# Spivak's Calculus On Manifolds: Solutions Manual

Thomas Hughes

August 2017

# Chapter 1

# Functions on Euclidean Space

1.1 Prove that $|x| \leq \sum_{i=1}^{n} |x_i|$

*Proof.* If $\{e_1, e_2, \ldots, e_n\}$ is the usual basis on $\mathbb{R}^n$, then we can write

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n$$

and thus

$$|x| = \left|\sum_{i=1}^{n} x_ie_i\right| \leq \sum_{i=1}^{n} |x_ie_i| = \sum_{i=1}^{n} |x_i||e_i| = \sum_{i=1}^{n} |x_i|$$

□

1.2 When does equality hold in Theorem 1-1(3)?

*Proof.* Notice in the proof that we get

$$|x+y|^2 = \sum_{i=1}^{n}(x_i)^2 + \sum_{i=1}^{n}(y_i)^2 + 2\sum_{i=1}^{n} x_iy_i \leq |x|^2 + |y|^2 + 2|x||y|$$

and so we have equality precisely when $\sum_{i=1}^{n} x_iy_i = |\sum_{i=1}^{n} x_iy_i|$ and $x$ and $y$ are dependent. That is, when $x$ and $y$ are dependent and $sgn(x_i) = sgn(y_i)$ for all $i$. That is, when one is a non-negative multiple of the other. □

1.3 Prove that $|x - y| \leq |x| + |y|$. When does equality hold?

*Proof.*

$$|x-y| = |x + (-y)| \leq |x| + |-y| = |x| + |y|$$

Conditions for equality are the same as in 1.2, for $x$ and $-y$. □

1.4 Prove that $||x| - |y|| \leq |x - y|$.

*Proof.* Notice

$$|x| = |x - y + y| \leq |x - y| + |y|$$

Thus

$$|x| - |y| \leq |x - y|$$

Similarly,

$$|y| = |y - x + x| \leq |y - x| + |x| = |x - y| + |x|$$

Thus

$$\begin{aligned} |y| - |x| &\leq |x - y| \\ -|x - y| &\leq |x| - |y| \end{aligned}$$

So, combining these results yields

$$-|x - y| \leq |x| - |y| \leq |x - y|$$

which implies

$$||x| - |y|| \leq |x - y|$$

as desired. □

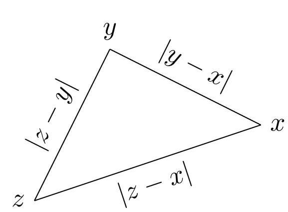
1.5 The quantity $|y - x|$ is called the **distance** between $x$ and $y$. Prove and interpret geometrically the "triangle inequality":

$$|z - x| \leq |z - y| + |y - x|$$

*Proof.*

$$|z - x| = |z - y + y - x| \leq |z - y| + |y - x|$$

Geometrically, we have




□

1.6 Let $f$ and $g$ be integrable on $[a, b]$.

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

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

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

*Proof.* (a) One way to prove this would be to observe that 1-1(2) implies that

$$\left|\sum_{k=1}^{n} f(t_k)g(t_k)\Delta x_k\right| \leq \left(\sum_{k=1}^{n}(f(t_k))^2\Delta x_k\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n}(g(t_k))^2\Delta x_k\right)^{\frac{1}{2}}$$

Notice, by integrability, all the sums can be considered functions of the tagged partition $\dot{\mathcal{P}}$, such that each will approach $\int_a^b f \cdot g$, $\int_a^b f^2$ and $\int_a^b g^2$, respectively, when $\left|\left|\dot{\mathcal{P}}\right|\right| \to 0$. Thus, by continuity of the square root function and the absolute value, taking the limit as $\left|\left|\dot{\mathcal{P}}\right|\right| \to 0$ will give us the desired result.

However, following Spivak's hint, we observe that either there exists $\lambda \in \mathbb{R}$ such that

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

or, since $(f - \lambda g)^2$ is nonnegative, for all $\lambda \in \mathbb{R}$

$$0 < \int_a^b (f - \lambda g)^2$$

Notice from Lebesgue's theorem of Riemann-integrability that in the first case we must have that $f = \lambda g$ almost everywhere on $[a, b]$, and therefore

$$\int_a^b f \cdot g = \lambda \int_a^b g^2$$

and

$$\int_a^b f^2 = \int_a^b \lambda^2 g^2$$

Which would give us that

$$\left|\int_a^b f \cdot g\right| = \left|\lambda \int_a^b g^2\right| = \sqrt{\lambda^2 \left(\int_a^b g^2\right)^2} = \sqrt{\int_a^b \lambda^2 g^2 \int_a^b g^2} = \sqrt{\int_a^b f^2 \int_a^b g^2}$$

which can be rewritten as

$$\left|\int_a^b f \cdot g\right| = \left(\int_a^b f^2\right)^{\frac{1}{2}} \left(\int_a^b g^2\right)^{\frac{1}{2}}$$

In the second case, we have that

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

On the right, we have a quadratic in $\lambda$ which has no real roots since the inequality holds for all $\lambda \in \mathbb{R}$. Thus

$$4 \left( \int_a^b f \cdot g \right)^2 - 4 \int_a^b f^2 \int_a^b g^2 < 0$$

which implies

$$\left( \int_a^b f \cdot g \right)^2 < \int_a^b f^2 \int_a^b g^2$$

which finally gives us

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

Thus, together, the results from both cases imply that

$$\left| \int_a^b f \cdot g \right| \leq \left( \int_a^b f^2 \right)^{\frac{1}{2}} \cdot \left( \int_a^b g^2 \right)^{\frac{1}{2}}$$

as desired.

(b) No, in general, we may have $f \neq \lambda g$ and still have $\left| \int_a^b f \cdot g \right| = \left( \int_a^b f^2 \right)^{\frac{1}{2}} \left( \int_a^b g^2 \right)^{\frac{1}{2}}$. For example, take

$$f = \begin{cases} 1, & x = 1 \\ 0, & elsewhere \end{cases}$$

and

$$g = \begin{cases} 1, & x = 0 \\ 0, & elsewhere \end{cases}$$

Then $\left| \int_0^1 f \cdot g \right| = \left( \int_0^1 f^2 \right)^{\frac{1}{2}} \left( \int_0^1 g^2 \right)^{\frac{1}{2}} = 0$ while $f \neq \lambda g$. However, if $f$ and $g$ are continuous, then equality holds if and only if $f = \lambda g$ for some $\lambda \in \mathbb{R}$. This follows from the theorem which states that if $f$ is continuous on $[a, b]$ and $f \geq 0$ with $f(x_0) > 0$ for some $x_0 \in [a, b]$ then $\int_a^b f > 0$. The proof of this can be found in Exercise 7.4.4 from the solutions manual for Abbott's *Understanding Analysis*.

(c) Divide $[a, b]$ into $n$ sub-intervals so that on the $i^{th}$ sub-interval, $\left[a + \frac{b-a}{n}(i-1), a + \frac{b-a}{n}i\right)$, we define $f$ to be

$$f(x) = \frac{x_i}{\sqrt{\frac{b-a}{n}}}$$

and $g$ to be

$$g(x) = \frac{y_i}{\sqrt{\frac{b-a}{n}}}$$

Then

$$\int_a^b f \cdot g = \sum_{i=1}^{n} x_i y_i$$
$$\int_a^b f^2 = \sum_{i=1}^{n} x_i^2$$
$$\int_a^b g^2 = \sum_{i=1}^{n} y_i^2$$

which, by (a), gives us the desired inequality.

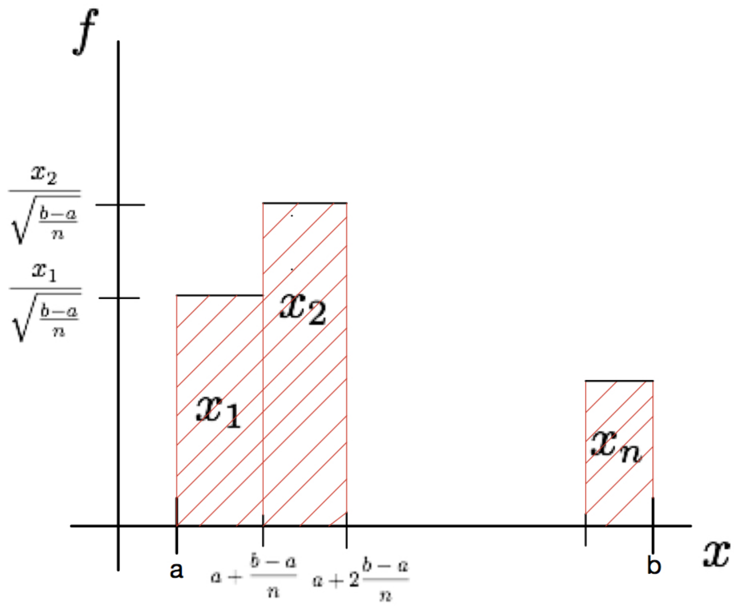



Table 1.1: A poorly drawn graph of $f$

□

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

(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 injective and $T^{-1}$ is of the same sort.

*Proof.* (a) Suppose that $T$ is norm preserving. Then, by the polarization identity, we have

$$\langle x, y\rangle = \frac{|x+y|^2 - |x-y|^2}{4}$$

and since $T$ is norm preserving, we have

$$= \frac{|T(x+y)|^2 - |T(x-y)|^2}{4}$$
$$= \frac{\langle T(x+y), T(x+y)\rangle - \langle T(x-y), T(x-y)\rangle}{4}$$

and by linearity of $T$, we have

$$= \frac{\langle T(x)+T(y), T(x)+T(y)\rangle - \langle T(x)-T(y), T(x)-T(y)\rangle}{4}$$

and by bilinearity of the inner product, we have

$$= \langle Tx, Ty\rangle$$

Now, suppose that $T$ is inner product preserving. Then

$$|T(x)| = \sqrt{\langle T(x), T(x)\rangle} = \sqrt{\langle x, x\rangle} = |x|$$

(b) Suppose that $T$ is norm preserving (and therefore, also inner product preserving) and that

$$T(x) = T(y)$$

then

$$\begin{aligned} T(x) - T(y) &= 0 \\ T(x-y) &= 0 \\ |T(x-y)| &= 0 \\ |x-y| &= 0 \end{aligned}$$

which implies that

$$\begin{aligned} x - y &= 0 \\ x &= y \end{aligned}$$

Therefore, $T$ is injective. Since injectivity of a linear transformation is equivalent to surjectivity, it follows that $T^{-1}$ exists. To show that $T^{-1}$ is of the same sort, we show that it is norm preserving. Observe that

$$\left|T^{-1}(y)\right| = |x| = |T(x)| = |y|$$

□

1.8 If $x, y \in \mathbb{R}^n$ are non-zero, the **angle** between $x$ and $y$, denoted $\angle(x, y)$, is defined as $\arccos\left(\frac{\langle x,y\rangle}{|x|\cdot|y|}\right)$, which makes sense by Theorem 1-1(2). The linear transformation $T$ is **angle preserving** if $T$ is injective, and for $x, y \neq 0$ 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 if and only if all $|\lambda_i|$ are equal.

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

*Proof.* (a) If $T$ is norm preserving, then it is also inner product preserving and injective. Thus, if $x, y \neq 0$,

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

so that $T$ is also angle preserving.

(b) This is not true. Take $((-1, -1), (1, 0)) = (x_1, x_2)$ as a basis for $\mathbb{R}^2$. If

$$T(x_1) = 3x_1$$
$$T(x_2) = -3x_2$$

Then $T$ is injective and has that $|\lambda_i|$ are all equal. However,

$$\angle(x_1, x_2) = \frac{3\pi}{4}$$

while

$$\angle(Tx_1, Tx_2) = \frac{\pi}{4}$$

□

1.9 If $0 \le \theta < \pi$, let $T : \mathbb{R}^2 \to \mathbb{R}^2$ have the matrix

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

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

*Proof.* Observe that

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1\cos\theta + x_2\sin\theta \\ -x_1\sin\theta + x_2\cos\theta \end{bmatrix}$$

and

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1\cos\theta + y_2\sin\theta \\ -y_1\sin\theta + y_2\cos\theta \end{bmatrix}$$

Now, some rather tedious calculations show that

$$\langle Tx, Ty\rangle = x_1y_1 + x_2y_2 = \langle x, y\rangle$$

and

$$|Tx| = |x|$$

and

$$|Ty| = |y|$$

which, together, imply that $\angle(Tx, Ty) = \angle(x, y)$ so that $T$ is angle preserving. Again, a rather tedious calculation shows that

$$\langle x, Tx\rangle = |x|^2 \cos\theta$$

which implies that, for $x \neq 0$

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

□

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 $h \in \mathbb{R}^m$.

**Lemma.** *Let $T \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$. Then there exists $M > 0$ such that for all $x$ admitting $|x| = 1$, we have*

$$|T(x)| \leq M$$

*Proof of Lemma.* Let $|x| = 1$. Given the standard basis, $\{e_1, \ldots, e_m\}$, on $\mathbb{R}^m$ we may write

$$x = \alpha_1 e_1 + \ldots + \alpha_m e_m$$

where $|\alpha_i| \le 1$ for all $1 \le i \le m$. Thus

$$\begin{aligned} |T(x)| &= |T(\alpha_1 e_1 + \ldots + \alpha_m e_m)| \\ &= |\alpha_1 T(e_1) + \ldots + \alpha_m T(e_m)| \\ &\le |\alpha_1|\,|T(e_1)| + \ldots + |\alpha_m|\,|T(e_m)| \\ &\le |T(e_1)| + \ldots + |T(e_m)| \\ &= M \end{aligned}$$

□

*Proof of 1.10.* From the Lemma, we get that there exists $M > 0$ such that, if $h \neq 0$, then

$$\begin{aligned} \left| T\left(\frac{h}{|h|}\right) \right| &\le M \\ \left| \frac{1}{|h|} T(h) \right| &\le M \\ |T(h)| &\le M\,|h| \end{aligned}$$

□

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 $|(x,z)| = \sqrt{|x|^2 + |z|^2}$. Note that $(x,z)$ and $(y,w)$ denote points in $\mathbb{R}^{n+m}$.

*Proof.* Observe

$$(x,z) = (x_1, \ldots, x_n, z_1, \ldots, z_m)$$

and

$$(y,w) = (y_1, \ldots, y_n, w_1, \ldots, w_m)$$

and thus

$$\langle (x,z),(y,w)\rangle = \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{m} z_i w_i = \langle x,y\rangle + \langle z,w\rangle$$

and

$$|(x,z)| = \sqrt{\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{m} z_i^2} = \sqrt{|x|^2 + |z|^2}$$

□

1.12* Let $(\mathbb{R}^n)^*$ denote the dual space of the vector space $\mathbb{R}^n$. If $x \in \mathbb{R}^n$, define $\phi_x \in (\mathbb{R}^n)^*$ by $\phi_x(y) = \langle x,y\rangle$. Define $T : \mathbb{R}^n \to (\mathbb{R}^n)^*$ by $T(x) = \phi_x$. Show that $T$ is an injective linear transformation and conclude that every $\phi \in (\mathbb{R}^n)^*$ is $\phi_x$ for a unique $x \in \mathbb{R}^n$.

*Proof.* The linearity of the inner product implies the linearity of $T$. To show that $T$ is injective, by linearity of $T$, it is sufficient to show that $ker(T) = \{0\}$. To that end, suppose that

$$T(x) = 0$$

then

$$\phi_x = 0$$

so that

$$\phi_x(x) = 0$$
$$\langle x, x \rangle = 0$$

if and only if

$$x = 0$$

Thus $ker(T) = \{0\}$ and $T$ is injective. Moreover, since $\dim(\mathbb{R}^n)^* = \dim \mathbb{R}^N$, injectivity of $T$ implies surjectivity. Thus $T$ is a bijection between $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$. Thus each $\phi \in (\mathbb{R}^n)^*$ corresponds to a $\phi_x$ for some unique $x \in \mathbb{R}^n$. □

1.13* If $x, y \in \mathbb{R}^n$, then $x$ and $y$ are called **perpendicular** (or **orthogonal**) if $\langle x, y \rangle = 0$. If $x$ and $y$ are perpendicular, prove that

$$|x + y|^2 = |x|^2 + |y|^2$$

*Proof.*

$$|x + y|^2 = \langle x+y, x+y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle = \langle x, x \rangle + \langle y, y \rangle = |x|^2 + |y|^2$$

□

1.14* Prove that the union of any (even infinite) number of open sets is open. Prove that the intersection of two (and hence infinitely many) open sets is open. Give a counterexample for infinitely many open sets.

*Proof.* Let $x \in \cup_\lambda O_\lambda$ with each $O_\lambda$ open. Then there exists $\lambda$ such that $x \in O_\lambda$. Since $O_\lambda$ is open, there exists an open rectangle, $A$, such that $x \in A \subset O_\lambda \subset \cup_\lambda O_\lambda$. Thus the arbitrary union of open sets is open. A simple but tedious argument could be made to show that the intersection of two open rectangles is, again, an open rectangle. Then the intersection of two open sets being, again open is a simple corollary. Notice that

$$\bigcap_{n=1}^{\infty} \left( -\frac{1}{n}, \frac{1}{n} \right) = \{0\}$$

Clearly, $\{0\}$ does not contain any open rectangles, and is therefore not open. □

1.15 Prove that $\{x \in \mathbb{R}^n : |x-a| < r\}$ is open.

**Lemma.** *There exists an open rectangle, $U$, such that $U \subset B_0(r) = \{x \in \mathbb{R}^n : |x| < r\}$*

*Proof.* Let $U = \Pi_{i=1}^n(-\frac{\epsilon}{\sqrt{n}}, \frac{r}{\sqrt{n}})$. Then

$$|x|^2 \leq \sum_{i=1}^n |x_i|^2 < \sum_{i=1}^n \left(\frac{r}{\sqrt{n}}\right)^2 = r^2$$

implying $|x| < r$. Therefore $U \subset B_0(r)$. □

**Corollary.** There exists an open rectangle, $U$, such that $U \subset B_a(r) = \{x \in \mathbb{R}^n : |x-a| < r\}$ is open.

*Proof.* By the transformation $T(x) = x - a$ and the open rectangle $U$ from the Lemma. □

**Corollary.** $B_a(r)$ is open.

*Proof.* Let $x \in B_a(r)$. Then, from the Corollary, we get that there exists an open rectangle, $U$, such that

$$U \subset B_x(r - |x-a|) \subset B_a(r)$$

Therefore, $B_a(r)$ is open. □

It is not to hard to apply the same idea backwards: every open rectangle contains an open ball. Thus a set, $A$, is open if and only if for each $x \in A$ there exists $B_x(r)$ such that $B_x(r) \subset A$. This characterization of open sets is often easier to work with than the characterization provided in the book. We will consider this as given for future exercises.

1.16 Find the interior, exterior, and boundary of the sets.

$$\{x \in \mathbb{R}^n : |x| \leq 1\}$$
$$\{x \in \mathbb{R}^n : |x| = 1\}$$
$$\{x \in \mathbb{R}^n : \text{ each } x_i \text{ is rational.}\}$$

*Proof.* For the first, we have that the interior is $\{x \in \mathbb{R}^n : |x| < 1\}$, the exterior is $\{x \in \mathbb{R}^n : |x| > 1\}$, and the boundary is $\{x \in \mathbb{R} : |x| = 1\}$.

For the second, we have that the interior is empty (every open rectangle in $\mathbb{R}^n$ will contain some points outside this set), the exterior is $\{x \in \mathbb{R}^n : |x| \neq 1\}$, and the boundary is the set itself.

For the third, we have that both the interior and exterior are empty, and the boundary is again the set itself. □

1.17 Construct a set $A \subset [0,1] \times [0,1]$ such that $A$ contains at most one point on each horizontal and each vertical line but boundary $A = [0,1] \times [0,1]$.

*Proof.* Let $A$ be the set of all points with rational coordinates in $(0,1) \times (0,1)$. Then clearly $boundary(A) \subset [0,1] \times [0,1]$. If $x \in [0,1] \times [0,1]$ then $x = (x_1, x_2)$ where $0 \leq x_1, x_2 \leq 1$. Thus, if $B$ is an open set containing $x$, it follows that there exist $0 \leq \epsilon_1, \epsilon_2$ such that $[x_1 - \epsilon_1, x_1 + \epsilon_1] \times [x_2 - \epsilon_2, x_2 + \epsilon_2] \subset B \cap [0,1] \times [0,1]$. Density of $\mathbb{Q}$ in $\mathbb{R}$ and the existence of an irrational between any two reals implies that $[0,1] \times [0,1] \subset boundary(A)$. Thus $boundary(A) = [0,1] \times [0,1]$. □

1.18 If $A \subset [0,1]$ is the union of open intervals $(a_i, b_i)$ such that each rational number in $(0,1)$ is contained in some $(a_i, b_i)$, show that $boundary(A) = [0,1] \setminus A$.

*Proof.* Notice, by openness of $A$, $x \in boundary(A)$ implies $x \notin A$. Furthermore, by openness of $[0,1]^c$, $x \in boundary(A)$ implies $x \notin [0,1]^c$. If $x \in [0,1] \setminus A$ and if $B$ is open with $x \in B$, then $B \cap A^c \neq \emptyset$. Furthermore, by density of $\mathbb{Q}$ in $\mathbb{R}$ we also get that there exists a rational $r \in B \cap A$, so that $B \cap A \neq \emptyset$. Thus $boundary(A) = [0,1] \setminus A$. □

1.19* If $A$ is a closed set that contains every rational number $r \in [0,1]$, show that $[0,1] \subset A$

*Proof.* Suppose, to the contrary, that there exists $x \in [0,1] \cap A^c$. Since $A$ is closed, it follows that $A^c$ is open and that therefore, there exists an open $B$ such that $x \in B \subset A^c$. But, by density of $\mathbb{Q}$ in $\mathbb{R}$, we know that there exists a rational $r \in A \cap B$ contradicting with $B \subset A^c$. Thus $[0,1] \subset A$. □

1.20 Show that a compact set in $\mathbb{R}^n$ is closed and bounded.

*Proof.* If $K \subset \mathbb{R}^n$ is compact then clearly it is bounded, for if it were unbounded, then $\{\Pi_1^n(-k,k) : k \in \mathbb{N}\}$ would be an open cover of $K$ which admits no finite subcover. To show that $K$ is closed, let $x \in K^c$ and consider the open cover of $K$ given by

$$\left\{ B_k \left( \frac{|x-k|}{2} \right) : k \in K \right\}$$

By compactness, there exists a finite subcover

$$\left\{ B_{k_i} \left( \frac{|x-k_i|}{2} \right) : 1 \leq i \leq n \right\}$$

Thus $B_x \left( \frac{\min_i |x-k_i|}{2} \right) \subset K^c$. Therefore, $K^c$ is open. □

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

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

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

*Proof.* (a) Let $A$ be closed and $x \in A^c$. Since $A$ is closed, there exists an open $B \subset A^c$ with $x \in B$. So there exists an open ball $B_x(r) \subset A^c$. Thus, if $y \in A$, then $y \in (B_x(r))^c$. Thus

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

(b) Let $A$ be closed and $B$ be compact with $A \cap B = \emptyset$. Since $A$ is closed and $A$ and $B$ are disjoint, by (a), we know that for every $x \in B$ there exists $d_x > 0$ such that for all $y \in A$ we have $|y - x| \geq d_x$. So

$$\left\{ B_x\left(\frac{d_x}{2}\right) : x \in B \right\}$$

forms an open cover of $B$. By compactness of $B$, there exists a finite subcover

$$\left\{ B_{x_1}\left(\frac{d_{x_1}}{2}\right), \ldots, B_{x_m}\left(\frac{d_{x_m}}{2}\right) \right\}$$

So for all $y \in A$ and $1 \leq i \leq m$

$$|y - x_i| \leq |y - x| + |x - x_i|$$

Thus, if $x \in B$ then $x \in B_{x_i}\left(\frac{d_{x_i}}{2}\right)$ for some $i$, and the above inequality implies

$$\min_{1 \leq i \leq m} \frac{d_{x_i}}{2} \leq \frac{d_{x_i}}{2} = d_{x_i} - \frac{d_{x_i}}{2} \leq |y - x_i| - |x - x_i| \leq |y - x|$$

(c) If

$$A = \{(0, n) : n \in \mathbb{N}\}$$

and

$$B = \left\{ (0, n - \frac{1}{n}) : n \in \mathbb{N} \right\}$$

then it is clear that both $A$ and $B$ are closed, but not bounded and therefore not compact. Furthermore, $A \cap B = \emptyset$ and for all $n \in \mathbb{N}$ we can find $y \in A$ and an $x \in B$ such that

$$|y - x| < \frac{1}{n}$$

$\square$

1.22* If $U$ is open and $C \subset U$ is compact, show that there is a compact set $D$ such that $C \subset interior(D)$ and $D \subset U$.

*Proof.* Let $U$ be open and $C \subset U$ be compact. It follows that $U^c$ is closed and that $U^c \cap C = \emptyset$. By the previous exercise, we know then that there exists a $d > 0$ such that

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

for all $x \in C$ and $y \in U^c$. We observe that since

$$\{B_x(d/2) : x \in C\}$$

forms an open cover of $C$, there is a finite subcover $\{B_{x_1}, \ldots, B_{x_m}\}$. So if we denote

$$\overline{B_x(d/2)} = \left\{ z : |x - z| \leq \frac{d}{2} \right\}$$

then

$$\bigcup_{i=1}^{m} \overline{B_{x_i}(d/2)} = D$$

is compact and

$$C \subset D \subset U$$

If $x \in C$ then it follows $x \in B_x(d/2) \subset D$ which implies $C \subset Int(D)$ $\square$

1.23 If $f : A \to \mathbb{R}^m$ and $a \in A$, show that $\lim_{x \to a} f(x) = b$ if and only if $\lim_{x \to a} f_i(x) = b_i$ for $1 \leq i \leq m$.

*Proof.* We first show sufficiency by contrapositive. Suppose that there is $1 \leq k \leq m$ such that

$$\lim_{x \to a} f_k(x) \neq b_k$$

Then there is a sequence $(x_r) \subset A \setminus \{a\}$ and $\epsilon > 0$ such that $x_r \to a$ and for all $r \in \mathbb{N}$ we have

$$|f_k(x_r) - b_k| \geq \epsilon$$

which, in turn, implies

$$(f_k(x_r) - b_k)^2 \geq \epsilon^2$$

Thus, for the same $\epsilon$ we have that for every $\delta > 0$ there is $r \in \mathbb{N}$ such that $0 < |x_r - a| < \delta$ while

$$|f(x_r) - b|^2 = \sum_{i=1}^{m} (f_k(x_r) - b_k)^2 \geq \sum_{\substack{i=1 \\ i \neq k}}^{m} (f_k(x_r) - b_k)^2 + \epsilon^2 \geq \epsilon^2$$

which implies that for all $\delta > 0$ there exists $r \in \mathbb{N}$ such that $0 < |x_r - a| < \delta$ while

$$|f(x_r) - b| \geq \epsilon$$

so that $\lim_{x\to a} f(x) \neq b$. To demonstrate necessity, let

$$\lim_{x\to a} f_i(x) = b_i$$

for every $1 \leq i \leq m$. Given $\epsilon > 0$, we may choose $\delta > 0$ so that $0 < |x-a| < \delta$ implies

$$|f_i(x) - b_i| < \frac{\epsilon}{m}$$

for every $1 \leq i \leq m$. Thus for $0 < |x-a| < \delta$ we have

$$|f(x) - b| \leq \sum_{i=1}^{m} |f_i(x) - b_i| \leq \sum_{i=1}^{m} \frac{\epsilon}{m} = \epsilon$$

Therefore,

$$\lim_{x\to a} f(x) = b$$

□

1.24 Prove that $f : A \to \mathbb{R}^m$ is continuous at $a$ if and only if each $f_i$ is.

*Proof.* Notice $f$ is continuous at $a$ if and only if $\lim_{x\to a} f(x) = f(a)$ if and only if $\lim_{x\to a} f_i(x) = f_i(a)$ for every $1 \leq i \leq m$. □

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

*Proof.* From 1.10 we know that there exists an $M > 0$ such that

$$|T(h)| \leq M\,|h|$$

So, if $\epsilon > 0$ is given, and $h \in \mathbb{R}^n$ then whenever $0 < |x-h| < \frac{\epsilon}{M}$ we get that

$$|T(x) - T(h)| = |T(x-h)| \leq M\,|x-h| < M\frac{\epsilon}{M} = \epsilon$$

Therefore $T$ is continuous. □

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 $A^c$.

(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 $(0,0)$.

*Proof.* (a) Let $s$ be a straight line through $(0,0)$. If $s$ is just the vertical line through $x = 0$ then $s \subset A^c$. Similarly, if $s(x) = mx$ with $m \leq 0$ then $s \subset A^c$. If $s(x) = mx$ with $m > 0$ , then $\{(x, mx) : x \leq m\} \subset s \cap A^c$.

(b) First, we demonstrate that $f$ is not continuous at $(0,0)$. Observe that $f((0,0)) = 0$. Notice that the sequence $\left(\frac{1}{n}, \frac{1}{n}\right) \to (0,0)$ and yet $f\left(\left(\frac{1}{n}, \frac{1}{n}\right)\right) = 1$ for all $n$ implying that $f$ is not continuous at $(0,0)$. Now, if $h$ is of the form $(0, h_2)$ then $g_h(t) = 0$ and is therefore continuous at 0. If $h$ is not of the form $(0, h_2)$ then $g_h(0) = f(0) = 0$. Notice that $th = t(h_1, h_2)$ is simply the line $s(x) = \frac{h_2}{h_1}x$. By (a), we know that for $\delta$ small enough $f(th) = 0$ whenever $|t| < \delta$. Therefore $g_h$ is continuous at 0.

□

1.27 Prove that $B_a(r)$ is open by considering the function $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) = |x - a|$.

**Lemma.** $|x| : \mathbb{R}^n \to \mathbb{R}$ is continuous.

*Proof of Lemma.* Let $\epsilon > 0$ be given and $c \in \mathbb{R}^n$. Then $|x - c| < \epsilon$, along with 1.4, imply that

$$||x| - |c|| \leq |x - c| < \epsilon$$

□

**Corollary.** $f(x) = |x - a|$ is continuous.

*Proof of 1.27.* Notice $B_a(r) = f^{-1}\left(\{f(x) \in \mathbb{R} : f(x) < r\}\right)$. Clearly $\{f(x) \in \mathbb{R} : 0 < f(x) < r\}$ is open, so by continuity of $f$ and definition of the norm, we have that $B_a(r) \setminus \{a\}$ is open. Furthermore, since $a \in B_a(r) \subset B_a(r)$ it follows $B_a(r)$ is open. □

1.28 If $A \subset \mathbb{R}^n$ is not closed, show that there is a continuous function $f : A \to \mathbb{R}$ which is unbounded.

*Proof.* If $A$ is not closed then $A^c$ is not open and so there exists a sequence $(a_m) \subset A$ such that $a_m \to x \in A^c$. Thus $f : A \to \mathbb{R}$ defined by

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

is continuous and unbounded. □

1.29 If $A$ is compact, prove that every continuous function $f : A \to \mathbb{R}$ takes on a maximum and minimum value.

*Proof.* From compactness of $A$ and continuity of $f$, it follows $f(A)$ is compact. Thus $f(A)$ is closed and bounded. Thus $inf f(A), sup f(A) \in f(A)$ implying $f$ takes on a maximum and minimum value. □

1.30 Let $f : [a,b] \to \mathbb{R}$ be an increasing function. If $x_1, x_2, \ldots, x_n \in [a,b]$ are distinct, show that $\sum_{i=1}^{n} o(f, x_i) < f(b) - f(a)$.

*Proof.* Let $x_1 < x_2 < \ldots < x_n \in [a,b]$ and let$\delta = \frac{\min\{|x_i - x_j| : i \neq j, 1 \leq i,j \leq n\}}{4}$. If $1 \leq i < n$ then, by our choice of $\delta$, we have

$$x_i + \delta < x_i + 2\delta < x_{i+1} - \delta$$

Since $f$ is increasing, it follows

$$M(x_i, f, \delta) \leq f(x_i + \delta) < f(x_i + 2\delta) < f(x_{i+1} - \delta) \leq m(x_{i+1}, f, \delta)$$

which establishes that

$$M(x_i, f, \delta) < m(x_{i+1}, f, \delta)$$

Furthermore, we observe the following

$$\begin{aligned} M(x_n, f, \delta) &\leq f(b) \\ f(a) &\leq m(x_1, f, \delta) \end{aligned}$$

and, for all $1 \leq i \leq n$

$$o(f, x_i) \leq M(x_i, f, \delta) - m(x_i, f, \delta)$$

Thus it follows that

$$\sum_1^n o(f, x_i) \leq \sum_1^n M(x_i, f, \delta) - m(x_i, f, \delta) < f(b) - m(x_n, f, \delta) - \sum_1^{n-1} m(x_{i+1}, f, \delta) - m(x_i, f, \delta)$$

which reduces to

$$\sum_1^n o(f, x_i) < f(b) - m(x_1, f, \delta) \leq f(b) - f(a)$$

which was to be shown. $\square$

# Chapter 2

# Differentiation

2.1* Prove that if $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a$, then it is continuous at $a$.

*Proof.* We shall present two proofs. Let $\Delta f_a(h) = f(a+h) - f(a)$. For the first proof, let $\epsilon > 0$ be given. We observe that differentiability at $a$ implies that

$$\Delta f_a(h) = Df_a(h) + r(h)$$

where $\lim_{h\to 0} \frac{|r(h)|}{|h|} = 0$. Note that this then implies $\lim_{h\to 0} |r(h)| = 0$. So then

$$|\Delta f_a(h)| = |Df_a(h) + r(h)| \leq |Df_a(h)| + |r(h)|$$

From linearity of $Df_a$ and exercise 1.10, we know that there exists $\lambda$ such that, for all $h$, $|Df_a(h)| \leq \lambda |h|$. Thus by choosing $\delta$ small enough, $|h| < \delta$ implies

$$|\Delta f_a(h)| \leq |Df_a(h)| + |r(h)| < \lambda |h| + |r(h)| < \lambda \frac{\epsilon}{2\lambda} + \frac{\epsilon}{2} = \epsilon$$

which implies that

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

so that $f$ is continuous at $a$.

For the second proof, we observe that given $\epsilon > 0$ we can choose $\delta > 0$ so that $0 < |h| < \delta$ implies

$$||\Delta f_a(h)| - |Df_a(h)|| < |\Delta f_a(h) - Df_a(h)| < \epsilon |h|$$

Thus

$$-\epsilon |h| < |\Delta f_a(h)| - |Df_a(h)| < \epsilon |h|$$

so that

$$|Df_a(h)| - \epsilon |h| < |\Delta f_a(h)| < |Df_a(h)| + \epsilon |h|$$

Linearity of $Df_a$ along with exercise 1.10 implies that, by taking the limit as $h \to 0$ and applying the squeeze theorem, we get that $f$ is continuous at $a$. $\square$

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

*Proof.* If $f$ is independent of the second variable, let $x \in \mathbb{R}$ and set

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

Thus $f(x, y) = g(x)$. On the other hand, if $g : \mathbb{R} \to \mathbb{R}$ is a function such that $f(x, y) = g(x)$, then

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

implying that $f$ is independent of the second variable. Now, assuming that $f$ is differentiable at $(a, b)$, we claim

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

To demonstrate, we see that there is a unique linear $\lambda : \mathbb{R}^2 \to \mathbb{R}$ such that

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

which implies then that $g'(a)$ exists and

$$\lambda(h, k) = g'(a) \cdot h = \begin{bmatrix} g'(a) & 0 \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}$$

Thus

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

$\square$

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

*Proof.* A function $f : \mathbb{R}^2 \to \mathbb{R}$ is independent of the first variable if and only if there is a $g : \mathbb{R} \to \mathbb{R}$ such that $g(y) = f(x, y)$. As before, it can be shown that, if $f$ is differentiable at a point $(a, b)$, then

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

Functions which are independent of the first and the second variable are constant functions. $\square$

2.4 Let $g$ be a continuous real-valued function on the unit circle $\{x \in \mathbb{R}^2 : |x| = 1\}$ 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 $h : \mathbb{R} \to \mathbb{R}$ is defined by $h(t) = f(tx)$, show that $h$ is differentiable.

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

*Proof.* (a) If $x = 0 \in \mathbb{R}^2$ then there is nothing to show. So, let $x \neq 0$, then

$$\lim_{s\to 0} \frac{h(t+s) - h(t)}{s} = \lim_{s\to 0} \frac{|(t+s)x|\, g\left(\frac{(t+s)x}{|(t+s)x|}\right) - |tx|\, g\left(\frac{tx}{|tx|}\right)}{s}$$

since $t > 0$ and $s$ can always be chosen small enough, we get that

$$= \left( \lim_{s\to 0} \frac{|t+s|\, g\left(\frac{(t+s)x}{|(t+s)x|}\right) - |t|\, g\left(\frac{tx}{|tx|}\right)}{s} \right) |x|$$

$$= |x|\, g\left(\frac{x}{|x|}\right)$$

To justify the last equality, we observe that

$$t > 0 \Rightarrow s \text{ can be chosen so that } t + s > 0$$

which will produce the last equality.

$$t < 0 \Rightarrow s \text{ can be chosen so that } t + s < 0$$

which will produce the last equality

$$t = 0 \Rightarrow \text{regardless of } sgn(s) \text{ we get } \lim_{s\to 0} \frac{sg\left(\frac{x}{|x|}\right)}{s}$$

which will produce the last equality.

(b) Suppose that $Df(0,0)$ exists. Then there is $\lambda \in \mathcal{L}(\mathbb{R}^2, \mathbb{R})$ such that

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

Approaching along the $x$-axis and $y$-axis separately, along with the linearity of $\lambda$, implies that $\lambda(x, y) = 0$. That is, that $Df(0, 0) = 0$. Thus

$$\begin{aligned} 0 &= \lim_{(h,k)\to(0,0)} \frac{|f(h,k)|}{|(h,k)|} \\ &= \lim_{(h,k)\to(0,0)} \left| g\left( \frac{(h,k)}{|(h,k)|} \right) \right| \end{aligned}$$

Now, let $\epsilon > 0$ be given and let $(x, y) \in S^1$. For $|\alpha| > 0$ small enough, we will have that

$$|g(x,y)| = \left| g\left( \frac{\alpha(x,y)}{|\alpha|\,|(x,y)|} \right) \right| < \epsilon$$

So $|g(x, y)| < \epsilon$ for every $\epsilon > 0$. Thus $g(x, y) = 0$. Therefore $g = 0$.

□

*Note: Continuity of $g$ was not needed in 2.4*

2.5 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

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

Show that $f$ is a function of the kind considered in 2.4, so that $f$ is not differentiable at $(0, 0)$.

*Proof.* Set $g(x, y) = x\,|y|$. Note then that $g$ is defined on $S^1$. It is simple to show that

$$g(0,1) = g(1,0) = 0 \text{ and } g(-x) = -g(x) \text{ and } g \neq 0$$

We also see that for $(x, y) \neq (0, 0)$ we have that

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

Thus, by 2.4, $f$ is not differentiable at $(0, 0)$. □

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.* Notice, for $(x, y) \neq (0, 0)$

$$|(x,y)|\,f\left( \frac{(x,y)}{|(x,y)|} \right) = \sqrt{|xy|} = f(x,y)$$

Thus, we have

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

Furthermore, the conclusion in 2.4 (b) still holds when $g(-x) = g(x)$. Thus, observing that

$$f(0,1) = f(1,0) = 0 \text{ and } f(-(x,y)) = f(x,y) \text{ and } f \neq 0$$

it follows from 2.4 (b) that $f$ is not differentiable at $(0,0)$. $\square$

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 0.

*Proof.* We claim that $Df(0,0) = 0$. First, observe that

$$0 \leq |f(0)| \leq |0|^2 = 0$$

implying that $f(0) = 0$. Now, let $\epsilon > 0$ be given. Then $0 < |h| < \epsilon$ implies

$$\frac{|f(h) - f(0) - \lambda(h)|}{|h|} = \frac{|f(h)|}{|h|} \leq |h| < \epsilon$$

$\square$

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.*

$$\begin{aligned} 0 &= \lim_{h \to 0} \frac{|f(a+h) - f(a) - \lambda(h)|}{|h|} \\ &= \lim_{h \to 0} \frac{|(f_1(a+h) - f_1(a) - \lambda_1(h), f_2(a+h) - f_2(a) - \lambda_2(h))|}{|h|} \\ &= \lim_{h \to 0} \left| \left( \frac{f_1(a+h) - f_1(a) - \lambda_1(h)}{h}, \frac{f_2(a+h) - f_2(a) - \lambda_2(h)}{h} \right) \right| \end{aligned}$$

if and only if

$$(0,0) = \lim_{h \to 0} \left( \frac{f_1(a+h) - f_1(a) - \lambda_1(h)}{h}, \frac{f_2(a+h) - f_2(a) - \lambda_2(h)}{h} \right)$$

by Exercise 1.23, if and only if

$$0 = \lim_{h \to 0} \frac{f_1(a+h) - f_1(a) - \lambda_1(h)}{h}$$

and

$$0 = \lim_{h \to 0} \frac{f_2(a+h) - f_2(a) - \lambda_2(h)}{h}$$

This establishes that $Df(a)$ exists if and only if

$$f'(a) = \begin{bmatrix} (f_1)'(a) \\ (f_2)'(a) \end{bmatrix}$$

$\square$

2.9 Two functions $f, g : \mathbb{R} \to \mathbb{R}$ are equal up to $n^{th}$ order at $a$ if

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

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

(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 $n^{th}$ order at $a$.

*Proof.* (a) $\Rightarrow$: If $f$ is differentiable at $a$ then simply set $g(x) = f(a) + c(x-a)$ where $f'(a) = c$.

$\Leftarrow$: We disagree, and offer a counter-example:

$$f(x) = \begin{cases} x & , x \neq 0 \\ 1 & , x = 0 \end{cases}$$

and $g(x) = x$. Then

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

Yet $f$ is not continuous at 0, and so, a fortiori, is not differentiable at 0. However, if we insist that $f$ be continuous at $a$, then the proof follows from the fact that

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

implies that

$$\lim_{h \to 0} f(a+h) = a_0$$

which, together with continuity, implies

$$f(a) = a_0$$

which gives the desired result.

(b)

$\square$

2.10 Use theorems of this section to find $f'$ for the following:

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

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

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

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

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

*Proof.* (a) Notice $x^y = e^{\ln(x^y)}$. Thus $f(x, y, z) = exp\{\pi_2 \ln(\pi_1)\}(x, y, z)$. So, if

$$g(x) = e^x$$

and

$$h(x, y, z) = \pi_2 \ln(\pi_1)(x, y, z)$$

then $f(x, y, z) = (g \circ h)(x, y, z)$. It easily follows that

$$g'(h(a, b, c)) = g'(b \cdot ln(a)) = \left[a^b\right]$$

On the other hand

$$\begin{aligned} Dh(a, b, c) &= \left[\ln(\pi_1) D\pi_2 + \pi_2 D \ln(\pi_1)\right](a, b, c) \\ &= ln(a) D\pi_2(a, b, c) + bD \ln(\pi_1)(a, b, c) \\ &= ln(a)\pi_2 + bD \ln(\pi_1)(a, b, c) \end{aligned}$$

and

$$\begin{aligned} D\ln(\pi_1)(a,b,c) &= D\ln(\pi_1(a,b,c)) \circ D\pi_1(a,b,c) \\ &= D\ln(a) \circ \pi_1 \\ &= \frac{1}{a} \circ \pi_1 \\ &= \frac{1}{a}\pi_1 \end{aligned}$$

Thus

$$h'(a,b,c) = \left[\begin{array}{ccc} 0 & \ln(a) & 0 \end{array}\right] + \left[\begin{array}{ccc} \frac{b}{a} & 0 & 0 \end{array}\right] = \left[\begin{array}{ccc} \frac{b}{a} & \ln(a) & 0 \end{array}\right]$$

Therefore,

$$f'(a,b,c) = g'(a,b,c) \cdot h'(a,b,c) = \left[a^b\right] \cdot \left[\begin{array}{ccc} \frac{b}{a} & \ln(a) & 0 \end{array}\right] = \left[\begin{array}{ccc} ba^{b-1} & a^b\ln(a) & 0 \end{array}\right]$$

(b) From part (a) we know that

$$\left[x^y\right]'(a,b,c) = \left[\begin{array}{ccc} ba^{b-1} & a^b\ln(a) & 0 \end{array}\right]$$

By linearity, we also have

$$[z]'(a,b,c) = \left[\begin{array}{ccc} 0 & 0 & 1 \end{array}\right]$$

and so

$$f'(a,b,c) = \left[\begin{array}{ccc} ba^{b-1} & a^b\ln(a) & 0 \\ 0 & 0 & 1 \end{array}\right]$$

(c) Let $f(x,y) = \sin(x\sin(y))$ then, if

$$g(x) = \sin(x)$$

and

$$h(x,y) = x\sin(y) = \pi_1\sin(\pi_2)$$

then $f(x,y) = (g \circ h)(x,y)$. So then

$$g'(h(a,b)) = [\cos(a\sin(b))]$$

and

$$Dh(a,b) = \sin(b)D\pi_1(a,b) + aD\sin(\pi_2)(a,b) = \sin(b)\pi_1 + a\cos(b)\pi_2$$

so that

$$f'(a,b) = \left[\begin{array}{cc} \sin(b)\cos(a\sin(b)) & a\cos(b)\cos(a\sin(b)) \end{array}\right]$$

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

$$g(x) = \sin(x)$$

and

$$h(x,y,z) = x\sin(y\sin(z)) = \pi_1 \sin\circ(\pi_2(\sin\circ\pi_3))(x,y,z)$$

Thus

$$g_1'(g_2(a,b,c)) = [\cos(a\sin(b\sin(c)))]$$

and

$$Dh(a,b,c) = [\sin\circ(\pi_2(\sin\circ\pi_3))]\,(a,b,c)D\pi_1(a,b,c) + \pi_1(a,b,c)D\sin\circ(\pi_2(\sin\circ\pi_3))(a,b,c)$$

and so by (c) and by an argument similar to those in exercises (2.2) and (2.3) we get

$$g_2'(a,b,c) = \sin(b\sin(c))\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} + a\begin{bmatrix} 0 & \sin(c)\cos(b\sin(c)) & b\cos(c)\cos(b\sin(c)) \end{bmatrix}$$

so that

$$g_2'(a,b,c) = \begin{bmatrix} \sin(b\sin(c)) & a\sin(c)\cos(b\sin(c)) & ab\cos(c)\cos(b\sin(c)) \end{bmatrix}$$

Therefore,

$$\begin{aligned} f'(a,b,c) &= g_1'(g_2(a,b,c)) \cdot g_2'(a,b,c) \\ &= [\cos(a\sin(b\sin(c)))]\begin{bmatrix} \sin(b\sin(c)) & a\sin(c)\cos(b\sin(c)) & ab\cos(c)\cos(b\sin(c)) \end{bmatrix} \end{aligned}$$

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

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

and

$$\begin{aligned} h(x,y,z) &= (x, y^z, 1) \\ &= (\pi_1(x,y,z), \pi_2^{\pi_3}(x,y,z), 1) \end{aligned}$$

Thus

$$(g\circ h)(x,y,z) = g(h(x,y,z)) = g((x,y^z,0)) = x^{(y^z)^1} = x^{y^z} = f(x,y,z)$$

Thus

$$Df(a,b,c) = Dg(h(a,b,c)) \circ Dh(a,b,c)$$

Recalling from part (a)

$$g'(h(a,b,c)) = g'((x,y^z,1)) = \begin{bmatrix} b^c a^{b^c-1} & a^{b^c}\ln(a) & 0 \end{bmatrix}$$

and

$$\begin{aligned} Dh(a,b,c) &= (D\pi_1(a,b,c), D\pi_2^{\pi_3}(a,b,c), D1(a,b,c)) \\ &= (\pi_1, D\pi_2^{\pi_3}(a,b,c), 0) \end{aligned}$$

Again, from part (a), we get that

$$(\pi_2^{\pi_3})'(a,b,c) = \begin{bmatrix} 0 & cb^{c-1} & b^c \ln(b) \end{bmatrix}$$

which gives us that

$$h'(a,b,c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cb^{c-1} & b^c \ln(b) \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore,

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

$\square$

2.11 Find $f'$ for the following (where $g : \mathbb{R} \to \mathbb{R}$ is continuous):

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

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

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

*Remark*: It's not too hard to show that if

$$f(x) = \int_a^{h(x)} g$$

with $g$ continuous, then

$$Df(c) = g(h(c)) \circ Dh(c)$$

*Proof.* (a) Let $h(x,y) = x + y = \pi_1 + \pi_2$. Then by our remark, we get

$$Df(c,d) = g(c+d) \circ D(\pi_1 + \pi_2)(c,d)$$

which implies

$$f'(c,d) = [g(c+d)] \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} g(c+d) & g(c+d) \end{bmatrix}$$

(b) Let $h(x, y) = xy = [\pi_1 \cdot \pi_2](x, y)$. Then by our remark, we get

$$Df(c, d) = g(cd) \circ D(\pi_1 \pi_2)(c, d)$$

which implies

$$f'(c, d) = \begin{bmatrix} dg(cd) & cg(cd) \end{bmatrix}$$

(c) Let $h(x, y, z) = \sin(x \sin(y \sin(z)))$. Then by our remark, we get

$$Df(a, b, c) = g(\sin(a \sin(b \sin(c)))) \circ Dh(a, b, c)$$

whose Jacobian is not worth writing out (see 2.10 (d)).

$\square$

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

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

(a) Prove that if $f$ is bilinear, then

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

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

(c) Show that the formula for $Dp(a, b)$ in Theorem 2-3 is a special case of (b).

**Lemma.** *If $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ is bilinear, then there is $M > 0$ such that for all $(a, b) \in \mathbb{R}^n \times \mathbb{R}^m$ with $|(a, b)| \le 1$*

$$|f(a, b)| \le M$$

*Proof of Lemma.* Let $\{e_1, \ldots, e_n\}$ and $\{s_1, \ldots, s_m\}$ be the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and let $(a, b)$ be such that $|(a, b)| \le 1$. Bilinearity of $f$ yields

$$|f(a, b)| = \left| f\left( \sum_{i=1}^n \alpha_i e_i, \sum_{j=1}^m \beta_j s_j \right) \right| \le \sum_{i=1}^n \sum_{j=1}^m |\alpha_i| \, |\beta_j| \, |f(e_i, s_j)|$$

Setting $M' = \max_{i,j}\{|f(e_i, s_j)|\}$ gives us that

$$|f(a, b)| \le \sum_{i=1}^n \sum_{j=1}^m |\alpha_i| \, |\beta_j| \, M'$$

We observe that $|a|, |b| \leq |(a,b)| \leq 1$, which implies that $|\alpha_i|, |\beta_j| \leq 1$ for all $i$ and $j$. So we have

$$|f(a,b)| \leq \sum_{i=1}^{n} \sum_{j=1}^{m} M' = n \cdot m \cdot M' = M$$

□

*Proof of Exercise.* (a) Let $\epsilon > 0$ be given and $0 < |(h,k)| < \frac{\epsilon}{M}$ where $M$ is as defined in the Lemma. Let $\gamma = |(h,k)|$. Then

$$\begin{aligned}
\frac{|f(h,k)|}{\gamma} &= \frac{\gamma^2 \left| f\left(\frac{1}{\gamma}(h,k)\right)\right|}{\gamma} \\
&= \gamma \left| f\left(\frac{1}{\gamma}(h,k)\right)\right| \\
&\leq \gamma M \\
&< \epsilon
\end{aligned}$$

(b) Notice that

$$f(a+h, b+k) - f(a,b) - (f(a,k) + f(h,b)) = f(a+h, b+k) - f(a,b) - f(a,k) - f(h,b)$$

which, by bilinearity, yields

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

which, by (a), will produce the result.

(c) It is a simple matter to show that $p : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by $p(x,y) = x \cdot y$ is bilinear. Part (b) in the above implies then that

$$Dp(a,b)(x,y) = p(a,y) + p(x,b) = ay + bx = bx + ay$$

as desired.

□

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

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

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

(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.

*Proof.* (a) It is a simple matter to show that $IP$ is bilinear. So it follows from the previous exercise that

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

which implies

$$(IP)'(a,b) = \begin{bmatrix} b & a \end{bmatrix}$$

where $b \cdot x = IP(b,x)$ and $a \cdot y = IP(a,y)$

(b) Note that $f(t) = (f_1(t), f_2(t), \ldots, f_n(t))$ and $(g_1(t), g_2(t), \ldots, g_n(t))$. So

$$h(t) = \langle f(t), g(t)\rangle = \sum_{i=1}^{n} f_i(t)g_i(t)$$

which implies

$$h'(a) = \sum_{i=1}^{n}(f_i \cdot g_i)'(a) = \sum_{i=1}^{n} [g_i(a)]\,(f_i)'(a) + [f_i(a)]\,(g_i)'(a)$$

and so, considering $[f'(a)]^T$ and $[g'(a)]^T$ as elements of $\mathbb{R}^n$, then we get

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

(c) Notice that $|f(t)| = 1$ for all $t$ implies that $\langle f(t), f(t)\rangle = 1$ for all $t$. Then if $g(t) = \langle f(t), f(t)\rangle$, by Theorem 2-3 and (b) from this exercise, we get that for all $a$

$$0 = g'(a) = \langle f'(a)^T, f(a)\rangle + \langle f(a), f'(a)^T\rangle = 2\langle f'(a), f(a)\rangle$$

which implies that $\langle f'(t)^T, f(t)\rangle = 0$.

(d) $f(x) = x^3$ is such a function.

□

1. Let $E_i, i = 1, 2, \ldots, k$ 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 $j \neq i$, show that for $h = (h_1, h_2, \ldots, h_k)$, with $h_l \in E_l$, we have

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

(b) Prove that

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

*Proof.* (a) It is a simple matter to show that if $g(x, y)$ is defined by

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

then $g$ is bilinear. Thus, by Exercise 2-12, we have

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

(b) First, a simple inductive argument establishes that for $n \geq 2$ we will have

$$\lim_{h \to 0} \frac{|f(h_1, \ldots, h_n)|}{|(h_1, \ldots, h_n)|} = 0$$

To show this, we first observe that

$$\begin{aligned} f(h_1, \ldots, h_i, \ldots, h_n) &= f\left(h_1, \ldots, \sum_{k=1}^{d} \alpha_k e_k, \ldots, h_n\right) \\ &= \sum_{k=1}^{d} \alpha_k f(h_1, \ldots, e_i, \ldots, h_n) \end{aligned}$$

where $\alpha_k(h_i) \to 0$ as $h_i \to 0$. So, if the claim holds for $n$ then, by the inductive hypothesis along with our previous observation, we have

$$\lim_{h \to 0} \frac{|f(h_1, \ldots, h_i, \ldots, h_{n+1})|}{|(h_1, \ldots, h_i, \ldots, h_{n+1})|} \leq \lim_{h \to 0} \sum_{k=1}^{d} |\alpha_k| \frac{|f(h_1, \ldots, e_k, \ldots, h_{n+1})|}{|(h_1, \ldots, e_k, \ldots, h_{n+1})|} = 0$$

Now, it's not too hard to show that the difference

$$\begin{aligned} &= f(a_1 + h_1, \ldots, a_k + h_k) - f(a_1, \ldots, a_k) - \sum_{i=1}^{k} f(a_1, \ldots, h_i, \ldots, a_k) \\ &= f(a_1, \ldots, a_k) + \sum_{i=1}^{k} f(a_1, \ldots, h_i, \ldots, a_k) + \sigma - f(a_1, \ldots, a_k) \\ &\quad - \sum_{i=1}^{k} f(a_1, \ldots, h_i, \ldots, a_k) \\ &= \sigma(h_1, \ldots, h_k) \end{aligned}$$

where

$$\sigma = \sum f(a_1, \dots, h_{i_1}, \dots, h_{i_2}, \dots, a_k) +$$
$$\sum f(a_1, \dots, h_{i_1}, \dots, h_{i_2}, \dots, h_{i_3} \dots, a_k) + \dots + f(h_1, \dots, h_k)$$

To be clear, $\sigma$ is simply the sum of all images by $f$ where there are two or more $h$'s in the argument. Thus,

$$= \lim_{h \to 0} \frac{\left| f(a_1 + h_1, \dots, a_k + h_k) - f(a_1, \dots, a_k) - \sum_{i=1}^{k} f(a_1, \dots, h_i, \dots, a_k) \right|}{|h|}$$

$$= \lim_{h \to 0} \frac{|\sigma|}{|h|}$$

and by our previous considerations, we get

$$= 0$$

□

2.15 Regard an $n \times n$ matrix as a point in the $n$-fold product $\mathbb{R}^n \times \dots \times \mathbb{R}^n$ by considering each row as a member of $\mathbb{R}^n$.

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

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

(b) If $a_{ij} : \mathbb{R} \to \mathbb{R}$ are differentiable and $f(t) = det(a_{ij}(t))$, show that

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

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

$$\sum_{j=1}^{n} a_{ji}(t) s_j(t) = b_i(t) \quad i = 1, 2, \dots, n$$

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

*Proof.* (a) It is well-known that the determinant, $det$, is multilinear in both the rows and the columns. Thus, the result follows immediately from the previous exercise part (b).

(b) Let $g : \mathbb{R} \to \mathbb{R}^n \times \ldots \times \mathbb{R}^n$ be defined by

$$g(t) = \begin{bmatrix} a_{11}(t) & \ldots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{n1}(t) & \ldots & a_{nn}(t) \end{bmatrix} = \begin{bmatrix} g_1(t) \\ \vdots \\ g_n(t) \end{bmatrix}$$

Then

$$f = det \circ g$$

and thus

$$Df(k) = Ddet(g(k)) \circ Dg(k)$$

Now, we know

$$Ddet(g(k))(x_1, \ldots, x_n) = \sum_{i=1}^{n} det \begin{bmatrix} g_1(k) \\ \vdots \\ x_i \\ \vdots \\ g_n(k) \end{bmatrix}$$

which, by Theorem 2-3, implies that

$$Ddet(g(k)) \circ D(g(k)) = \sum_{i=1}^{n} det \begin{bmatrix} g_1(k) \\ \vdots \\ Dg_i(k) \\ \vdots \\ g_n(k) \end{bmatrix}$$

which, by $Ddet(c) = det'(c)$ and, again, Theorem 2-3, implies that

$$f'(t) = \sum_{i=1}^{n} det \begin{bmatrix} g_1(k) \\ \vdots \\ g_i'(k) \\ \vdots \\ g_n(k) \end{bmatrix} = \sum_{i=1}^{n} det \begin{bmatrix} a_{11}k & \ldots & a_{1n}(k) \\ \vdots & & \vdots \\ a_{i1}'(k) & \ldots & a_{in}'(k) \\ \vdots & & \vdots \\ a_{n1}(k) & \ldots & a_{nn}(k) \end{bmatrix}$$

(c) The first part follows from Cramer's rule. To find $s'(t)$ it is enough to apply (b) and the quotient rule.

$\square$

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 $(f^{-1})'(a) = \left[f'(f^{-1}(a))\right]^{-1}$.

*Proof.*

$$f \circ f^{-1}(x) = I(x) = x$$

implies

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

which implies

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

□

2.17 Find the partial derivatives of the following functions:

2.18 Find the partial derivatives of the following functions (where $g : \mathbb{R} \to \mathbb{R}$ is continuous):

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

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

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

*Proof.* (a) For $1 \leq i \leq 2$

$$\begin{aligned} D_i f(c, d) &= \lim_{h \to 0} \frac{\int_a^{c+d+h} g - \int_a^{c+d} g}{h} \\ &= \lim_{h \to 0} \frac{G(c + d + h) - G(c + d)}{h} \\ &= g(c + d) \end{aligned}$$

(b)

$$\begin{aligned} D_1 f(c, d) &= D_1[G(x) - G(y)](c, d) \\ &= g(c) \\ D_2 f(c, d) &= D_2[G(x) - G(y)](c, d) \\ &= -g(d) \end{aligned}$$

(c)

$$\begin{aligned} D_1 f(c, d) &= 0 \\ D_2 f(c, d) &= g(G(d) - G(b)) \cdot g(d) \end{aligned}$$

□

2.19 If $f(x, y) = x^{x^{x^{x^y}}} + \log(x)(\arctan(\arctan(\arctan(\sin(\cos xy) - \log(x+y)))))$ find $D_2 f(1, y)$.

*Proof.* $f(1, y) = 1$ implying $D_2 f(1, y) = 0$. ☐

2.20 Find the partial derivatives of $f$ in terms of the derivatives of $g$ and $h$ if

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

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

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

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

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

*Proof.* (a) $D_1 f(a, b) = h(b)g'(a)$ and $D_2 f(a, b) = g(a)h'(b)$.

(b) $D_1 f(a, b) = h(b)g(a)^{h(b)} g'(a)$ and $D_2 f(a, b) = g(a)^{h(b)} \ln(g(a))h'(b)$

(c) $D_1 f(a, b) = g'(a)$ and $D_2 f(a, b) = 0$

(d) $D_1 f(a, b) = 0$ and $D_2 f(a, b) = h'(b)$

(e) $D_1 f(a, b) = g'(a + b) = D_2 f(a, b)$

☐

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 so 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 $D_2 f(x, y) = x$.

*Proof.* (a) Let us define

$$f_1(x) = \int_0^x g_1(t, 0)dt$$
$$f_2(y) = \int_0^y g_2(x, t)dt$$

Then $f(x, y) = f_1(x) + f_2(y)$ which implies that

$$D_2 f(x, y) = f_2'(y) = g_2(x, y)$$

(b) If $f(x, y) = \int_0^x g_1(t, y)dt + \int_0^y g_2(0, t)dt$ then $D_2 f(x, y) = g_1(x, y)$.

(c) Let $f(x,y) = \frac{x^2}{2} + \frac{y^2}{2}$. Then

$$D_1 f(x,y) = x \text{ and } D_2 f(x,y) = y$$

If $f(x,y) = xy$ then

$$D_1 f(x,y) = y \text{ and } D_2 f(x,y) = x$$

□

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.* Let $f : \mathbb{R}^2 \to \mathbb{R}$ such that $D_2 f = 0$. Given $x$, if we set $g(y) = f(x,y)$ then $g'(y) = D_2 f(x,y) = 0$ for all $y$, implying, by MVT, that $g(y)$ is constant. But then

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

so that $f$ is independent of the second variable. Now, if $D_1 f = D_2 f = 0$ then $D_1 f = 0$ implies that $f$ is independent of the first variable and $D_2 f = 0$ implies that $f$ is independent of the second variable. So

$$f(x_1, y_1) = f(x_2, y_2)$$

implying that $f$ is constant. □

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.

*Proof.* (a) Let $(x_1, y_1), (x_2, y_2) \in A$. Notice there exists a sequence of lines, each parallel to one of the axes, going from $(x_1, y_1)$ to $(x_2, y_2)$ so that no line passes through the origin. So, if $(x^{(i)}, y^{(i)})$ is the appropriate endpoint of the $i^{th}$ line, then from Exercise 2.22, we have

$$f(x_1, y_2) = f(x', y') = f(x'', y'') = \ldots = f(x_2, y_2)$$

so that $f$ is constant.

(b) Notice that

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

is such a function.

□

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

*Proof.* (a) It is a simple matter to show that

$$D_2 f(x,0) = \lim_{h \to 0} \frac{f(x,h) - f(x,0)}{h} = x$$

and

$$D_1 f(0,y) = \lim_{h \to 0} \frac{f(h,y) - f(0,y)}{h} = -y$$

(b) It is a simple matter to show that

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

□

2.25* Define $f : \mathbb{R} \to \mathbb{R}$ by

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

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

*Proof.* □

2.26* Too bored. Will return later.

*Proof.* □

2.27 Too bored. Will return later.

*Proof.* □

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

*Proof.* (a) Let us define

$$q_1(x,y) = g(x)k(y)$$
$$q_2(x,y) = g(x) + h(y)$$

Then

$$q_1'(x,y) = \begin{bmatrix} k(y)g'(x) & g(x)k'(y) \end{bmatrix}$$
$$q_2'(x,y) = \begin{bmatrix} g'(x) & h'(y) \end{bmatrix}$$

and

$$F(x,y) = f(q_1(x,y), q_2(x,y))$$

By Theorem 2.9, we get that

$$D_1F(a,b) = D_1f(q_1(a,b), q_2(a,b))k(b)g'(a) + D_2f(q_1(a,b), q_2(a,b))g'(a)$$
$$D_2F(a,b) = D_1f(q_1(a,b), q_2(a,b))g(a)k'(b) + D_2f(q_1(a,b), q_2(a,b))h'(b)$$

(b) We see that

$$(g \circ (\pi_1 + \pi_2))'(x,y,z) = \begin{bmatrix} g'(x+y) & g'(x+y) & 0 \end{bmatrix}$$

and

$$(h \circ (\pi_2 + \pi_3))'(x,y,z) = \begin{bmatrix} 0 & h'(y+z) & h'(y+z) \end{bmatrix}$$

and so if $r = (g(a+b), h(b+c))$, then

$$D_1F(a,b,c) = D_1f(r)g'(a+b)$$
$$D_2F(a,b,c) = D_1f(r)g'(a+b) + D_2f(r)h'(b+c)$$
$$D_3F(a,b,c) = D_2f(r)h'(b+c)$$

(c) Skipping the details, if $r = (a^b, b^c, c^a)$ then

$$D_1F(a,b,c) = D_1f(r)ba^{b-1} + D_3f(r)\ln(c)c^a$$
$$D_2F(a,b,c) = D_1f(r)\ln(a)a^b + D_2f(r)cb^{c-1}$$
$$D_3F(a,b,c) = D_2f(r)\ln(b)b^c + D_3f(r)ac^{a-1}$$

(d) Skipping the details, if $r = (a, g(a), h(a,b))$ then

$$D_1F(a,b) = D_1f(r) + D_2f(r)g'(a) + D_3f(r)D_1h(a,b)$$
$$D_2F(a,b) = D_3f(r)D_2h(a,b)$$

$\square$

2.29 Let $f : \mathbb{R}^n \to \mathbb{R}$. For $x \in \mathbb{R}^n$, the limit

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

if it exists, is denoted $D_x f(a)$, and called the **directional derivative** of $f$ at $a$, in the direction of $x$.

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

(b) Show that $D_{tx} f(a) = t D_x f(a)$.

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

*Proof.* (a) Notice that $a + he_i = (a_1, \ldots, a_i + h, \ldots, a_n)$. Specifically,

$$\begin{aligned}
D_{e_i} f(a) &= \lim_{h\to 0} \frac{f(a+he_i) - f(a)}{h} \\
&= \lim_{h\to 0} \frac{f(a_1, \ldots, a_i + h, \ldots, a_n) - f(a_1, \ldots, a_n)}{h} \\
&= D_i f(a)
\end{aligned}$$

(b)

$$\begin{aligned}
D_{\alpha x} f(a) &= \lim_{t\to 0} \frac{f(a + t\alpha x) - f(a)}{t} \\
&= \lim_{t\to 0} \frac{f(a + \frac{t}{\alpha}\alpha x) - f(a)}{\frac{t}{\alpha}} \\
&= \alpha \lim_{t\to 0} \frac{f(a+tx) - f(a)}{t} \\
&= \alpha D_x f(a)
\end{aligned}$$

(c) Let $f$ be differentiable at $a$. If $g(t) = a + tx$ then

$$\begin{aligned}
D_x f(a) &= \lim_{t\to 0} \frac{f(a+tx) - f(a)}{t} \\
&= \lim_{t\to 0} \frac{(f\circ g)(t) - (f \circ g)(0)}{t} \\
&= D(f\circ g)(0) \\
&= Df(g(0)) \circ Dg(0) \\
&= Df(a)(x)
\end{aligned}$$

By linearity of $Df(a)$ we get the desired result.

$\square$

2.30 Let $f$ be defined as in Problem 2.4. Show that $D_x f(0,0)$ exists for all $x$, but if $g \neq 0$, then $D_{x+y} f(0,0) = D_x f(0,0) + D_y f(0,0)$ is not true for all $x$ and $y$.

*Proof.* We see from 2.4(a) that $D_x f(0,0) = h'(0) = |x|\, g\left(\frac{x}{|x|}\right)$ for $x \neq 0$ and $D_{(0,0)} f(0,0) = 0$. If $g \neq 0$, then there exists $z \neq 0$ such that $g\left(\frac{z}{|z|}\right) \neq 0$. If $x = (\pi_1(z), 0)$ and $y = (0, \pi_2(z))$, then

$$D_{x+y} f(0,0) = D_z f(0,0) = |z|\, g\left(\frac{z}{|z|}\right) \neq 0$$

while, by our assumptions on $g$, we have

$$D_x f(0,0) + D_y f(0,0) = 0 + 0 = 0$$

which shows that

$$D_{x+y} f(0,0) \neq D_x f(0,0) + D_y f(0,0)$$

$\square$

2.31 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined as in Problem 1.26. Show that $D_x f(0,0)$ exists for all $x$, although $f$ is not even continuous at $(0,0)$.

*Proof.* Recall that from 1.26(a), we have that every straight line through $(0,0)$ contains an interval about $(0,0)$ which is contained in $\mathbb{R}^2 \setminus A$, where $A$ is as defined in 1.26. This means $D_x f(0,0)$ exists for all $x$ and

$$D_x f(0,0) = 0$$

$\square$

2.32 (a) Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

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

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

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

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

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

*Proof.* (a) We see that

$$f'(0) = 0$$

and that, for $x \neq 0$

$$f'(x) = 2x \sin(1/x) - \cos(1/x)$$

which is clearly not continuous at 0.

(b) We see that

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

so that $Df(0,0)=0$. With a little calculation, we can see that for $(x,y)\neq(0,0)$ we get

$$D_if(x,y)=2\cdot\pi_i(x,y)\cdot\sin\left(\frac{1}{\sqrt{x^2+y^2}}\right)-\pi_i(x,y)\cdot\frac{\cos\left(\frac{1}{\sqrt{x^2+y^2}}\right)}{\sqrt{x^2+y^2}}$$

It is not too hard to show that when $(x,y)\to(0,0)$ the first term tends to zero but that the term on the end is pathological (Check $\lim_{x\to 0^+} D_1f(x,0)$ for example). Therefore, $D_if$ is not continuous at $(0,0)$.

□

2.33 Show that continuity of $D_1f_i$ at $a$ may be eliminated from the hypothesis of Theorem 2-8.

*Proof.* This is because existence of $D_1f_j(a)$ will imply that

$$\lim_{h\to 0}\frac{|f_j(a+h)-f_j(a)-\sum_{i=1}D_if_j(a)(h_i)|}{|h|}=$$
$$\lim_{h\to 0}\frac{|f_j(a_1+h_1,a_2\dots,a_n)-f_j(a_1,\dots,a_n)-D_1f_j(a)(h_1)|}{|h|}+\frac{|\sum_{i=2}^n[D_if_j(c_i)-D_if_j(a)]h_i|}{|h|}$$

so that continuity of only the other partials is need to obtain the desired limit. □

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

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

*Proof.* If we let $g(t)=f(tx)$ and $h(t)=tx$, then

$$g(t)=(f\circ h)(t)$$

From this it is not too difficult to show that

$$g'(1)=\sum_{i=1}^n x_iD_if(x)$$

Now, if $f$ is homogeneous of degree $m$, then

$$g(t) = f(tx) = t^m f(x)$$

so that

$$g'(t) = mf(x)t^{m-1}$$

and

$$g'(1) = mf(x)$$

□

2.35 If $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable 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.* The author's hint is very revealing. Let $h_x(t) = f(tx)$. So then

$$\int_0^1 h'_x(t)dt = h_x(1) - h_x(0) = f(x) - f(0) = f(x)$$

However, we can see that

$$h'_x(t) = \sum_{i=1}^{n} x_i D_i f(tx)$$

which implies that

$$\begin{aligned} f(x) &= \int_0^1 h'_x(t)dt \\ &= \sum_{i=1}^{n} x_i \int_0^1 D_i f(tx)dt \end{aligned}$$

which implies that if $g_i(x) = \int_0^t D_i f(tx)dt$, we obtain the desired result. □

2.36 Let $A \subset \mathbb{R}^n$ be an open set and $f : A \to \mathbb{R}^n$ a continuously differentiable 1-1 function such that $det f'(x) \neq 0$ 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 $B \subset A$.

*Proof.* Note that the assumptions on $f$ and $A$ imply that the Inverse Function Theorem holds on $f$ for all $x \in A$. So, given $y \in f(A)$ there are open $W_y \subset f(A)$ and $V_x \subset A$ and a function $\tilde{f}^{-1} : W_y \to V_x$ which is continuous and differentiable. But by injectivity of $f$, we know that $f^{-1} : f(A) \to A$ exists and $f^{-1}(z) = \tilde{f}^{-1}(z)$ for $z \in W_y$. So $f^{-1}$ is continuous and differentiable at $y$. Therefore, $f^{-1}$ is differentiable and continuous on $f(A)$. Continuity of $f^{-1}$ implies both that $f(A)$ is open, and that $f(B)$ is open whenever $B \subset A$ is (Willard). □

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

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

*Proof.* (a) Assuming $f$ is continuously differentiable (though we could use a weaker assumption than this) and injective, it follows that $f(x, y)$ is differentiable on $\mathbb{R}^2$ and therefore continuous on $\mathbb{R}^2$. So $f(x, 0)$ and $f(0, y)$ are continuous and injective on the interval $[-1, 1] \subset \mathbb{R}$. By connectedness of $[-1, 1]$ and continuity of $f(x, 0)$ and $f(0, y)$, the images of $[-1, 1]$ by $f(x, 0)$ and $f(0, y)$ are both connected. $f(0, 0)$ being a maximum or minimum of either function on $[-1, 1]$ contradicts with $f$ being injective. If $f(0, 0)$ is not a maximum or a minimum, then it follows by connectedness that both images contain some open interval about $f(0, 0)$, which, again, contradicts with injectivity of $f$.

(b) Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $m < n$, be continuously differentiable. Suppose for contradiction that $f$ is injective. If we define $g : \mathbb{R}^n \to \mathbb{R}^n$ by

$$g(x_1, \ldots, x_n) = (x_1, x_2, \ldots, x_{n-m}, f(x_1, \ldots, x_n))$$

then we know that

$$g'(x) = \left[\begin{array}{c|c} Id & 0 \\ \hline * & \frac{\partial f}{\partial \vec{y}}(x) \end{array}\right]$$

where

$$\frac{\partial f}{\partial \vec{y}}(x) = \begin{bmatrix} D_{n-m+1}f_1(x) & \ldots & D_n f_1(x) \\ \vdots & & \vdots \\ D_{n-m+1}f_m(x) & \ldots & D_n f_m(x) \end{bmatrix}$$

and $*$ is unimportant since $det g'(x) \neq 0$ iff $det \frac{\partial f}{\partial \vec{y}}(x) \neq 0$. Now, if $det \frac{\partial f}{\partial \vec{y}}(x) = 0$ then there exists $v \in \mathbb{R}^m$ such that $v \neq 0$ and

$$\left[\frac{\partial f}{\partial \vec{y}}(x)\right] v = 0$$

which implies that if $v^* \in \mathbb{R}^n$ admitting $v^* = (0, \ldots, 0, v)$ then we get that

$$D_{v^*} f(x) = Df(x)(v^*) = \left[\frac{\partial f}{\partial \vec{y}}(x)\right] v = \vec{0} \in \mathbb{R}^m$$

□

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 not injective (on all $\mathbb{R}$).

(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 $(x,y) \in \mathbb{R}^2$ but $f$ is not injective.

*Proof.* (a) This is a simple consequence of the Mean Value Theorem.

(b) We see that

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

so that

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

and yet $f(0,0) = (1,0) = f(0, 2\pi)$.

□

2.39 Use the function $f : \mathbb{R} \to \mathbb{R}$ defined by

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

To show that continuity of the derivative cannot be eliminated from the hypotheses in theorem 2-11.

*Proof.* Observe that

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

and, for $x \neq 0$

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

so that $f'(0) \neq 0$ but $f'(x)$ is continuously differentiable only when $x \neq 0$. One could argue by relatively elementary means that

$$f\left(\frac{1}{(2n+1)\pi}\right) < f\left(\frac{1}{(n+1)2\pi}\right) < f\left(\frac{1}{n2\pi}\right)$$

which implies that $f$ is not injective on any interval about 0. □

2.40 Use the implicit function theorem to re-do Problem 2-15(c).

*Proof.* The assumptions in 2-15(c) are not strong enough to enable us to use the Implicit Function Theorem. If we assume that each $a_{ij}(t)$ and $b_i(t)$ are continuously differentiable, then letting $[A(t)]_{ij} = a_{ij}(t)$, $s = (s_1, \ldots, s_n) \in \mathbb{R}^n$ and $b : \mathbb{R} \to \mathbb{R}^n$ with $b(t) = (b_1(t), \ldots, b_n(t))$, we can define $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ by

$$f(t,s) = A(t)^T s - b(t)$$

By non-singularity of $A(t)$, we know that for every $t$ there is $s$ such that $f(t,s) = 0$. Furthermore, by $a_{ij}(t)$ and $b_i(t)$ being continuously differentiable, we know that $D_1 f(t,s)$ exists and is continuous. It is easy to show that for all $i \geq 2$, $D_i f(t,s)$ exists and is continuous since $f(t,s)$ is affine in $s$. Indeed, if $M$ is as defined in the theorem, then $M = A$ and so is non-singular for all $t$. Thus, by the Implicit Function Theorem, we will have that for all $t$, we can find $s(t)$ so that $f(t, s(t)) = 0$ with $s$ differentiable at $t$. $\square$

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

(a) If $D_{2,2} f(x,y) \neq 0$ 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

$$\begin{aligned} D_{2,1} f(x,y) &= 0 \\ D_2 f(x,y) &= 0 \end{aligned}$$

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

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

*Proof.* (a) If we just assume that $D_2 f(x,y)$ is continuously differentiable, then we will get that, for $M$ as defined in Theorem 2-12, $M = D_{2,2} f(x, c(x))$ so that $det\ M \neq 0$ for all $x \in \mathbb{R}$. The Implicit Function Theorem then implies that $c'(x)$ exists for all $x \in \mathbb{R}$. To compute $c'(x)$ we simply see that if we let $h(x) = (x, c(x))$, then $[D_2 f] \circ h = 0$. This implies

$$\begin{aligned} 0 &= (D_2 f \circ h)'(x) \\ &= \left[ D_{2,1} f(x, c(x)) \;\; D_{2,2} f(x, c(x)) \right] \begin{bmatrix} 1 \\ c'(x) \end{bmatrix} \\ &= D_{2,1} f(x, c(x)) + D_{2,2} f(x, c(x)) c'(x) \end{aligned}$$

which implies the desired result.

(b) Note that $g_x'(y) = D_2 f(x,y)$. So, if $c'(x) = 0$, then, letting $y = c(x)$ yields

$$g_x'(y) = D_2 f(x,y) = 0 = -\frac{D_{2,1} f(x,y)}{D_{2,2} f(x,y)}$$

(c) Maybe some other time.

$\square$

# Chapter 3

# Integration

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

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

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

*Proof.* We see that for any partition $P = (P_1, P_2)$ that admits

$$\left\{0, \frac{1}{2}, 1\right\} \subset P_1$$

We have

$$L(f,P) = U(f,P) = \frac{1}{2}$$

Thus, if we have an arbitrary partition, $P$, then, with the proper refinement and by Corollary 3-2, we obtain

$$\int_{[0,1]\times[0,1]} f = \frac{1}{2}$$

□

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

*Proof.* Let us denote those values of $x$ where $g$ and $f$ differ by $x_1, \dots, x_m$. We will show that $\forall U(f,P)$, $\exists_{P'} U(g,P') \leq U(f,P)$, and similar for the lower sums. This will imply that $\forall P$, $\exists P'$ such that

$$L(f,P) \leq L(g,P') \leq U(g,P') \leq U(f,P)$$

which, in turn, implies that $g$ is integrable and that $\int_A g = \int_A f$. So, given a partition, $P$ we can first refine it by adding enough subrectangles so that each subrectangle contains at most one $x_i$. Let us denote this partition $P^*$. We can refine $P^*$ by including subrectangles around each $x_i$ with sufficiently small volume. This refinement we denote by, $P'$. Now, if $S$ is a subrectangle in $P^*$ that does not contain any $x_i$ then

$$M_S(f)v(S) = M_S(g)v(S)$$

If $R$ is a subrectangle in $P^*$ that does include an $x_i$, then, in $P'$, we have $R = \cup_j R_j$ where $R_j$ is a subrectangle in $P'$ and $x_i \in R_k$ for some $k$. Note $R_j$ does not contain any $x_i$ for $j \neq k$. But then $M_{R_j}(g) \leq M_R(f)$ for all $j \neq k$, and $\sum_{j \neq k} v(R_j) < v(R)$ so that

$$\sum_{j \neq k} M_{R_j}(g)v(R_j) < M_R(f)v(R)$$

Since $v(R_k)$ was sufficiently small, in the sense that

$$M_{R_k}(g)v(R_k) < M_R(f)v(R) - \sum_{j \neq k} M_{R_i}(g)v(R_i)$$

we get

$$0 < M_R(f)v(R) - \sum_j M_{R_i}(g)v(R_i)$$

But then $U(f, P^*) \leq U(f, P)$ and

$$\begin{aligned} U(f, P^*) - U(g, P') &= \sum_{S \cup R} M_{S \cup R}(f)v(S \cup R) - \sum_{S \cup R'} M_{S \cup R'}(g)v(S \cup R') \\ &= \sum_R M_R(f)v(R) - \sum_{R'} M_{R'}(g)v(R') \\ &> 0 \end{aligned}$$

which implies

$$U(g, P') \leq U(f, P)$$

Similar reasoning produces an analogous result for the lower sums, which implies the conclusion. $\square$

3.3 *Proof.* (a) Notice that for all $x \in S$ we have

$$m_S(f) + m_S(g) \leq f(x) + g(x) \quad \text{and} \quad f(x) + g(x) \leq M_S(f) + M_S(g)$$

which implies that

$$\begin{aligned} L(f,P)+L(g,P) &= \sum_S m_S(f)v(S) + \sum_S m_S(g)v(S) \\ &= \sum_S (m_S(f)+m_S(g))v(S) \\ &\le \sum_S m_S(f+g)v(S) \\ &= L(f+g,P) \end{aligned}$$

and similar for the upper sums.

(b) Note that, for $\epsilon > 0$ given, for well-chosen $P$, we have

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

which implies that $\int_A f+g$ exists. Now, if for some $P$

$$U(f+g,P) < \int_A f + \int_A g$$

then there is $P'$ such that

$$U(f+g,P) < L(f,P') + L(g,P') < \int_A f + \int_A g$$

which would imply $U(f+g,P) < L(f+g,P')$ which is absurd. Combined with the above inequality, we get that for all $\epsilon > 0$, there exists $P$ such that

$$\int_A f + \int_A g < U(f+g,P) < \int_A f + \int_A g + \epsilon$$

Therefore, $\int_A f+g = \int_A f + \int_A g$.

(c) Two cases, too tedious.

□

3.4 Let $f : A \to \mathbb{R}$ and let $P$ be a partition of $A$. Show that $f$ is integrable if and only if for each subrectangle $S$ the function $f|S$ is integrable, and that in this case

$$\int_A f = \sum_S \int_S f|S$$

*Proof.* If $P_S$ is the portion of the partition $P$ which produces subrectangle $S$, then

$$\sum_S U(f|S, P_S) = U(f, P)$$
$$\sum_S L(f|S, P_S) = L(f, P)$$

So, there exists a refinement, $P'$, of $P$ such that

$$U(f, P') - L(f, P') < \epsilon$$

which implies

$$\sum_S U(f|S, P'_S) - L(f|S, P'_S) < \epsilon$$

which implies $\sum_S \int_S f|S$ exists. Furthermore

$$0 < U(f, P') - \int_A f < \epsilon$$

which implies

$$0 < \sum_S U(f|S, P'_S) - \int_A f < \epsilon$$

which implies that

$$\sum_S \int_S f|S = \int_A f$$

□

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.* Notice that for any $P$

$$U(f, P) \leq U(g, P)$$

By integrability of both $f$ and $g$, we get

$$\int_A f = \inf_P U(f, P) \leq \inf_P U(g, P) = \int_A g$$

□

3.6 If $f : A \to \mathbb{R}$ is integrable, show that $|f|$ is integrable and $\left|\int_A f\right| \leq \int_A |f|$.

*Proof.* We wish to show that for any partition $P$,

$$M_S(|f|) - m_S(|f|) \leq M_S(f) - m_S(f)$$

Notice that this is trivially true if $\forall_{x \in S} f(x) \geq 0$ or $f(x) \leq 0$. Otherwise, we observe that

(a) $m_S(f) < 0 \leq m_s(|f|) \leq M_S(f)$

(b) Either $M_S(|f|) = M_S(f)$ or $M_S(|f|) = -m_S(f)$.

If $M_S(|f|) = M_S(f)$ then

$$M_S(|f|) - m_S(|f|) = M_S(f) - m_S(|f|) \leq M_S(f) - m_S(f)$$

where the inequality comes from (a). If $M_S(|f|) = -m_S(f)$ then

$$M_S(|f|) - m_S(|f|) \leq M_S(|f|) = -m_S(f) \leq M_S(f) - m_S(f)$$

Thus

$$M_S(|f|) - m_S(|f|) \leq M_S(f) - m_S(f)$$

which implies that $|f|$ is integrable and

$$\left| \int_A f \right| \leq \int_A |f|$$

follows from the fact that $|M_S(f)| \leq M_S(|f|)$, the integrability of $|f|$, and the triangle inequality. □

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

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

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

*Proof.* □

3.8 Prove that $[a_1, b_1] \times \ldots \times [a_n, b_n]$ does not have content zero if $a_i < b_i$ for each $i$.

*Proof.* Similar to 3-5. □

3.9 (a) Show that an unbounded set cannot have content 0.

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

*Proof.* (a) Notice that given $\{U_1, \ldots, U_n\}$ we will get that $\cup_{i=1}^n U_i$ is bounded and therefore cannot cover an unbounded set.

(b) $\mathbb{N}$ is closed but unbounded and therefore cannot have content 0. However, given $0 < \epsilon < 1$, if we take $U_i$ to be the rectangle, centered at $i$, such that $v(U_i) = (\frac{\epsilon}{2})^i$ then

$$\sum_{i=1}^{\infty} v(U_i) = \sum_{i=1}^{\infty} \left(\frac{\epsilon}{2}\right)^i = \frac{\epsilon}{2-\epsilon} < \epsilon$$

□

3.10 (a) If $C$ is a set of content 0, show that the boundary of $C$ has content 0.

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

*Proof.* (a) If $\{U_1, \ldots, U_n\}$ is a cover of $C$ then $\cup_{i=1}^{n} U_i$ is closed and contains $C$. Therefore, it contains the boundary of $C$. So $\{U_1, \ldots, U_n\}$ is also a cover for the boundary of $C$. Therefore, the boundary also has content 0.

(b) Notice that $\mathbb{Q} \cap [0,1]$ is countable and therefore has measure 0. On the other hand the boundary is $[0,1]$ which does not have measure 0.

□

3.11 Let $A$ be the set in problem 1.18. If $\sum_{i=1}^{\infty}(b_i - a_i) < 1$, show that the boundary of $A$ does not have measure 0.

*Proof.* Suppose to the contrary. From 1.18, we know that boundary$A = [0,1] \setminus A$. So $[0,1] = A \cup \text{boundary}A$. Since we can use open rectangles, we have a cover of $A$, $\{(a_i, b_i) : i \in \mathbb{N}\}$, and a cover of boundary$A$, $\{U_1, U_2, \ldots\}$, such that

$$\sum_{i=1}^{\infty} v(U_i) < 1 - \sum_{i=1}^{\infty}(b_i - a_i)$$

But then $\{(a_i, b_i) : i \in \mathbb{N}\} \cup \{U_1, U_2, \ldots\}$ forms a cover of $[0,1]$. But

$$\sum_{i=1}^{\infty} v(U_i) + \sum_{i=1}^{\infty}(b_i - a_i) < 1$$

which would contradict with it being a cover of $[0,1]$. □

3.12 Let $f : [a,b] \to \mathbb{R}$ be an increasing function. Show that $\{x : f(x) \text{ is discontinuous at } x\}$ has measure 0.

*Proof.* First we will show that $\{x : o(f,x) > \frac{1}{n}\}$ is finite for any given $n$. To this end, we suppose to the contrary, that there exists an $n \in \mathbb{N}$ so that we have $\{x : o(f,x) > \frac{1}{n}\}$ is infinite. Then we can select a countable sequence of distinct elements from it:

$$x_0 < x_1 < x_2 < \dots$$

So, by $f$ increasing, we get that for every $k$

$$f(x_k) - f(x_0) = \sum_{i=1}^{k} f(x_i) - f(x_{i-1}) \geq \sum_{i=1}^{k} \frac{1}{n} = \frac{k}{n}$$

which would imply that $f$ is not bounded from above on $[a,b]$ contradicting with $f(x) \leq f(b)$ for all $x \in [a,b]$. Thus

$$\left\{x : \exists n \text{ such that } o(f,x) > \frac{1}{n}\right\}$$

is at most countable, and therefore has measure 0. But, by Theorem 1.10, $f$ is continuous at $x$ iff $o(f,x) = 0$. Therefore, $\{x : f(x) \text{ is discontinuous at } x\}$ has measure 0. □

3.13

3.14 Show if $f, g : A \to \mathbb{R}$ are integrable, then so is $f \cdot g$.

*Proof.* Notice that $(f \cdot g)(x)$ is continuous wherever both $f$ and $g$ are. Since both $f$ and $g$ are continuous, it follows that both $\{x : f \text{ is not continuous at } x\}$ and $\{x : g \text{ is not continuous at } x\}$ have measure zero. This implies $\{x : f \cdot g \text{ is not continuous at } x\}$ has measure zero. Therefore $f \cdot g$ is integrable. □

3.15 Show that if $C$ has content 0, then $C \subset A$ for some closed rectangle $A$ and $C$ is Jordan-measurable and $\int_A \chi_C = 0$.

*Proof.* Since $C$ has content zero, $C \subset \cup_{i=1}^{n} U_i \subset A$ for some rectangle $A$, where the latter inclusion follows from boundedness of the union. Notice that $\cup_{i=1}^{n} U_i$ contains the closure of $C$ and therefore its boundary. Therefore, the boundary of $C$ has content 0 and so $C$ is Jordan-measurable. Finally, notice that if $P$ is a partition containing the cover $\{U_i\}$, then $U(f,P) = \sum v(U_i) < \epsilon$, implying $\int_A \chi_C = 0$. □

3.16 Give an example of a bounded set $C$ of measure 0 such that $\int_A \chi_C$ does not exist.

*Proof.* If $C = \mathbb{Q} \cap [0,1]$ then clearly $C$ is bounded. Moreover, since $C \subset \mathbb{Q}$ and $\mathbb{Q}$ has measure 0, it follows $C$ has measure 0. Furthermore, boundary $C = [0,1]$ which clearly does not have measure 0. Therefore, $\chi_C$ is not integrable, and $\int_A \chi_C$ does not exist. □

3.17 If $C$ is a bounded set of measure 0 and $\int_A \chi_C$ exists, show that $\int_A \chi_C = 0$.

*Proof.* We claim that if $U$ is a (non-degenerate) rectangle, then $U \not\subset C$, for otherwise, by problem 3.8, $C$ would not have measure 0. Therefore, it follows that if $U$ is a rectangle then $U$ contains points not in $C$. Therefore $L(\chi_C, P) = 0$ for every partition $P$ which, by integrability of $\chi_C$, implies that $\int_A \chi_C = 0$. □

3.18 If $f : A \to \mathbb{R}$ is nonnegative and $\int_A f = 0$, show that $\{x : f(x) \neq 0\}$ has measure 0.

*Proof.* Observe that our hypotheses on $f$ imply that for any subrectangle, $S$, we have that $m_S(f) = 0$. In order to obtain a contradiction, let us suppose that there is $c \in A$ such that $f(c) \neq 0$ and $f$ is continuous at $c$. But then there exists some subrectangle $S$ such that $x \in S$ implies

$$|f(x) - f(c)| < \frac{f(c)}{2}$$

which further implies that $0 < \frac{f(c)}{2} \leq f(x)$ for all $x \in S$, contradicting with $m_S(f) = 0$. Therefore $\{x : f(x) \neq 0\} \subset \{x : f \text{ is not continuous at } x\}$. Integrability of $f$ and Theorem 3-8 imply the result. □

3.19 Let $U$ be the open set of Problem 3.11. Show that if $f = \chi_U$ except on a set of measure 0, then $f$ is not integrable on $[0,1]$.

3.20 Show that an increasing function $f : [a,b] \to \mathbb{R}$ is integrable on $[a,b]$.

*Proof.* Let $B$ be defined as in Theorem 3.8. Notice, by $f$ increasing, we have

$$f(a) + \sum_{i=1}^{\infty} o(f, \beta_i) \leq f(b)$$

for any countable subcollection $\{\beta_i : \beta \in B, i \in \mathbb{N}\}$. But then

$$B_n = \left\{ \beta : o(f, \beta) \geq \frac{1}{n} \right\}$$

is finite for every $n \geq 2$. But then

$$B = \bigcup_{n=2}^{\infty} B_n$$

is at most countable. Therefore, by Theorem 3.8, it follows that $f$ is integrable. □

3.21 If $A$ is a closed rectangle, show that $C \subset A$ is Jordan-measurable if and only if for every $\epsilon > 0$ there exists a partition $P$ of $A$ such that $\sum_{S \in \mathcal{S}_1} v(S) - \sum_{S \in \mathcal{S}_2} < \epsilon$ where $\mathcal{S}_1$ consists of all subrectangles intersecting $C$ and $\mathcal{S}_2$ all subrectangles contained in $C$.

*Proof.* Notice that

$$\mathcal{S}_1 = \mathcal{S}_2 \cup \mathcal{S}_3$$

where $\mathcal{S}_3$ is the collection of subrectangles that intersect with $C$ but are not contained in $C$. Thus the expression

$$\sum_{S \in \mathcal{S}_1} v(S) - \sum_{S \in \mathcal{S}_2} v(S) < \epsilon$$

is equivalent to

$$\sum_{S \in \mathcal{S}_2 \cup \mathcal{S}_3} v(S) - \sum_{S \in \mathcal{S}_2} v(S) = \sum_{S \in \mathcal{S}_2} v(S) + \sum_{S \in \mathcal{S}_3} v(S) - \sum_{S \in \mathcal{S}_2} v(S) = \sum_{S \in \mathcal{S}_3} v(S) < \epsilon$$

which is equivalent to $C$ having boundary with measure zero, and therefore, Jordan-measurable. □

3.22 If $A$ is a Jordan-measurable set and $\epsilon > 0$, show that there is a compact Jordan-measurable set $C \subset A$ such that $\int_{A \setminus C} 1 < \epsilon$.

*Proof.* Set $C = \mathcal{S}_2$, where $\mathcal{S}_2$ is as defined in problem 3.21. □

3.23 Let $C \subset A \times B$ be a set of content 0. Let $A' \subset A$ be the set of all $x \in A$ such that $\{y \in B : (x, y) \in C\}$ is *not* of content 0. Show that $A'$ is a set of measure 0.