Chapter 9: Functions of Several Variables

Author: Meng-Gen Tsai Email: plover@gmail.com

Exercise 9.1. If S is a nonempty subset of a vector space X, prove (as asserted in Section 9.1) that the span of S is a vector space.

Denote the span of S by span(S).

Proof.

- (1) Since $S \neq \emptyset$, there is $\mathbf{z} \in S$. So $1\mathbf{z} = \mathbf{z} \in \text{span}(S) \neq \emptyset$. (In fact, $\text{span}(S) \supseteq S$.)
- (2) If $\mathbf{x}, \mathbf{y} \in \text{span}(S)$, then there exist elements $\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{y}_1, \dots, \mathbf{y}_n \in S$ and scalars $a_1, \dots, a_m, b_1, \dots, b_n$ such that

$$\mathbf{x} = a_1 \mathbf{x}_1 + \dots + a_m \mathbf{x}_m,$$

$$\mathbf{y} = b_1 \mathbf{y}_1 + \dots + b_n \mathbf{y}_n.$$

Then

$$\mathbf{x} + \mathbf{y} = a_1 \mathbf{x}_1 + \dots + a_m \mathbf{x}_m + b_1 \mathbf{y}_1 + \dots + b_n \mathbf{y}_n$$

is a linear combination of the elements of S. For any scalar c,

$$c\mathbf{x} = (ca_1)\mathbf{x}_1 + \dots + (ca_m)\mathbf{x}_m$$

is again linear combination of the elements of S.

(3) By (1)(2), span(S) is a vector space.

Note. Any subspace of X that contains S must also contain span(S).

Exercise 9.2. Prove (as asserted in Section 9.6) that BA is linear if A and B are linear transformations. Prove also that A^{-1} is linear and invertible if A is invertible.

Proof. Use the notation in Definitions 9.6.

(1) Show that BA is linear if A and B are linear transformations. Let X, Y, Z be vector spaces, $A \in L(X, Y)$ and $B \in L(Y, Z)$.

(a) Given any $\mathbf{x}_1, \mathbf{x}_2 \in X$.

$$(BA)(\mathbf{x}_1 + \mathbf{x}_2) = B(A(\mathbf{x}_1 + \mathbf{x}_2))$$

= $B(A\mathbf{x}_1 + A\mathbf{x}_2)$ (A is a linear transformation)
= $B(A\mathbf{x}_1) + B(A\mathbf{x}_2)$ (B is a linear transformation)
= $(BA)\mathbf{x}_1 + (BA)\mathbf{x}_2$.

(b) For any $\mathbf{x} \in X$ and scalar c,

$$(BA)(c\mathbf{x}) = B(A(c\mathbf{x}))$$

= $B(cA\mathbf{x})$ (A is a linear transformation)
= $cB(A\mathbf{x})$ (B is a linear transformation)
= $c(BA)\mathbf{x}$.

By (a)(b), $BA \in L(X, Z)$.

- (2) Show that A^{-1} is linear if A is invertible.
 - (a) Given any $\mathbf{y}_1, \mathbf{y}_2 \in X$. Since A is surjective, there exist $\mathbf{x}_1, \mathbf{x}_2 \in X$ such that

$$\mathbf{y}_1 = A\mathbf{x}_1$$
$$\mathbf{y}_2 = A\mathbf{x}_2.$$

So

$$A^{-1}\mathbf{y}_1 = A^{-1}(A\mathbf{x}_1) = \mathbf{x}_1$$

 $A^{-1}\mathbf{y}_2 = A^{-1}(A\mathbf{x}_2) = \mathbf{x}_2$

(by Definitions 9.4). Hence

$$A^{-1}(\mathbf{y}_1 + \mathbf{y}_2) = A^{-1}(A\mathbf{x}_1 + A\mathbf{x}_2)$$

$$= A^{-1}(A(\mathbf{x}_1 + \mathbf{x}_2)) \quad (A \text{ is a linear transformation})$$

$$= \mathbf{x}_1 + \mathbf{x}_2 \qquad (Definitions 9.4)$$

$$= A^{-1}\mathbf{y}_1 + A^{-1}\mathbf{y}_2.$$

(b) For any $\mathbf{y} \in X$ and scalar c, there is a corresponding $\mathbf{x} \in X$ such that $\mathbf{y} = A\mathbf{x}$ since A is surjective. So $A^{-1}\mathbf{y} = \mathbf{x}$ by Definition 9.4. Hence

$$A^{-1}(c\mathbf{y}) = A^{-1}(cA\mathbf{x})$$

= $A^{-1}(A(c\mathbf{x}))$ (A is a linear transformation)
= $c\mathbf{x}$ (Definitions 9.4)
= $cA^{-1}\mathbf{y}$.

By (a)(b), $A^{-1} \in L(X)$.

- (3) Show that A^{-1} is invertible if A is invertible. It suffices to show that A^{-1} is injective and surjective.
 - (a) Show that A^{-1} is injective. Given any $\mathbf{y}_1, \mathbf{y}_2 \in X$. Since A is surjective, there exist $\mathbf{x}_1, \mathbf{x}_2 \in X$ such that

$$\mathbf{y}_1 = A\mathbf{x}_1$$
$$\mathbf{y}_2 = A\mathbf{x}_2.$$

Suppose $A^{-1}\mathbf{y}_1 = A^{-1}\mathbf{y}_2$. So $A^{-1}(A\mathbf{x}_1) = A^{-1}(A\mathbf{x}_2)$, or $\mathbf{x}_1 = \mathbf{x}_2$, or $\mathbf{y}_1 = A\mathbf{x}_1 = A\mathbf{x}_2 = \mathbf{y}_2$.

(b) Show that A^{-1} is surjective. For any $\mathbf{x} \in X$, there exists $A\mathbf{x} \in X$ such that $A^{-1}(A\mathbf{x}) = \mathbf{x}$ by Definitions 9.4.

Exercise 9.3. Assume $A \in L(X,Y)$ and $A\mathbf{x} = \mathbf{0}$ only when $\mathbf{x} = \mathbf{0}$. Prove that A is then 1-1.

Proof. Suppose $A\mathbf{x} = A\mathbf{y}$. Since A is a linear transformation, $A(\mathbf{x} - \mathbf{y}) = A\mathbf{x} - A\mathbf{y} = \mathbf{0}$. By assumption, $\mathbf{x} - \mathbf{y} = \mathbf{0}$ or $\mathbf{x} = \mathbf{y}$. \square

Exercise 9.4. Prove (as asserted in Section 9.30) that null spaces and ranges of linear transformations are vector spaces.

Proof. Use the notation in Definitions 9.30. Suppose X, Y are vector spaces, and $A \in L(X,Y)$, as in Definition 9.6.

- (1) Show that $\mathcal{N}(A)$ is a vector space in X.
 - (a) Note that $\mathbf{0} \in X$. Since $A\mathbf{0} = \mathbf{0}$, $\mathbf{0} \in \mathcal{N}(A) \neq \emptyset$.
 - (b) Suppose $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{N}(A)$. Then

$$\begin{split} A(\mathbf{x}_1+\mathbf{x}_2) &= A\mathbf{x}_1 + A\mathbf{x}_2 & \quad (A \text{ is a linear transformation}) \\ &= \mathbf{0} + \mathbf{0} & \quad (\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{N}(A)) \\ &= \mathbf{0}. \end{split}$$

So $\mathbf{x}_1 + \mathbf{x}_2 \in \mathcal{N}(A)$.

(c) Suppose $\mathbf{x} \in \mathcal{N}(A)$ and c is a scalar. Then

$$A(c\mathbf{x}) = cA\mathbf{x}$$
 (A is a linear transformation)
= $c\mathbf{0}$ ($\mathbf{x} \in \mathcal{N}(A)$)
= $\mathbf{0}$.

So $c\mathbf{x} \in \mathcal{N}(A)$.

By (a)(b)(c), $\mathcal{N}(A)$ is a vector space.

- (2) Show that $\mathcal{R}(A)$ is a vector space in Y.
 - (a) Note that $\mathbf{0} \in X$. So $A\mathbf{0} = \mathbf{0} \in \mathcal{R}(A) \neq \emptyset$.
 - (b) Suppose $\mathbf{y}_1, \mathbf{y}_2 \in \mathcal{R}(A)$. Then there exist $\mathbf{x}_1, \mathbf{x}_2 \in X$ such that $A\mathbf{x}_1 = \mathbf{y}_1$ and $A\mathbf{x}_2 = \mathbf{y}_2$. Hence

$$\mathbf{y}_1 + \mathbf{y}_2 = A\mathbf{x}_1 + A\mathbf{x}_2$$

= $A(\mathbf{x}_1 + \mathbf{x}_2)$ (A is a linear transformation).

So $\mathbf{y}_1 + \mathbf{y}_2 \in \mathscr{R}(A)$.

(c) Suppose $\mathbf{y} \in \mathcal{R}(A)$ and c is a scalar. Then there exists $\mathbf{x} \in X$ such that $A\mathbf{x} = \mathbf{y}$. Hence

$$c\mathbf{y} = cA\mathbf{x}$$

= $A(c\mathbf{x})$ (A is a linear transformation).

So $c\mathbf{y} \in \mathcal{R}(A)$.

By (a)(b)(c), $\mathcal{R}(A)$ is a vector space.

Exercise 9.5. Prove that to every $A \in L(\mathbb{R}^n, \mathbb{R}^1)$ corresponds a unique $\mathbf{y} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{x} \cdot \mathbf{y}$. Prove also that $||A|| = |\mathbf{y}|$. (Hint: Under certain conditions, equality holds in the Schwarz inequality.)

Proof.

- (1) Recall that $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is the standard basis of \mathbb{R}^n (Definitions 9.1). Given any $\mathbf{x} \in \mathbb{R}^n$, write $\mathbf{x} = (x_1, \dots, x_n)$ as $\mathbf{x} = \sum x_i \mathbf{e}_i$.
- (2) Show that y exists. Since A is a linear transformation,

$$A\mathbf{x} = A\left(\sum x_j \mathbf{e}_j\right)$$

$$= \sum x_j A \mathbf{e}_j$$

$$= (x_1, \dots, x_n) \cdot (A \mathbf{e}_1, \dots, A \mathbf{e}_n)$$

$$= \mathbf{x} \cdot \sum (A \mathbf{e}_j) \mathbf{e}_j.$$

Define $\mathbf{y} = \sum (A\mathbf{e}_i)\mathbf{e}_i \in \mathbb{R}^n$ so that $A\mathbf{x} = \mathbf{x} \cdot \mathbf{y}$.

(3) Show that **y** is unique. Suppose there exists some $\mathbf{z} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{x} \cdot \mathbf{z}$. So

$$0 = A\mathbf{x} - A\mathbf{x}$$
$$= \mathbf{x} \cdot \mathbf{y} - \mathbf{x} \cdot \mathbf{z}$$
$$= \mathbf{x} \cdot (\mathbf{y} - \mathbf{z})$$

for any $\mathbf{x} \in \mathbb{R}^n$. In particular, take $\mathbf{x} = \mathbf{y} - \mathbf{z} \in \mathbb{R}^n$ to get

$$0 = (\mathbf{y} - \mathbf{z}) \cdot (\mathbf{y} - \mathbf{z}) = |\mathbf{y} - \mathbf{z}|^2$$

or y - z = 0 or y = z.

(4) Show that $||A|| = |\mathbf{y}|$. By the Schwarz inequality (Theorem 1.37(d)),

$$|A\mathbf{x}| = |\mathbf{x} \cdot \mathbf{y}| \le |\mathbf{x}||\mathbf{y}| \le |\mathbf{y}|$$

as $|\mathbf{x}| \leq 1$. Take the sup over all $|\mathbf{x}| \leq 1$ to get

$$||A|| \leq |\mathbf{y}|.$$

If $\mathbf{y} = \mathbf{0}$, then $||A|| = |\mathbf{y}| = 0$. If $\mathbf{y} \neq \mathbf{0}$, then the equality holds when $\mathbf{x} = \frac{\mathbf{y}}{|\mathbf{y}|} \in \mathbb{R}^n$. (Here $|\mathbf{x}| = 1$.)

Exercise 9.6. If f(0,0) = 0 and

$$f(x,y) = \frac{xy}{x^2 + y^2}$$
 if $(x,y) \neq (0,0)$,

prove that $(D_1f)(x,y)$ and $(D_2f)(x,y)$ exist at every point of \mathbb{R}^2 , although f is not continuous at (0,0).

Proof.

(1) Show that

$$(D_1 f)(x,y) = \begin{cases} 0 & \text{if } (x,y) = (0,0), \\ \frac{y(y^2 - x^2)}{(x^2 + y^2)^2} & \text{if } (x,y) \neq (0,0). \end{cases}$$

Write

$$(D_1 f)(x,y) = \lim_{t \to 0} \frac{f((x,y) + t(1,0)) - f(x,y)}{t}$$
$$= \lim_{t \to 0} \frac{f(x+t,y) - f(x,y)}{t}.$$

If (x, y) = (0, 0),

$$(D_1 f)(0,0) = \lim_{t \to 0} \frac{f(t,0) - f(0,0)}{t} = \lim_{t \to 0} \frac{0 - 0}{t} = 0.$$

If $(x, y) \neq (0, 0)$,

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

$$= \lim_{t \to 0} \frac{\frac{(x + t)y}{(x + t)^2 + y^2} - \frac{xy}{x^2 + y^2}}{t}$$

$$= \lim_{t \to 0} \frac{y(y^2 - x^2) - txy}{((x + t)^2 + y^2)(x^2 + y^2)}$$

$$= \frac{y(y^2 - x^2)}{(x^2 + y^2)^2}.$$

(2) Show that

$$(D_2 f)(x, y) = \begin{cases} 0 & \text{if } (x, y) = (0, 0), \\ \frac{x(x^2 - y^2)}{(x^2 + y^2)^2} & \text{if } (x, y) \neq (0, 0). \end{cases}$$

Similar to (1).

(3) Show that f is not continuous at (0,0). Note that

$$\lim_{n \to \infty} f\left(\frac{1}{n}, \frac{1}{n}\right) = \lim_{n \to \infty} \frac{\frac{1}{n} \cdot \frac{1}{n}}{\frac{1}{n^2} + \frac{1}{n^2}} = \lim_{n \to \infty} \frac{1}{2} = \frac{1}{2}$$

and

$$\lim_{n\to\infty} f\left(\frac{1}{n},0\right) = \lim_{n\to\infty} \frac{0}{\frac{1}{n^2}+0} = \lim_{n\to\infty} 0 = 0.$$

Hence the limit $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Exercise 9.7. Suppose that f is a real-valued function defined in an open set $E \subseteq \mathbb{R}^n$, and that the partial derivatives $D_1 f, \ldots, D_n f$ are bounded in E. Prove that f is continuous in E. (Hint: Proceed as in the proof of Theorem 9.21.)

Proof.

- (1) Since $D_j f$ is bounded in E, there is a real number M_j such that $|D_j f| \le M_j$ in E. Take $M = \max_{1 \le j \le n} M_j$ so that $|D_j f| \le M$ in E for all $1 \le j \le n$.
- (2) Fix $\mathbf{x} \in E$ and $\varepsilon > 0$. Since E is open, there is an open neighborhood

$$B(\mathbf{x}; r) = {\mathbf{x} + \mathbf{h} \in E : |\mathbf{h}| < r} \subseteq E$$

with

$$0 < r < \frac{\varepsilon}{n(M+1)}.$$

(3) Write $\mathbf{h} = \sum h_j \mathbf{e}_j$, $|\mathbf{h}| < r$, put $\mathbf{v}_0 = \mathbf{0}$, and $\mathbf{v}_k = h_1 \mathbf{e}_1 + \cdots + h_k \mathbf{e}_k$ for $1 \le k \le n$. Then

$$f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = \sum_{j=1}^{n} [f(\mathbf{x} + \mathbf{v}_j) - f(\mathbf{x} + \mathbf{v}_{j-1})].$$

Since $|\mathbf{v}_k| < r$ for $1 \le k \le n$ and since $B(\mathbf{x}; r)$ is convex, the open interval with end points $\mathbf{x} + \mathbf{v}_{j-1}$ and $\mathbf{x} + \mathbf{v}_j$ lie in $B(\mathbf{x}; r)$. Since $\mathbf{v}_j = \mathbf{v}_{j-1} - h_j \mathbf{e}_j$, the mean value theorem (Theorem 5.10) show that

$$f(\mathbf{x} + \mathbf{v}_j) - f(\mathbf{x} + \mathbf{v}_{j-1}) = h_j(D_j f)(\mathbf{x} + \mathbf{v}_{j-1} + \theta_j h_j \mathbf{e}_j)$$

for some $\theta_i \in (0,1)$.

(4) Note that $|h_j| \leq |\mathbf{h}| < r < \frac{\varepsilon}{n(M+1)}$. Hence

$$|f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})| \le \sum_{j=1}^{n} |f(\mathbf{x} + \mathbf{v}_j) - f(\mathbf{x} + \mathbf{v}_{j-1})|$$

$$= \sum_{j=1}^{n} |h_j| |(D_j f)(\mathbf{x} + \mathbf{v}_{j-1} + \theta_j h_j \mathbf{e}_j)|$$

$$\le \sum_{j=1}^{n} \frac{\varepsilon}{n(M+1)} \cdot M$$

$$< \varepsilon$$

as $|\mathbf{h}| < r < \frac{\varepsilon}{n(M+1)}$. Hence f is continuous at all $\mathbf{x} \in E$.

Exercise 9.8. Suppose that f is a differentiable real function in an open set $E \subseteq \mathbb{R}^n$, and that f has a local maximum at a point $\mathbf{x} \in E$. Prove that $f'(\mathbf{x}) = 0$.

Proof (Theorem 5.8).

(1) Apply Theorem 5.8 to each $D_j f$ for $1 \leq j \leq n$. Since f has a local maximum at a point $\mathbf{x} \in E$, there is an open neighborhood $B(\mathbf{x}; r)$ of \mathbf{x} in E such that

$$f(\mathbf{y}) \le f(\mathbf{x})$$

for all $\mathbf{y} \in B(\mathbf{x}; r)$. Therefore,

$$f(\mathbf{x} + t\mathbf{e}_j) \le f(\mathbf{x})$$

for all |t| < r and $1 \le j \le n$, or $t \mapsto f(\mathbf{x} + t\mathbf{e}_j)$ has a local maximum at a point $t = 0 \in (-r, r)$.

(2) Since f is a differentiable in E, each partial derivatives $D_j f$ exist (Theorem 9.21). Hence Theorem 5.8 implies that $(D_j f)(\mathbf{x}) = 0$ for all $1 \le j \le n$. So

$$f'(\mathbf{x}) = [(D_1 f)(\mathbf{x}) \cdots (D_k f)(\mathbf{x})] = [0 \cdots 0] = 0$$

(as the zero matrix).

Exercise 9.9. If **f** is a differentiable mapping of a connected open set $E \subseteq \mathbb{R}^n$, and if $\mathbf{f}'(\mathbf{x}) = 0$ for every $\mathbf{x} \in E$, prove that **f** is a constant in E.

Proof.

- (1) Show that \mathbf{f} is locally constant. Given any $\mathbf{x} \in E$. Since E is open, there exists an open neighborhood $B(\mathbf{x};r)$ of \mathbf{x} such that $B(\mathbf{x};r) \subseteq E$ and r > 0. Corollary to Theorem 9.19 implies that \mathbf{f} is a constant on $B(\mathbf{x};r)$, that is, \mathbf{f} is locally constant.
- (2) Show that **f** is constant if **f** is locally constant in a connected set $E \subseteq \mathbb{R}^n$. Might assume that $E \neq \emptyset$. (Otherwise there is nothing to do.) Take some $\mathbf{x}_0 \in E$.
 - (a) Let

$$U = \{ \mathbf{y} \in E : \mathbf{f}(\mathbf{y}) = \mathbf{f}(\mathbf{x}_0) \}.$$

- (b) U is open since \mathbf{f} is locally constant (by (1)). (Take any $\mathbf{y} \in U$. Since \mathbf{f} is locally constant, there is an open neighborhood $B(\mathbf{y}) \subseteq E$ of \mathbf{y} such that $f(\mathbf{z}) = f(\mathbf{y}) = f(\mathbf{x}_0)$ whenever $\mathbf{z} \in B(\mathbf{y})$. So that $B(\mathbf{y}) \subseteq U$, or U is open.)
- (c) Besides, since \mathbf{f} is continuous (Remarks 9.13(c)), the set U is closed. (The proof is the same as Proof (Definition 2.18(d)) in Exercise 4.3.)
- (d) So U is open and closed. Write $E = U \cup (E U)$. Here U and E U are both open and closed. Hence $U \cap \overline{E U} = U \cap (E U) = \emptyset$ and $\overline{U} \cap (E U) = U \cap (E U) = \emptyset$. Note that $\mathbf{x}_0 \in U \neq \emptyset$. By the connectedness of E, $E U = \emptyset$, or E = U, or \mathbf{f} is constant on E.

Note. The only subsets of a connected set E which are both open and closed are E and \varnothing .

Exercise 9.10. If f is a real function defined in a convex open set $E \subseteq \mathbb{R}^n$, such that $(D_1 f)(\mathbf{x}) = 0$ for every $\mathbf{x} \in E$, prove that $f(\mathbf{x})$ depends only on x_2, \ldots, x_n . Show that the convexity of E can be replaced by a weaker condition, but that some condition is required. For example, if n = 2 and E is shaped like

a horseshoe, the statement may be false.

Proof.

(1) It suffices to show that

$$f(a, x_2, \dots, x_n) = f(b, x_2, \dots, x_n)$$

whenever $\mathbf{x} = (a, x_2, \dots, x_n) \in E$ and $\mathbf{y} = (b, x_2, \dots, x_n) \in E$ if $(D_1 f)(\mathbf{x}) = 0$ in the convex open set E.

(2) Might assume that a < b. Since $g: t \mapsto f(t, x_2, \dots, x_n)$ is a real continuous function on [a, b] (by the openness of E) and differentiable in (a, b) (by the existence of $D_1 f$),

$$g(b) - g(a) = (b - a)g'(\xi)$$

for some $\xi \in (a, b)$. Note that

$$g'(\xi) = (D_1 f)(\xi, x_2, \dots, x_n) = 0$$

by assumption. g(b) = g(a) or $f(a, x_2, \dots, x_n) = f(b, x_2, \dots, x_n)$.

(3) (2) shows that the convexity of E can be replaced by a weaker condition that $E \subseteq \mathbb{R}^n$ is convex in the first coordinate, say E is open and

$$\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} = (\lambda a + (1 - \lambda)b, x_2, \dots, x_n) \in E$$

whenever $\mathbf{x} = (a, x_2, ..., x_n) \in E$, $\mathbf{y} = (b, x_2, ..., x_n) \in E$, and $0 < \lambda < 1$.

(4) Show that the convexity of E or some weaker condition is required. Define $f(x,y) = \operatorname{sgn}(x)$ on $E = \{(x,y) \in \mathbb{R}^2 : x \neq 0\}$. E is open and $(D_1f)(x,y) = 0$ in E. Note that f(1989,0) = 1 and f(-64,0) = -1, and thus f(x,y) does not depend only on y = 0.

Exercise 9.11. If f and g are differentiable real functions in \mathbb{R}^n , prove that

$$\nabla(fg) = f\nabla g + g\nabla f$$

and that

$$\nabla\left(\frac{1}{f}\right) = -\frac{1}{f^2}\nabla f$$

whenever $f \neq 0$.

Proof. Recall Example 9.18:

$$(\nabla(f))(\mathbf{x}) = \sum_{i=1}^{n} (D_i f)(\mathbf{x}) \mathbf{e}_i.$$

(1) Show that $\nabla(fg) = f\nabla g + g\nabla f$. For any $\mathbf{x} \in \mathbb{R}^n$,

$$(\nabla(fg))(\mathbf{x}) = \sum_{i=1}^{n} (D_i(fg))(\mathbf{x})\mathbf{e}_i$$

$$= \sum_{i=1}^{n} (g(D_if) + f(D_ig))(\mathbf{x})\mathbf{e}_i \qquad (\text{Theorem 5.3(b)})$$

$$= \sum_{i=1}^{n} [g(\mathbf{x})(D_if)(\mathbf{x}) + f(\mathbf{x})(D_ig)(\mathbf{x})] \mathbf{e}_i$$

$$= g(\mathbf{x}) \sum_{i=1}^{n} (D_if)(\mathbf{x})\mathbf{e}_i + f(\mathbf{x}) \sum_{i=1}^{n} (D_ig)(\mathbf{x})\mathbf{e}_i$$

$$= g(\mathbf{x})(\nabla f)(\mathbf{x}) + f(\mathbf{x})(\nabla g)(\mathbf{x})$$

$$= (f\nabla g + g\nabla f)(\mathbf{x}).$$

(2) Show that

$$\nabla\left(\frac{1}{f}\right) = -\frac{1}{f^2}\nabla f$$

whenever $f \neq 0$. Note that $\nabla(1) = 0$ since

$$\nabla(1)(\mathbf{x}) = \sum (D_i 1)(\mathbf{x})\mathbf{e}_i = \sum (0)(\mathbf{x})\mathbf{e}_i = \sum 0\mathbf{e}_i = 0.$$

Hence as $f \neq 0$, we have

$$0 = \nabla(1)$$

$$= \nabla \left(f \frac{1}{f} \right) \qquad (f \neq 0)$$

$$= f \nabla \left(\frac{1}{f} \right) + \frac{1}{f} \nabla f \qquad ((1)),$$

or
$$\nabla \left(\frac{1}{f}\right) = -\frac{1}{f^2} \nabla f$$
.

Exercise 9.12. ...

Proof.

- (1)
- (2)

Exercise 9.13. Suppose \mathbf{f} is a differentiable mapping of \mathbb{R}^1 into \mathbb{R}^3 such that $|\mathbf{f}(t)| = 1$ for every t. Prove that $\mathbf{f}'(t) \cdot \mathbf{f}(t) = 0$. Interpret this result geometrically.

Proof.

(1) Write $\mathbf{f} = (f_1, f_2, f_3)$ as a vector-valued function. By Remarks 5.16, \mathbf{f} is differentiable if and only if each f_1, f_2, f_3 is differentiable. So $\mathbf{f}' = (f'_1, f'_2, f_3)'$. Hence

$$|\mathbf{f}(t)| = 1 \text{ for every } t$$

$$\iff \mathbf{f}(t) \cdot \mathbf{f}(t) = 1$$

$$\iff f_1(t)^2 + f_2(t)^2 + f_3(t)^2 = 1$$

$$\iff 2f_1(t)f_1'(t) + 2f_2(t)f_2'(t) + 2f_3(t)f_3'(t) = 0$$

$$\iff f_1(t)f_1'(t) + f_2(t)f_2'(t) + f_3(t)f_3'(t) = 0$$

$$\iff (f_1(t), f_2(t), f_3(t)) \cdot (f_1'(t), f_2'(t), f_3'(t)) = 0$$

$$\iff \mathbf{f}(t) \cdot \mathbf{f}'(t) = \mathbf{f}'(t) \cdot \mathbf{f}(t) = 0.$$

(2) The vector $\mathbf{f}'(t)$ is called the **tangent vector** (or **velocity vector**) of \mathbf{f} at t. Geometrically, given any mapping \mathbf{f} lying on the sphere S^2 , its tangent vector at t is lying on the tangent plane of S^2 at t.

Exercise 9.14. Define f(0,0) = 0 and

$$f(x,y) = \frac{x^3}{x^2 + y^2}$$
 if $(x,y) \neq (0,0)$.

- (a) Prove that D_1f and D_2f are bounded functions in \mathbb{R}^2 . (Hence f is continuous.)
- (b) Let **u** be any unit vector in \mathbb{R}^2 . Show that the directional derivative $(D_{\mathbf{u}}f)(0,0)$ exists, and that its absolute value is at most 1.
- (c) Let γ be a differentiable mapping of \mathbb{R}^1 into \mathbb{R}^2 (in other words, γ is a differentiable curve in \mathbb{R}^2), with $\gamma(t) = (0,0)$ and $\gamma'(t) \neq (0,0)$ for any $t \in \mathbb{R}^1$. Put $g(t) = f(\gamma(t))$ and prove that g is differentiable for every $t \in \mathbb{R}^1$. If $\gamma \in \mathscr{C}'$, prove that $g \in \mathscr{C}'$.
- (d) In spite of this, prove that f is not differentiable at (0,0).

Proof of (a).

(1) Show that

$$(D_1 f)(x,y) = \begin{cases} 1 & \text{if } (x,y) = (0,0), \\ \frac{x^2 (x^2 + 3y^2)}{(x^2 + y^2)^2} & \text{if } (x,y) \neq (0,0). \end{cases}$$

If (x, y) = (0, 0),

$$(D_1 f)(0,0) = \lim_{t \to 0} \frac{f(t,0) - f(0,0)}{t} = \lim_{t \to 0} \frac{t - 0}{t} = 1.$$

If $(x, y) \neq (0, 0)$,

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

$$= \lim_{t \to 0} \frac{\frac{(x+t)^3}{(x+t)^2 + y^2} - \frac{x^3}{x^2 + y^2}}{t}$$

$$= \lim_{t \to 0} \frac{x^2 (x^2 + 3y^2) + tx(2x^2 + 3y^2) + t^2(x^2 + y^2)}{((x+t)^2 + y^2)(x^2 + y^2)}$$

$$= \frac{x^2 (x^2 + 3y^2)}{(x^2 + y^2)^2}.$$

(Or differentiate directly.)

(2) Show that $(D_1 f)(x, y)$ is bounded. It suffices to show that $(D_1 f)(x, y)$ is bounded if $(x, y) \neq (0, 0)$. Write $x = r \cos \theta$ and $y = r \sin \theta$ in the polar coordinates. (Here r > 0.) Hence

$$(D_1 f)(x, y) = \frac{x^2(x^2 + 3y^2)}{(x^2 + y^2)^2} = \cos^2 \theta(\cos^2 \theta + 3\sin^2 \theta)$$

is bounded by $1 \cdot (1+3) = 4$.

(3) Show that

$$(D_2 f)(x,y) = \begin{cases} 0 & \text{if } (x,y) = (0,0), \\ \frac{-2x^3y}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0). \end{cases}$$

If (x, y) = (0, 0),

$$(D_2 f)(0,0) = \lim_{t \to 0} \frac{f(0,t) - f(0,0)}{t} = \lim_{t \to 0} \frac{0-0}{t} = 0.$$

If $(x, y) \neq (0, 0)$,

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

$$= \lim_{t \to 0} \frac{\frac{x^3}{x^2 + (y+t)^2} - \frac{x^3}{x^2 + y^2}}{t}$$

$$= \lim_{t \to 0} \frac{-2x^3y - tx^3}{(x^2 + (y+t)^2)(x^2 + y^2)}$$

$$= \frac{-2x^3y}{(x^2 + y^2)^2}.$$

(Or differentiate directly.)

- (4) Show that $(D_2f)(x,y)$ is bounded. Similar to (2).
- (5) Show that f is continuous. Apply Exercise 9.7 to (2)(4).

Proof of (b).

(1) Write $\mathbf{u} = (u_1, u_2)$. The formula

$$(D_{\mathbf{u}}f)(0,0) = (D_1f)(0,0)u_1 + (D_2f)(0,0)u_2 = u_1$$

might be false since we don't know if f is differentiable or not. Actually, we will show that $(D_{\mathbf{u}}f)(0,0) = u_1^3 \neq u_1$.

(2)

$$(D_{\mathbf{u}}f)(0,0) = \lim_{t \to 0} \frac{f(tu_1, tu_2) - f(0,0)}{t}$$

$$= \lim_{t \to 0} \frac{\frac{t^3 u_1^3}{t^2 u_1^2 + t^2 u_2^2} - 0}{t}$$

$$= \lim_{t \to 0} u_1^3 \qquad (|\mathbf{u}| = 1)$$

$$= u_1^3.$$

Also $|(D_{\mathbf{u}}f)(0,0)| = |u_1|^3 \le 1$ since $|\mathbf{u}| = 1$.

Proof of (c).

(1) Given any $t \in \mathbb{R}^1$.

$$g'(t) = \lim_{x \to t} \frac{g(x) - g(t)}{x - t} = \lim_{x \to t} \frac{f(\gamma(x)) - f(\gamma(t))}{x - t}.$$

Write $\gamma(t) = (\gamma_1(t), \gamma_2(t)).$

(2) Suppose that $\gamma(t) \neq (0,0)$. Since γ is differentiable, γ is continuous. So there exists an open neighborhood $B(t) \subseteq \mathbb{R}^1$ of t such that $\gamma(x) \neq (0,0)$ whenever $x \in B(t)$. Hence

$$g'(t) = \lim_{x \to t} \frac{\frac{\gamma_1(x)^3}{\gamma_1(x)^2 + \gamma_2(x)^2} - \frac{\gamma_1(t)^3}{\gamma_1(t)^2 + \gamma_2(t)^2}}{x - t}$$

$$= \frac{d}{dt} \left(\frac{\gamma_1(t)^3}{\gamma_1(t)^2 + \gamma_2(t)^2} \right)$$

$$= \frac{3\gamma_1(t)^2 \gamma_1'(t)}{\gamma_1(t)^2 + \gamma_2(t)^2} - \frac{\gamma_1(t)^3 (2\gamma_1(t)\gamma_1'(t) + 2\gamma_2(t)\gamma_2'(t))}{(\gamma_1(t)^2 + \gamma_2(t)^2)^2}.$$

exists since γ_1 and γ_2 are differentiable

(3) Suppose that $\gamma(t) = (0,0)$ and thus $\gamma'(t) \neq (0,0)$. So

$$g'(t) = \lim_{x \to t} \frac{f(\gamma(x))}{x - t}$$

Note that $\gamma(x) \neq (0,0)$ in some open neighborhood of t since

$$\lim_{\substack{x \to t \\ \gamma(x) = (0,0)}} \frac{\gamma(x) - \gamma(t)}{x - t} = (0,0),$$

contrary to the assumption that $\gamma'(t) \neq (0,0)$. Note that $\gamma_1(t) = \gamma_2(t) = 0$. So

$$g'(t) = \lim_{x \to t} \frac{f(\gamma(x))}{x - t}$$

$$= \lim_{x \to t} \frac{\gamma_1(x)^3}{\gamma_1(x)^2 + \gamma_2(x)^2} \cdot \frac{1}{x - t}$$

$$= \lim_{x \to t} \frac{(\gamma_1(x) - \gamma_1(t))^3}{(\gamma_1(x) - \gamma_1(t))^2 + (\gamma_2(x) - \gamma_2(t))^2} \cdot \frac{1}{x - t}$$

$$= \lim_{x \to t} \frac{\left(\frac{\gamma_1(x) - \gamma_1(t)}{x - t}\right)^3}{\left(\frac{\gamma_1(x) - \gamma_1(t)}{x - t}\right)^2 + \left(\frac{\gamma_2(x) - \gamma_2(t)}{x - t}\right)^2}$$

$$= \frac{\gamma'_1(t)^3}{\gamma'_1(t)^2 + \gamma'_2(t)^2}$$

since $\gamma'(t) \neq (0,0)$.

(4) By (2)(3), g'(t) exists and

$$g'(t) = \begin{cases} \frac{3\gamma_1(t)^2\gamma_1'(t)}{\gamma_1(t)^2 + \gamma_2(t)^2} - \frac{\gamma_1(t)^3(2\gamma_1(t)\gamma_1'(t) + 2\gamma_2(t)\gamma_2'(t))}{(\gamma_1(t)^2 + \gamma_2(t)^2)^2} & \text{if } \gamma(t) \neq (0,0), \\ \frac{\gamma_1'(t)^3}{\gamma_1'(t)^2 + \gamma_2'(t)^2} & \text{if } \gamma(t) = (0,0). \end{cases}$$

(5) Now suppose $\gamma \in \mathscr{C}'$. To show $g' \in \mathscr{C}'$, it suffices to show that

$$\lim_{x \to t} g'(x) = g'(t)$$

if $\gamma(t)=(0,0)$ since g'(t) is always continuous if $\gamma(t)\neq(0,0)$. Here all $\gamma_1,\gamma_2,\gamma_1',\gamma_2'$ are continuous and $\gamma_1(t)^2+\gamma_2(t)^2\neq0$ by assumption. So

$$\lim_{x \to t} \frac{3\gamma_1(x)^2 \gamma_1'(x)}{\gamma_1(x)^2 + \gamma_2(x)^2}$$

$$= \lim_{x \to t} \frac{3\left(\frac{\gamma_1(x) - \gamma_1(t)}{x - t}\right)^2 \gamma_1'(x)}{\left(\frac{\gamma_1(x) - \gamma_1(t)}{x - t}\right)^2 + \left(\frac{\gamma_2(x) - \gamma_2(t)}{x - t}\right)^2}$$

$$= \frac{3\gamma_1'(t)^2 \cdot \gamma_1'(t)}{\gamma_1'(t)^2 + \gamma_2'(t)^2}$$

$$= \frac{3\gamma_1'(t)^3}{\gamma_1'(t)^2 + \gamma_2'(t)^2}$$

and similarly

$$\lim_{x \to t} \frac{\gamma_1(t)^3 (2\gamma_1(t)\gamma_1'(t) + 2\gamma_2(t)\gamma_2'(t))}{(\gamma_1(t)^2 + \gamma_2(t)^2)^2}$$

$$= \lim_{x \to t} \frac{\left(\frac{\gamma_1(x) - \gamma_1(t)}{x - t}\right)^3 \left(2\frac{\gamma_1(x) - \gamma_1(t)}{x - t}\gamma_1'(t) + 2\frac{\gamma_2(x) - \gamma_2(t)}{x - t}\gamma_2'(t)\right)}{\left(\left(\frac{\gamma_1(x) - \gamma_1(t)}{x - t}\right)^2 + \left(\frac{\gamma_2(x) - \gamma_2(t)}{x - t}\right)^2\right)^2}$$

$$= \frac{\gamma_1'(t)^3 \cdot (2\gamma_1'(t)\gamma_1'(t) + 2\gamma_2'(t)\gamma_2'(t))}{(\gamma_1'(t)^2 + \gamma_2'(t)^2)^2}$$

$$= \frac{2\gamma_1'(t)^3}{\gamma_1'(t)^2 + \gamma_2'(t)^2}.$$

Hence

$$\lim_{x \to t} g'(x) = \frac{3\gamma_1'(t)^3}{\gamma_1'(t)^2 + \gamma_2'(t)^2} - \frac{2\gamma_1'(t)^3}{\gamma_1'(t)^2 + \gamma_2'(t)^2} = g'(t).$$

Proof of (d). (Reductio ad absurdum) If f were differentiable, then

$$(D_{\mathbf{u}}f)(0,0) = (D_1f)(0,0)u_1 + (D_2f)(0,0)u_2 = u_1$$

(Formula (40) in Chapter 9), contrary to (b) if we take $\mathbf{u} = \left(\frac{1}{64}, \frac{\sqrt{4095}}{64}\right)$.

Exercise 9.15. Define f(0,0) = 0, and put

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

if $(x, y) \neq (0, 0)$.

(a) Prove, for all $(x, y) \in \mathbb{R}^2$, that

$$4x^4y^2 < (x^4 + y^2)^2$$
.

Conclude that f is continuous.

(b) For $0 \le \theta \le 2\pi$, $-\infty < t < \infty$, define

$$g_{\theta}(t) = f(t\cos\theta, t\sin\theta).$$

Show that $g_{\theta}(0) = 0$, $g'_{\theta}(0) = 0$, $g''_{\theta}(0) = 2$. Each g_{θ} has therefore a strict local minimum at t = 0. In other words, the restriction of f to each line through (0,0) has a strict local minimum at (0,0).

(c) Show that (0,0) is nevertheless not a local minimum for f, since $f(x,x^2) = -x^4$.

Proof.

- (1)
- (2)

Exercise 9.16. ...

Proof.

- (1)
- (2)

Exercise 9.17. Let $\mathbf{f} = (f_1, f_2)$ be the mapping of \mathbb{R}^2 into \mathbb{R}^2 given by

$$f_1(x,y) = e^x \cos y,$$
 $f_2(x,y) = e^x \sin y.$

(a) What is the range of \mathbf{f} ?

- (b) Show that the Jacobian of \mathbf{f} is not zero at any point of \mathbb{R}^2 . Thus every point of \mathbb{R}^2 has a neighborhood in which \mathbf{f} is one-to-one. Nevertheless, \mathbf{f} is not one-to-one on \mathbb{R}^2 .
- (c) Put $\mathbf{a} = (0, \frac{\pi}{3})$, $\mathbf{b} = \mathbf{f}(\mathbf{a})$, let \mathbf{g} be the continuous inverse of \mathbf{f} , defined in a neighborhood of \mathbf{b} such that $\mathbf{g}(\mathbf{b}) = \mathbf{a}$. Find an explicit formula for \mathbf{g} , compute $\mathbf{f}'(\mathbf{a})$ and $\mathbf{g}'(\mathbf{b})$, and verify the formula

$$\mathbf{g}'(\mathbf{y}) = {\{\mathbf{f}'(\mathbf{g}(\mathbf{y}))\}}^{-1} \qquad (\mathbf{y} \in \mathbb{R}^2 - {\{(0,0)\}}).$$

(d) What are the images under **f** of lines parallel to the coordinate axes?

Proof of (a).

- (1) The range of **f** is $\mathbb{R}^2 \{(0,0)\}$.
- (2) If $(a, b) \neq (0, 0)$, then **f**: $(\log(\sqrt{a^2 + b^2}), \operatorname{atan}(a, a)) \mapsto (a, b)$ where

$$\operatorname{atan2}(b,a) = \begin{cases} \arctan\left(\frac{b}{a}\right) & \text{if } a > 0, \\ \arctan\left(\frac{b}{a}\right) + \pi & \text{if } a < 0 \text{ and } b \geq 0, \\ \arctan\left(\frac{b}{a}\right) - \pi & \text{if } a < 0 \text{ and } b < 0, \\ \frac{\pi}{2} & \text{if } a = 0 \text{ and } b > 0, \\ -\frac{\pi}{2} & \text{if } a = 0 \text{ and } b < 0, \\ \operatorname{undefined} & \text{if } a = 0 \text{ and } b = 0. \end{cases}$$

(Or apply Theorem 8.7(d).)

(3) If (a,b) = (0,0), then for any $(x,y) \in \mathbb{R}^2$ we have $f_1(x,y)^2 + f_2(x,y)^2 = e^{2x} \neq 0$. So that there is no (x,y) such that $\mathbf{f}: (x,y) \mapsto (0,0)$.

Exercise 9.18. ...

Proof.

- (1)
- (2)

Exercise 9.19. ...

Proof.

(1)	
(2)	
Exercise 9.20	
Proof.	
(1)	
(2)	
Exercise 9.21	
Proof.	
(1)	
(2)	
Exercise 9.22	
Proof.	
(1)	
(2)	
Exercise 9.23	
Proof.	
(1)	
(2)	

Exercise 9.24
Proof.
(1)
(2)
Exercise 9.25
Proof.
(1)
(2)
Exercise 9.26
Proof.
(1)
(2)
Exercise 9.27
Proof.
(1)
(1)
(1) (2)

Proof.

(1)	
(2)	
Exercise 9.29.	
Proof.	
(1)	
(2)	
Exercise 9.30.	
Proof.	
(1)	
(2)	
Exercise 9.31.	
Proof.	
(1)	
(2)	