frac{2\delta^2}{3}}} = \frac{1}{\sqrt{2}} \geq \varepsilon$$

so no δ works and f is not differentiable. □

Exercise 2-7 Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function such that $|f(x)| \leq |x|^2$. Show that f is differentiable at $\mathbf{0}$.

Proof. Let $\lambda : \mathbb{R}^n \to \mathbb{R}$ be the zero transformation. Then let $\varepsilon > 0$ be arbitrary, and set $\delta = \varepsilon$. Whenever $0 < |x| < \delta$, by assumption we have

$$\frac{|f(x)|}{|x|} \leq |x|$$

In particular, $|f(0)| \leq |0|^2 = 0$ so $f(0) = 0$. Thus

$$\frac{|f(x) - f(0) - \lambda(x)|}{|x|} = \frac{|f(x)|}{|x|} \leq |x| < \delta = \varepsilon$$

so f is differentiable at 0 with derivative $Df(0) = \lambda$ the zero transformation. □

Exercise 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

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

Proof. ($\implies$) Suppose f is differentiable at $a \in \mathbb{R}$. Then let $\varepsilon > 0$ be arbitrary. Since f is differentiable, there exists $\delta > 0$ such that whenever $0 < |h| < \delta$ we have

$$\frac{|f(a+h) - f(a) - Df(a)(h)|}{|h|} < \varepsilon$$

If we suppose that $Df(a)(h)$ has matrix representation given by

$$f'(a) = \begin{bmatrix} b \\ c \end{bmatrix}$$

then it is the case that

$$Df(a)(h) = \begin{bmatrix} bh \\ ch \end{bmatrix}$$

Now if we write for convenience $(x, y) = f(a+h) - f(a) - Df(a)(h)$, then we know that $|x| \leq |(x, y)|$, so whenever $0 < |h| < \delta$

$$\frac{|f^1(a+h) - f(a) - bh|}{|h|} = \frac{|x|}{|h|} \leq \frac{|(x, y)|}{|h|} = \frac{|f(a+h) - f(a) - Df(a)(h)|}{|h|} < \varepsilon$$

so f^1 is differentiable at a. The proof for f^2 is similar. Moreover, this proves that in this case $bh = (f^1)'(a)$ and $ch = (f^2)'(a)$, so that

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

($\Longleftarrow$) Now suppose that f^1 and f^2 are differentiable at a. Let $\varepsilon > 0$ be arbitrary. Then there exist $\delta_1, \delta_2 > 0$ such that whenever $0 < |h| < \delta_1$ we have

$$\frac{|f^1(a+h) - f^1(a) + (f^1)'(a)(h)|}{|h|} < \frac{\varepsilon}{2}$$

and whenever $0 < |h| < \delta_2$ we have

$$\frac{|f^2(a+h) - f^2(a) + (f^2)'(a)(h)|}{|h|} < \frac{\varepsilon}{2}$$

Let $\delta = \min\{\delta_1, \delta_2\}$. Let $\lambda : \mathbb{R} \to \mathbb{R}^2$ have the matrix

$$[\lambda] = \begin{bmatrix} (f^1)'(a) \\ (f^2)'(a) \end{bmatrix}$$

Then whenever $0 < |h| < \delta$,

$$\frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = \frac{\left| \begin{bmatrix} x \\ y \end{bmatrix} \right|}{|h|}$$

where

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f^1(a+h) - f^1(a) - (f^1)'(a)(h) \\ f^2(a+h) - f^2(a) - (f^2)'(a)(h) \end{bmatrix}$$

Then

$$\frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} = \frac{\left| \begin{bmatrix} x \\ y \end{bmatrix} \right|}{|h|} \leq \frac{|x|}{|h|} + \frac{|y|}{|h|} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

so f is differentiable at a, and once again we have

$$f'(a) = [\lambda] = \begin{bmatrix} (f^1)'(a) \\ (f^2)'(a) \end{bmatrix}$$

□

Exercise 2-9 Two functions $f, g : \mathbb{R} \to \mathbb{R}$ are **equal up to nth order** at $a \in \mathbb{R}$ if

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

(a) Show that a continuous function 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. (**Note:** Spivak did not assume continuity in the original exercise, but it is required in the if direction, and continuity in the only if direction follows from differentiability).

(b) If $f'(a), \ldots, f^{(n)}(a)$ 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 nth order at a. (This is the nth degree Taylor polynomial of f expanded about a).

(a) *Proof.* ($\Longrightarrow$) Suppose f is differentiable at a. Then define

$$g(x) = f(a) + f'(a)(x-a)$$

We have

$$\lim_{h \to 0} \frac{f(a+h) - f(a+h)}{h} = \lim_{h \to 0} \frac{f(a+h) - f(a) - f'(a)(h)}{h} = 0$$

since f is differentiable, so f and g are equal up to first order.

($\Longleftarrow$) Now suppose $g(x) = a_0 + a_1(x-a)$ is equal to f up to first order. Since f (and g) are continuous,

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

so $f(a) = g(a) = a_0$. Thus we have

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

so f is differentiable at a. □

(b) *Proof.* We induct on n. Suppose that for any function f, whenever $f'(a), \ldots, f^{(n-1)}(a)$ exist, then

$$f(x) \overset{n-1}{\sim} \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i$$

where $\overset{n-1}{\sim}$ represents equality up to order $n-1$. Now suppose that $f'(a), \ldots, f^{(n)}(a)$ all exist. Then we have

$$\begin{aligned}
\lim_{h\to 0} \frac{f(a+h)-g(a+h)}{h^n} &= \lim_{h\to 0} \frac{f(a+h) - \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(a+h-a)^i}{h^n} \\
&= \lim_{h\to 0} \frac{f(a+h) - \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}h^i}{h^n}
\end{aligned}$$

Note that since f and g are continuous (where f is continuous since it is differentiable), we have

$$\lim_{h\to 0} f(a+h) - \sum_{i=0}^{n} \frac{f^{(i)}}{i!}h^i = f(a) - \frac{f^{(0)}(a)}{0!} - \sum_{i=1}^{n} \frac{f^{(i)}}{i!}0^i = f(a) - f(a) = 0$$

Clearly g is differentiable and so is f, so $f(a+h) - g(a+h)$ is differentiable, and thus L'Hopital's Rule applies. So

$$\begin{aligned}
\lim_{h\to 0} \frac{f(a+h) - \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}h^i}{h^n} &\overset{\text{LH}}{=} \lim_{h\to 0} \frac{f'(a+h) - \sum_{i=1}^{n} \frac{f^{(i)}(a)}{(i-1)!}h^{i-1}}{nh^{n-1}} \\
&= \lim_{h\to 0} \frac{f'(a+h) - \sum_{i=0}^{n-1} \frac{f^{(i+1)}(a)}{i!}h^i}{h^{n-1}} \\
&= \lim_{h\to 0} \frac{f'(a+h) - \sum_{i=0}^{n-1} \frac{(f')^{(i)}(a)}{i!}h^i}{h^{n-1}}
\end{aligned}$$

Since $f''(a), \ldots, f^{(n)}(a)$ all exist, $(f')'(a), \ldots, (f')^{(n-1)}(a)$ all exist, so the inductive hypothesis applies and

$$f'(x) \overset{n-1}{\sim} \sum_{i=0}^{n-1} \frac{(f')^{(i)}(a)}{i!}(x-a)^i$$

so

$$\lim_{h\to 0} \frac{f'(a+h) - \sum_{i=0}^{n-1} \frac{(f')^{(i)}(a)}{i!}h^i}{h^{n-1}} = 0$$

Thus

$$\lim_{h\to 0} \frac{f(a+h)-g(a+h)}{h^n} = \lim_{h\to 0} \frac{f'(a+h) - \sum_{i=0}^{n-1} \frac{(f')^{(i)}(a)}{i!}h^i}{h^{n-1}} = 0$$

so f and g are equal up to nth order. $\square$

Exercise 2-10 Use the theorems of this section [Section 2.2] 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) $f(x, y, z) = \sin(x \sin(y \sin z))$.

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

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

(a) We write
$$f = [\pi^1]^{[\pi^2]} = (e^{\ln \circ [\pi^1]})^{[\pi^2]} = e^{\pi^2 \cdot \ln \circ \pi^1}$$
. Then
$$\begin{aligned}
f'(a, b, c) &= (e^{\pi^2 \cdot \ln \circ \pi^1})'(a, b, c) \\
&= e^{b \ln a}(\pi^2 \cdot \ln \circ \pi^1)'(a, b, c) \\
&= a^b(\ln a(\pi^2)'(a, b, c) + b(\ln \circ \pi^1)'(a, b, c) \\
&= a^b(\ln a \pi^2 + b\frac{1}{a}(\pi^1)'(a, b, c) \\
&= a^b(\ln a \pi^2 + \frac{b}{a}\pi^1) \\
&= (ba^{b-1}, a^b \ln a, 0)
\end{aligned}$$

(b) Following easily from part (a) we have:
$$\begin{aligned}
f'(a, b, c) &= \begin{bmatrix} - & (x^y)'(a, b, c) & - \\ - & (\pi^3)'(a, b, c) & - \end{bmatrix} \\
&= \begin{bmatrix} ba^{b-1} & a^b \ln a & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

(c) Similarly to the example, we have $f = \sin\circ(\pi^1 \cdot \sin\circ\pi^2)$. Thus,

$$\begin{aligned}
f'(a,b) &= (\sin\circ(\pi^1 \cdot \sin\circ\pi^2))'(a,b) \\
&= \cos(a\sin b)(\pi^1 \cdot \sin\circ\pi^2)'(a,b) \\
&= \cos(a\sin b)(\sin b(\pi^1)'(a,b) + a(\sin\circ\pi^2)'(a,b)) \\
&= \cos(a\sin b)\sin b\pi^1 + a\cos(a\sin b)\cos b\pi^2 \\
&= (\cos(a\sin b)\sin b, a\cos(a\sin b)\cos b)
\end{aligned}$$

(d) As above, we have

$$f = \sin\circ(\pi^1 \cdot (\sin\circ(\pi^2 \cdot (\sin\circ\pi^3))))$$

so

$$\begin{aligned}
f'(a,b,c) &= (\sin\circ(\pi^1 \cdot (\sin\circ(\pi^2 \cdot (\sin\circ\pi^3)))))'(a,b,c) \\
&= \cos(a\sin(b\sin c))(\pi^1 \cdot (\sin\circ(\pi^2 \cdot (\sin\circ\pi^3))))'(a,b,c) \\
&= \cos(a\sin(b\sin c))(\sin(b\sin c)\pi^1 + a\cos(b\sin c)(\pi^2 \cdot (\sin\circ\pi^3))'(a,b,c)) \\
&= \cos(a\sin(b\sin c))(\sin(b\sin c)\pi^1 + a\cos(b\sin c)(\sin c\pi^2 + b\cos c\pi^3)) \\
&= \cos(a\sin(b\sin c)) * (\sin(b\sin c), a\cos(b\sin c)\sin c, ab\cos(b\sin c)\cos c)
\end{aligned}$$

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

$$f(x,y,z) = g(x, g(y,z))$$

so that

$$f = g\circ(\pi^1, g\circ(\pi^2,\pi^3))$$

Using our result from part (a),

$$\begin{aligned}
f'(a,b,c) &= g'(a, g(b,c)) \begin{bmatrix} - & (\pi^1)'(a,b,c) & - \\ - & (g\circ(\pi^2,\pi^3))'(a,b,c) & - \end{bmatrix} \\
&= \begin{bmatrix} b^c a^{b^c-1} & a^{b^c}\ln a \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & cb^{c-1} & b^c\ln b \end{bmatrix} \\
&= \begin{bmatrix} b^c a^{b^c-1} & a^{b^c} cb^{c-1}\ln a & a^{b^c} b^c \ln a \ln b \end{bmatrix}
\end{aligned}$$

(f) Letting g be as defined in part (e), we have

$$f = g\circ(\pi^1, \pi^2+\pi^3)$$

Thus

$$\begin{aligned}
f'(a,b,c) &= (g\circ(\pi^1,\pi^2+\pi^3))'(a,b,c) \\
&= g'(a, b+c) \begin{bmatrix} - & (\pi^1)'(a,b,c) & - \\ - & (\pi^2+\pi^3)'(a,b,c) & - \end{bmatrix} \\
&= \begin{bmatrix} (b+c)a^{b+c-1} & a^{b+c}\ln a \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} (b+c)a^{b+c-1} & a^{b+c}\ln a & a^{b+c}\ln a \end{bmatrix}
\end{aligned}$$

(g) Again letting g be as in part (e), we have

$$f = g \circ (\pi^1 + \pi^2, \pi^3)$$

so that

$$\begin{aligned} f'(a,b,c) &= (g \circ (\pi^1 + \pi^2, \pi^3))'(a,b,c) \\ &= g'(a+b,c) \begin{bmatrix} - & (\pi^1+\pi^2)'(a,b,c) & - \\ - & (\pi^3)'(a,b,c) & \end{bmatrix} \\ &= \begin{bmatrix} c(a+b)^{c-1} & (a+b)^c \ln(a+b) \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} c(a+b)^{c-1} & c(a+b)^{c-1} & (a+b)^c \ln(a+b) \end{bmatrix} \end{aligned}$$

(h) We can straightforwardly write this as

$$f = \sin \circ (\pi^1 \cdot \pi^2)$$

Then

$$\begin{aligned} f'(a,b) &= (\sin \circ (\pi^1 \cdot \pi^2))'(a,b) \\ &= \cos(ab)(b\pi^1 + a\pi^2) \\ &= (b\cos(ab), a\cos(ab)) \end{aligned}$$

(i) Using the same definition of g,

$$f = g \circ (\sin \circ (\pi^1 \cdot \pi^2), \cos 3)$$

Since $\cos 3$ is constant,

$$\begin{aligned} f'(a,b) &= (g \circ (\sin \circ (\pi^1 \cdot \pi^2), \cos 3))'(a,b) \\ &= g'(\sin(ab), \cos 3) \begin{bmatrix} - & (\pi^1 \cdot \pi^2)'(a,b) & - \\ - & (\cos 3)'(a,b) & - \end{bmatrix} \\ &= \begin{bmatrix} \cos 3[\sin(ab)]^{\cos 3} & [\sin(ab)]^{\cos 3} \ln \sin(ab) \end{bmatrix} \begin{bmatrix} b & a \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} b\cos 3[\sin(ab)]^{\cos 3} & a\cos 3[\sin(ab)]^{\cos 3} \end{bmatrix} \end{aligned}$$

(j) From parts (h), (c), and (a), respectively, we already know that

$$\begin{gathered} (\sin(xy))'(a,b) = \begin{bmatrix} b\cos(ab) & a\cos(ab) \end{bmatrix} \\ (\sin(x \sin y))'(a,b) = \begin{bmatrix} \cos(a \sin b)\sin b & a\cos(a\sin b)\cos b \end{bmatrix} \\ (x^y)'(a,b) = \begin{bmatrix} ba^{b-1} & a^b \ln a \end{bmatrix} \end{gathered}$$

Then f' is simply given by putting each of these matrices in as row vectors, such that

$$f'(a,b,c) = \begin{bmatrix} - & (\sin(xy))'(a,b) & - \\ - & (\sin(x\sin y))'(a,b) & - \\ - & (x^y)'(a,b) & - \end{bmatrix} = \begin{bmatrix} b\cos(ab) & a\cos(ab) \\ \cos(a\sin b)\sin b & a\cos(a\sin b)\cos b \\ ba^{b-1} & a^b \ln a \end{bmatrix}$$

Exercise 2-11 Find f' for the following (where $g : \mathbb{R} \to \mathbb{R}$ is continuous, and $s \in \mathbb{R}$ is fixed):

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

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

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

(a) Define $F : \mathbb{R} \to \mathbb{R}$ by

$$F(x) = \int_s^x g(t)\,\mathrm{d}t$$

Since g is continuous, the fundamental theorem of calculus tells us that

$$F'(x) = g(x)$$

Then we can here write f as

$$f = F \circ (\pi^1 + \pi^2)$$

so that

$$\begin{aligned} f'(a, b) &= (F \circ (\pi^1 + \pi^2))'(a, b) \\ &= F'(a + b)(\pi^1 + \pi^2)'(a, b) \\ &= g(a + b)(\pi^1 + \pi^2) \\ &= (g(a + b), g(a + b)) \end{aligned}$$

(b) Similarly, write

$$f = F \circ (\pi^1 \cdot \pi^2)$$

Then

$$\begin{aligned} f'(a, b) &= (F \circ (\pi^1 \cdot \pi^2))'(a, b) \\ &= (bg(ab), ag(ab)) \end{aligned}$$

(c) First note that we can pick any $s \in \mathbb{R}$ and separate this integral:

$$f(x, y, z) = \int_{x^y}^{\sin(x \sin(y \sin z))} g = \int_s^{\sin(x \sin(y \sin z))} g + \int_{x^y}^s g = \int_s^{\sin(x \sin(y \sin z))} g - \int_s^{x^y} g$$

Then using the same method as parts (a) and (b) of this problem, and using the results from parts (d) and (a) of Exercise 2-10, the Jacobian of the first term, evaluated at (a, b, c), is given by

$$\begin{bmatrix} g(\sin(a \sin(b \sin c))) \cos(a \sin(b \sin c)) \sin(b \sin c) \\ ag(\sin(a \sin(b \sin c))) \cos(a \sin(b \sin c)) \cos(b \sin c) \sin c \\ abg(\sin(a \sin(b \sin c))) \cos(a \sin(b \sin c)) \cos(b \sin c) \cos c \end{bmatrix}^T$$

and the Jacobian of the second by

$$\begin{bmatrix} g(a^b)ba^{b-1} \\ g(a^b)a^b \ln a \\ 0 \end{bmatrix}$$

Thus we have

$$f'(a,b,c) = \begin{bmatrix} g(\sin(a\sin(b\sin c)))\cos(a\sin(b\sin c))\sin(b\sin c) - g(a^b)ba^{b-1} \\ ag(\sin(a\sin(b\sin c)))\cos(a\sin(b\sin c))\cos(b\sin c)\sin c - g(a^b)a^b\ln a \\ abg(\sin(a\sin(b\sin c)))\cos(a\sin(b\sin c))\cos(b\sin c)\cos c \end{bmatrix}^T$$

Exercise 2-12 A function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ is **bilinear** if for $x, x_1, x_2 \in \mathbb{R}^n$, $y, y_1, y_2 \in \mathbb{R}^m$ and $a \in \mathbb{R}$ we have

$$\begin{aligned} f(ax,y) = af(x,y) &= f(x,ay) \\ f(x_1+x_2,y) &= f(x_1,y)+f(x_2,y) \\ f(x,y_1+y_2) &= f(x,y_1)+f(x,y_2) \end{aligned}$$

(a) Prove that if f is bilinear, then

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

(b) Prove that $Df(a,b)(x,y) = f(a,y) + f(x,b)$.

(c) Show that the formula for $Dp(a,b)$ in Section 2.2 is a special case of (b).

(a) *Proof.* Suppose f is bilinear, and suppose $h = (h_1,\ldots,h_n)$, $k = (k_1,\ldots,k_m)$. Then we can write

$$\begin{aligned} \lim_{(h,k)\to\mathbf{0}} \frac{|f(h,k)|}{|(h,k)|} &= \lim_{(h,k)\to\mathbf{0}} \frac{|f(\sum_{i=1}^n h_i, \sum_{j=1}^m k_j)}{|(h,k)|} \\ &= \lim_{(h,k)\to\mathbf{0}} \frac{|\sum_{i=1}^n \sum_{j=1}^m h_i k_j f(e_i,e_j)|}{|(h,k)|} \\ &\leq \sum_{i=1}^n \sum_{j=1}^m f(e_i,e_j) \lim_{(h_i,k_j)\to\mathbf{0}} \frac{|h_i k_j|}{|(h,k)|} \\ &\leq \sum_{i=1}^n \sum_{j=1}^m f(e_i,e_j) \lim_{(h_i,k_j)\to\mathbf{0}} \frac{|h_i k_j|}{|(h_i,k_j)|} \end{aligned}$$

Now we proved in the proof of $Dp(a,b)$ that

$$\lim_{(h_i,k_j)\to\mathbf{0}} \frac{|h_i k_j|}{|(h_i,k_j)|} = 0$$

so we have

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

(b) *Proof.* Note that

$$\begin{aligned}
f(a+x,b+y) - f(a,b) - f(a,y) - f(x,b) &= f(a+x,b+y) - f(a,b+y) - f(x,b) \\
&= f(a+x,b+y) - f(a,b+y) - f(x,b) - f(x,y) + f(x,y) \\
&= f(a+x,b+y) - f(a,b+y) - f(x,b+y) + f(x,y) \\
&= f(a+x,b+y) - f(a+x,b+y) + f(x,y) \\
&= f(x,y)
\end{aligned}$$

Then we have

$$\lim_{(x,y)\to\mathbf{0}} \frac{|f(a+x,b+y) - f(a,b) - f(a,y) - f(x,b)|}{|(x,y)|} = \lim_{(x,y)\to\mathbf{0}} \frac{|f(x,y)|}{|(x,y)|}$$

and by part (a) we know this limit is 0. $\square$

(c) *Proof.* Note that our work in part (a) implies that f is completely determined by its values on the various pairs (e_i, e_j). So $Dp(a,b)$ is simply the case where $n = m = 1$ and $f(1,1) = 1$. $\square$

Exercise 2-13 Define $IP : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by $IP(x,y) = \langle x, y \rangle$.

(a) Find $D(IP)(a,b)$ and $(IP)'(a,b)$.

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

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

(Note that $f'(a)$ is an $n \times 1$ matrix; its transpose $f'(a)^T$ is a $1 \times n$ matrix, which we consider as a member of $\mathbb{R}^n$.)

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

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

(a) Since the (real) inner product is bilinear by definition, we can apply Exercise 2-12 to conclude that

$$D(IP)(a,b)(x,y) = IP(a,y) + IP(x,b) = \langle a, y \rangle + \langle x, b \rangle = \langle a, y \rangle + \langle b, x \rangle = \langle (b,a), (x,y) \rangle$$

Moreover, we can rewrite this to be

$$D(IP)(a,b)(x,y) = (b,a)(x,y)^T$$

from which we can conclude that

$$(IP)'(a,b) = (b,a)$$

where (b,a) is the $1 \times 2n$ matrix given by concatenating the row vectors b and a.

(b) *Proof.* Directly from the definition of h, we have

$$h = IP \circ (f,g)$$

so the chain rule says that

$$\begin{aligned}
h'(a) &= IP'(f(a),g(a)) \begin{bmatrix} | \\ f'(a) \\ | \\ \\ | \\ g'(a) \\ | \end{bmatrix} \\
&= \begin{bmatrix} -g(a)- & & -f(a)- \end{bmatrix} \begin{bmatrix} | \\ f'(a) \\ | \\ \\ | \\ g'(a) \\ | \end{bmatrix} \\
&= \begin{bmatrix} -g(a)- \end{bmatrix} \begin{bmatrix} | \\ f'(a) \\ | \end{bmatrix} + \begin{bmatrix} -f(a)- \end{bmatrix} \begin{bmatrix} | \\ g'(a) \\ | \end{bmatrix} \\
&= \left\langle g(a), f'(a)^T \right\rangle + \left\langle f(a), g'(a)^T \right\rangle \\
&= \left\langle f'(a)^T, g(a) \right\rangle + \left\langle f(a), g'(a)^T \right\rangle
\end{aligned}$$

□

(c) *Proof.* Define

$$h(t) := \langle f(t), f(t) \rangle = \sqrt{|f(t)|}$$

Then by part (b),

$$h'(t) = 2\left\langle f'(t)^T, f(t) \right\rangle$$

But the assumption that $|f(t)|$ is identically 1 means that h is constant, and thus

$$\left\langle f'(t)^T, f(t) \right\rangle = \frac{h'(t)}{2} = 0$$

□

(d) The identity function satisfies this, since $x \mapsto |x|$ is not differentiable at $x = 0$.

Exercise 2-14 Let E_i, $i = 1, \ldots, k$ be Euclidean spaces of various dimensions. A function $f : E_1 \times \ldots \times E_k \to \mathbb{R}^p$ is called **multilinear** if for each choice of $x_j \in E_j$, $j \neq i$, the function $g : E_i \to \mathbb{R}^p$ defined by $g(x) = f(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_k)$ is a linear transformation.

(a) If f is multilinear and $i \neq j$, show that for $h = (h_1, \ldots, h_k)$ with $h_l \in E_l$, we have

$$\lim_{h \to \mathbf{0}} \frac{|f(a_1, \ldots, h_i, \ldots, h_j, \ldots, a_k)|}{|h|} = 0$$

(b) Prove that

$$Df(a_1, \ldots, a_k)(x_1, \ldots, x_k) = \sum_{i=1}^{k} f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_k)$$

(a) *Proof.* Suppose that $\dim E_i = k_1$ and $\dim E_j = k_2$. Then define the function $g : \mathbb{R}^{k_1} \times \mathbb{R}^{k_2} \to \mathbb{R}^p$ by

$$g(x, y) = f(a_1, \ldots, x, \ldots, y, \ldots, a_k)$$

Then we need to prove that

$$\lim_{(h_i, h_j) \to \mathbf{0}} \frac{|g(h_i, h_j)|}{|(h_i, h_j)|} = 0$$

To do this, we first prove that g is bilinear. Using multilinearity, we have that

$$g(ax, y) = f(a_1, \ldots, ax, \ldots, y, \ldots, a_k) = af(a_1, \ldots, x, \ldots, y, \ldots, a_k) = ag(x, y)$$

and

$$\begin{aligned} g(x_1 + x_2, y) &= f(a_1, \ldots, x_1 + x_2, \ldots, y, \ldots, a_k) \\ &= f(a_1, \ldots, x_1, \ldots, y, \ldots, a_k) + f(a_1, \ldots, x_2, \ldots, y, \ldots, a_k) \\ &= g(x_1, y) + g(x_2, y) \end{aligned}$$

The last property is similar. So g is bilinear, and Exercise 2-12 part (a) tells us that

$$\lim_{h \to \mathbf{0}} \frac{|f(a_1, \ldots, h_i, \ldots, h_j)|}{|h|} = \lim_{(h_i, h_j) \to \mathbf{0}} \frac{|g(h_i, h_j)|}{|(h_i, h_j)|} = 0 \qquad \square$$

(b) *Proof.* For notational convenience, we define the following. Given a set of distinct indices $i_1, \ldots, i_n \in [1, k]$, and vectors $\overrightarrow{a} = (a_1, \ldots, a_k)$, $\overrightarrow{h} = (h_1, \ldots, h_k)$, we write

$$f_{\{i_1, \ldots, i_n\}}(\overrightarrow{a}, \overrightarrow{h}) = f(a_1, \ldots, h_{i_1}, \ldots, h_{i_2}, \ldots, h_{i_n}, \ldots, a_k)$$

In other words, if $S \subseteq [1, k]$, then $f_S(\overrightarrow{a}, \overrightarrow{h})$ passes in h_i if $i \in S$ and a_i otherwise.

Now, we prove an extension of part (a), namely, that for any k-linear function f, if we pick $n \leq k$ indices $i_1, \ldots, 1_n$, then

$$\lim_{\overrightarrow{h} \to \mathbf{0}} \frac{|f_{\{i_1,\ldots,i_n\}}(\overrightarrow{a}, \overrightarrow{h})|}{|\overrightarrow{h}|} = 0$$

We skip the proof that $f_{\{i_1,\ldots,i_n\}}$ is n-linear, so this reduces to simply showing that for any multilinear function ($n > 1$) we have

$$\lim_{\overrightarrow{h} \to \mathbf{0}} \frac{|f(\overrightarrow{h})|}{|\overrightarrow{h}|} = 0$$

Let $d_i = \dim E_i$ for each i. Suppose also that $h_i = (h_{i,1}, \ldots, h_{i,d_i})$. Then

$$\begin{aligned}
\lim_{\overrightarrow{h} \to \mathbf{0}} \frac{|f(h_1, \ldots, h_k)|}{\left|\overrightarrow{h}\right|} &= \lim_{\overrightarrow{h} \to \mathbf{0}} \frac{\left|f(\sum_{j_1=1}^{d_1} h_{i,j_1}, \ldots, \sum_{j_k=1}^{d_k} h_{i,j_k})\right|}{\left|\overrightarrow{h}\right|} \\
&= \lim_{\overrightarrow{h} \to \mathbf{0}} \frac{\left|\sum_{j_1=1}^{d_1} \ldots \sum_{j_k=1}^{d_k} h_{1,j_1} \ldots h_{k,j_k} f(e_{j_1}, \ldots, e_{j_k})\right|}{\left|\overrightarrow{h}\right|} \\
&\leq \sum_{j_i=1}^{d_i} \ldots \sum_{j_k=1}^{d_k} |f(e_{j_1}, \ldots, e_{j_k})| \lim_{(h_{1,j_1},\ldots,h_{k,j_k}) \to \mathbf{0}} \frac{|h_{1,j_1} \ldots h_{k,j_k}|}{\left|\overrightarrow{h}\right|} \\
&\leq \sum_{j_i=1}^{d_i} \ldots \sum_{j_k=1}^{d_k} |f(e_{j_1}, \ldots, e_{j_k})| \lim_{(h_{1,j_1},\ldots,h_{k,j_k}) \to \mathbf{0}} \frac{|h_{1,j_1} \ldots h_{k,j_k}|}{|(h_{1,j_1}, \ldots, h_{k,j_k})|} \\
&= \sum_{j_i=1}^{d_i} \ldots \sum_{j_k=1}^{d_k} |f(e_{j_1}, \ldots, e_{j_k})| \cdot 0 \\
&= 0
\end{aligned}$$

Thus we have shown that any multilinear function satisfies

$$\lim_{\overrightarrow{h} \to \mathbf{0}} \frac{|f(\overrightarrow{h})|}{|\overrightarrow{h}|} = 0$$

Now, I claim that

$$f(a_1 + x_1, \ldots, a_k + x_k) = f(\overrightarrow{a} + \overrightarrow{x}) = \sum_{S \in \mathcal{P}([1,k])} f_S(\overrightarrow{a}, \overrightarrow{x})$$

where $\mathcal{P}([1,k])$ represents the set of all subsets of $[1,k]$. We prove this by induction. Supposing it is true for $k-1$, we can then partition $\mathcal{P}([1,k])$ into X, consisting of those subsets which contain k, and A, consisting of those subsets which do not. Then

$$\sum_{S \in \mathcal{P}([1,k])} f_S(\overrightarrow{a}, \overrightarrow{x}) = \sum_{S \in X} f_S(\overrightarrow{a}, \overrightarrow{x}) + \sum_{S \in A} f_S(\overrightarrow{a}, \overrightarrow{x})$$

Now, the inductive hypothesis applies, and we have

$$\sum_{S \in X} f_S(\overrightarrow{a}, \overrightarrow{x}) = f(a_1 + x_1, \ldots, a_{k-1} + x_{k-1}, x_k)$$

and

$$\sum_{S \in A} f_S(\overrightarrow{a}, \overrightarrow{x}) = f(a_1 + x_1, \ldots, a_{k-1} + x_{k-1}, a_k)$$

and by applying multilinearity, we conclude that

$$\begin{aligned} \sum_{S \in \mathcal{P}([1,k])} f_S(\overrightarrow{a}, \overrightarrow{x}) &= \sum_{S \in X} f_S(\overrightarrow{a}, \overrightarrow{x}) + \sum_{S \in A} f_S(\overrightarrow{a}, \overrightarrow{x}) \\ &= f(a_1 + x_1, \ldots, a_{k-1} + x_{k-1}, x_k) + f(a_1 + x_1, \ldots, a_{k-1} + x_{k-1}, a_k) \\ &= f(\overrightarrow{a} + \overrightarrow{x}) \end{aligned}$$

Lastly, we have

$$\begin{aligned} &\lim_{\overrightarrow{h} \to \mathbf{0}} \frac{\left| f(\overrightarrow{a} + \overrightarrow{h}) - f(\overrightarrow{a}) - \sum_{i=1}^{k} f_{\{i\}}(\overrightarrow{a}, \overrightarrow{h}) \right|}{\left| \overrightarrow{h} \right|} \\ &= \lim_{\overrightarrow{h} \to \mathbf{0}} \frac{\left| \sum_{S \in \mathcal{P}([1,k])} f_S(\overrightarrow{a}, \overrightarrow{h}) - f(\overrightarrow{a}) - \sum_{i=1}^{k} f_{\{i\}}(\overrightarrow{a}, \overrightarrow{h}) \right|}{|\overrightarrow{h}|} \end{aligned}$$

Now, after cancelling, the numerator will be left only with terms of the form $f_S(\overrightarrow{a}, \overrightarrow{h})$ where S contains at least two elements, and f_S is therefore n-linear for $n > 1$. Thus the first part of this proof shows that the quotient goes to 0. $\square$

Exercise 2-15 Regard an $n \times n$ matrix as a point in the n-fold product $\mathbb{R}^n \times \ldots \times \mathbb{R}^n$ by considering each column as a member of $\mathbb{R}^n$. (**Note:** Spivak considers the rows as elements of $\mathbb{R}^n$, but we use columns here for convention.)

(a) Prove that $\det : \mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}$ is differentiable and

$$D(\det)(a_1, \ldots, a_n)(x_1, \ldots, x_n) = \sum_{i=1}^{n} \det \begin{bmatrix} | & & | & & | \\ a_1 & \ldots & x_i & \ldots & a_n \\ | & & | & & | \end{bmatrix}$$

(b) If $a_{ij} : \mathbb{R} \to \mathbb{R}$ are differentiable, let $A(t)$ be the matrix such that $A(t)_{ij} = a_{ij}(t)$. If $f(t) = \det(A(t))$, show that

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

(c) If $\det(A(t)) \neq 0$ for all t and $b_1, \ldots, b_n : \mathbb{R} \to \mathbb{R}$ are differentiable, let $s_1, \ldots, s_n : \mathbb{R} \to \mathbb{R}$ be the functions such that $s_1(t), \ldots, s_n(t)$ are solutions of the equations

$$\sum_{j=1}^{n} a_{ij}(t) s_j(t) = b_i(t)$$

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

(a) *Proof.* We take it for granted that det is multilinear, as this is one possible definition of the determinant, and otherwise can easily be concluded from Laplace expansion along various columns. Then det is differentiable by Exercise 2-14 part (b), and moreover the result from that problem shows that

$$D(\det)(a_1, \ldots, a_n)(x_1, \ldots, x_n) = \sum_{i=1}^{n} \det \begin{bmatrix} | & & | & & | \\ a_1 & \ldots & x_i & \ldots & a_n \\ | & & | & & | \end{bmatrix} \qquad \square$$

(b) *Proof.* Note that $f'(t)$ is just a number, so we ignore the distinction between $Df(t)$

and $f'(t)$. By the chain rule, and using the result from part (a),

$$
\begin{aligned}
Df(t) &= D\left(\det \circ \left(\begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix}, \dots, \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix}\right)\right)(t) \\
&= D(\det)\left(\begin{bmatrix} a_{11}(t) \\ \vdots \\ a_{n1}(t) \end{bmatrix}, \dots, \begin{bmatrix} a_{1n}(t) \\ \vdots \\ a_{nn}(t) \end{bmatrix}\right)\left(\begin{bmatrix} a'_{11}(t) \\ \vdots \\ a'_{n1}(t) \end{bmatrix}, \dots, \begin{bmatrix} a'_{1n}(t) \\ \vdots \\ a'_{nn}(t) \end{bmatrix}\right) \\
&= \sum_{i=1}^{n} \det \begin{bmatrix} | & & | & & | \\ a_1(t) & \dots & a'_i(t) & \dots & a_n(t) \\ | & & | & & | \end{bmatrix}
\end{aligned}
$$

□

(c) For any fixed t, we essentially have the condition that

$$
\begin{bmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}(t) & \dots & a_{nn}(t) \end{bmatrix} \begin{bmatrix} s_1(t) \\ \vdots \\ s_n(t) \end{bmatrix} = \begin{bmatrix} b_1(t) \\ \vdots \\ b_n(t) \end{bmatrix}
$$

or more concisely, we can write

$$
A(t)\overrightarrow{s}(t) = \overrightarrow{b}(t)
$$

Since we are given that $\det A(t) \neq 0$, we know that $A(t)$ is invertible. Then by Cramer's Rule,

$$
s_i(t) = \frac{\det(A_i(t))}{\det(A(t))}
$$

where

$$
A_i(t) = \begin{bmatrix} a_{11}(t) & \dots & b_1(t) & \dots & a_{1n}(t) \\ \vdots & & \vdots & & \vdots \\ a_{n1}(t) & \dots & b_n(t) & \dots & a_{1n}(t) \end{bmatrix}
$$

Then $s_i(t)$ is differentiable as the quotient of differentiable functions. To calculate $s'_i(t)$, we have

$$
s'_i(t) = \frac{\det(A(t))D(\det \circ A_i)(t) - \det(A_i(t))D(\det \circ A)(t)}{[\det(A(t))]^2}
$$

Define the following matrices for convenience:

$$A^j(t) = \begin{bmatrix} | & & | & & | \\ a_1(t) & \dots & a_j'(t) & \dots & a_n(t) \\ | & & | & & | \end{bmatrix}$$

$$A_i^j(t) = \begin{cases} \begin{bmatrix} | & & | & & | & & | \\ a_1(t) & \dots & a_j'(t) & \dots & b_i(t) & \dots & a_n(t) \\ | & & | & & | & & | \end{bmatrix}, & i \neq j \\ \begin{bmatrix} | & & | & & | \\ a_1(t) & \dots & b_i'(t) & \dots & a_n(t) \\ | & & | & & | \end{bmatrix}, & i = j \end{cases}$$

Then the results from part (b), and the quotient rule,

$$s_i'(t) = \frac{\det(A(t)) \sum_{j=1}^n \det A_i^j(t) - \det(A_i(t)) \sum_{j=1}^n \det A^j(t)}{[\det(A(t))]^2}$$

Exercise 2-16 Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is differentiable and has a differentiable inverse $f^{-1} : \mathbb{R}^n \to \mathbb{R}^n$. Show that

$$(f^{-1})'(a) = [f'(f^{-1}(a))]^{-1}$$

Proof. By definition,

$$f \circ f^{-1} = \mathrm{Id}$$

Since both f and f^{-1} are differentiable, we can apply the chain rule in matrix form:

$$I_n = f'(f^{-1}(a)) \cdot (f^{-1})'(a)$$

Since both $f'(f^{-1}(a))$ and $(f^{-1})'(a)$ are $n \times n$ matrices, being single sided inverses is equivalent to being inverses, so we conclude that

$$(f^{-1})'(a) = [f'(f^{-1}(a))]^{-1} \qquad \square$$

Exercise 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) $f(x,y,z) = \sin(x \sin(y \sin z))$

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

(a)

$$\begin{aligned} D_1 f(x,y,z) &= y x^{y-1} \\ D_2 f(x,y,z) &= x^y \ln x \\ D_3 f(x,y,z) &= 0 \end{aligned}$$

(b)

$$\begin{aligned} D_1 f(x,y,z) &= 0 \\ D_2 f(x,y,z) &= 0 \\ D_3 f(x,y,z) &= 1 \end{aligned}$$

(c)

$$\begin{aligned} D_1 f(x,y) &= \sin y \cos(x \sin y) \\ D_2 f(x,y) &= x \cos y \cos(x \sin y) \end{aligned}$$

(d)

$$\begin{aligned} D_1 f(x,y,z) &= \sin(y \sin z) \cos(x \sin(y \sin z)) \\ D_2 f(x,y,z) &= x \sin z \cos(y \sin z) \cos(x \sin(y \sin z)) \\ D_3 f(x,y,z) &= xy \cos z \cos(y \sin z) \cos(x \sin(y \sin z)) \end{aligned}$$

(e)

$$\begin{aligned} D_1 f(x,y,z) &= y^z x^{y^z - 1} \\ D_2 f(x,y,z) &= z y^{z-1} x^{y^z} \ln x \\ D_3 f(x,y,z) &= y^z x^{y^z} \ln x \ln y \end{aligned}$$

(f)

$$\begin{aligned}
D_1 f(x,y,z) &= (y+z)x^{y+z-1} \\
D_2 f(x,y,z) &= x^z x^y \ln x^{y+z} \ln x \\
D_3 f(x,y,z) &= x^{y+z} \ln x
\end{aligned}$$

(g)

$$\begin{aligned}
D_1 f(x,y,z) &= z(x+y)^{z-1} \\
D_2 f(x,y,z) &= z(x+y)^{z-1} \\
D_3 f(x,y,z) &= (x+y)^z \ln(x+y)
\end{aligned}$$

(h)

$$\begin{aligned}
D_1 f(x,y) &= y\cos(xy) \\
D_2 f(x,y) &= y\cos(xy)
\end{aligned}$$

(i)

$$\begin{aligned}
D_1 f(x,y) &= y\cos 3[\sin(xy)]^{\cos 3-1}\cos(xy) \\
D_2 f(x,y) &= x\cos 3[\sin(xy)]^{\cos 3-1}\cos(xy)
\end{aligned}$$

Exercise 2-18 If $g : \mathbb{R} \to \mathbb{R}$ is continuous, find the partial derivatives of each of the following functions:

(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) $f(x,y) = \int_a^{(\int_b^y g)} g$

(a) By the fundamental theorem of calculus,

$$\begin{aligned}
D_1 f(x,y) &= g(x+y) \\
D_2 f(x,y) &= g(x+y)
\end{aligned}$$

(b) Let $a \in \mathbb{R}$. Then

$$\int_y^x g = \int_a^x g - \int_a^y g$$

so

$$\begin{aligned}
D_1 f(x,y) &= g(x) \\
D_2 f(x,y) &= -g(y)
\end{aligned}$$

(c)

$$D_1 f(x,y) = yg(xy)$$
$$D_2 f(x,y) = xg(xy)$$

(d)

$$D_1 f(x,y) = 0$$
$$D_2 f(x,y) = g\left(\int_b^y g\right) g(y)$$

Exercise 2-19 If

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

Find $D_2 f(1,y)$.

Since we are calculating D_2, we treat x as constant, and in particular, we can substitute in $x = 1$. So we have

$$g_2(y) = f(1,y) = \underbrace{1^{1^{1^{1^y}}}}_{=1} + \underbrace{(\ln 1)}_{=0}(\arctan(\arctan(\arctan(\sin(\cos y) - \ln(1+y)))))$$

So $g_2(y) = 1$ for all y, and thus $g_2'(y) = D_2 f(1,y) = 0$.

Exercise 2-20 Find the partial derivatives of f in terms of g, h, g', h'.

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

(a)

$$D_1 f(x,y) = h(y)g'(x)$$
$$D_2 f(x,y) = g(x)h'(y)$$

(b)

$$D_1 f(x,y) = h(y)g(x)^{h(y)-1}$$
$$D_2 f(x,y) = g(x)^{h(y)} \ln(g(x))$$

(c)

$$D_1 f(x,y) = g'(x)$$
$$D_2 f(x,y) = 0$$

(d)

$$D_1 f(x,y) = 0$$
$$D_2 f(x,y) = g'(y)$$

(e)

$$D_1 f(x,y) = g'(x+y)$$
$$D_2 f(x,y) = g'(x+y)$$

Exercise 2-21 Let $g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$ be continuous. Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

$$f(x,y) = \int_0^x g_1(t,0)dt + \int_0^y g_2(x,t)dt$$

(a) Show that $D_2 f(x,y) = g_2(x,y)$.

(b) How should f be defined such that $D_1 f(x,y) = g_1(x,y)$?

(c) Find a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $D_1 f(x,y) = x$ and $D_2 f(x,y) = y$. Find one such that $D_1 f(x,y) = y$ and $D_2 f(x,y) = x$.

(a) *Proof.* Define

$$h_2(y) \coloneqq f(x,y)$$

Then

$$D_2 f(x,y) = h_2'(y) = \frac{\mathrm{d}}{\mathrm{d}y} \int_0^x g_1(t,0)dt + \frac{\mathrm{d}}{\mathrm{d}y} \int_0^y g_2(x,t)dt$$

Since the first integral is constant with respect to y,

$$\frac{\mathrm{d}}{\mathrm{d}y} \int_0^x g_1(t,0)dt = 0$$

By the fundamental theorem of calculus,

$$\frac{\mathrm{d}}{\mathrm{d}y} \int_0^y g_2(x,t)dt = g_2(x,y)$$

Thus

$$D_2 f(x,y) = h_2'(y) = g_2(x,y)$$ ◻

(b) Define

$$f(x, y) = \int_0^x g_1(t, y)dt + \int_0^y g_2(0, t)dt$$

Then by a similar argument as above, $D_1 f(x, y) = g_1(x, y)$.

(c) The function $f(x, y) = \frac{x^2}{2} + \frac{y^2}{2}$ satisfies

$$D_1 f(x, y) = x, D_2 f(x, y) = y$$

The function $f(x, y) = xy$ satisfies

$$D_1 f(x, y) = y, D_2 f(x, y) = x$$

Exercise 2-22 If $f : \mathbb{R}^2 \to \mathbb{R}$ and $D_2 f = 0$, show that f is independent of the second variable. If $D_1 f = D_2 f = 0$, show that f is constant.

Proof. Fix some $x \in \mathbb{R}$, and define $h_x(y) = f(x, y)$. Since $D_2 f = 0$, $h'_x(y) = 0$ everywhere, so h_x is constant. Thus for any $y_1, y_2 \in \mathbb{R}$,

$$f(x, y_1) = h_x(y_1) = h_x(y_2) = f(x, y_2)$$

and thus f is independent of the second variable.

When $D_1 f = 0$, f is independent of the first variable as well. Moreover, we showed in Exercise 2-3 that functions which are independent of both variables are constant, so f is constant. $\square$

Exercise 2-23 Let $A = \{(x, y) \in \mathbb{R}^2 : x < 0, \text{ or } x \geq 0 \text{ and } y \neq 0\}$.

(a) If $f : A \to \mathbb{R}$ and $D_1 f = D_2 f = 0$, show that f is constant.

(b) Find a function $f : A \to \mathbb{R}$ such that $D_2 f = 0$ but f is not independent of the second variable.

Note: The set A as defined here is the plane excluding the nonnegative x-axis.

(a) *Proof.* Let $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ be arbitrary. Suppose $y_1 \neq 0$ and $y_2 \neq 0$. Define $g_x(y) = f(x, y)$ and $h_y(x) = f(x, y)$. Pick some $a < 0$. Then

$$\begin{aligned} f(x_1, y_1) - f(x_2, y_2) &= f(x_1, y_1) - f(a, y_1) + f(a, y_1) - f(a, y_2) + f(a, y_2) - f(x_2, y_2) \\ &= h_{y_1}(x_1) - h_{y_1}(a) + g_a(y_1) - g_a(y_2) + h_{y_2}(a) - h_{y_2}(x_2) \end{aligned}$$

Since $y_1 \neq 0$, h_{y_1} is defined on all of $\mathbb{R}$ and h'_{y_1} is identically, h_{y_1} is constant. Similarly, h_{y_2} is constant, and g_a is also constant since $a < 0$. Thus

$$f(x_1, y_1) - f(x_2, y_2) = \underbrace{h_{y_1}(x_1) - h_{y_1}(a)}_{=0} + \underbrace{g_a(y_1) - g_a(y_2)}_{=0} + \underbrace{h_{y_2}(a) - h_{y_2}(x_2)}_{=0} = 0$$

The case where $y_1 = 0$ or $y_2 = 0$ is proved similarly. (Geometrically, we have connected the points (x_1, y_1) and (x_2, y_2) using three segments, but this can be adjusted to use only two or one if either y-coordinate is 0.) Thus $f(x_1, y_1) = f(x_2, y_2)$ for all points, and thus f is constant. □

(b) Define $f : A \to \mathbb{R}$ by

$$f(x,y) = \begin{cases} 1, & x = 0, y > 0 \\ 0, & \text{otherwise} \end{cases}$$

Pick some point (x, y). Then there exists an interval $(y - \varepsilon, y + \varepsilon) \subseteq A$. Moreover, f is constant on this interval. Thus $D_2 f(x, y) = 0$ everywhere, but f is not constant.

Exercise 2-24 Define $f : \mathbb{R}^2 \to \mathbb{R}$ by

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

(a) Show that $D_2 f(x, 0) = x$ for all x and $D_1 f(0, y) = -y$ for all y.

(b) Show that $D_{1,2} f(0, 0) \neq D_{2,1} f(0, 0)$.

(a) *Proof.* Define $g_x(y) = g(x, y)$ and $h_y(x) = f(x, y)$. Then

$$\begin{aligned} D_2 f(x,0) &= g_x'(0) \\ &= \frac{\mathrm{d}}{\mathrm{d}y}\left(xy\frac{x^2-y^2}{x^2+y^2}\right)\bigg|_{y=0} \\ &= \left(x\frac{x^2-y^2}{x^2+y^2} + xy\frac{-2x^2y - 2y^3 - 2x^2y + 2y^3}{(x^2+y^2)^2}\right)\bigg|_{y=0} \\ &= x\frac{x^2}{x^2} \end{aligned}$$

And

$$\begin{aligned} D_1 f(0,y) &= h_y'(0) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left(xy\frac{x^2-y^2}{x^2+y^2}\right)\bigg|_{x=0} \\ &= \left(y\frac{x^2-y^2}{x^2+y^2} + xy\frac{2x^3 + 2xy^2 - 2x^3 + 2xy^2}{(x^2+y^2)^2}\right)\bigg|_{x=0} \\ &= y\frac{-y^2}{y^2} \\ &= -y \end{aligned}$$

□

(b) Taking the derivative of the functions we computed in part (a),

$$D_{1,2}f(0,0) = \frac{\mathrm{d}}{\mathrm{d}y}D_1f(0,y) = \frac{\mathrm{d}}{\mathrm{d}y}(-y) = -1$$
$$D_{2,1}f(0,0) = \frac{\mathrm{d}}{\mathrm{d}x}D_2f(x,0) = \frac{\mathrm{d}}{\mathrm{d}x}x = 1$$

so

$$D_{1,2}f(0,0) = -1 \neq 1 = D_{2,1}f(0,0)$$

Exercise 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 C^∞, and $f^{(i)}(0) = 0$ for all i.

Proof. For points $x \neq 0$, we have

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

and

$$f''(x) = \frac{-6}{x^4}e^{-x^{-2}} + \frac{4}{x^6}e^{-x^{-2}}$$

Claim: In general, for any $i > 0$ and $x \neq 0$, $f^{(i)}(x)$ is composed of terms of the form

$$\frac{a}{x^b}e^{-x^{-2}}, \qquad a \in \mathbb{Z}, b \in \mathbb{Z}_{\geq 0}$$

We prove this by induction. As shown, we already know this is true for $i = 1, 2$. Now suppose it is true for $i = k$. Then for $k+1$, it is sufficient to show that each term of the above form differentiates into further terms of that form. Differentiating,

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{a}{x^b}e^{-x^{-2}} = \frac{-ab}{x^{b+1}}e^{-x^{-2}} + \frac{2a}{x^{b+3}}e^{-x^{-2}}$$

and the two terms are also of the form requested. Thus the claim is proved. This shows that $f^{(i)}(x)$ exists for all i when $x \neq 0$.

For $x = 0$, we use L'Hopital's rule:

$$\begin{aligned} f'(0) &= \lim_{h\to 0}\frac{e^{-h^{-2}}}{h} \\ &= \lim_{h\to 0}\frac{\frac{1}{h}}{e^{h^{-2}}} \\ (LH) &= \lim_{h\to 0}\frac{-\frac{1}{h^2}}{-2h^{-3}e^{h^{-2}}} \\ &= \lim_{h\to 0}\frac{h}{2e^{h^{-2}}} \\ &= 0 \end{aligned}$$

Similarly, for higher derivatives, we can apply the claim proved above to write

$$f^{(i)}(x) = \sum_{j=1}^{n} \frac{a_j}{x^{b_j}} e^{-x^{-2}}, \qquad a_j \in \mathbb{Z}, b_j \in \mathbb{Z}_{\geq 0}$$

for some finite n. Then

$$\begin{aligned}
f^{(i+1)}(0) &= \lim_{h \to 0} \frac{f^{(i)}(h)}{h} \\
&= \lim_{h \to 0} \frac{\sum_{j=1}^{n} \frac{a_j}{h^{b_j}} e^{-h^{-2}}}{h} \\
&= \sum_{j=1}^{n} \left(\lim_{h \to 0} \frac{a_j}{h^{b_j+1}} e^{-h^{-2}} \right) \\
&= \sum_{j=1}^{n} \left(a_j \lim_{h \to 0} \frac{\frac{1}{h^{b_j+1}}}{e^{h^{-2}}} \right) \\
(LH) &= \sum_{j=1}^{n} \left(a_j \lim_{h \to 0} \frac{\frac{-(b_j+1)}{h^{b_j+2}}}{\frac{-2}{h^3} e^{h^{-2}}} \right) \\
&= \sum_{j=1}^{n} \left(a_j \frac{b_j+1}{2} \lim_{h \to 0} \frac{e^{x^{-2}}}{h^{b_j-1}} \right) \\
&\vdots \\
&= 0
\end{aligned}$$

Thus $f^{(i)}(x)$ exists for all i, x, so f is C^∞, and $f^{(i)}(0) = 0$ for all i. $\square$

Exercise 2-26 Let

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

(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 $s : \mathbb{R} \to [0,1]$ such that $s(x) = 0$ for $x \leq 0$ and $s(x) = 1$ for $x \geq \varepsilon$.

(c) If $a = (a_1, \dots, a_n) \in \mathbb{R}^n$, define $g_a : \mathbb{R}^n \to \mathbb{R}$ by

$$g_a(x) = g_a(x_1, \dots, x_n) = f(\frac{x_1 - a_1}{\varepsilon}) \cdot \ldots \cdot f(\frac{x_n - a_n}{\varepsilon})$$

Show that g_a is a C^∞ function which is positive on

$$(a_1 - \varepsilon, a_1 + \varepsilon) \times \ldots \times (a_n - \varepsilon, a_n + \varepsilon)$$

(d) If $A \subseteq \mathbb{R}^n$ is open and $C \subseteq A$ is compact, show that there is a nonnegative C^∞ function $h : A \to \mathbb{R}$ 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 h so that $h : A \to [0,1]$ and $h(x) = 1$ for $x \in C$.

(a) *Proof.* By definition, f is 0 outside of $(-1,1)$, and it must be positive on $(-1,1)$ since each of the exponential factors are positive.

To show that f is C^∞, define $f_1, f_2 : (-1,1) \to \mathbb{R}$ by

$$f_1(x) = e^{-(x-1)^{-2}}$$
$$f_2(x) = e^{-(x+1)^{-2}}$$

We proved in Exercise 2-25 that both f_1, f_2 are C^∞, so

$$f'(x) = f_1(x) f_2'(x) + f_1'(x) f_2(x)$$

and higher order derivatives will in general be sums of products of $f_1^{(i)}(x)$ and $f_2^{(j)}(x)$, which all exist and are continuous. Thus f is C^∞. ☐

(b) *Proof.* Fix $\varepsilon > 0$. Then define

$$s(x) = \begin{cases} 1, & x \geq 2\varepsilon \\ \frac{f(1-\frac{x}{\varepsilon})}{f(\frac{x}{\varepsilon}) + f(1-\frac{x}{\varepsilon})}, & -\varepsilon < x < 2\varepsilon \\ 0, & x \leq -\varepsilon \end{cases}$$

By definition, $s(x) = 0$ for $x \leq -\varepsilon$ and $s(x) = 1$ for $x \geq 2\varepsilon$.

On the interval $(-\varepsilon, 0]$, $1 - \frac{x}{\varepsilon} \geq 1$, so $f(1 - \frac{x}{\varepsilon}) = 0$ and thus $s = 0$. So $s = 0$ for any $x \leq 0$.

Similarly, for the interval $[\varepsilon, 2\varepsilon)$, $\frac{x}{\varepsilon} \geq 1$, so $f(\frac{x}{\varepsilon}) = 0$ and

$$s(x) = \frac{f(1 - \frac{x}{\varepsilon})}{f(1 - \frac{x}{\varepsilon})} = 1$$

So $s = 1$ for any $x \geq \varepsilon$.

To prove that s is C^∞, we can obvioulsy ignore the constant regions.

On $(0, \varepsilon)$, at least one of $f(\frac{x}{\varepsilon}), f(1 - \frac{x}{\varepsilon})$ will be positive, so the quotient rule says that $s'(x)$ exists. In general, we can continue to apply the quotient rule, since the quotient will never be zero, and f is smooth. Thus $s^{(i)}(x)$ exists and is continuous for all i and $x \in (0, \varepsilon)$, and we conclude that s is C^∞. □

(c) *Proof.* The fact that g_a is positive follows from the fact that for each i,

$$\left| \frac{x_i - a_i}{\varepsilon} \right| < 1$$

so $f(\frac{x_i - a_i}{\varepsilon}) > 0$. Thus their product g_a is positive.

To show that g_a is C^∞, we need to prove that the mixed partials of all orders exist. Here, we can actually prove a more general result:

Lemma

If $f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}$ are C^∞, then $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x_1, \ldots, x_n) \coloneqq f_1(x_1) \cdot \ldots \cdot f_n(x_n)$$

is C^∞.

Proof. To prove that derivatives of all orders exist and are continuous, pick any index i. Then define

$$g_i(x_i) = f_i(x_i) \left(\prod_{j \neq i} f_j(x_j) \right)$$

Then $g_i(x_i)$ is just a constant multiple of $f_i(x_i)$, so $g_i'(x_i)$ exists. Moreover, f_i' is also C^∞, so the function

$$D_i f(x_1, \ldots, x_n) = f_1(x_1) \ldots f_i'(x_i) \ldots f_n(x_n)$$

satisfies the hypotheses of this lemma and we can differentiate it again using the above method. So derivatives of all orders exist and are continuous. Thus f is C^∞. □

We can then apply the above lemma to conclude that g_a is C^∞. □

(d) *Proof.* For each $x = (x_1, \dots, x_n) \in C$, there exists ε_x such that the rectangle

$$R_x = (x_1 - \varepsilon_x, x_1 + \varepsilon_x) \times \dots \times (x_n - \varepsilon_x, x_n + \varepsilon_x) \subseteq A$$

In fact, we may choose ε_x small enough such that the closed rectangle is contained in A as well. Let $\mathcal{O}$ be the collection of R_x for $x \in C$. Since C is compact, we pick a finite subcover $\mathcal{O}' = \{R_{x_i}\}_{i=1}^m$. Then define $h : A \to \mathbb{R}$ by

$$h(x_1, \dots, x_n) = \sum_{i=1}^m g_{x_i}(x_1, \dots, x_n)$$

h is C^∞ since it is the product of C^∞ functions (by the lemma in part (d)). For any $y \in C$, $\mathcal{O}'$ covers C, so $y \in R_{x_i}$ for some x_i. Then $g_{x_i} > 0$, and each other g_{x_j} is at least nonnegative, so $h(y) > 0$.

Now let $\overline{R}_{x_i}$ be the closed rectangle about x_i.

$$B = \bigcup_{i=1}^m \overline{R}_{x_i}$$

We showed that we can pick ε small enough that $\overline{R}_{x_i} \subseteq A$. Thus B is a closed set contained in A. Moreover, if $y \notin B$, then $y \notin R_{x_i}$ for any i, and hence $h(y) = 0$. So h is 0 on outside of a closed set contained in A. □

(e) *Proof.* Since h is C^∞, it is continuous, and hence achieves a minimum value on C. Since h is positive on C, this minimum value $\varepsilon = \min_{x \in C} h(x)$ is positive. Let $s_\varepsilon : \mathbb{R} \to [0, 1]$ be as defined in part (b). Then the function

$$s_\varepsilon \circ h : \mathbb{R}^n \to [0, 1]$$

is still C^∞ (since the composition of C^∞ functions is C^∞ using repeated applications of the chain rule, similarly to the lemma in part (c)). Letting B be as defined previously, if $y \notin B$ then $h(y) = 0$, so $s_\varepsilon(h(y)) = s_\varepsilon(0) = 0$. Thus $s_\varepsilon \circ h$ is still of the form in part (d).

Moreover, whenever $x \in C$, $h(x) \geq \varepsilon$ so $s_\varepsilon(h(x)) = 1$. □

Exercise 2-27 Define $g, h : \{x \in \mathbb{R}^2 : |x| \leq 1\} \to \mathbb{R}^3$ by

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

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

Proof. Let $D_2 = \{x \in \mathbb{R}^2 : |x| \leq 1\}$ and $C_3 = \{x \in \mathbb{R}^3 : |x| = 1\}$. Then supposing that f has a maximum $m = \max_{x \in D_2} f(x)$, then there exists at least one point $x = (x_1, x_2, x_3)$ such that $f(x) = m$. Then we have the cases $x_3 \geq 0$ and $x_3 < 0$.

Case 1: Since $|x| = 1$, $x_1^2 + x_2^2 + x_3^2 = 1$, and hence

$$x_3 = \sqrt{1 - x_1^2 - x_2^2}$$

Thus we have $g(x_1, x_2) = x$, so $(f \circ g)(x_1, x_2) = m$. $(f \circ g)$ certainly cannot achieve a higher value, or else it would contradict m being the maximum of f, so m is also the maximum of g.

Case 2: Similar to Case 1, but we use $h(x_1, x_2)$ instead, and we find that $(f \circ h)$ achieves the maximum m.

Thus we see that m is the maximum of at least one of $f \circ g$ or $f \circ h$ on D_2. □

Exercise 2-28 Find expressions for the partial derivatives of the following functions:

(a) $F(x, y) = f(g(x)k(y), g(x) + h(y))$

(b) $F(x, y, z) = f(g(x + y), h(y + z))$

(c) $F(x, y, z) = f(x^y, y^z, z^x)$

(d) $F(x, y) = f(x, g(x), h(x, y))$

(a) Let $f(*) = f(g(x)k(y), g(x) + h(y))$. Using the chain rule for partial derivatives,

$$\begin{aligned}
D_1F(x, y) &= D_1f(*)D_x[g(x)k(y)] + D_2f(*)D_x[g(x) + h(y)] \\
&= k(y)g'(x)D_1f(*) + g'(y)D_2f(*) \\
D_2F(x, y) &= D_1f(*)D_y[g(x)k(y)] + D_2f(*)D_y[g(x) + h(y)] \\
&= g(x)k'(y)D_1f(*) + h'(y)D_2f(*)
\end{aligned}$$

(b) Let $f(*) = f(g(x + y), h(y + z))$. Then

$$\begin{aligned}
D_1F(x, y, z) &= D_1f(*)D_xg(x + y) + D_2f(*)D_xh(y + z) \\
&= g'(x + y)D_1f(*) \\
D_2F(x, y, z) &= D_1f(*)D_yg(x + y) + D_2f(*)D_yh(y + z) \\
&= g'(x + y)D_1f(*) + h'(y + z)D_2f(*) \\
D_3F(x, y, z) &= D_1f(*)D_zg(x + y) + D_2f(*)D_zh(y + z) \\
&= h'(y + z)D_2f(*)
\end{aligned}$$

(c) Let $f(*) = f(x^y, y^z, z^x)$. Omitting zero terms,

$$\begin{aligned}
D_1F(x, y, z) &= yx^{y-1}D_1f(*) + z^x \ln z D_3f(*) \\
D_2F(x, y, z) &= x^y \ln x D_1f(*) + zy^{z-1}D_2f(*) \\
D_3F(x, y, z) &= y^z \ln y D_2f(*) + xz^{x-1}D_3f(*)
\end{aligned}$$

(c) Let $f(*) = f(x, g(x), h(x,y))$. Then

$$\begin{aligned} D_1 F(x,y) &= D_1 f(*) + g'(x) D_2 f(*) + D_1 h(x,y) D_3 f(*) \\ D_2 F(x,y) &= D_2 h(x,y) D_3 f(*) \end{aligned}$$

Exercise 2-29 Let $f : \mathbb{R}^n \to \mathbb{R}$. For $\overrightarrow{x} \in \mathbb{R}^n$, if the limit

$$\lim_{t \to 0} \frac{f(a + t\overrightarrow{x}) - f(a)}{t}$$

exists, it is called the **directional derivative** of f at a in the direction $\overrightarrow{x}$, denoted $D_{\overrightarrow{x}}(a)$.

(a) Show that $D_{e_i} f(a) = D_i f(a)$.

(b) Show that $D_{t\overrightarrow{x}} f(a) = t D_{\overrightarrow{x}} f(a)$.

(c) If f is differentiable at a, show that $D_{\overrightarrow{x}} f(a) = Df(a)(\overrightarrow{x})$ and therefore $D_{\overrightarrow{x}+\overrightarrow{y}} f(a) = D_{\overrightarrow{x}} f(a) + D_{\overrightarrow{y}} f(a)$.

(a) *Proof.* Immediate from the definitions. ☐

(b) *Proof.* Fix $t \in \mathbb{R}$. Then

$$\begin{aligned} \lim_{s \to 0} \frac{f(a + s(t\overrightarrow{x})) - f(a)}{s} &= t \lim_{s \to 0} \frac{f(a + st\overrightarrow{x}) - f(a)}{st} \\ &= t \lim_{st \to 0} \frac{f(a + (st)\overrightarrow{x}) - f(a)}{(st)} \\ &= t D_{\overrightarrow{x}}(a) \end{aligned}$$ ☐

(c) *Proof.* Since the derivative exists, we know that

$$\lim_{t\overrightarrow{x} \to 0} \frac{f(a + t\overrightarrow{x}) - f(a) - Df(a)(t\overrightarrow{x})}{t|\overrightarrow{x}|} = 0$$

We can multiply both sides by $|\overrightarrow{x}|$ to clear the denominator, and apply linearity of $Df(a)$ to see that

$$\begin{aligned} &\lim_{t \to 0} \frac{f(a + t\overrightarrow{x}) - f(a) - tDf(a)(\overrightarrow{x})}{t} = 0 \\ \Longrightarrow\ &\lim_{t \to 0} \frac{f(a + t\overrightarrow{x}) - f(a)}{t} = Df(a)(\overrightarrow{x}) \end{aligned}$$

and thus $Df(a)(\overrightarrow{x}) = D_{\overrightarrow{x}} f(a)$. Since $Df(a)$ is linear,

$$\begin{aligned} D_{\overrightarrow{x}+\overrightarrow{y}} f(a) &= Df(a)(\overrightarrow{x} + \overrightarrow{y}) \\ &= Df(a)(\overrightarrow{x}) + Df(a)(\overrightarrow{y}) \\ &= D_{\overrightarrow{x}} f(a) + D_{\overrightarrow{y}} f(a) \end{aligned}$$ ☐

Exercise 2-30 Let f be defined as in Exercise 2-4. Show that $D_{\overrightarrow{x}} f(0,0)$ exists for all x, but if $g \neq 0$, then $D_{\overrightarrow{x}+\overrightarrow{y}} f(0,0) = D_{\overrightarrow{x}} f(0,0) + D_{\overrightarrow{y}} f(0,0)$ is not true for all x, y.

Proof. The result of Exercise 2-4 part (a) says that for $x \in \mathbb{R}^2$, defining $h_x(t) = f(tx)$ means that h_x is differentiable at $(0,0)$. This means that $D_{\overrightarrow{x}} f(0,0)$ exists for all $\overrightarrow{x}$. Similarly, as the result in part (b) shows, $D_{e_1} f(0) = D_{e_2} f(0) = 0$. However, if g is nonzero, then we can take a directional derivative in some direction which is a linear combination of e_1 and e_2, so the linearity condition fails. □

Exercise 2-31 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined as in Exercise 1-26. Show that $D_x f(0,0)$ exists for all x, even though f is not continuous at $(0,0)$.

Proof. As we showed in the proof of Exercise 1-26 part (b), f is 0 in an interval about $(0,0)$ in each direction, and is thus differentiable. □

Exercise 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

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

Show that f is differentiable at $(0,0)$ but $D_i f$ is not continuous at $(0,0)$.

(a) *Proof.* Let $\varepsilon > 0$. Then whenever $|x - 0| < \delta = \varepsilon$, we have

$$\left| \frac{f(x) - f(0)}{x} - 0 \right| = \left| \frac{f(x)}{x} \right| = \left| x \sin \frac{1}{x} \right| < \varepsilon$$

Thus f is differentiable at 0 with $f'(0) = 0$.

If we differentiate f elsewhere, we find that

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

But

$$\lim_{x\to 0} 2x\sin\frac{1}{x} - \cos\frac{1}{x} = -\lim_{x\to 0}\cos\frac{1}{x}$$

which doesn't exist. Thus f' is not continuous at 0 (it has an oscillating discontinuity). □

(b) *Proof.* Let $\varepsilon > 0$. Then whenever $|(x,y)| = \sqrt{x^2+y^2} < \delta = \varepsilon$, we have

$$\left|\frac{f(x,y)-f(0,0)}{|(x,y)|}\right| = \left|\frac{f(x,y)}{\sqrt{x^2+y^2}}\right| = \left|\sqrt{x^2+y^2}\sin\frac{1}{\sqrt{x^2+y^2}}\right| \leq \sqrt{x^2+y^2} < \varepsilon$$

Thus

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

so $Df(0,0)$ exists and is the zero transformation. But in the directions e_1, e_2, f is simply the single variable case considered in part (a), so we know $D_i f$ is not continuous at $(0,0)$. □

Exercise 2-33 Show that the continuity of $D_1 f^j$ at a may be eliminated from the hypothesis of Theorem 2-8.

Proof. In the proof of Theorem 2-8, we attempted to prove that

$$\lim_{\vec{h}\to 0} \frac{\left|f\left(\vec{a}+[\vec{h}]^j\right) - f\left(\vec{a}+[\vec{h}]^{j-1}\right) - D_j f(\vec{a})h_j\right|}{\left|\vec{h}\right|} = 0$$

for all j. We did this by using the continuity of $D_j f$ at a to extend its differentiability nearby. However, in the case of the first partial derivative $D_1 f$, the continuous differentiability condition already shows us that

$$\lim_{\vec{h}\to 0} \frac{|f(\vec{a}+h_1 e_1) - f(\vec{a}) - D_1 f(\vec{a})h_1|}{\left|\vec{h}\right|} = 0$$

so we can omit continuity. (Obviously, any other direction would also work.) □

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

$$\sum_{i=1}^{n} x^i D_i f(x) = mf(x)$$

Proof. Define $g(t) = f(tx)$. Then $D_x f(x) = g'(1)$. Moreover, we showed in Exercise 2-30 that D_* is linear, so

$$D_x f(x) = \sum_{i=1}^{n} x_i D_i f(x)$$

At the same time, we know that $g(t) = f(tx) = t^m f(x)$. Differentiating with respect to t,

$$g'(x) = mt^{m-1} f(x)$$

so

$$\sum_{i=1}^{n} x_i D_i f(x) = D_x f(x) = g'(1) = mf(x)$$

□

Exercise 2-35 If $f : \mathbb{R}^n \to \mathbb{R}$ is diferentiable and $f(0) = 0$, prove that there exist $g_i : \mathbb{R}^n \to \mathbb{R}$ such that

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

Proof. Since f is differentiable, the directional derivative $D_x f(tx)$ exists for all t, x. Define $h_x(t) = f(tx)$. Then $h_x'(t) = D_x f(tx)$. Thus h_x is differentiable. Then by the fundamental theorem of calculus,

$$f(x) = f(1x) = \int_0^1 h_x'(t)dt = \int_0^1 D_x f(tx)dt$$

Since D_* is linear with respect to direction, we then have

$$f(x) = \int_0^1 \sum_{i=1}^{n} x_i D_1 f(tx)dt = \sum_{i=1}^{n} x_i \int_0^1 D_1 f(tx)dt$$

Then defining $g_i(x) = \int_0^1 D_1 f(tx)dt$, we have found g_i satisfying

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

□

Exercise 2-36 Let $A \subseteq \mathbb{R}^n$ be an open set and $f : A \to \mathbb{R}^n$ a continuously differentiable one-to-one function such that $\det f'(x) \neq 0$ for all 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 $B \subseteq A$.

Proof. Let $y \in f(A)$. Then since f is one-to-one, there exists a unique $x \in A$ such that $f(x) = y$. Since f is continuously differentiable at x and $\det f'(x) \neq 0$, the Inverse Function Theorem tells us there exist open sets $V \subseteq A$ containing x and $W \subseteq \mathbb{R}^n$ such that $f : V \to W$

has an inverse. Thus $W \subseteq f(A)$ and $y \in W$, so $f(A)$ is open. Moreover, the Inverse Function Theorem also says f^{-1} is differentiable at y. But this is true for every $y \in f(A)$, so f^{-1} is differentiable. Lastly, let $B \subseteq A$ be open. Then the restriction $\overline{f} : B \to \mathbb{R}^n$ is also continuously differentiable and one-to-one, so $f(B)$ is open. $\square$

Exercise 2-37

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

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

(a) *Proof.* If $D_1 f(x,y) = 0$ for all $(x,y) \in \mathbb{R}^2$, then f is independent of the first variable and is not one-to-one. So suppose there exists some $(x_1, y_1) \in \mathbb{R}^2$ with $D_1 f(x_1, y_1) \neq 0$. Since f is continuously differentiable, there exists an open set A containing (x_1, y_1) such that $D_1 f(x,y) \neq 0$ for any $(x,y) \in A$. Then define $g : A \to \mathbb{R}^2$ by $g(x,y) = (f(x,y), y)$. Then the derivative is given by

$$g'(x,y) = \begin{bmatrix} D_1 f(x,y) & D_2 f(x,y) \\ 0 & 1 \end{bmatrix} \implies \det g'(x,y) = D_1 f(x,y) \neq 0$$

In particular, $\det g'(x_1, y_1) \neq 0$. Then by the Inverse Function Theorem, there exists an open set V containing (x_1, y_1) and an open set W containing $(f(x_1, y_1), y_1)$ such that $g : V \to W$ has a continuous, differentiable inverse $g^{-1} : W \to V$. Then pick some $y_2 \neq y_1$ such that $(f(x_1, y_1), y_2) \in W$. Then we have

$$g(g^{-1}(f(x_1, y_1), y_2)) = (f(x_1, y_1), y_2)$$

but by definition,

$$g(g^{-1}(f(x_1, y_1), y_2)) = (f(g^{-1}(f(x_1, y_1), y_2)), g_2^{-1}(f(x_1, y_1), y_2))$$

So

$$f(x_1, y_1) = f(g^{-1}(f(x_1, y_1), y_2))$$

While the x coordinate of $g^{-1}(f(x_1, y_1), y_2)$ is unknown, the y coordinate is certainly y_2. Thus we have

$$f(x_1, y_1) = f(*, y_2)$$

But we mandated that $y_1 \neq y_2$, so $(x_1, y_1) \neq (*, y_2)$. So f is not one-to-one. $\square$

(b) *Proof.* $\square$

Exercise 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 one-to-one (on all of $\mathbb{R}$).

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

(a) *Proof.* Suppose without loss of generality that $f'(a) > 0$ for some $a \in \mathbb{R}$. One can prove in single variable analysis that if $g = f'$ for some function f, then g satisfies the intermediate value property. If $f'(b) < 0$ for some $b \in \mathbb{R}$, then there exists c between a and b such that $f'(c) = 0$, contradicting the assumption. So we must have $f'(x) > 0$ for all x. Thus f is strictly increasing (or decreasing), so it is one-to-one. $\square$

(b) *Proof.* The Jacobian matrix is given by

$$f'(x,y) = \begin{bmatrix} e^x \cos y & e^x \sin y \\ -e^x \sin y & e^x \cos y \end{bmatrix}$$

so

$$\det f'(x,y) = e^x(\sin^2 y + \cos^2 y) = e^x \neq 0$$

But for any (x,y), we have

$$f(x,y) = f(x, y+2\pi)$$

so f is not one-to-one. $\square$

Exercise 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 the Inverse Function Theorem.

First, we verify that f is differentiable at 0. We have

$$\lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} \frac{1}{2} + h \sin \frac{1}{h} = \frac{1}{2} + \lim_{h \to 0} h \sin \frac{1}{h} = \frac{1}{2}$$

and by the formula f is clearly differentiable everywhere else. So f is differentiable in an open set around 0. However, I claim that for any open set V around 0, f is **not** injective onto $f(V)$.

To see this, let V be an open set aroud 0. Then pick n large enough that

$$a = \frac{1}{2\pi n} \in V$$

Now, we have

$$f'(a) = \frac{1}{2} + 2a \sin \frac{1}{a} - \cos \frac{1}{a} = \frac{1}{2} - 1 = -\frac{1}{2} < 0$$

Thus there exists $b < a$ with $f(b) > f(a)$ and $b > 0$. Now, pick m large enough that

$$c = \frac{1}{2\pi m} < b$$

Then we have

$$f(c) = \frac{c}{2} < \frac{a}{2} = f(a)$$

So $f(c) < f(a) < f(b)$, and $b \in [c, a]$. Pick some y with $f(a) < y < f(b)$. By the Intermediate Value Theorem, there exists $x_1 \in (c, b)$ with $f(x_1) = y$, and $x_2 \in (b, a)$ with $f(x_2) = y$, so f is not one-to-one onto $f(V)$. Thus the Inverse Function Theorem is false for f.

Exercise 2-40 Use the implicit function theorem to redo Problem 2-15 (c). For reference, this problem is reprinted here:

If $a_{ij} : \mathbb{R} \to \mathbb{R}$ are differentiable, let $A(t)$ be the matrix such that $A(t)_{ij} = a_{ij}(t)$. If $\det(A(t)) \neq 0$ for all t and $b_1, \ldots, b_n : \mathbb{R} \to \mathbb{R}$ are differentiable, let $s_1, \ldots, s_n : \mathbb{R} \to \mathbb{R}$ be the functions such that $s_1(t), \ldots, s_n(t)$ are the solutions of the equations

$$\sum_{j=1}^{n} a_{ji}(t) s_j(t) = b_i(t)$$

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

Proof. Define $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ such that the component functions are given by

$$F^i(t, x) = -b_i(t) + \sum_{j=1}^{n} a_{ji}(t) x_j$$

Then F^i can alternately be written as

$$F^i = -b_i \circ (\pi^2, \ldots, \pi^n) + \sum_{j=1}^{n} (a_{ji} \circ \pi^1) \pi^j$$

which makes it clear that it can be written as sums, products, and compositions of differentiable functions. If we assume that the a_{ij} and b_i are additionally continuously differentiable, then F is also continuously differentiable.

Now, fix t_1. Let $M(t, x)$ be the matrix with ijth entry given by $D_{j+1} F^i(t, x)$. To calculate the matrix of partial derivatives, for $k \geq 2$ we have

$$\begin{aligned} D_k F^i(t_1, x) &= D_k \left(\sum_{j=1}^{n} a_{ji}(t_1) x_j \right) \\ &= \sum_{j=1}^{n} a_{ji}(t_1) e_j \delta_{ik} \\ &= a_{ki}(t_1) e_k \end{aligned}$$

Thus $M(t_1, x)$ is simply the matrix $[A(t_1)]^T$, where $A(t_1)$ has ij-th entry given by $a_{ij}(t_1)$. By assumption, $\det[A(t_1)]^T = \det A(t_1) \neq 0$, so $\det M(t_1, x) \neq 0$ and the Implicit Function Theorem applies. Then there exists an open set $A \subseteq \mathbb{R}$ containing t and a function $g : A \to \mathbb{R}^n$ such that

$$F(t, g(t)) = 0$$

But this happens precisely when each component function is zero, so for each component we have

$$-b_i(t) + \sum_{j=1}^{n} a_{ji}(t) g^j(t) = 0 \iff \sum_{j=1}^{n} a_{ji}(t) g^j(t) = b_i(t)$$

Thus we may let $s_j = g^j$. Since $\det A(t) \neq 0$ for all t we are able to "patch" the local definitions of g^j into a global function without issue. Moreover, the Implicit Function Theorem tells us that g is differentiable at t_1, so each s_j is everywhere.

To calculate s_i', we know that $F^i(t, \overrightarrow{s}(t)) = 0$. Taking partial derivatives on both sides, we have

$$\begin{aligned} D_1 F^i(t, \overrightarrow{s}(t)) &= 0 \\ D_2 F^i(t, \overrightarrow{s}(t)) s_1'(t) &= 0 \\ D_3 F^i(t, \overrightarrow{s}(t)) s_2'(t) &= 0 \\ &\vdots \\ D_{n+1}(t, \overrightarrow{s}(t)) s_n'(t) &= 0 \end{aligned}$$

which we can combine as

$$D_1 F^i(t, \overrightarrow{s}(t)) + \sum_{j=1}^{n} D_{j+1} F^i(t, \overrightarrow{s}(t)) s_j'(t) = 0$$

Consider the system of equations this forms. We can rewrite it in matrix-vector multiplication using our definition of $M(t, x)$ from above as

$$M(t, \overrightarrow{s}(t)) s'(t) = -(D_1 F^i(t, \overrightarrow{s}(t)))$$

Moreover, the ith coordinate of the vector $(D_1 F^i(t, \overrightarrow{s}(t)))$ is given by

$$-b_i'(t) + \sum_{j=1}^{n} a_{ji}'(t) s_j(t)$$

Since $M(t, \overrightarrow{s}(t))$ is invertible by assumption, we find that

$$s'(t) = [M(t, \overrightarrow{s}(t))]^{-1} \begin{bmatrix} b_1'(t) - \sum_{j=1}^{n} a_{j1}'(t) s_j(t) \\ \vdots \\ b_n'(t) - \sum_{j=1}^{n} a_{jn}'(t) s_j(t) \end{bmatrix} \qquad \square$$

Exercise 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 $g_x'(y)$. Then let $c(x)$ be this y.

(a) If $D_{2,2}f(x, y) \neq 0$ for all (x, y), show that c is differentiable and

$$c'(x) = -\frac{D_{2,1}f(x, c(x))}{D_{2,2}f(x, c(x))}$$

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

$$\begin{aligned} D_{2,1}f(x, y) &= 0 \\ D_2f(x, y) &= 0 \end{aligned}$$

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

$$\max_{\frac{1}{2}\leq x\leq 2}\left(\min_{\frac{1}{3}\leq y\leq 1} f(x, y)\right)$$

Note: Spivak does not include this, but we must assume that f is twice continuously differentiable.

(a) *Proof.* Note that by our definition, $D_2f(x, y) = g_x'(y)$. So $y = c(x)$ precisely when $D_2f(x, y) = 0$. Note that D_2f is a function $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, and the matrix $M = (D_{j+1}(D_2f)^i(x, y))$ is simply the matrix with sole entry $D_{2,2}f(x, y)$. By assumption, $D_{2,2}f(x, y) \neq 0$, so $\det M \neq 0$ and the Implicit Function Theorem applies to D_2f, and we conclude that c is differentiable.

Now, the function $x \mapsto D_2f(x, c(x))$ is a function $\mathbb{R} \to \mathbb{R}$ and is 0 everywhere, so we can differentiate it:

$$D_{2,1}f(x, c(x)) + D_{2,2}f(x, c(x))c'(x) = 0$$

which we can rearrange as

$$c'(x) = -\frac{D_{2,1}f(x, c(x))}{D_{2,2}f(x, c(x))}$$

□

(b) *Proof.* Pick $y = c(x)$. Then by definition, $g_x'(c(x)) = 0$, and $D_2f(x, c(x)) = g_x'(c(x))$, so $D_2f(x, c(x)) = 0$. Moreover, from part (a),

$$D_{2,1}f(x, c(x)) = -c'(x)D_{2,2}f(x, c(x)) = 0$$

so this choice of y works. □

(c) For any fixed x,

$$\min_{\frac{1}{3}\le y\le 1} f(x,y) = \min_{\frac{1}{3}\le y\le 1} g_x(y)$$

We already know that $g_x'(c(x)) = 0$, so it is a critical point. If we calculate $g_x''(y) = D_{2,2}f(x,y)$ for any y, we get

$$\begin{aligned} D_2 f(x,y) &= x(\ln y + 1 - 1) - \ln x = x\ln y - \ln x \\ D_{2,2}f(x,y) &= \frac{x}{y} \end{aligned}$$

which is strictly positive (as both x, y must be positive for this function to be defined). Thus g_x is concave upward, and the critical point at $c(x)$ is in fact a global minimum.[1] So if $c(x) \in [\frac{1}{3}, 1]$, then the minimum is at $c(x)$. If $c(x) < \frac{1}{3}$, then the minimum is at $\frac{1}{3}$, and if $c(x) > 1$, then the minimum is at 1.

If we explicitly calculate $c(x)$, we use the fact that $D_2 f(x, c(x)) = 0$ to find

$$\ln c(x) = \frac{\ln x}{x} \implies c(x) = e^{\frac{\ln x}{x}} = \sqrt[x]{x}$$

and the derivative of this is positive, so c is strictly increasing. Thus there exists a unique α with $c(\alpha) = \frac{1}{3}$, and $x < \alpha \implies c(x) < \frac{1}{3}$. Similarly, $x > 1 \implies c(x) > 1$. So we can explicitly find the minimum of g_x:

$$\begin{aligned} \min_{\frac{1}{3}\le y\le 1} g_x(y) &= \begin{cases} f(x, \frac{1}{3}), & x < \alpha \\ f(x, c(x)), & \alpha \le x \le 1 \\ f(x,1), & x > 1 \end{cases} \\ &= \begin{cases} x(\frac{\ln\frac{1}{3}}{3} - \frac{1}{3}) - \frac{\ln x}{3}, & x < \alpha \\ x(\sqrt[x]{x}\frac{\ln x}{x} - \sqrt[x]{x}) - \sqrt[x]{x}\ln x, & \alpha \le x \le 1 \\ -x - \ln x, & x > 1 \end{cases} \\ &= \begin{cases} \frac{-x\ln 3 - x - \ln x}{3}, & x < \alpha \\ -x\sqrt[x]{x}, & \alpha \le x \le 1 \\ -x - \ln x, & x > 1 \end{cases} \end{aligned}$$

Call the above function $h(x)$. Then

$$\begin{aligned} h'(x) &= \begin{cases} \frac{-\ln 3 - 1}{3} - \frac{1}{3x}, & x < \alpha \\ \frac{d}{dx}(-xc(x)), & \alpha < x < 1 \\ -1 - \frac{1}{x}, & x > 1 \end{cases} \\ &= \begin{cases} \frac{-\ln 3 - 1}{3} - \frac{1}{3x}, & x < \alpha \\ -c(x) - xc'(x), & \alpha < x < 1 \\ -1 - \frac{1}{x}, & x > 1 \end{cases} \end{aligned}$$

[1] Credit for work past this part to the solution presented here

Now, since $D_{2,2}f(x,y) \neq 0$ for all x, y, part a) applies and

$$c'(x) = -\frac{D_{2,1}f(x,c(x))}{D_{2,2}f(x,c(x))} = -\frac{\ln c(x) - \frac{1}{x}}{\frac{x}{c(x)}} = -\frac{\frac{\ln x - 1}{x}}{\frac{x}{c(x)}} = -c(x)\frac{\ln x - 1}{x^2}$$

Thus

$$h'(x) = \begin{cases} \frac{-\ln 3 - 1}{3} - \frac{1}{3x}, & x < \alpha \\ -c(x)\frac{x+1-\ln x}{x}, & \alpha < x < 1 \\ -1 - \frac{1}{x}, & x > 1 \end{cases}$$

Note that $x > \ln x$, so $\frac{x+1-\ln x}{x} > 0$ and $c(x) > 0$, so $h'(x)$ is negative everywhere (except possibly the boundary points $\alpha, 1$, but it is continuous there). Thus the minimum of h on $[\frac{1}{2}, 2]$ is given when $x = \frac{1}{2}$. To check whether $\frac{1}{2} < \alpha$, simply note that $c(\frac{1}{2}) = \frac{1}{4} < \frac{1}{3}$, so $\frac{1}{2} < \alpha$. Thus

$$\max_{\frac{1}{2} \leq x \leq 2} \left(\min_{\frac{1}{2} \leq y \leq 2} f(x,y) \right) = h\left(\frac{1}{2}\right) = \frac{-\ln 3 - 1 - 2\ln\frac{1}{2}}{6} = \frac{\ln\frac{3}{4} - 1}{6}$$

1.3 Chapter 3 Exercises

Exercise 3-1 Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} 0, & x \in [0, \frac{1}{2}) \\ 1, & x \in [\frac{1}{2}, 1] \end{cases}$$

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

Proof. Let $\varepsilon > 0$. Choose a partition $\mathcal{P}$ with subrectangles given by

$$A = \left[0, \frac{1}{2} - \frac{\varepsilon}{2}\right] \times [0,1]$$
$$B = \left[\frac{1}{2} - \frac{\varepsilon}{2}, \frac{1}{2} + \frac{\varepsilon}{2}\right] \times [0,1]$$
$$C = \left[\frac{1}{2} + \frac{\varepsilon}{2}, 1\right] \times [0,1]$$

Then

$$m_A(f) = M_A(f) = 0$$
$$m_B(f) = 0, M_B(f) = 1$$
$$m_C(f) = M_C(f) = 1$$

and

$$v(B) = \varepsilon$$

So

$$\begin{aligned} U(f,\mathcal{P}) - L(f,\mathcal{P}) &= v(B)(M_B(f) - m_B(f)) \\ &= v(B) = \varepsilon \end{aligned}$$

So f is integrable by the alternate criterion for integrability. Moreover,

$$U(f,\mathcal{P}) = v(A)M_A(f) + v(B)M_B(f) + v(C)M_C(f) = v(B \sqcup C) = \frac{1}{2} + \frac{\varepsilon}{2}$$

and similarly

$$L(f,\mathcal{P}) = \frac{1}{2} - \frac{\varepsilon}{2}$$

So $L \geq \frac{1}{2}$ and $U \leq \frac{1}{2}$, but we know that $U = L$ so $\int_A f = \frac{1}{2}$. $\square$

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

Proof. Refer to Exercise 3-3. Its proof does not depend on this problem, and we will use the fact that $\int_A f + g = \int_A f + \int_A g$ when f, g are integrable.

Let $\varepsilon > 0$ be arbitrary. We aim to show that $g - f$ is integrable with $\int_A g - f = 0$. Since $g \neq f$ at only finitely many points, it is bounded. Let $\mu = \max\{|f - g|\}$. Let $p_1, \ldots, p_k$ be those points where $g - f \neq 0$. Let $S_1, \ldots, S_k$ be the subrectangles they are in for a given partition (pick them small enough that they are distinct). Then choose $\mathcal{P}$ such that

$$\sum_{i=1}^{k} v(S_i) < \varepsilon$$

Then

$$\begin{aligned} U(g-f,\mathcal{P}) - L(g-f,\mathcal{P}) &= \sum_{S \in \mathcal{P}} [M_S(g-f) - m_S(g-f)]v(S) \\ &= \sum_{i=1}^{k} [M_{S_i}(g-f) - m_{S_i}(g-f)]v(S_i) \\ &= \sum_{i=1}^{k} v(S_i) \\ &< \varepsilon \end{aligned}$$

So $g - f$ is integrable and a similar argument shows $\int_A g - f = 0$. So $\int_A g = \int_A g - f + f = \int_A g - f + \int_A f = \int Af$. $\square$

Exercise 3-3 Let $f, g : A \to \mathbb{R}$ be integrable.

(a) For any partition $\mathcal{P}$ of A and subrectangle $S \in \mathcal{P}$, show that

$$m_S(f) + m_S(g) \leq m_S(f+g)$$

and

$$M_S(f+g) \leq M_S(f) + M_S(g)$$

so that

$$L(f, \mathcal{P}) + L(g, \mathcal{P}) \leq L(f+g, \mathcal{P})$$

and

$$U(f+g, \mathcal{P}) \leq U(f, \mathcal{P}) + U(g, \mathcal{P})$$

(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 $\int_A cf = c\int_A f$.

(a) *Proof.* Let $S \in \mathcal{P}$. Then for any point $x \in S$, we have

$$(f+g)(x) = f(x) + g(x) \geq m_S(f) + m_S(g)$$

Thus

$$m_S(f) + m_S(g) \leq m_S(f+g)$$

Similarly,

$$M_S(f+g) \leq M_S(f) + M_S(g)$$

Thus we have

$$\begin{aligned} L(f, \mathcal{P}) + L(g, \mathcal{P}) &= \sum_{S \in \mathcal{P}} v(S)[m_S(f) + m_S(g)] \\ &\leq \sum_{S \in \mathcal{P}} v(S) m_S(f+g) \\ &= L(f+g, \mathcal{P}) \end{aligned}$$

Similarly,

$$U(f+g, \mathcal{P}) \leq U(f, \mathcal{P}) + U(g, \mathcal{P})$$

□

(b) *Proof.* Let $\varepsilon > 0$ be arbitrary. Pick $\mathcal{P}_1, \mathcal{P}_2$ such that

$$\begin{aligned} U(f_1, \mathcal{P}_1) - L(f_1, \mathcal{P}_1) &< \frac{\varepsilon}{2} \\ U(f_2, \mathcal{P}_2) - L(f_2, \mathcal{P}_2) &< \frac{\varepsilon}{2} \end{aligned}$$

Let $\mathcal{Q}$ be the common refinement of $\mathcal{P}_1, \mathcal{P}_2$. Then

$$\begin{aligned}
U(f_1+f_2,\mathcal{Q}) - L(f_1+f_2,\mathcal{Q}) &= \sum_{S\in\mathcal{Q}} v(S)[M_S(f_1+f_2) - m_S(f_1+f_2)] \\
&\le \sum_{S\in\mathcal{Q}} v(S)[M_S(f_1) + M_S(f_2) - m_S(f_1) - m_S(f_2)] \\
&= U(f_1,\mathcal{Q}) + U(f_2,\mathcal{Q}) - L(f_1,\mathcal{Q}) - L(f_2,\mathcal{Q}) \\
&\le U(f_1,\mathcal{P}_1) - L(f_1,\mathcal{P}_1) + U(f_2,\mathcal{P}_2) - L(f_2,\mathcal{P}_2) \\
&< \varepsilon
\end{aligned}$$

So f_1+f_2 is integrable and a similar argument shows $\int_A f_1 + f_2 = \int_A f_1 + \int_A f_2$. $\square$

(c) *Proof.* Let $\mathcal{P}$ be a partition and let $S \in \mathcal{P}$. Since S is a closed rectangle, it is compact, so there exists $x \in S$ with $f(x) = M_S(f)$. Then $(cf)(x) = cM_S(f)$ so $M_S(cf) \ge cM_S(f)$. But for any $y \in S$, we also have $(cf)(y) = cf(y) \le cM_S(f)$ so $M_S(cf) = cM_S(f)$. Similarly, $m_S(cf) = cm_S(f)$.

Now, let $\varepsilon > 0$. Then there exists a partition $\mathcal{P}$ with

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) < \frac{\varepsilon}{c}$$

Then we have

$$\begin{aligned}
U(cf,\mathcal{P}) - L(cf,\mathcal{P}) &= \sum_{S\in\mathcal{P}} v(S)[M_S(cf) - m_S(cf)] \\
&= \sum_{S\in\mathcal{P}} cv(S)[M_S(f) - m_S(f)] \\
&= c[U(f,\mathcal{P}) - L(f,\mathcal{P})] \\
&< \varepsilon
\end{aligned}$$

So that cf is integrable. Now, let $\varepsilon > 0$ be arbitrary. Then there exists a partition $\mathcal{P}$ such that

$$U(f,\mathcal{P}) \le \int_A f + \frac{\varepsilon}{c}$$

Then we have

$$U(cf,\mathcal{P}) \le c\int_A +\varepsilon$$

So $\int_A cf = c\int_A f$. $\square$

Exercise 3-4 Let $f : A \to \mathbb{R}$ and let $\mathcal{P}$ be a partition of A. Show that f is integrable if and only if, for each subrectangle $S \in \mathcal{P}$ the restriction $f|_S$ of f to S is integrable, and in this case $\int_A f = \sum_S \int_S f|_S$.

Proof. ($\implies$) Suppose that f is integrable on A, and let $\mathcal{P}$ be given. Let $\varepsilon > 0$. Then there exists a partition $\mathcal{P}'$ of A with

$$U(f, \mathcal{P}') - L(f, \mathcal{P}') < \varepsilon$$

Now let $\mathcal{Q}$ be the common refinement of $\mathcal{P}$ and $\mathcal{P}'$. Then each subrectangle of $\mathcal{Q}$ is entirely contained within a subrectangle of $\mathcal{P}$. In other words, for any $S \in \mathcal{P}$, we may enumerate $S_1, \ldots, S_k \in \mathcal{Q}$ such that $S_1 \sqcup \ldots \sqcup S_k = S$, which means that $\mathcal{S} = \{S_1, \ldots, S_k\}$ is a partition of S. Thus

$$\begin{aligned}
U(f|_S, \mathcal{S}) - L(f|_S, \mathcal{S}) &= \sum_{S' \in \mathcal{S}} v(S')[M_{S'}(f|_S) - m_{S'}(f|_S)] \\
&\leq v(S)[M_S(f) - m_S(f)] \\
&\leq \sum_{S'' \in \mathcal{P}} v(S'')[M_{S''}(f) - m_{S''}(f)] \\
&= U(f, \mathcal{P}) - L(f, \mathcal{P}) \\
&< \varepsilon
\end{aligned}$$

So $f|_S$ is integrable on S.

($\impliedby$) Let $\mathcal{P}$ be given, and suppose each $f|_S$ is integrable on the resepctive S. Let $\varepsilon > 0$. Then let N be the number of subrectangles in the partition $\mathcal{P}$. For each S, pick a partition $\mathcal{P}^S$ such that

$$U(f|_S, \mathcal{P}^S) - L(f|_S, \mathcal{P}^S) < \frac{\varepsilon}{N}$$

Now, suppose that $\mathcal{P}^S = (\mathcal{P}_1^S, \ldots, \mathcal{P}_n^S)$. Then

$$\mathcal{Q}_1 := \bigcup_{S \in \mathcal{P}} \mathcal{P}_1^S$$

is a partition of $[a_1, b_1]$. Let $\mathcal{Q} := (\mathcal{Q}_1, \ldots, \mathcal{Q}_n)$. Then $\mathcal{Q}$ is a refinement of $\mathcal{P}$, and moreover, for any $S \in \mathcal{P}$, $\mathcal{Q}^S$ (which is the collection of subrectangles in $\mathcal{Q}$ which are contained in S) is a refinement of $\mathcal{P}^S$. Thus

$$\begin{aligned}
U(f, \mathcal{Q}) - L(f, \mathcal{Q}) &= \sum_{S' \in \mathcal{Q}} v(S')[M_{S'}(f) - m_{S'}(f)] \\
&= \sum_{S \in \mathcal{P}} \sum_{S'' \in \mathcal{Q}^S} v(S'')[M_{S''}(f) - m_{S''}(f)] \\
&\leq \sum_{S \in \mathcal{P}} \sum_{S'' \in \mathcal{P}^S} v(S'')[M_{S''}(f) - m_{S''}(f)] \\
&= \sum_{S \in \mathcal{P}} [U(f|_S, \mathcal{P}^S) - L(f|_S, \mathcal{P}^S)] \\
&< \sum_{S \in \mathcal{P}} \frac{\varepsilon}{N} \\
&= \varepsilon
\end{aligned}$$

So f is integrable on A. A similar argument shows that $\int_A f = \sum_S \int_S f|_S$. □

Exercise 3-5 Let $f, g : A \to \mathbb{R}$ be integrable and suppose $f \leq g$. Show that $\int_A f \leq \int_A g$.

Proof. Let $\mathcal{P}$ be a partition of A. Then for any $S \in \mathcal{P}$, $M_S(f) \leq M_S(g)$. Thus

$$U(f, \mathcal{P}) = \sum_{S \in \mathcal{P}} v(S) M_S(f) \leq \sum_{S \in \mathcal{P}} v(S) M_S(g) = U(g, \mathcal{P})$$

Since we know f and g are integrable, we conclude that

$$\int_A f = \inf U(f, \mathcal{P}) \leq \inf U(g, \mathcal{P}) = \int_A g \qquad \square$$

Exercise 3-6 If $f : A \to \mathbb{R}$ is integrable, show that $|f|$ is integrable and $|\int_A f| \leq \int_A |f|$.

Proof. Let $\varepsilon > 0$. Let $\mathcal{P}$ be a partition such that

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon$$

Let $S \in \mathcal{P}$. If $M_S(f) \geq m_S(f) \geq 0$, then $M_S(|f|) = M_S(f)$ and $m_S(|f|) = m_S(f)$. If $m_S(f) \leq M_S(f) \leq 0$, then $M_S(|f|) = -m_S(f)$ and $m_S(|f|) = -M_S(f)$. If $M_S(f) > 0$ and $m_S(f) < 0$, then I claim that $M_S(|f|) \leq \max\{|M_S(f)|, |m_S(f)|\}$.

To see this, note that for any $x \in S$, if $f(x) < 0$ then $|f(x)| = -f(x) \leq -m_S(f) = |m_S(f)|$. If $f(x) > 0$, then $|f(x)| = f(x) \leq M_S(f) = |M_S(f)|$. So $M_S(|f|) \leq \max\{|M_S(f)|, |m_S(f)|\}$. Using the fact that $m_S(|f|) \geq 0$, we have

$$\begin{aligned}
M_S(|f|) - m_S(|f|) &\leq M_S(|f|) \\
&\leq \max\{|M_S(f)|, |m_S(f)|\} \\
&= \begin{cases} M_S(f), & |M_S(f)| \geq |m_S(f)| \\ -m_S(f), & |m_S(f)| > |M_S(f)| \end{cases} \\
&\leq M_S(f) - m_S(f)
\end{aligned}$$

As a result, we have the following:

$$\begin{aligned}
M_S(|f|) - m_S(|f|) &\leq \begin{cases} M_S(f) - m_S(f), & M_S(f) \geq m_S(f) \geq 0 \\ -m_S(f) - (-M_S(f)), & m_S(f) \leq M_S(f) \leq 0 \\ M_S(f) - m_S(f), & M_S(f) > 0, m_S(f) < 0 \end{cases} \\
&= M_S(f) - m_S(f)
\end{aligned}$$

Thus, we have

$$\begin{aligned} U(|f|,\mathcal{P}) - L(|f|,\mathcal{P}) &= \sum_{S\in\mathcal{P}} v(S)[M_S(|f|) - m_S(|f|)] \\ &\leq \sum_{S\in\mathcal{P}} v(S)[M_S(f) - m_S(f)] \\ &= U(f,\mathcal{P}) - L(f,\mathcal{P}) \\ &< \varepsilon \end{aligned}$$

So $|f|$ is integrable.

For any partition $\mathcal{P}$, and any $S \in \mathcal{P}$, we showed that $M_S(|f|) \leq \max\{|M_s(f)|, |m_S(f)|\}$. However, we can make a stronger statement, that $M_S(|f|) = \max\{|M_S(f)|, |m_S(f)|\}$. Indeed, since S is compact there exists $x, y \in S$ with $f(x) = M_S(f)$ and $f(y) = m_S(f)$. Then $|f|(x) = |M_S(f)|$ and $|f|(y) = |m_S(f)|$ so $|f|$ attains the value of $\max\{|M_S(f)|, |m_S(f)|\}$. Thus $|M_S(f)| \leq M_S(|f|)$. So

$$\left|\int_A f\right| \leq |U(f,\mathcal{P})| = \left|\sum_{S\in\mathcal{P}} v(S)M_S(f)\right| \leq \sum_{S\in\mathcal{P}} v(S)|M_S(f)| \leq \sum_{S\in\mathcal{P}} v(S)M_S(|f|) = U(|f|,\mathcal{P})$$

So for any partition $\mathcal{P}$, $U(|f|,\mathcal{P}) \geq |\int_A f|$ so $\int_A |f| \geq |\int_A f|$. $\square$

Exercise 3-7 Let $f : [0,1] \times [0,1] \to \mathbb{R}$ be defined by

$$f(x,y) = \begin{cases} 0, & x \notin \mathbb{Q} \text{ or } y \notin \mathbb{Q} \\ \frac{1}{q}, & x \in \mathbb{Q}, y = \frac{p}{q} \in \mathbb{Q} \end{cases}$$

where we assume that $y = \frac{p}{q}$ is given in lowest terms. Show that f is integrable and $\int_{[0,1]\times[0,1]} f = 0$.

Proof. First, note that for any partition $\mathcal{P}$ the density of $\mathbb{Q}$ implies that $L(f,\mathcal{P}) = 0$. So it suffices to show that $U = 0$.

Let $\varepsilon > 0$. Pick a partition $\mathcal{P}$ as follows: Choose N large enough that

$$\frac{1}{N} < \frac{\varepsilon}{2}$$

Then there are finitely many $y = p/q \in \mathbb{Q}$ such that $q < N$. Denote them by $y_1, \ldots, y_k$. Then pick intervals $I_1, \ldots, I_k$ about each such that the total length of the intervals is less than $\varepsilon/2$ (and such that the I_i are disjoint). Let $\mathcal{P}_2$ be the partition of $[0,1]$ given by these intervals, with the gaps filled in appropriately.

Let $\mathcal{P}_1$ be the single partition $\{0,1\}$. Then $\mathcal{P} = (\mathcal{P}_1, \mathcal{P}_2)$ consists of subrectangles of the form $[0,1] \times I$, where I is either one of the I_i we defined previously, or it is not (in this case, it is a gap between them). Let $\mathcal{L}$ denote the set of all subrectangles of the form $[0,1] \times I_i$, and let $\mathcal{R}$ denote the set of all other subrectangles. Then

$$U(f,\mathcal{P}) = \sum_{S\in\mathcal{P}} v(S)M_S(f) = \sum_{S\in\mathcal{L}} v(S)M_S(f) + \sum_{S\in\mathcal{R}} v(S)M_S(f)$$

Now, if $S \in \mathcal{L}$, then f attains a value of at most 1 on S, so $M_S(f) \leq 1$. But if $M_S(f) \in \mathcal{R}$, then by construction there is no point $(x,y) \in S$ with $y = p/q$ and $q < N$. Thus

$$f(x,y) = \frac{1}{q} < \frac{1}{N} < \frac{\varepsilon}{2}$$

so $M_S(f) \leq \frac{\varepsilon}{2}$. Thus

$$\sum_{S\in\mathcal{L}} v(S)M_S(f) + \sum_{S\in\mathcal{R}} v(S)M_S(f) \leq \sum_{S\in\mathcal{L}} v(S) + \frac{\varepsilon}{2}\sum_{S\in\mathcal{R}} v(S) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Thus $U = 0$. So f is integrable and $\int_{[0,1]\times[0,1]} f = 0$. ☐

Exercise 3-8 Prove that $A = [a_1, b_1] \times \ldots \times [a_n, b_n]$ does not have content zero if $a_i < b_i$ for each i.

Proof. Let $\mathcal{O}$ be a finite cover of A by closed rectangles. Without loss of generality we may assume that each rectangle is contained within A. Then let $T_i = \{t_0^i, \ldots, t_{k_i}^i\}$ be the set of endpoints of the rectangles in the ith direction (that is, if $O = \{[c_1, d_1] \times \ldots \times [c_n, d_n]\} \in \mathcal{O}$, then $c_1, d_1 \in T_1$ and $c_i, d_i \in T_i$ for any i). Without loss of generality we may order them so that $a_i = t_0^i \leq \ldots \leq t_{k_i}^i = b_i$. Then each $v(O_i)$ for $O_i \in \mathcal{O}$ is the sum of $v(A_i)$ for A_i of the form $[t_{j_1-1}^1, t_{j_1}^1] \times \ldots \times [t_{j_n-1}^n, t_{j_n}^n]$. Moreover, each of those rectangles is contained within some O_i. So

$$\sum_{i=1}^{n} v(O_i) \geq \sum_{j=1}^{k_1\cdot\times\cdot k_n} v(A_i) = \prod_{j=1}^{n}(b_j - a_j)$$

So A does not have content zero. ☐

Exercise 3-9

(a) Show that an unbounded set cannot have content zero.

(b) Give an example of a closed set of measure zero which does not have content zero.

(a) *Proof.* Let A be an unboudned set and $\mathcal{O}$ a finite cover of A by closed rectangles. Then there exists a closed rectangle M such that

$$\bigcup_{O\in\mathcal{O}} O \subseteq M$$

But since A is unbounded it contains points outside M. So $\mathcal{O}$ cannot be a cover of A, contradiction. Thus A is in fact not covered by any finite set of closed (or open) rectangles, so it cannot have content zero. ☐

(b) *Proof.* $\mathbb{Q}$ is closed and has measure zero (this follows from the fact that it is countable). However, it is unbounded, and thus does not have content zero by part a). □

Exercise 3-10

(a) If C is a set of content zero, show that the boundary of C has content zero.

(b) Give an example of a bounded set C of measure zero such that the boundary of C does not have measure zero.

(a) *Proof.* Let $\mathcal{O}$ be a finite cover of C by closed rectangles. I claim that $\mathcal{O}$ contains ∂C. To see this, suppose that there exists a point $x \in \partial C$ such that $x \notin O$ for each $O \in \mathcal{O}$. Then

$$x \in \bigcap_{i=1}^{k} \mathbb{R}^n \setminus O_i$$

But since each O_i is closed, $\mathbb{R}^n \setminus O_i$ is open, and this is a finite intersection of open sets, which is open. Then since $x \in \partial C$, there exists a point $y \in C$ with

$$y \in \bigcap_{i=1}^{k} \mathbb{R}^n \setminus O_i$$

But this contradicts the assumption that $\mathcal{O}$ is a cover of C. Thus $\mathcal{O}$ covers ∂C. So any closed cover of C is a cover of ∂C. Then let $\varepsilon > 0$. We may produce a finite cover of C by closed rectangles with total volume less than ε. This cover works for ∂C as well. Thus ∂C has content zero. □

(b) Pick $\mathbb{Q} \cap [0,1]$. This is a bounded set of measure zero. But $\partial(\mathbb{Q} \cap [0,1]) = [0,1]$, which does not have measure zero.

Exercise 3-11 Let A be 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), as in Exercise 1-18. If

$$\sum_{i=1}^{\infty} b_i - a_i < 1$$

show that ∂A does not have measure zero.

Proof. Suppose that ∂A has measure zero. Pick a cover $\mathcal{O}$ of ∂A by open intervals such that

$$\sum_{O \in \mathcal{O}} v(O) < 1 - \sum_{i=1}^{\infty} b_i - a_i$$

which we rewrite as

$$1 > \sum_{O \in \mathcal{O}} + \sum_{i=1}^{\infty} b_i - a_i$$

From Exercise 1-18, we know that $\partial A = [0,1] \setminus A$. So the collection of intervals in $\mathcal{O}$ combined with the open intervals which make up A form a cover of $[0,1]$ by open intervals. Call this cover $\mathcal{O}'$. Then we know

$$\sum_{O \in \mathcal{O}'} v(O) \geq 1$$

But we also have

$$\sum_{O \in \mathcal{O}} v(O) + \sum_{i=1}^{\infty} b_i - a_i \geq \sum_{O \in \mathcal{O}'} v(O)$$

So

$$1 > \sum_{O \in \mathcal{O}} v(O) + \sum_{i=1}^{\infty} b_i - a_i \geq \sum_{O \in \mathcal{O}'} v(O) \geq 1$$

and we conclude that $1 > 1$, contradiction. So ∂A does not have measure zero. $\square$

Exercise 3-12 Let $f : [a,b] \to \mathbb{R}$ be an increasing function. Show that $\{x : f \text{ is discontinuous at } x\}$ has measure zero.

Proof. I claim that for any n, there are at most $n(f(b) - f(a))$ points with $o(f,x) > \frac{1}{n}$.

To prove this, suppose there are more than $n(f(b) - f(a))$ such points, $x_1, \ldots, x_k$. Then we may pick $y_0, \ldots, y_k$ with $a = y_0 < x_1 < y_1 < \ldots < x_k < y_k = b$. Then because f is increasing, for each x_i we have

$$o(f, x_i) \leq f(y_i) - f(y_{i-1})$$

Then by a telescoping argument,

$$\sum_{i=1}^{k} o(f, x_i) \leq f(y_k) - f(y_0) = f(b) - f(a)$$

But we also have

$$\sum_{i=1}^{k} o(f, x_i) \geq \sum_{i=1}^{k} \frac{1}{n} = \frac{k}{n} > \frac{n(f(b) - f(a))}{n} = f(b) - f(a)$$

contradiction. Thus there are at most $n(f(b) - f(a))$ such points. Recall that f is discontinuous at x precisely when $o(f,x) > 0$. But

$$\{x : o(f,x) > 0\} = \bigcup_{n=1}^{\infty} \{x : o(f,x) > \frac{1}{n}\}$$

So $\{x : f \text{ is discontinuous at } x\}$ is the countable union of finite sets and thus has measure zero. $\square$

Exercise 3-13

(a) Show that the collection of all rectangles $[a_1, b_1] \times \ldots \times [a_n, b_n]$ with all a_i and b_i rational can be arranged in a sequence.

(b) If $A \subseteq \mathbb{R}^n$ is any set and $\mathcal{O}$ is an open cover of A, show that there is a sequence $O_1, O_2, \ldots$ of members of $\mathcal{O}$ which also cover A.

(a) *Proof.* This collection may be placed in bijection with $\mathbb{Q}^{2n}$, which is a finite Cartesian product of countable sets, so it is countable. ☐

(b) *Proof.* For each point $x \in A$, $x \in O$ for some $O \in \mathcal{O}$, and O is open, so there exists an open rectangle $R_x \subseteq O$ containing x. Moreover, we demand that each endpoint of R_x is rational. Then the set of $R = \{R_x : x \in A\}$ is a subset of the set of all rectangles with rational endpoints, which we showed is countable. Thus R is countable, so we may order its elements as $R_1, R_2, \ldots$.

We then pick a countable subcover $\mathcal{O}'$ of $\mathcal{O}$ by picking $\mathcal{O}'_1$ such that $R_1 \subseteq \mathcal{O}'_1$, and so on. We may skip terms if R_i is already contained in a previously chosen open set. This gives a countable subcover of R, and R covers A, so this is a countable subcover of A. ☐

Exercise 3-14 Show that if $f, g : A \to \mathbb{R}$ are integrable, then fg is as well.

Proof. Since f and g are both integrable, they are discontinuous on sets $C_1, C_2 \subseteq A$ of measure zero. For any x such that $x \notin C_1$ and $x \notin C_2$, f, g are both continuous at x so fg is continuous at x. Thus C_3, the set of points where fg is continuous, is a subset of $C_1 \cup C_2$ and has measure zero. So fg is integrable. ☐

Exercise 3-15 Show that if C has content zero, then $C \subseteq A$ for some closed rectangle A and C is Jordan measurable with $\int_A \chi_C = 0$

Proof. We showed in Exercise 3-9 part a) that any unbounded set does not have content zero. So $C \subseteq A$ for a closed rectangle A. We showed in Exercise 3-10 part a) that ∂C has content zero whenever C has content zero. So C is Jordan-measurable.

Now pick a partition $\mathcal{P}$ of A. For every subrectangle S of $\mathcal{P}$, we cannot have $S \subseteq C$, since otherwise C would not have content zero. So $m_S(\chi_C) = 0$ for each S and thus $L(f, \mathcal{P}) = 0$. This is true for all partitions $\mathcal{P}$, so

$$\int_A \chi_C = L = 0$$ ☐

Exercise 3-16 Give an example of a bounded set C of measure zero such that $\int_A \chi_C$ does not exist.

Set $C = \mathbb{Q} \cap [0,1]$. Then χ_C is the Dirichlet function, which is discontinuous on $[0,1]$ (since both irrationals and rationals are dense in $[0,1]$). So χ_C is not discontinuous on a set of measure zero, so $\int_A \chi_C$ does not exist.

Exercise 3-17 If C is a bounded set of measure zero and $\int_A \chi_C$ exists, show that $\int_A \chi_C = 0$.

Proof. See the second paragraph of the argument from Exercise 3-15. □

Exercise 3-18 If $f : A \to \mathbb{R}$ is nonnegative and $\int_A f = 0$, show that $\{x : f(x) \neq 0\}$ has measure zero.

Proof. Consider the set $B_n = \{x : f(x) \geq \frac{1}{n}\}$ for any n. I claim that B_n has content zero. Suppose it does not. Then there exists $\varepsilon > 0$ such that any cover of B_n has total volume no less than ε. Then let $\mathcal{P}$ be any partition. If $\mathcal{S}$ is the collection of subrectangles which intersect B_n, then $M_S(f) \geq \frac{1}{n}$ for any $S \in \mathcal{S}$. So

$$U(f,\mathcal{P}) = \sum_{S \in \mathcal{P}} v(S) M_S(f) \geq \sum_{S \in \mathcal{S}} v(S) M_S(f) \geq \frac{1}{n} \sum_{S \in \mathcal{S}} v(S) \geq \frac{\varepsilon}{n}$$

So $U \geq \frac{\varepsilon}{n} > 0$, but this contradicts the assumption that $\int_A f = 0$. So B_n has content zero. Thus

$$\{x : f(x) \neq 0\} = \bigcup_{n=1}^{\infty} B_n$$

has measure zero. □

Exercise 3-19 Let U be 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), and

$$\sum_{i=1}^{\infty} b_i - a_i < 1$$

as in Exercise 3-11. Show that if $f = \chi_U$ except on a set of measure zero, then f is not integrable on $[0,1]$.

Proof. In Exercise 3-11 we showed that $\partial U = [0,1] \setminus U$ does not have measure zero. χ_U is discontinuous on ∂U, so it is discontinuous on a set that is not of measure zero, and thus not integrable. Then we need to show that f is also discontinuous on a set not of measure zero.

Let $x \in \partial U$, and suppose that $f(x) = \chi_U(x)$. Suppose for contradiction, suppose that f is continuous at x. Since $x \in \partial U$ and $\partial U = [0,1] \setminus U$, $x \notin U$. Thus $f(x) = \chi_U(x) = 0$. If f is continuous at x, then for any $\varepsilon > 0$ there exists a neighborhood around x such that $|f(y)| < \varepsilon$ for y in the neighborhood. We will show that this is not the case.

Let $\varepsilon = \frac{1}{2}$. Let V be any neighborhood around x contained in $[0,1]$. Then there exists a rational $q \in V$. $q \in U$ which is open, so there exists an open rectangle R containing q contained in $U \cap V$. So $\chi_U = 1$ on an open rectangle within V. So if $|f(y)| < \varepsilon$ for any $y \in V$, we must have $f \neq \chi_U$ on R. But R is not a set of measure zero, so this contradicts the assumption that $f = \chi_U$ on a set of measure zero. So f is not continuous at x.

We have shown that for any $x \in \partial U$ such that $f(x) = \chi_U(x)$, f is discontinuous at x. Then we must show that the set of $x \in \partial U$ with $f(x) = \chi_U(x)$ does not have measure zero.

Suppose that it does. Let $\varepsilon > 0$. Then there exists a cover $\mathcal{U}$ of $\{x \in \partial U : f(x) = \chi_U(x)\}$ by open intervals with total length less than $\varepsilon/2$. We also know that $\{x \in [0,1] : f(x) \neq \chi_U(x)\}$ has measure zero by assumption, so $\{x \in \partial U : f(x) \neq \chi_U(x)\}$ also has measure zero and we may cover it by an open cover $\mathcal{O}$ with total length less than $\varepsilon/2$.

Now for any $x \in \partial U$, we must have $f(x) = \chi_U(x)$ or $f(x) \neq \chi_U(x)$, so $\mathcal{U} \cup \mathcal{O}$ covers ∂U. Now we have

$$\sum_{(c_i,d_i)\in\mathcal{U}\cup\mathcal{O}} d_i - c_i \leq \sum_{(c_i,d_i)\in\mathcal{U}} d_i - c_i + \sum_{(c_i,d_i)\in\mathcal{O}} d_i - c_i < \varepsilon$$

So ∂U has measure zero. But in Exercise 3-11 we showed that this is not the case. So the assumption that $\{x \in \partial U : f(x) = \chi_U(x)\}$ has measure zero is incorrect. But we showed that f is discontinuous on this set, and it does not have measure zero, so f is not integrable on $[0,1]$. □

Exercise 3-20 Show that an increasing function $f : [a,b] \to \mathbb{R}$ is integrable on $[a,b]$.

Proof. In Exercise 3-12 we showed that f is discontinuous on a set of measure zero. So it is integrable on $[a,b]$. □

Exercise 3-21 If A is a closed rectangle, show that $C \subseteq A$ is Jordan-measurable if and only if for every $\varepsilon > 0$ there is a partition $\mathcal{P}$ of A such that

$$\sum_{S\in\mathcal{S}_1} v(S) - \sum_{S\in\mathcal{S}_2} v(S) < \varepsilon$$

where $\mathcal{S}_1$ consists of all subrectangles intersecting C and $\mathcal{S}_2$ all subrectangles contained in C.

We first prove the following fact:

Lemma

If $A \subseteq \mathbb{R}^n$ and $x \in \operatorname{int} A, y \in \operatorname{ext} A$, then there exists $z = tx + (1-t)y$ with $0 \leq t \leq 1$ such that $z \in \partial A$. (Intuitively, this z lies along the line segment between x and y).

Proof. To see this, first note for sufficiently small $t > 0$, $tx + (1-t)y \in A$ since $x \in \operatorname{int} A$. Thus the set $\{0 \le t \le 1 : tx + (1-t)y \in A\}$ is nonempty. Moreover, it is clearly bounded. Then let

$$t' = \sup\{0 \le t \le 1 : tx + (1-t)y \in A\}$$

Now, first note that $t' < 1$. This is because $y \in \operatorname{ext} A$, so there exists a ball around y entirely contained in $\mathbb{R}^n \setminus A$.

I claim that $z = t'x + (1-t')y \in \partial A$. To see this, let $B_r(z)$ be any open ball around z. $B_r(z)$ contains a point in A, as we can simply pick $tx + (1-t)y$ for $t \le t'$ such that $t' - t < r$. Then we need to show that $B_r(z)$ contains a point in $\mathbb{R}^n \setminus A$.

Let $z' = (t' + \varepsilon)x + (1 - t' - \varepsilon)y$, where $\varepsilon < r$ and $t' + \varepsilon < 1$ (possible because $t' < 1$). Then $|z - z'| = \varepsilon < r$, so $z' \in B_r(z)$. But $t' + \varepsilon > t'$, so $t' + \varepsilon \notin \{0 \le t \le 1 : tx + (1-t)y \in A\}$. Since we provided that $t' + \varepsilon < 1$, we conclude that $z' \notin A$. So $z \in \partial A$. $\square$

Now, continuing to the main proof:

Proof. ($\implies$) Suppose that $C \subseteq A$ is Jordan-measurable. ∂C has measure zero, and is compact, so we may pick a finite collection of closed rectangles $\mathcal{O}$ whose interiors cover ∂C with total volume is less than ε. Then apply Lemma **??** to pick a partition $\mathcal{P}$ such that every subrectangle $S \in \mathcal{P}$ is either contained in some $O \in \mathcal{O}$ or does not intersect ∂C. If $S \in \mathcal{P}$ and S intersects C but is not contained in C, I claim that there exists $z \in S$ with $z \in \partial C$.

Indeed, we can pick $x, y \in S$ such that $x \in C$ and $y \notin C$. Then if either of these points is in ∂C, then we are done. Otherwise, $x \in \operatorname{int} C$ and $y \in \operatorname{ext} C$. By the Lemma, there exists $z = tx + (1-t)y$ with $0 \le t \le 1$ such that $z \in \partial C$. Since S is convex, $z \in S$. So the claim is proved. Then S intersects ∂C, so we must have $S \subseteq O$ for some $O \in \mathcal{O}$. Thus

$$\sum_{S \in \mathcal{S}_1} v(S) - \sum_{S \in \mathcal{S}_2} v(S) \le \sum_{O \in \mathcal{O}} v(O) < \varepsilon$$

($\impliedby$) Suppose that $C \subseteq A$ satisfies the condition that for every $\varepsilon > 0$ there is a partition $\mathcal{P}$ such that

$$\sum_{S \in \mathcal{S}_1} v(S) - \sum_{S \in \mathcal{S}_2} v(S) < \varepsilon$$

I claim that $\mathcal{S}_1 \setminus \mathcal{S}_2$ covers ∂C. To see this, let $x \in \partial C$. Then $x \in S$ for some $S \in \mathcal{P}$. Then $S \in \mathcal{S}_1$ or $S \notin \mathcal{S}_1$. But if $S \notin \mathcal{S}_1$, then there exists an open rectangle (S) around x entirely contained in $\operatorname{ext} C$, contradicting $x \in \partial C$. So $S \in \mathcal{S}_1$. But similarly, if $S \in \mathcal{S}_2$ then that contradicts $x \in \partial C$. So $S \in \mathcal{S}_1 \setminus \mathcal{S}_2$.

Thus $\mathcal{S}_1 \setminus \mathcal{S}_2$ covers ∂C, and by assumption it can be made as small as required. So ∂C has measure zero and C is Jordan-measurable. $\square$

Exercise 3-22 If A is Jordan-measurable and $\varepsilon > 0$, show that there exists a compact Jordan-measurable set $C \subseteq A$ such that $\int_{A \setminus C} 1 < \varepsilon$.

Proof. Let A be Jordan-measurable and let $\varepsilon > 0$. Then by Exercise 3-21, we may pick a partition $\mathcal{P}$ such that

$$\sum_{S \in \mathcal{S}_1} v(S) - \sum_{S \in \mathcal{S}_2} v(S) < \varepsilon$$

where $\mathcal{S}_1$ is the collection of subrectangles intersecting A and $\mathcal{S}_2$ is the collection of subrectangles contained in A. Then $C = \bigcup \mathcal{S}_2$ is a union of finite closed rectangles and thus closed. Moreover, $C \subseteq A$. Since A is bounded, C is also bounded and thus compact. So we need to show that it is Jordan-measurable.

I claim that $\partial C \subseteq \bigcup_{S \in \mathcal{S}_2} \partial S$. Let $x \in \partial C$. Then consider the sequence of open balls (B_n), where $B_n = B_{1/n}(x)$. Then for each B_n, there exists some point $y_n \in C$. Each $y_n \in S$ for some $S \in \mathcal{S}_2$, but there are only finitely many such S, so there is some S' such that $y_n \in S'$ for infinitely many n. Moreover, each B_n contains a point not contained in C, which is thus also not contained in S'. So $x \in \partial S'$. Thus the claim is proved.

We take without proof the fact that a rectangle is Jordan-measurable. Then ∂S has measure zero for each $S \in \mathcal{S}_2$, so the finite union $\bigcup_{S \in \mathcal{S}_2} \partial S$ also has measure zero, and thus ∂C has measure zero and C is Jordan measurable

Now, because $C \subseteq A$, we have $\int_{A \setminus C} 1 = \int_A 1 - \int_C 1$. Moreover, $\mathcal{S}_1$ covers A. So

$$\int_A 1 \leq \int_{\bigcup \mathcal{S}_1} 1$$

and thus

$$\int_{A \setminus C} 1 = \int_A 1 - \int_C 1 \leq \int_{\bigcup \mathcal{S}_1} 1 - \int_C 1 = \sum_{S \in \mathcal{S}_1} v(S) - \sum_{S \in \mathcal{S}_2} v(S) < \varepsilon \qquad \square$$

Exercise 3-23 Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$. Let $C \subseteq A \times B$ be a set of $n+m$-dimensional content zero. Let $A' \subseteq A$ be the set of all $x \in A$ such that $\{y \in B : (x, y) \in C\}$ is not of m-dimensional content zero. Show that A' is a set of n-dimensional measure zero.

Proof. First, because C has content zero, ∂C has content zero so χ_C is integrable on $A \times B$ and $\int_{A \times B} \chi_C = 0$. Let $\mathcal{U}(x) = \mathbf{U} \int_B \chi_C(x, y) dy$. Then by Fubini's Theorem,

$$\int_{A \times B} \chi_C = \int_A \mathcal{U} = 0$$

Now, fix some $x \in A$, and let $\mathcal{P}$ be a partition of B.

If $x \in A'$, then there exists some $\varepsilon_x > 0$ such that any finite cover of $\{y \in B : (x, y) \in C\}$ by closed rectangles has total length at least ε_x. Let $\mathcal{S}_1$ be the collection of subrectangles S in $\mathcal{P}$ that intersect $\{y \in B : (x, y) \in C\}$. Because $\mathcal{S}_1$ is a finite cover of $\{y \in B : (x, y) \in C\}$,

$$U(\chi_C, \mathcal{P}) = \sum_{S \in \mathcal{S}_1} M_S(\chi_C) v(S) = \sum_{S \in \mathcal{S}_1} v(S) \geq \varepsilon_x$$

Then $\mathcal{U}(x) = \mathbf{U} \int_B \chi_C \geq \varepsilon_x$.

Now, $\mathcal{U}$ is clearly nonnegative, and we know that $\int_A \mathcal{U} = 0$. So by Exercise 3-18, $\{x : \mathcal{U}(x) \neq 0\}$ has measure zero. But we just showed that $A' \subseteq \{x : \mathcal{U}(x) \neq 0\}$, so A' has measure zero. $\square$

Exercise 3-24 Let $C \subseteq [0,1] \times [0,1]$ be the union of all $\{p/q\} \times [0, 1/q]$, where p/q is a rational in $[0,1]$ in lowest terms. Show that it is not true that the set A' in Exercise 3-23 has content zero.

Proof. First we show that C has content zero. Let $\varepsilon > 0$. Then let

$$R_0 = [0,1] \times \left[0, \frac{\varepsilon}{2}\right]$$

Then there a finite number of rationals p/q such that $\{p/q\} \times [0, 1/q]$ is not contained in R_0. Call these $r_1, \ldots, r_k = p_1/q_1, \ldots, p_k/q_k$. Then for $1 \leq i \leq k$, let

$$R_i = \left[\frac{p_i}{q_i} - \frac{q_i \varepsilon}{2^{i+1}}, \frac{p_i}{q_i} + \frac{q_i \varepsilon}{2^{i+1}}\right] \times \left[0, \frac{1}{q_i}\right]$$

Letting $\mathcal{R} = \{R_0, R_1, \ldots, R_k\}$, $\mathcal{R}$ is a finite cover of C by closed rectangles with

$$\sum_{R \in \mathcal{R}} v(R) = v(R_0) + \sum_{i=1}^{k} v(R_i) = \frac{\varepsilon}{2} + \sum_{i=1}^{k} \frac{\varepsilon}{2^{i+1}} \leq \frac{\varepsilon}{2} + \sum_{i=1}^{\infty} \frac{\varepsilon}{2^{i+1}} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

So C has content zero.

But for each rational $p/q \in [0,1]$, the set $\{y \in [0,1] : (p/q, y) \in C\}$ is simply the set $[0, 1/q]$, which does not have content zero. So $A' = \mathbb{Q} \cap [0,1]$, which does not have content zero. $\square$

Exercise 3-25 Use induction on n to show that $[a_1, b_1] \times \ldots \times [a_n, b_n]$ is not a set of measure zero (or content zero) if $a_i < b_i$.

Proof. In the base case, $n = 1$, let $\mathcal{U}$ be a cover of $[a_1, b_1]$ by open intervals. Since $[a_1, b_1]$ is compact, we can assume $\mathcal{U}$ is finite. From here the base case proceeds as in Exercise 3-8.

Now suppose the theorem is true for n, and we will prove it for $n+1$. Then $[a_1, b_1] \times \ldots [a_{n+1}, b_{n+1}] = ([a_1, b_1] \times \ldots \times [a_n, b_n]) \times [a_{n+1}, b_{n+1}]$. Let $A = [a_1, b_1] \times \ldots \times [a_n, b_n]$ and $B = [a_{n+1}, b_{n+1}]$. By Fubini's Theorem[2]

$$\int_{A \times B} 1 = \int_A \left(\int_B 1 \, \mathrm{d}y\right) \mathrm{d}x = \left(\int_A 1 \, \mathrm{d}x\right) \left(\int_B 1 \, \mathrm{d}y\right)$$

[2] Credit for work past this point to https://hidenori-shinohara.github.io/2019/12/23/measure-zero-ex-3-25.html

Now, the constant function 1 is a nonnegative function, and Exercise 3-18 shows that if $\int_A 1 \, \mathrm{d}x = 0$, then 1 is nonzero on a set of measure zero. But 1 is nonzero on A, which is not a set of measure zero by the inductive hypothesis. So

$$\int_A 1 \, \mathrm{d}x > 0$$

and similarly

$$\int_B 1 \, \mathrm{d}y > 0$$

so

$$\int_{A \times B} 1 > 0$$

Now $A \times B$ is bounded. If it has measure zero, then Exercise 3-18 says that $\int_{A \times B} \chi_{A \times B} = \int_{A \times B} 1 = 0$. But this is not the case, so $A \times B$ does not have measure zero. □

Exercise 3-26 Let $f : [a, b] \to \mathbb{R}$ be integrable and nonnegative and let $A_f = \{(x, y) : a \leq x \leq b, 0 \leq y \leq f(x)\}$. Show that A_f is Jordan-measurable and has area $\int_a^b f$.

Proof. Since $f : [a, b] \to \mathbb{R}$ is integrable and nonnegative, there exists $M > 0$ such that $M > f(x)$ for any x.

Claim 1.1

Let

$$\begin{aligned} B &= ([a, b] \times \{0\}) \\ C &= \{(x, f(x)) : x \in [a, b]\} \\ D &= \{a\} \times [0, M] \\ E &= \{b\} \times [0, M] \\ F &= \{x : f \text{ is discontinuous at } x\} \times [0, M] \end{aligned}$$

Then

$$\partial A_f \subseteq B \cup C \cup D \cup E \cup F$$

To prove this, note that any (x, y) satisfies exactly one of the following conditions[3]:

1. $(x, y) \notin [a, b]$.
2. $x = a$.
3. $x = b$.

[3]Strictly speaking, conditions 5 and 8 are both filled by $(x, 0)$ for $x : f(x) = 0$, but this does not detract from the overall argument.

4. $(x, y) \in (a, b), y < 0$.
5. $(x, y) \in (a, b), y = 0$.
6. $(x, y) \in (a, b), 0 < y < f(x)$, f is continuous at x.
7. $(x, y) \in (a, b), 0 < y < f(x)$, f is not continuous at x.
8. $(x, y) \in (a, b), y = f(x)$.
9. $(x, y) \in (a, b), y > f(x)$, f is continuous at x.
10. $(x, y) \in (a, b), f(x) < y \leq M$, f is not continuous at x.
11. $(x, y) \in (a, b), y > M$, f is not continuous at x.

For cases 2, 3, 5, 7, 8, 10, $(x, y) \in B \cup C \cup D \cup E \cup F$. Thus we must show that $(x, y) \notin \partial A_f$ whenever conditions 1, 4, 6, 9, or 11 are met.

Case 1: We can pick an open rectangle R containing (x, y) such that $(x_1, y_1) \in R \implies x_1 \notin [a, b]$. So $(x, y) \in \operatorname{ext} A_f$.

Case 4: We can pick an open rectangle R containing (x, y) such that $(x_1, y_1) \in R \implies y_1 < 0$. So $(x, y) \in \operatorname{ext} A_f$.

Case 6: Since f is continuous at x, there exists an interval $(x - \delta, x + \delta)$ such that $f(x_1) > y + \varepsilon$ whenever $x_1 \in (x - \delta, x + \delta)$, for $\varepsilon > 0$ sufficiently small (where δ is chosen small enough that this makes sense). Then the rectangle $R = (x - \delta, x + \delta) \times (0, y + \varepsilon)$ is an open rectangle containing (x, y) which is contained in A_f. So $(x, y) \in \operatorname{int} A_f$.

Case 9: Similarly to Case 4, since f is continuous at x, there exists an interval $(x - \delta, x + \delta)$ such that $f(x_1) < y - \varepsilon$ whenever $x_1 \in (x - \delta, x + \delta)$ for $\varepsilon > 0$ sufficiently small. Then the rectangle $R = (x - \delta, x + \delta) \times (y - \varepsilon, M)$ shows that $(x, y) \in \operatorname{ext} A$.

Case 11: Similarly to Case 2, we may pick an open rectangle R containing (x, y) such that $(x_1, y_1) \in R \implies y_1 > M \implies (x_1, y_1) \notin A_f$. So $(x, y) \in \operatorname{ext} A_f$.

Thus Claim 1 is proved.

Claim 1.2

The sets B, C, D, E, F each have measure zero.

The line interval $[a, b] \times \{0\}$ has measure zero, as for any $\varepsilon > 0$ we cover it by

$$R_\varepsilon = [a, b] \times \left[-\frac{\varepsilon}{2(b-a)}, \frac{\varepsilon}{2(b-a)} \right]$$

which has $v(R_\varepsilon) = \varepsilon$. So B has measure zero. A similar proof holds for the line segments D and E.

The set $\{x : f \text{ is discontinuous at } x\}$ has measure zero since f is integrable. Let $\varepsilon > 0$. Then we may pick a cover $\mathcal{I}$ of $\{x : f \text{ is discontinuous at } x\}$ by open intervals such that

$$\sum_{(c,d) \in \mathcal{I}} d - c < \frac{\varepsilon}{4M}$$

Then the collection $\mathcal{U}$ of rectangles of the form $(c,d) \times (-\frac{M}{2}, \frac{3M}{2})$ for $(c,d) \in \mathcal{I}$ forms a cover of $\{x : f \text{ is discontinuous at } x\} \times [0, M]$, Moreover, consider the remaining set

$$S = [a,b] \setminus \bigcup_{I \in \mathcal{I}} I$$

Since each I is open, S is closed. It is also bounded, so it is compact. Moreover, f is continuous at each $x \in S$. Since f is continuous on S compact, it is uniformly continuous. Thus we may pick $\delta > 0$ such that

$$x, y \in S, |x-y| < \delta \implies |f(x) - f(y)| < \frac{\varepsilon}{4(b-a)}$$

Moreover, pick δ such that $m\delta = b-a$ for some $m \in \mathbb{N}$. Now let $\delta_i = [a + (i-1)\delta, a + i\delta]$. Then the collection $\{\delta_i\}_{i=1}^m$ partitions the interval $[a,b]$. Now for each i, define the rectangle P_i as follows: if $S \cap \delta_i = \varnothing$, then let $P_i = \delta_i \times \{0\}$. Otherwise, pick $x_i \in S \cap \delta_i$. Then let

$$P_i = \delta_i \times \left[f(x_i) - \frac{\varepsilon}{4(b-a)}, f(x_i) + \frac{\varepsilon}{4(b-a)} \right]$$

Let $\mathcal{P} = \{P_1, \ldots, P_n\}$, and let $\overline{\mathcal{U}} = \{\overline{U} : U \in \mathcal{U}\}$. I claim that $\mathcal{P} \cup \overline{UUU}$ is a cover of $C \cup F$. Indeed, we already showed that $\mathcal{U}$ covers F, so $\overline{\mathcal{U}}$ does as well.

Now, for any $x \in [a,b]$, either $x \in S$ or $x \notin S$. If $x \notin S$, then $x \in I$ for some $I \in \mathcal{I}$ and thus $(x, f(x)) \in U$ for some $U \in \overline{\mathcal{U}}$. On the other hand, if $x \in S$, then $x \in \delta_i$ for some i (this does not require $x \in S$, just $x \in [a,b]$). Then $|x - x_i| < \delta_i$, so

$$|f(x) - f(x_i)| < \frac{\varepsilon}{4(b-a)}$$

so $(x, f(x)) \in P_i$. Thus $\mathcal{P} \cup \overline{\mathcal{U}}$ is a cover of $C \cup F$ by closed rectangles. Lastly, we have

$$\sum_{\overline{U} \in \overline{\mathcal{U}}} = \sum_{U \in \mathcal{U}} v(U) = \sum_{(c,d) \in \mathcal{I}} v((c,d) \times (-\frac{M}{2}, \frac{3M}{2})) = 2M \sum_{(c,d) \in \mathcal{I}} d - c < \frac{\varepsilon}{2}$$

and

$$\sum_{i=1}^m v(P_i) = \sum_{i=1}^m \delta \cdot \frac{\varepsilon}{2(b-a)} = \frac{\varepsilon}{2(b-a)} m\delta = \frac{\varepsilon}{2}$$

so the total volume of $\mathcal{P} \cup \overline{\mathcal{U}}$ is less than ε. Thus $C \cup F$ has measure zero, and C and F each do.

Thus Claim 2 is proved.

Now, by Claim 2, each of B, C, D, E, F has measure zero. So $B \cup C \cup D \cup E \cup F$ has measure zero, and by Claim 1 $\partial A_f \subseteq B \cup C \cup D \cup E \cup F$, so ∂A_f has measure zero. It is also bounded, so A_f is Jordan-measurable.

The last part of the proof is to show that $v(A_f) = \int_a^b f$. Since A_f is Jordan-measurable, χ_{A_f} is integrable on $[a,b] \times [0,M]$. So by Fubini's Theorem,

$$v(A_f) = \int_{[a,b] \times [0,M]} \chi_{A_f} = \int_a^b \left(\mathbf{L} \int_0^M \chi_{A_f}(x,y) \, \mathrm{d}y \right) \mathrm{d}x$$

For each fixed $x \in [a,b]$, $g_x = \chi_{A_f}(x,\cdot)$ is integrable as it is only discontinuous at $f(x)$. Thus

$$\mathbf{L}\int_0^M \chi_{A_f}(x,y)\,\mathrm{d}y = \int_0^M \chi_{A_f}(x,y)\,\mathrm{d}y$$

Moreover,

$$\int_0^M \chi_{A_f}(x,y)\,\mathrm{d}y = \int_0^f (x)1\,\mathrm{d}y = f(x)$$

So we have

$$v(A_f) = \int_a^b \left(\int_0^M \chi_{A_f}(x,y)\,\mathrm{d}y\right)\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x = \int_a^b f \qquad \square$$

Exercise 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. Define $C = \{(x,y) \in [a,b] : y \geq x\}$. Then C has boundary $\partial C = (\{a\} \times [a,b]) \cup ([a,b] \times \{b\}) \cup \{(x,x) : x \in [a,b]\}$ which are all line segments, and thus have measure zero. So C is Jordan-measurable and $\chi_C f$ is integrable on $[a,b] \times [a,b]$. By Fubini's Theorem, since f is continuous,

$$\int_{[a,b]\times[a,b]} \chi_C f = \int_a^b \int_a^b \chi_C(x,y) f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_a^b \int_x^b f(x,y)\,\mathrm{d}y\,\mathrm{d}x$$

But applying it in the opposite order,

$$\int_{[a,b]\times[a,b]} \chi_C f = \int_a^b \int_a^b \chi_C(x,y) f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_a^b \int_a^y f(x,y)\,\mathrm{d}x\,\mathrm{d}y \qquad \square$$

Exercise 3-28 Use Fubini's theorem to prove that $D_{1,2}f = D_{2,1}f$ if both are continuous.

Proof. Suppose that $D_{1,2}f$ and $D_{2,1}$ both exist and are continuous. Then $D_{1,2}f - D_{2,1}f$ is continuous. Suppose there exists a such that $D_{1,2}f(a) - D_{2,1}f(a) > 0$ (for the case < 0 the proof is analogous). Then there exists a rectangle $A = [a,b] \times [c,d]$ containing a such that

$$D_{1,2}f(x) - D_{2,1}f(x) > \varepsilon$$

for any $x \in A$ and $\varepsilon > 0$ smaller than $D_{1,2}f(a) - D_{2,1}f(a)$. Since $D_{1,2}f - D_{2,1}f$ is continuous, it is integrable on A. So

$$\int_A D_{1,2}f - D_{2,1}f \geq \int_A \varepsilon = \varepsilon \int_A 1 = \varepsilon v(A) > 0$$

But by Fubini's Theorem,

$$\begin{aligned}
\int_A D_{1,2}f &= \int_a^b \int_c^d D_{1,2}f(x,y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_a^b \left(\int_c^d \frac{\mathrm{d}}{\mathrm{d}y} D_1 f(x,y)\,\mathrm{d}y \right) \mathrm{d}x \\
&= \int_a^b D_1 f(x,d) - D_1 f(x,c)\,\mathrm{d}x \\
&= f(b,d) - f(b,c) - f(a,d) + f(a,c)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\int_A D_{2,1}f &= \int_c^d \int_a^b D_{2,1}f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_c^d D_2 f(b,y) - D_2 f(a,y)\,\mathrm{d}y \\
&= f(b,d) - f(a,d) - f(b,c) + f(a,c)
\end{aligned}$$

So

$$\begin{aligned}
\int_A D_{1,2}f - D_{2,1}f &= f(b,d) - f(b,c) - f(a,d) + f(a,c) - f(b,d) + f(a,d) + f(b,c) - f(a,c) \\
&= 0
\end{aligned}$$

contradiction. Thus $D_{1,2}f - D_{2,1}f = 0$ and $D_{1,2}f = D_{2,1}f$. ☐

Exercise 3-29 Use Fubini's theorem to derive an expression for the volume of a set of $\mathbb{R}^3$ obtained by revolving a Jordan-measurable set in the yz-plane about the z-axis.

Exercise 3-30 Let $C \subseteq [0,1] \times [0,1]$ contain at most one point on each horizontal and each vertical line, with $\partial C = [0,1] \times [0,1]$, as in Exercise 1-17. Show that

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

but

$$\int_{[0,1]\times[0,1]} \chi_C$$

does not exist.

Proof. Fix some $y \in [0,1]$. Then A intersects $[0,1] \times \{y\}$ at at most one point, so $h_y(x) = \chi_C(x,y)$ is zero everywhere except possibly one point. Thus it is nonzero at a finite number of points, so

$$\int_{[0,1]} \chi_C(x,y)\,\mathrm{d}x = 0$$

so

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

Similarly, for any $x \in [0,1]$, A intersects $\{x\} \times [0,1]$ at at most one point, so $g_x(y) = \chi_C(x,y)$ is nonzero at a finite number of points, so

$$\int_{[0,1]} \chi_C(x,y)\,\mathrm{d}y = 0$$

and

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

On the other hand, $\partial A = [0,1] \times [0,1]$ by assumption, which does not have measure zero and thus χ_C is not integrable on $[0,1] \times [0,1]$. □

Exercise 3-31 If $A = [a_1, b_1] \times \ldots \times [a_n, b_n]$ and $f : A \to \mathbb{R}$ is continuous, define $F : A \to \mathbb{R}$ by

$$F(x) = \int_{[a_1,x_1] \times \ldots \times [a_n,x_n]} f$$

What is $D_i F(x)$ for $x \in \operatorname{int} A$?

Define $G_1 : \mathbb{R} \to \mathbb{R}$ by

$$G_1(y) = F(y, x_2, \ldots, x_n) = \int_{[a_1,y] \times \ldots \times [a_n,x_n]} f$$

and $g_1 : \mathbb{R} \to \mathbb{R}$ by

$$g_1(y) = f(y, x_2, \ldots, x_n)$$

Since f is continuous, we may apply Fubini's theorem to write

$$G_1(y) = \int_{a_1}^{y} \left(\int_{[a_2,x_2] \times \ldots \times [a_n,x_n]} f(y, x^2, \ldots, x^n)\,\mathrm{d}x \right) \mathrm{d}y$$

(where x^i represents a variable being integrated against, as opposed to x_i which is the ith component of x). So by the Fundamental Theorem of Calculus,

$$G_1'(y) = \left(\int_{[a_2,x_2] \times \ldots \times [a_n,x_n]} f(y, x^2, \ldots, x^n)\,\mathrm{d}x \right) = \int_{a_2}^{x_2} \ldots \int_{a_n}^{x_n} f(y, x^2, \ldots, x^n)\,\mathrm{d}x^n \ldots \mathrm{d}x^2$$

We can make a similar argument for g_i for any i, so that

$$D_i F(x) = g_i'(y) = \int_{a_1}^{x_1} \ldots \cancel{\int_{a_i}^{x_i}} \ldots \int_{a_n}^{x_n} f(x^1, \ldots, x^{i-1}, x_i, x^{i+1}, \ldots, x^n)\,\mathrm{d}x^n \ldots \cancel{\mathrm{d}x_i} \ldots \mathrm{d}x^1$$

where the strikethroughs indicate that the ith variables is not integrated against (that is, we integrate against all other variables but hold x_i constant).

Exercise 3-32 Let $f : [a, b] \times [c, d] \to \mathbb{R}$ be continuous and suppose $D_2 f$ 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. Define $g_x(y) : [c, d] \to \mathbb{R}$ by

$$g_x(y) = f(x, y)$$

Then by definition,

$$g_x'(y) = D_2 f(x, y)$$

Since $D_2 f$ is continuous, by the Fundamental Theorem of Calculus,

$$f(x, y) = g_x(y) = g_x(c) + \int_c^y g_x'(t)\,\mathrm{d}t = f(x, c) + \int_c^y D_2 f(x, t)\,\mathrm{d}t$$

So

$$F(y) = \int_a^b \left(f(x, c) + \int_c^y D_2 f(x, t)\,\mathrm{d}t \right) \mathrm{d}x = \int_a^b f(x, c)\,\mathrm{d}x + \int_a^b \int_c^y D_2 f(x, t)\,\mathrm{d}t\,\mathrm{d}x$$

Now, by Fubini's Theorem we have

$$\int_a^b \int_c^y D_2 f(x, t)\,\mathrm{d}t\,\mathrm{d}x = \int_c^y \int_a^b D_2 f(x, t)\,\mathrm{d}x\,\mathrm{d}t$$

so

$$F'(y) = \frac{\mathrm{d}}{\mathrm{d}y} \int_a^b \int_c^y D_2 f(x, t)\,\mathrm{d}t\,\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}y} \int_c^y \int_a^b D_2 f(x, t)\,\mathrm{d}x\,\mathrm{d}t = \int_a^b D_2 f(x, y)\,\mathrm{d}x \qquad \square$$

Exercise 3-33 If $f : [a, b] \times [c, d] \to \mathbb{R}$ is continuous and $D_2 f$ is continuous, define

$$F(x, y) = \int_a^x f(t, y)\,\mathrm{d}t$$

(a) Find $D_1 F$ and $D_2 F$.

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

(a) Define $h_y(x) = f(x, y)$. Let $F_y(x) = F(x, y)$, so that $D_1 F(x, y) = F_y'(x)$. Then

$$F_y(x) = F(x, y) = \int_a^x f(t, y)\,\mathrm{d}t = \int_a^x h_y(t)\,\mathrm{d}t$$

so

$$D_1F(x,y) = F_y'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\int_a^x h_y(t)\,\mathrm{d}t = h_y(x) = f(x,y)$$

Now, define $H_x(y) = F(x,y) = \int_a^x f(t,y)\,\mathrm{d}t$, so that $D_2F(x,y) = H_x'(y)$. By Leibnitz's rule from Exercise 3-32,

$$D_2F(x,y) = H_x'(y) = \int_a^x D_2f(t,y)\,\mathrm{d}t$$

(b) Here we have $G(x) = F(g(x),x)$. By the Chain Rule,

$$G'(x) = D_1F(g(x),x)g'(x) + D_2F(g(x),x) = f(g(x),x)g'(x) + \int_a^x D_2f(t,x)\,\mathrm{d}t$$

Exercise 3-34 Let $g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$ be continuously differentiable and suppose $D_1g_2 = D_2g_1$. As in Exercise 2-21, let

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

Show that $D_1f(x,y) = g_1(x,y)$.

Proof. Differentiating term by term, the Fundamental Theorem of Calculus gives us

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^x g_1(t,0)\,\mathrm{d}t = g_1(x,0)$$

Now, since g_2 is continuously differentiable, it is also continuous, so by Leibnitz's Rule (considering D_1 rather than D_2),

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_0^y g_2(x,t)\,\mathrm{d}t = \int_0^y D_1g_2(x,t)\,\mathrm{d}t$$

By assumption, $D_1g_2 = D_2g_1$, so

$$\int_0^y D_1g_2(x,t)\,\mathrm{d}t = \int_0^y D_2g_1(x,t)\,\mathrm{d}t$$

Then by the Fundamental Theorem of Calculus,

$$D_1f(x,y) = \frac{\mathrm{d}}{\mathrm{d}x}\int_0^x g_1(t,0)\,\mathrm{d}t + \frac{\mathrm{d}}{\mathrm{d}x}\int_0^y g_2(x,t)\,\mathrm{d}t = g_1(x,0) + \int_0^y D_2g_1(x,t)\,\mathrm{d}t = g_1(x,y) \quad \square$$

Exercise 3-35

(a) Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation of one of the following types:

$$\begin{cases} g(e_i) = e_i, & i \neq j \\ g(e_j) = ae_j, \end{cases}$$

$$\begin{cases} g(e_i) = e_i, & i \neq j \\ g(e_j) = e_j + e_k, \end{cases}$$

$$\begin{cases} g(e_k) = e_k, & k \neq i, k \neq j \\ g(e_i) = e_j \\ g(e_j) = e_i \end{cases}$$

If U is a rectangle, show that $v(g(U)) = |\det g| v(U)$.

(b) Prove that $v(g(U)) = |\det g| v(U)$ for any linear transformation $g : \mathbb{R}^n \to \mathbb{R}^n$.

(a) *Proof.* First note that the scaling factor of g is scale invariant, for any of the above cases. For instance, let $U = [a_1, b_1] \times \ldots \times [a_n, b_n]$. Let $\mathbf{x} = (a_1, \ldots, a_n)$. Then let $y \in U$. Since g is linear,

$$g(y) = g(y - \mathbf{x} + \mathbf{x}) = g(y - \mathbf{x}) + g(\mathbf{x})$$

So $g(U) = g(U - \mathbf{x}) + g(\mathbf{x})$, and thus $g(U)$ is a translated version of $g(U - \mathbf{x})$, which has the same volume. Thus we may assume that $U = [0, b_1] \times \ldots \times [0, b_n]$.

Let $\overrightarrow{y}_i = b_i e_i$, so that $\overrightarrow{y}_1, \ldots, \overrightarrow{y}_n$ are the edges of U. Then $g(U)$ is the rectangle with edges given by $g(\overrightarrow{y}_1), \ldots, g(\overrightarrow{y}_n)$.

Case 1: We have

$$g(\overrightarrow{y}_i) = b_i g(e_i) = \begin{cases} b_i e_i, & i \neq j \\ a b_i e_i, & i = j \end{cases}$$

so $g(U) = [0, b_1] \times \ldots \times [0, ab_j] \times \ldots \times [0, b_n]$. Then

$$v(g(U)) = b_1 b_2 \ldots a b_j \ldots b_n = a(b_1 \ldots b_n) = av(U)$$

Now, the matrix of g is

$$[g] = \begin{bmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & 1 \end{bmatrix}$$

so

$$\det g = \det[g] = a$$

Case 2: Since g is linear, it is continuous. Assume without loss of generality that $j = 1$ and $k = 2$. Then $g(U) = V \times [0, b_3] \times \ldots \times [0, b_n]$, where

$$V \subseteq \mathbb{R}^2 = \{(x, y) : 0 \le x \le b_1, x \le y \le x + b_2\}$$

is a rhombus. Then by Fubini's Theorem, (letting M be any rectangle bounding $g(U)$)

$$\begin{aligned}
v(g(U)) &= \int_M \chi_{g(U)} \\
&= \int_0^{b_1} \int_x^{x+b_2} \left(\int_0^{b_3} \ldots \int_0^{b_n} \mathrm{d}x_n \ldots \mathrm{d}x_3 \right) \mathrm{d}y \, \mathrm{d}x \\
&= b_3 \ldots b_n \int_0^{b_1} \int_x^{x+b_2} \mathrm{d}y \, \mathrm{d}x \\
&= b_3 \ldots b_n \int_0^{b_1} b_2 \\
&= b_1 \ldots b_n \\
&= v(U)
\end{aligned}$$

The matrix of g is given by

$$[g] = \begin{bmatrix} 1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & 1 & \vdots \\ \vdots & \ddots & a & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & 1 \end{bmatrix}$$

(where the off-diagonal 1 is an arbitrary off-diagonal location), which has determinant 1.

Case 3: We have

$$g(U) = [0, b_1] \times \ldots \times \underbrace{[0, b_j]}_{i\text{th position}} \times \ldots \times \underbrace{[0, b_i]}_{j\text{th position}} \times \ldots \times [0, b_n]$$

which has $v(g(U)) = b_1 \ldots b_n = v(U)$. The matrix of g is simply the identity matrix with two columns switched, so $\det g = -1$ and $|\det g| = 1$. $\square$

(b) *Proof.* If $\det g = 0$, then $g(U)$ has volume zero for any U. If $\det g \neq 0$, then $\mathrm{RREF}([g]) = I_n$. Moreover, note that the elementary row operations correspond to the following matrices:

$$\left\{\begin{array}{ll}
\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & a & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}, & \text{scaling of a row} \\
\begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}, & \text{row swap} \\
\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & a & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix}, & \text{addition of rows}
\end{array}\right.$$

The first two ERO matrices directly correspond to Cases 1 and 3, respectively.

For the third matrix, suppose the ERO in question sends R_i to $R_i + aR_j$. Then this ERO matrix may be written as $[g_1][g_2][g_3]$, where g_3 scales R_j by a (Case 1), g_2 is a Case 2 transformation which sends e_i to $e_i + e_j$, and g_1 scales R_j by $1/a$ (Case 1).

Thus any invertible transformation has a matrix which may be written as

$$[g] = [g_1] \dots [g_k] \operatorname{RREF}([g]) = [g_1] \dots [g_k]$$

where each of the g_k is of one of the three types considered above. By the property of the determinant,

$$\det[g] = \det([g_1] \dots [g_k]) = \det([g_1]) \dots \det([g_k])$$

By applying part a), we have

$$\begin{aligned} v(g(U)) &= v(g_1(\dots(g_k(U)))) \\ &= |\det g_1| v(g_2(\dots(g_k(U)))) \\ &= |\det g_1| \dots |\det g_k| v(U) \\ &= |\det g_1 \dots \det g_k| v(U) \\ &= |\det g| v(U) \end{aligned}$$

□

Exercise 3-36 (Cavalieri's Principle) Let A and B be Jordan-measurable subsets of $\mathbb{R}^3$. Let $A_c = \{(x,y) : (x,y,c) \in A\}$ and define B_c similarly. Suppose each A_c and B_c are Jordan-measurable and have the same area. Show that A and B have the same volume.

Proof. Let $M = [a_1,b_1]\times[a_2,b_2]\times[a_3,b_3]$ be a closed rectangle which bounds both A and B. Since A is Jordan-measurable, χ_A is integrable on M, and so is χ_B. By Fubini's Theorem,

$$\int_M \chi_A = \int_{a_3}^{b_3} \left(\int_{[a_1,b_1]\times[a_2,b_2]} \chi_A(x,y)\,\mathrm{d}x \right) \mathrm{d}y$$

where our use of the integral sign is justified since A_c is Jordan measurable. Then we may write

$$\int_{[a_1,b_1]\times[a_2,b_2]} \chi_A(x,y)\,\mathrm{d}x = \int_{[a_1,b_1]\times[a_2,b_2]} \chi_{A_y}$$

This is precisely the area of A_y, which by assumption is the area of B_y. So

$$\begin{aligned}
\int_M \chi_A &= \int_{a_3}^{b_3} \left(\int_{[a_1,b_1]\times[a_2,b_2]} \chi_A(x,y)\,\mathrm{d}x \right) \mathrm{d}y \\
&= \int_{a_3}^{b_3} \left(\int_{[a_1,b_1]\times[a_2,b_2]} \chi_{A_y} \right) \mathrm{d}y \\
&= \int_{a_3}^{b_3} v(A_y) \\
&= \int_{a_3}^{b_3} v(B_y) \\
&= \int_{a_3}^{b_3} \left(\int_{[a_1,b_1]\times[a_2,b_2]} \chi_{B_y} \right) \mathrm{d}y \\
&= \int_{a_3}^{b_3} \left(\int_{[a_1,b_1]\times[a_2,b_2]} \chi_B(x,y)\,\mathrm{d}x \right) \mathrm{d}y \\
&= \int_M \chi_B
\end{aligned}$$

so $v(A) = v(B)$. ☐

Exercise 3-37

(a) Suppose that $f : (0,1) \to \mathbb{R}$ is a nonnegative continuous function. Show that

$$\operatorname{ext} \int_{(0,1)} f$$

exists if and only if

$$\lim_{\varepsilon \to 0^+} \int_{\varepsilon}^{1-\varepsilon} f$$

exists.

(b) Define

$$A_n := \left[1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}}\right]$$

Suppose that $f : \mathbb{R} \to \mathbb{R}$ satisfies

$$\int_{A_n} f = \frac{(-1)^n}{n}$$

and $f = 0$ outside of $\bigcup_{i=1}^{\infty} A_i$. Suppose also that f does not change sign on the interiors of any of the A_n. Show that

$$\operatorname{ext} \int_{(0,1)} f$$

does not exist, but

$$\lim_{\varepsilon \to 0^+} \operatorname{ext} \int_{(\varepsilon, 1-\varepsilon)} f = -\ln 2$$

Note: The hypothesis that f does not change sign is not included in Spivak's original exercise. Spivak's exercise is incorrect as written, but this is not the only possible hypothesis to rectify the issue.

(a) *Proof.* ($\Longrightarrow$) Suppose that

$$\operatorname{ext} \int_{(0,1)} f$$

exists. Let Φ be some partition of unity subordinate to an admissible open cover $\mathcal{O}$ of $(0,1)$. Now, let $\varepsilon > 0$. Then let Φ_ε be the finite collection of $\varphi \in \Phi$ which are nonzero on $[\varepsilon, 1-\varepsilon]$. Then we have

$$\int_{\varepsilon}^{1-\varepsilon} f = \int_{\varepsilon}^{1-\varepsilon} f \sum_{\varphi \in \Phi_\varepsilon} \varphi = \sum_{\varphi \in \Phi_\varepsilon} \int_{\varepsilon}^{1-\varepsilon} \varphi f$$

Now, since f is nonnegative, we have

$$\sum_{\varphi\in\Phi_\varepsilon}\int_\varepsilon^{1-\varepsilon}\varphi f \le \sum_{\varphi\in\Phi_\varepsilon}\int_{C_\varphi}\varphi f \le \sum_{\varphi\in\Phi}\int_{C_\varphi}\varphi f = \mathrm{ext}\int_{(0,1)} f$$

So $\int_\varepsilon^{1-\varepsilon} f$ is bounded above. Moreover, let $\varepsilon' < \varepsilon$. Since f is nonnegative, we have

$$\int_\varepsilon^{1-\varepsilon} f \le \int_{\varepsilon'}^{1-\varepsilon'} f$$

so

$$\lim_{\varepsilon\to 0^+}\int_\varepsilon^{1-\varepsilon} f$$

exists.

($\Longleftarrow$) Suppose that

$$\lim_{\varepsilon\to 0^+}\int_\varepsilon^{1-\varepsilon} f$$

exists. For any $n\in\mathbb{N}$, let

$$A_n := \left[\frac{1}{2^{n+1}}, \frac{1}{2^n}\right] \cup \left[1-\frac{1}{2^n}, 1-\frac{1}{2^{n+1}}\right]$$

By Exercise 2-26 there exists a C^∞ function φ_n such that $\varphi_n > 0$ on A_n but $\varphi_n = 0$ outside of some closed set contained in

$$\left(\frac{1}{2^{n+2}}, \frac{1}{2^{n-1}}\right) \cup \left(1-\frac{1}{2^{n-1}}, 1-\frac{1}{2^{n+2}}\right)$$

which can be smoothly extended to have domain $(-1,2)$. Now, $(0,1) = \bigcup_{i=1}^\infty A_i$, so for any $x\in(0,1)$ at least one φ_n is nonzero at x. Moreover, it is clear that only finitely many are nonzero at x. So

$$\sum_{i=1}^\infty \varphi_i(x) > 0$$

and we may define the C^∞ function $\psi_n : (-1,2)\to\mathbb{R}$ by

$$\psi_n(x) = \frac{\varphi_n(x)}{\sum_{i=1}^\infty \varphi_i(x)}$$

Then $\Psi = \{\psi_1, \psi_2, \ldots\}$ is a partition of unity subordinate to the open cover

$$\mathcal{O} = \left\{\left(\frac{1}{2^{n+2}}, \frac{1}{2^{n-1}}\right) \cup \left(1-\frac{1}{2^{n-1}}, 1-\frac{1}{2^{n+2}}\right)\right\}_{n=1}^\infty$$

Now, let S_k be the partial sum

$$S_k := \sum_{n=1}^k \int_{C_{\varphi_n}} \varphi_n|f| = \sum_{n=1}^k \int_{C_{\varphi_n}} \varphi_n|f|$$

For each φ_i we have

$$C_{\varphi_i} \subseteq \left(\frac{1}{2^{k+2}}, 1 - \frac{1}{2^{k+2}}\right)$$

so

$$S_k = \sum_{n=1}^{k} \int_{\frac{1}{2^{k+2}}}^{1-\frac{1}{2^{k+2}}} \varphi_i f = \int_{\frac{1}{2^{k+2}}}^{1-\frac{1}{2^{k+2}}} \sum_{i=1}^{k} \varphi_i f \leq \int_{\frac{1}{2^{k+2}}}^{1-\frac{1}{2^{k+2}}} f \leq \lim_{\varepsilon \to 0^+} \int_{\varepsilon}^{1-\varepsilon} f$$

where the last inequality follows since f is nonnegative. Moreover, since f is nonnegative we have

$$\int_{C_{\varphi_i}} \varphi_i f \geq 0$$

so we have an increasing, bounded above series which thus converges. So f is extended integrable on $(0,1)$. □

(b) *Proof.* To show that

$$\text{ext} \int_{(0,1)} f$$

does not exist, we will exhibit a partition of unity Φ subordinate to an admissible open cover $\mathcal{O}$ of $(0,1)$ such that

$$\underset{\Phi}{\text{ext}} \int_{(0,1)} f = \sum_{\varphi \in \Phi} \int_{C_\varphi} \varphi |f|$$

does not converge. Define

$$O_n = \left(\frac{1}{2^{n+2}}, \frac{1}{2^{n-1}}\right) \cup \left(1 - \frac{1}{2^{n-1}}, 1 - \frac{1}{2^{n+2}}\right)$$

for each n, and let $\mathcal{O} = \{O_n\}_{n \in \mathbb{N}}$ be our open cover. By Exercise 2-26, pick ψ_n so that $\psi_n > 0$ on A_n but $\psi_n = 0$ outside of some closed set contained in O_n. Then only finitely many (but at least one) ψ_i are nonzero at any given point $x \in (0,1)$, so write

$$\varphi_n(x) = \frac{\psi_n(x)}{\sum_{i=1}^{\infty} \psi_i(x)}$$

$\Phi = \{\varphi_1, \varphi_2, \ldots\}$ is our desired partition of unity subordinate to $\mathcal{O}$.

Since $\bigcup_{i=1}^{\infty} A_i = [1/2, 1)$ and $f = 0$ outside of $\bigcup_{i=1}^{\infty} A_i$, we have

$$\text{supp}(\varphi_n |f|) \subseteq \left(1 - \frac{1}{2^{n-1}}, 1 - \frac{1}{2^{n+2}}\right)$$

so that

$$\int_{C_{\varphi_n}} \varphi_n |f| = \int_{\text{supp}\, \varphi_n |f|} \varphi_n |f| = \int_{A_{n-1}} \varphi_n |f| + \int_{A_n} \varphi_n |f| + \int_{A_{n+1}} \varphi_n |f|$$

(for $n = 1$ the first term is omitted). Letting

$$S_k = \sum_{i=1}^{k} \int_{C_{\varphi_i}} \varphi_i |f|$$

we have

$$\begin{aligned}
\sum_{i=1}^{\infty} \int_{C_{\varphi_i}} \varphi_i |f| &\geq S_k \\
&= \sum_{i=1}^{k} \left(\int_{A_{i-1}} \varphi_i |f| + \int_{A_i} \varphi_i |f| + \int_{A_{i+1}} \varphi_i |f| \right) \\
&= \sum_{i=1}^{k-1} \int_{A_i} \varphi_{i+1} |f| + \sum_{i=1}^{k} \int_{A_i} \varphi_i |f| + \sum_{i=2}^{k+1} \int_{A_i} \varphi_{i-1} |f| \\
&\geq \sum_{i=1}^{k-1} \int_{A_i} \varphi_{i+1} |f| + \sum_{i=1}^{k} \int_{A_i} \varphi_i |f| + \sum_{i=2}^{k} \int_{A_i} \varphi_{i-1} |f| \\
= \int_{A_1} |f|(\varphi_2 + \varphi_1) &+ \sum_{i=2}^{k-1} \left(\int_{A_i} |f|(\varphi_{i+1} + \varphi_i + \varphi_{i-1}) \right) + \int_{A_k} |f|(\varphi_k + \varphi_{k-1}) \\
&\geq \int_{A_1} |f|(\varphi_2 + \varphi_1) + \sum_{i=2}^{k-1} \left(\int_{A_i} |f|(\varphi_{i+1} + \varphi_i + \varphi_{i-1}) \right)
\end{aligned}$$

Note that by construction, φ_1 and φ_2 are the only nonzero φ on A_1, and $\varphi_{i-1}, \varphi_i, \varphi_{i+1}$ are the only nonzero φ on A_i for $i \geq 2$. Thus this simplifies to

$$\int_{A_1} |f| + \sum_{i=2}^{k-1} \int_{A_i} |f| \geq \sum_{i=1}^{k-1} \left| \int_{A_i} f \right| = \sum_{i=1}^{k-1} \frac{1}{n}$$

so (S_k) is the sequence of partial sums of the harmonic series, which diverges. Thus $\operatorname{ext}_\Phi \int_{(0,1)} f$ does not exist.

But in contrast, we have

$$\operatorname{ext} \int_{(\varepsilon, 1-\varepsilon)} f = \sum_{i=1}^{M-1} \int_{A_i} f + \int_{(1-1/2^M, 1-\varepsilon)} f$$

where M is the largest integer such that $1 - 1/2^M \leq 1 - \varepsilon$. If M is even then we have

$$\sum_{i=1}^{M-1} \int_{A_i} f \leq \operatorname{ext} \int_{(\varepsilon, 1-\varepsilon)} f \leq \sum_{i=1}^{M} \int_{A_i} f$$

and if M is odd then

$$\sum_{i=1}^{M-1} \int_{A_i} f \geq \operatorname{ext} \int_{(\varepsilon, 1-\varepsilon)} f \geq \sum_{i=1}^{M} \int_{A_i} f$$

so

$$\lim_{\varepsilon\to 0} \operatorname{ext} \int_{(\varepsilon,1-\varepsilon)} f = \sum_{i=1}^{\infty} \int_{A_i} f = \sum_{i=1}^{M} \frac{(-1)^i}{i} = -\ln 2 \qquad \square$$

Exercise 3-38 Let A_n be a closed set contained in $(n, n+1)$. Suppose that $f : \mathbb{R} \to \mathbb{R}$ satisfies

$$\int_{A_i} f = \frac{(-1)^i}{i}$$

and $f = 0$ outside of $\bigcup_{i=1}^{\infty} A_i$. Find two partitions of unity Φ, Ψ for $\mathbb{R}$ such that

$$\sum_{\varphi\in\Phi} \int_{C_\varphi} \varphi f$$

and

$$\sum_{\psi\in\Psi} \int_{C_\psi} \psi f$$

both converge absolutely, but to different values.

Proof. First, pick C^∞ functions $g_1, g_2, \ldots : \mathbb{R} \to [0,1]$ such that $g_i = 1$ on A_i and $g_i = 0$ outside of a closed set contained in $(i, i+1)$. Now, let $\varphi_n = g_{2n-1} + g_{2n}$. Then the collection $\Phi = \{\varphi_1, \varphi_2, \ldots\}$, together with appropriately chosen functions, forms a partition of unity for $\mathbb{R}$. We have

$$\int_{C_{\varphi_n}} \varphi_n f = \int_{C_{g_{2n-1}}} f + \int_{C_{g_{2n}}} f = \int_{A_{2n-1}} f + \int_{A_{2n}} f = \frac{-1}{2n-1} + \frac{1}{2n} = -\frac{1}{4n^2 - 2n}$$

Thus

$$\operatorname*{ext}_{\Phi} \int_{\mathbb{R}} f = \sum_{i=1}^{\infty} \int_{C_{\varphi_i}} \varphi_i f = \sum_{i=1}^{\infty} -\frac{1}{4n^2-2n} = -\ln 2$$

If we instead pick $\psi_1 = g_1$ and $\psi_n = g_{2n} + g_{2n+1}$, then $\Psi = \{\psi_1, \psi_2, \ldots\}$ (with appropriately chosen functions) forms a partition of unity and we similarly have

$$\operatorname*{ext}_{\Psi} \int_{\mathbb{R}} f = \int_{A_1} f + \sum_{i=2}^{\infty} \left(\int_{A_{2n}} f + \int_{A_{2n+1}} f \right) = -1 + \sum_{i=2}^{\infty} \frac{1}{4n^2+2n} = -\frac{1}{6} - \ln 2$$

Both of the series indicated converge absolutely since they converge, and do not change sign. $\square$

Exercise 3-39 Prove Theorem **??** without the assumption $\det u'(x) \neq 0$ using Sard's Theorem.

Proof. Suppose $u : A \to \mathbb{R}^n$ is injective and continuously differentiable, with A open. Let C be the set of points $x \in A$ such that $\det u'(x) = 0$. $\det u'(x)$ is composed of products and sums of the partial derivatives, which are continuous, so $x \mapsto \det u'(x)$ is continuous. So C is a closed set in A, which means that $A \setminus C$ is open in A and thus in $\mathbb{R}^n$. Then the restriction of u to $A \setminus C$ is an injective, continuously differentiable function defined on an open set with $\det u'(x) \neq 0$ for $x \in A \setminus C$. By Theorem **??**, we have

$$\text{ext} \int_{u(A\setminus C)} f = \text{ext} \int_{A\setminus C} (f \circ u)|\det u'|$$

Since u is injective, $u(A \setminus C) = u(A) \setminus u(C)$. By Sard's Theorem, $u(C)$ has measure zero so

$$\text{ext} \int_{u(A)} f = \text{ext} \int_{u(A)\setminus u(C)} f + \text{ext} \int_{u(C)} f = \text{ext} \int_{u(A)\setminus u(C)} f$$

Now, since $(f \circ u)|\det u'| = 0$ on C, and

$$\text{ext} \int_{A\setminus C} |\det u'| = \text{ext} \int_{A} |\det u'|$$

By Sard's Theorem, $u(C)$ has measure zero. So we have

$$\text{ext} \int_{u(A)} 1 = \text{ext} \int_{u(A)\setminus u(C)} 1 = \text{ext} \int_{u(A\setminus C)} 1 = \text{ext} \int_{A\setminus C} |\det u'| = \text{ext} \int_{A} |\det u'|$$

□

Exercise 3-40

(a) If $g : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable and $\det g'(a) \neq 0$, prove that in some open set containing a we can write $g = T \circ g_n \circ \ldots \circ g_1$, where g_i is of the form

$$g_i(x) = (x_1, \ldots, f_i(x), \ldots, x_n)$$

for some $f_i : \mathbb{R}^n \to \mathbb{R}$, and where T is a linear transformation.

Note: Spivak failed to require that g be C^1.

(b) Show that if f_i does not depend on $x_j, i \neq j$, then we can take $T = I$ if and only if $g'(a)$ is diagonal.

Note: Spivak's original question does not include the stipulation that f_i does not depend on the other variables, but it is incorrect as stated.

(a) *Proof.* First note that it suffices to prove the case $g'(a) = I$. In the general case, we would consider $(Dg(a))^{-1} \circ g$, and then g may be written as $Dg(a)$ composed with the representation produced in the identity case.

Recursively define the following:

$$\begin{aligned}
g_1(x) &= (g^1(x), x_2, \ldots, x_n) \\
g_2(x) &= (x_1, g^2(g_1^{-1}(x)), x_3, \ldots, x_n) \\
&\vdots \\
g_n(x) &= (x_1, \ldots, x_{n-1}, g^n(g_1^{-1}(\ldots(g_{n-1}^{-1}(x)))))
\end{aligned}$$

The fact that each g_i^{-1} exists is by the Inverse Function Theorem, since each has $g_i'(a) = I$ and thus there is an open set around a where all g_i are invertible. It follows that

$$g = g_n \circ \ldots \circ g_1$$

$\square$

(b) ($\Longrightarrow$) Suppose $T = I$. Then if $j \neq i$, we have

$$\begin{aligned} D_j g_i(a) &= D_j(g^i \circ (g_1^{-1} \circ \ldots \circ g_{i-1}^{-1})(a)) \\ &= \underbrace{D_j g^i(g_1^{-1} \circ \ldots \circ g_{i-1}^{-1})(a)}_{=0} D_j(g_1 \circ \ldots \circ g_{i-1})(a) \\ &= 0 \end{aligned}$$

so $g'(a)$ is diagonal.

($\Longleftarrow$) Suppose

$$g'(a) = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}$$

where each a_i is nonzero. Then $g \circ [Dg(a)]^{-1}$ satisfies

$$(g \circ [Dg(a)]^{-1})'(a) = g'(Dg(a)^{-1}(a))[g'([Dg(a)]^{-1}(a))]^{-1} = I$$

So we have $g = g_n \circ \ldots \circ g_1 \circ Dg(a)$. Since $Dg(a)$ is of the form

$$Dg(a) = f_1 \circ \ldots \circ f_n$$

we can write

$$g = g_n \circ \ldots \circ g_1 \circ f_1 \circ \ldots \circ f_n$$

Since f_i only depends on and changes the ith coordinate, and the same is true for g_i, we can freely interchange them so long as the relative order of g_i, f_i is preserved for each i. So this becomes

$$g = (g_n \circ f_n) \circ \ldots \circ (g_1 \circ f_1)$$

Define $f : \{r : r > 0\} \times (0, 2\pi) \to \mathbb{R}^2$ by $f(r, \theta) = (r\cos\theta, r\sin\theta)$.

(a) Show that f is injective, compute $f'(r, \theta)$, and show that $\det f'(r, \theta) \neq 0$ for all (r, θ). Show that $f(\{r : r > 0\} \times (0, 2\pi))$ is the set $A = \{x < 0 \text{ or } x \geq 0, y \neq 0\}$, as in Exercise 2-23.

(b) If $P = f^{-1}$, show that $P(x, y) = (r(x, y), \theta(x, y))$, where

$$r(x, y) = \sqrt{x^2 + y^2}$$
$$\theta(x, y) = \begin{cases} \arctan \frac{y}{x}, & x > 0, y > 0 \\ \pi + \arctan \frac{y}{x}, & x < 0 \\ 2\pi + \arctan \frac{y}{x}, & x > 0, y < 0 \\ \frac{\pi}{2}, & x = 0, y > 0 \\ \frac{3\pi}{2}, & x = 0, y < 0 \end{cases}$$

Find $P'(x, y)$. P is called the **polar coordinate system** on A.

(c) Let $C \subseteq A$ be the region between the circles of radius r_1 and r_2 and the half-lines through 0 which make angles of θ_1 and θ_2 with the x-axis. If $h : C \to \mathbb{R}$ is integrable and $h(x, y) = g(r(x, y), \theta(x, y))$, show that

$$\int_C h = \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} r g(r, \theta) \, \mathrm{d}\theta \, \mathrm{d}r$$

If $B_r = \{(x, y) : x^2 + y^2 \leq r^2\}$, show that

$$\int_{B_r} h = \int_0^r \int_0^{2\pi} r g(r, \theta) \, \mathrm{d}\theta \, \mathrm{d}r$$

(c) If $C_r = [-r, r] \times [-r, r]$, show that

$$\int_{B_r} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y = \pi(1 - e^{-r^2})$$

and

$$\int_{C_r} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y = \left(\int_{-r}^{r} e^{-x^2} \, \mathrm{d}x \right)^2$$

(e) Prove that

$$\lim_{r \to \infty} \int_{B_r} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y = \lim_{r \to \infty} \int_{C_r} e^{-(x^2+y^2)} \, \mathrm{d}x \, \mathrm{d}y$$

and conclude that

$$\int_{-\infty}^{\infty} e^{-x^2} = \sqrt{\pi}$$

(a) *Proof.* Suppose $r_1 \cos\theta_1 = r_2 \cos\theta_2$ and $r_1 \sin\theta_1 = r_2 \sin\theta_2$. Then

$$r_1^2 = r_1^2(\cos^2\theta_1 + \sin^2\theta_1) = r_2^2(\cos^2\theta_2 + \sin^2\theta_2) = r_2^2$$

so $r_1 = r_2$. So $\sin\theta_1 = \sin\theta_2$ and $\cos\theta_1 = \cos\theta_2$, and we conclude that $\theta_1 = \theta_2$. We have

$$\det f'(r,\theta) = \det \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = r\cos^2\theta + r\sin^2\theta = r > 0$$

$f(r,\theta) \in \mathbb{R}^2 \setminus A$ only if $y = 0$ and $x \geq 0$, which implies $\sin\theta = 0$ and $\cos\theta > 0$ and thus $\theta = 0$, or $\sin\theta = \cos\theta = 0$ which is impossible. So $f(\{r : r > 0\} \times (0, 2\pi)) \subseteq A$. Let $A = (x, y)$. Then take

$$r = \sqrt{x^2 + y^2}$$

$$\theta = \begin{cases} \arctan\frac{y}{x}, & x > 0, y > 0 \\ \pi + \arctan\frac{y}{x}, & x < 0 \\ 2\pi + \arctan\frac{y}{x}, & x > 0, y < 0 \\ \frac{\pi}{2}, & x = 0, y > 0 \\ \frac{3\pi}{2}, & x = 0, y < 0 \end{cases}$$

So $A \subseteq f(\{r : r > 0\} \times (0, 2\pi))$ and we have equality. $\square$

(b) *Proof.* It suffices to check that $r(f(r,\theta)) = r$ and $\theta(f(r,\theta)) = \theta$. The first equality is easy:

$$r(f(r,\theta)) = \sqrt{r^2\cos^2\theta + r^2\sin^2\theta} = r$$

For the second:

$$\begin{cases} 0 < \theta < \frac{\pi}{2} \implies \cos\theta > 0, \sin\theta > 0 \\ \frac{\pi}{2} < \theta < \frac{3\pi}{2} \implies \cos\theta < 0 \\ \frac{3\pi}{2} < \theta < 2\pi \implies \cos\theta > 0, \sin\theta < 0 \\ \theta = \frac{\pi}{2} \implies \cos\theta = 0, \sin\theta > 0 \\ \theta = \frac{3\pi}{2} \implies \cos\theta = 0, \sin\theta < 0 \end{cases}$$

Since $r > 0$, all of the following remain true when $\cos\theta$ is replaced by f_1 and $\sin\theta$ by f_2. So we have

$$\begin{cases} 0 < \theta < \frac{\pi}{2} \implies \theta(f(r,\theta)) = \arctan\tan\theta = \theta \\ \frac{\pi}{2} < \theta < \frac{3\pi}{2} \implies \theta(f(r,\theta)) = \pi + \arctan\tan\theta = \theta \\ \frac{3\pi}{2} < \theta < 2\pi \implies \theta(f(r,\theta)) = 2\pi + \arctan\tan\theta = \theta \\ \theta = \frac{\pi}{2} \implies \theta(f(r,\theta)) = \frac{\pi}{2} = \theta \\ \theta = \frac{3\pi}{2} \implies \theta(f(r,\theta)) = \frac{3\pi}{2} = \theta \end{cases}$$

We have

$$\begin{aligned}
D_1P^1(x,y) = D_1r(x,y) &= \frac{x}{\sqrt{x^2+y^2}} \\
D_2P^1(x,y) = D_2r(x,y) &= \frac{y}{\sqrt{x^2+y^2}} \\
D_1P^2(x,y) = D_1\theta(x,y) &= \frac{1}{1+\frac{y^2}{x^2}}\left(-\frac{y}{x^2}\right) = -\frac{y}{x^2+y^2} \\
D_2P^2(x,y) = D_2\theta(x,y) &= \begin{cases} \frac{1}{1+\frac{y^2}{x^2}}\frac{1}{x} = \frac{x}{x^2+y^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}
\end{aligned}$$

so

$$P'(x,y) = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{bmatrix}$$

□

(c) *Proof.* Let $C' = P(C) = (r_1, r_2) \times (0, 2\pi)$, so that $C = P^{-1}C$. Note also that $h = g \circ P$. P^{-1} is continuously differentiable by the Inverse Function Theorem, so by the Change of Variables theorem,

$$\int_C h = \int_{C'} (h \circ P^{-1})|\det(P^{-1})'| = \int_{C'} (h \circ P^{-1})\frac{1}{|\det P'|} = \int_{C'} g\frac{1}{|\det P'|}$$

We can calculate,

$$\det P'(x,y) = \frac{x^2+y^2}{\sqrt{x^2+y^2}^3} = \frac{1}{\sqrt{x^2+y^2}} = \frac{1}{r}$$

So

$$\int_{C'} rg = \int_{(r_1,r_2)\times(0,2\pi)} rg$$

By Fubini's Theorem, this becomes

$$\int_C h = \int_{r_1}^{r_2}\int_0^{2\pi} rg(r,\theta)\,\mathrm{d}\theta\,\mathrm{d}r$$

Similarly, let $B_r' = P(B_r) = (0,r) \times (0,2\pi)$. By similar logic,

$$\int_{B_r} h = \int_{B_r'} (h \circ P^{-1})\frac{1}{|\det P'|} = \int_{B_r'} gr = \int_{(0,r)\times(0,2\pi)} gr = \int_0^r\int_0^{2\pi} rg(r,\theta)\,\mathrm{d}\theta\,\mathrm{d}r \quad \square$$

(d) *Proof.* Using the result from part c),

$$\int_{B_r} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = \int_0^r\int_0^{2\pi} re^{-r^2}\,\mathrm{d}\theta\,\mathrm{d}r = \int_0^r 2\pi re^{-r^2}\,\mathrm{d}r = -\pi e^{-r^2}|_0^r = \pi(1-e^{-r^2})$$

By Fubini's Theorem,

$$\begin{aligned}\int_{C_r} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y &= \int_{-r}^{r}\left(\int_{-r}^{r} e^{-x^2}e^{-y^2}\,\mathrm{d}y\right)\mathrm{d}x\\ &= \int_{-r}^{r} e^{-x^2}\left(\int_{-r}^{r} e^{-y^2}\,\mathrm{d}y\right)\mathrm{d}x\\ &= \left(\int_{-r}^{r} e^{-x^2}\,\mathrm{d}x\right)^2\end{aligned}$$

□

(e) *Proof.* The quantity $e^{-(x^2+y^2)}$ is positive everywhere. So for any r, there exists $r' > r$ such that $C_r \subseteq B_{r'}$, giving

$$\int_{C_r} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y \le \int_{B_{r'}} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y$$

But we can also pick r'' so that $B_r \subseteq C_{r''}$ so that the other direction is true. This shows that

$$\lim_{r\to\infty}\int_{B_r} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = \lim_{r\to\infty}\int_{C_r} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y$$

Then we have

$$\begin{aligned}\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x &= \lim_{r\to\infty}\int_{-r}^{r} e^{-x^2}\,\mathrm{d}x\\ &= \lim_{r\to\infty}\sqrt{\int_{C_r} e^{-x^2+y^2}\,\mathrm{d}x\,\mathrm{d}y}\\ &= \sqrt{\lim_{r\to\infty}\int_{C_r} e^{-x^2+y^2}\,\mathrm{d}x\,\mathrm{d}y}\\ &= \sqrt{\lim_{r\to\infty}\int_{B_r} e^{-x^2+y^2}\,\mathrm{d}x\,\mathrm{d}y}\\ &= \sqrt{\lim_{r\to\infty}\pi(1-e^{-r^2})}\\ &= \sqrt{\pi}\end{aligned}$$

□

Definitions