MA 410 | Multivariable Calculus

Complete Homework Solutions

Sumanta Das
Teaching Assistant

February 10, 2026

Contents

Homework 1

Theorem 1. *If $T\colon \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation, there exists a number $M \geqslant 0$ such that*

$$\|T(v)\| \leqslant M\|v\| \quad \text{for all } v \in \mathbb{R}^m.$$

Problem 1

Proof. Let $\{e_1, e_2, \ldots, e_m\}$ be the standard basis for $\mathbb{R}^m$. Any vector $v \in \mathbb{R}^m$ can be expressed as a linear combination of these basis vectors:

$$v = \sum_{i=1}^{m} v_i e_i,$$

where v_i are the scalar components of v. Since T is linear, we can distribute it across the sum:

$$T(v) = T\left(\sum_{i=1}^{m} v_i e_i\right) = \sum_{i=1}^{m} v_i T(e_i).$$

Using the triangle inequality and the homogeneity property of the norm ($\|cv\| = |c|\|v\|$), we obtain:

$$\|T(v)\| = \left\|\sum_{i=1}^{m} v_i T(e_i)\right\| \leqslant \sum_{i=1}^{m} \|v_i T(e_i)\| = \sum_{i=1}^{m} |v_i|\|T(e_i)\|.$$

Recall that for the standard Euclidean norm, $|v_i| \leqslant \|v\|$ for each i. Substituting this into the inequality:

$$\|T(v)\| \leqslant \sum_{i=1}^{m} \|v\|\|T(e_i)\| = \|v\| \left(\sum_{i=1}^{m} \|T(e_i)\|\right).$$

Let $M = \sum_{i=1}^{m} \|T(e_i)\|$. Since the basis is finite, M is a finite constant independent of v. Thus, we conclude:

$$\|T(v)\| \leqslant M\|v\|. \qquad \square$$

Theorem 2 (Jordan–von Neumann, real case). *Let V be a real vector space with a norm $\|\cdot\|$. The norm is induced by an inner product if and only if it satisfies the parallelogram law*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \textit{for all } x, y \in V.$$

Proof. (⇒) Suppose the norm is induced by an inner product $\langle\cdot,\cdot\rangle$, so $\|x\|^2 = \langle x, x\rangle$. Then

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + 2\langle x,y\rangle + \langle y,y\rangle,\\ \|x-y\|^2 &= \langle x-y, x-y\rangle = \langle x,x\rangle - 2\langle x,y\rangle + \langle y,y\rangle.\end{aligned}$$

Adding these identities gives

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2,$$

which is the parallelogram law.

(⇐) Now suppose the norm satisfies the parallelogram law. Define a function $\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{R}$ by

$$\langle x, y\rangle := \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2\big).$$

We verify that this is an inner product.

Symmetry. Clearly,

$$\langle y, x\rangle = \frac{1}{4}\big(\|y+x\|^2 - \|y-x\|^2\big) = \langle x, y\rangle.$$

Positive definiteness. Taking $x = y$, we obtain

$$\langle x, x\rangle = \frac{1}{4}\|2x\|^2 = \|x\|^2 \geqslant 0,$$

with equality if and only if $x = 0$.

Additivity in the first variable. Using the parallelogram law, one checks that for all $x, y, z \in V$,

$$\|x+y+z\|^2 - \|x+y-z\|^2 = \big(\|x+z\|^2 - \|x-z\|^2\big) + \big(\|y+z\|^2 - \|y-z\|^2\big).$$

Dividing by 4 yields

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle.$$

Homogeneity. For integers n, additivity gives $\langle nx, y \rangle = n\langle x, y \rangle$. By symmetry, this holds in the second variable as well. Extending by continuity (using the continuity of the norm) yields

$$\langle \lambda x, y \rangle = \lambda \langle x, y \rangle \quad \text{for all } \lambda \in \mathbb{R}.$$

Thus $\langle \cdot, \cdot \rangle$ is a real inner product on V.

Finally, $\|x\|^2 = \langle x, x \rangle$, so the given norm is induced by this inner product. □

Theorem 3. *The ∞-norm on $\mathbb{R}^2$ is not induced by an inner product.*

Problem 2

Proof. By Theorem 2, a norm is induced by an inner product, then it must satisfy the parallelogram law:

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2 \quad \text{for all } u, v \in \mathbb{R}^2.$$

Consider the vectors $u = (1, 0)$ and $v = (0, 1)$ in $\mathbb{R}^2$. For the infinity norm defined by $\|(x, y)\|_\infty = \max(|x|, |y|)$, we have:

$$\|u\|_\infty = 1 \quad \text{and} \quad \|v\|_\infty = 1.$$

Thus, the left-hand side of the Parallelogram Law is:

$$2\|u\|_\infty^2 + 2\|v\|_\infty^2 = 2(1)^2 + 2(1)^2 = 4.$$

Now consider the sum and difference:

$$u + v = (1, 1) \implies \|u + v\|_\infty = 1,$$

$$u - v = (1, -1) \implies \|u - v\|_\infty = 1.$$

The right-hand side is:

$$\|u + v\|_\infty^2 + \|u - v\|_\infty^2 = 1^2 + 1^2 = 2.$$

Since $4 \neq 2$, the Parallelogram Law does not hold, and therefore $\|\cdot\|_\infty$ is not induced by an inner product. □

Theorem 4. *If A is a closed subset of $\mathbb{R}$ containing every rational number $r \in [0,1]$, then $[0,1] \subset A$.*

Problem 3

Proof. Let x be any real number such that $x \in [0,1]$. Since the set of rational numbers $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists a sequence of rational numbers $(q_n)_{n=1}^{\infty}$ such that $q_n \in [0,1]$ for all n, and

$$\lim_{n\to\infty} q_n = x.$$

By the hypothesis, A contains every rational number in $[0,1]$. Therefore, the sequence (q_n) is contained entirely in A (i.e., $q_n \in A$ for all n).

Since A is a closed set, it is sequentially closed. This means that if a sequence in A converges, its limit must also belong to A. Consequently,

$$x = \lim_{n\to\infty} q_n \in A.$$

Since x was an arbitrary element of $[0,1]$, we conclude that $[0,1] \subseteq A$. $\square$

Theorem 5. *If $Q = [a,b] \times [c,d] \subset \mathbb{R}^2$, then Q is the closure of its interior (i.e., $Q = \overline{\operatorname{int}(Q)}$).*

Problem 4

Proof. We first determine the interior of Q. Recall that a point $(x,y) \in Q$ is an interior point if there exists $\varepsilon > 0$ such that the open ball

$$B_\varepsilon(x,y) = \{(u,v) \in \mathbb{R}^2 : \|(u,v) - (x,y)\| < \varepsilon\}$$

is contained in Q.

Suppose $a < x < b$ and $c < y < d$. Define

$$\varepsilon = \min\{x-a,\ b-x,\ y-c,\ d-y\} > 0.$$

Then $B_\varepsilon(x,y) \subset (a,b) \times (c,d) \subset Q$, so (x,y) is an interior point of Q.

If $x = a$ or $x = b$, or $y = c$ or $y = d$, then every open ball centered at (x,y) contains points not belonging to Q. Hence such points are not interior points.

Therefore,

$$\text{int}(Q) = (a, b) \times (c, d).$$

Next, since $\text{int}(Q) \subset Q$ and Q is closed in $\mathbb{R}^2$ (being a product of closed intervals), we have

$$\overline{\text{int}(Q)} \subset Q.$$

To prove the reverse inclusion, let $(x, y) \in Q$ be arbitrary. Fix $\varepsilon > 0$. Because $a \leqslant x \leqslant b$ and $c \leqslant y \leqslant d$, we may choose

$$x' \in (a, b) \cap (x - \varepsilon/2,\ x + \varepsilon/2), \qquad y' \in (c, d) \cap (y - \varepsilon/2,\ y + \varepsilon/2).$$

Then $(x', y') \in (a, b) \times (c, d) = \text{int}(Q)$ and

$$\|(x', y') - (x, y)\| < \varepsilon.$$

Hence every open ball centered at (x, y) intersects $\text{int}(Q)$, which implies

$$(x, y) \in \overline{\text{int}(Q)}.$$

Thus $Q \subset \overline{\text{int}(Q)}$. Combining both inclusions, we conclude that

$$Q = \overline{\text{int}(Q)}. \qquad \square$$

Theorem 6. *Let A be a closed subset of $\mathbb{R}^n$ and $x \in \mathbb{R}^n$ such that $x \notin A$. Then there exists $d > 0$ such that $\|x - y\| > d$ for every $y \in A$.*

Problem 5

Proof. Since A is a closed set, its complement $A^c = \mathbb{R}^n \backslash A$ is an open set. Given that $x \notin A$, we have $x \in A^c$.

By the definition of an open set, there exists an $\epsilon > 0$ such that the open ball centered at x with radius ϵ is contained entirely in A^c:

$$B(x, \epsilon) \subset A^c.$$

This implies that $B(x, \epsilon) \cap A = \varnothing$. Consequently, for any point $y \in A$, y is not in the ball $B(x, \epsilon)$, which means:

$$\|x - y\| \geqslant \epsilon.$$

Let us choose $d = \frac{\epsilon}{2}$. Clearly $d > 0$, and for all $y \in A$:

$$\|x - y\| \geqslant \epsilon > \frac{\epsilon}{2} = d.$$

Thus, the condition is satisfied. $\square$

Theorem 7. *Let A be a closed subset of $\mathbb{R}^n$ and B be a compact subset of $\mathbb{R}^n \backslash A$. Then there exists $d > 0$ such that $\|x - y\| > d$ for every $x \in B$ and $y \in A$.*

Problem 6

Proof. Suppose for the sake of contradiction that such a $d > 0$ does not exist. This implies that

$$\inf\{\|x - y\| : x \in B, y \in A\} = 0.$$

Therefore, for every $n \in \mathbb{N}$, there exist points $x_n \in B$ and $y_n \in A$ such that:

$$\|x_n - y_n\| < \frac{1}{n}.$$

Since (x_n) is a sequence in the compact set B, it contains a convergent subsequence (x_{n_k}) that converges to a point $x^* \in B$.

Now consider the corresponding subsequence (y_{n_k}) in A. By the triangle inequality:

$$\|y_{n_k} - x^*\| \leqslant \|y_{n_k} - x_{n_k}\| + \|x_{n_k} - x^*\|.$$

As $k \to \infty$, the first term $\|y_{n_k} - x_{n_k}\| < \frac{1}{n_k} \to 0$, and the second term $\|x_{n_k} - x^*\| \to 0$ by the definition of convergence. Consequently, $y_{n_k} \to x^*$.

Since A is closed and (y_{n_k}) is a sequence in A converging to x^*, we must have $x^* \in A$. Thus, $x^* \in A \cap B$. However, this contradicts the hypothesis that $B \subset \mathbb{R}^n \backslash A$ (i.e., $A \cap B = \varnothing$).

Therefore, the assumption was false, and there exists $d > 0$ such that $\|x - y\| > d$ for all $x \in B, y \in A$. $\square$

Problem 8. *Show that the property of A being closed is necessary in Theorem 6 and that the property of B being compact is necessary in Theorem 7.*

Problem 7

Solution. If A is not closed, we can choose x to be a limit point of A that is not contained in A. In this case, there are points in A arbitrarily close to x, making the infimum distance zero.

Counterexample: Consider $A = (0, 1) \subset \mathbb{R}$ and $x = 0$.

- A is not closed.
- $x \notin A$.

For any proposed $d > 0$, by the Archimedean property, there exists a point $y \in A$ (specifically $y = \min(1/2, d/2)$) such that:

$$|x - y| = |0 - y| = y < d.$$

Thus, no such $d > 0$ exists satisfying $|x - y| > d$ for all $y \in A$.

If B is closed but not compact (specifically, if B is unbounded), the distance between the disjoint closed sets A and B can be zero.

Counterexample: Consider subsets of $\mathbb{R}^2$. Let A be the x-axis and B be the graph of the hyperbola $y = 1/x$ for $x > 0$.

$$A = \{(x, y) \in \mathbb{R}^2 : y = 0\}$$

$$B = \{(x, y) \in \mathbb{R}^2 : xy = 1, x > 0\}$$

Both sets are closed in $\mathbb{R}^2$, and they are clearly disjoint ($xy = 1$ implies $y \neq 0$). However, B is not compact because it is unbounded.

Consider the sequence of points $x_n = (n, 1/n) \in B$ and $y_n = (n, 0) \in A$ for $n \in \mathbb{N}$. The distance is:

$$\|x_n - y_n\| = \frac{1}{n}.$$

As $n \to \infty$, the distance approaches 0. Thus, there is no strictly positive d such that $\|x - y\| > d$ for all $x \in B, y \in A$. $\square$

Theorem 9. *Let U be an open subset of $\mathbb{R}^n$ and $C \subset U$ be a compact set. Then there exists a compact set D such that $C \subset \operatorname{int}(D)$ and $D \subset U$.*

Problem 8

Proof. Since U is open and $C \subset U$, for each $x \in C$, there exists $\epsilon_x > 0$ such that the open ball $B(x, \epsilon_x) \subseteq U$. For each x, let $r_x = \epsilon_x/2$. The closure of this smaller ball satisfies:

$$\overline{B}(x, r_x) \subset B(x, \epsilon_x) \subseteq U.$$

The collection $\{B(x, r_x)\}_{x \in C}$ is an open cover of C. Since C is compact, there exists a finite subcover corresponding to points $x_1, \ldots, x_k \in C$. That is:

$$C \subset \bigcup_{i=1}^{k} B(x_i, r_i).$$

Let us define the set D as the union of the closures of these balls:

$$D = \bigcup_{i=1}^{k} \overline{B}(x_i, r_i).$$

We now verify the required properties for D:

1. D is compact: Each closed ball $\overline{B}(x_i, r_i)$ is a compact set in $\mathbb{R}^n$. Since a finite union of compact sets is compact, D is compact.

2. $D \subset U$: By construction, each $\overline{B}(x_i, r_i) \subset U$. Therefore, their union $D \subset U$.

3. $C \subset \operatorname{int}(D)$: Let $V = \bigcup_{i=1}^{k} B(x_i, r_i)$. Since V is a union of open sets, V is open. We have $C \subset V$ and clearly $V \subset D$. Since V is an open subset of D, it must be contained in the interior of D. Thus, $C \subset \operatorname{int}(D)$.

This completes the proof. $\square$

Theorem 10. *Every linear transformation $T \colon \mathbb{R}^n \to \mathbb{R}^m$ is continuous.*

Problem 9

Proof. Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. By Theorem 1, there exists a constant $M > 0$ such that for all $v \in \mathbb{R}^n$:

$$\|T(v)\| \leqslant M\|v\|.$$

To prove continuity, fix any point $x_0 \in \mathbb{R}^n$. Let $\epsilon > 0$ be given. We choose $\delta = \frac{\epsilon}{M}$.

For any $x \in \mathbb{R}^n$ satisfying $\|x - x_0\| < \delta$, we examine the distance between their images:

$$\|T(x) - T(x_0)\|.$$

Using the linearity of T, we have $T(x) - T(x_0) = T(x - x_0)$. Applying the boundedness property:

$$\|T(x) - T(x_0)\| = \|T(x - x_0)\| \leqslant M\|x - x_0\|.$$

Substituting the bound on $\|x - x_0\|$:

$$\|T(x) - T(x_0)\| < M\delta = M\left(\frac{\epsilon}{M}\right) = \epsilon.$$

Since the choice of x_0 was arbitrary, T is continuous on all of $\mathbb{R}^n$. □

Theorem 11. *If $A \subset \mathbb{R}^n$ is not closed, there exists a continuous function $f\colon A \to \mathbb{R}$ that is unbounded.*

Problem 10

Proof. Since A is not closed, it does not contain all of its limit points. Therefore, there exists a point $x_0 \in \mathbb{R}^n$ such that x_0 is a limit point of A (i.e., $x_0 \in \overline{A}$), but $x_0 \notin A$.

Define the function $f\colon A \to \mathbb{R}$ by:

$$f(x) = \frac{1}{\|x - x_0\|}.$$

First, we observe that f is continuous on A. Since $x_0 \notin A$, the denominator $\|x - x_0\|$ is non-zero for all $x \in A$. The function $x \mapsto \|x - x_0\|$ is continuous, and the reciprocal of a non-zero continuous function is continuous.

Next, we show that f is unbounded. Since x_0 is a limit point of A, there exists a sequence (a_k) in A such that $a_k \to x_0$ as $k \to \infty$. Consequently, $\|a_k - x_0\| \to 0$. It follows that:

$$\lim_{k\to\infty} f(a_k) = \lim_{k\to\infty} \frac{1}{\|a_k - x_0\|} = \infty.$$

Thus, f is unbounded on A, as required. $\square$

Homework 2

Theorem 12. *Let $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at a point $a \in \mathbb{R}^n$. Then f is continuous at a.*

Problem 1

Proof. Since f is differentiable at a, by definition there exists a linear map $Df(a)\colon \mathbb{R}^n \to \mathbb{R}^m$ such that

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

This means that there exists a function $r\colon \mathbb{R}^n \to \mathbb{R}^m$ satisfying

$$f(a+h) - f(a) = Df(a)h + r(h),$$

where

$$\lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = 0.$$

By Theorem 1, there exists a constant $C > 0$ such that

$$\|Df(a)h\| \leqslant C\|h\| \quad \text{for all } h \in \mathbb{R}^n.$$

Since $\lim_{h\to 0} \frac{\|r(h)\|}{\|h\|} = 0$, there exists $\delta_1 > 0$ such that

$$\|h\| < \delta_1 \implies \|r(h)\| \leqslant \|h\|.$$

Let $\varepsilon > 0$ be given. Now choose

$$\delta = \min\left\{\delta_1,\ \frac{\varepsilon}{C+1}\right\}.$$

Then for $\|h\| < \delta$, we have

$$\begin{aligned}
\|f(a+h) - f(a)\| &\leqslant \|Df(a)h\| + \|r(h)\| \\
&\leqslant C\|h\| + \|h\| \\
&= (C+1)\|h\| \\
&< (C+1)\delta \\
&\leqslant \varepsilon.
\end{aligned}$$

Therefore,

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

which shows that f is continuous at a. $\square$

Theorem 13. *Let $f, g \colon \mathbb{R}^n \to \mathbb{R}$ be functions differentiable at a point $a \in \mathbb{R}^n$. Then the functions $f \pm g$ are also differentiable at a. Moreover,*

$$D(f \pm g)(a) = Df(a) \pm Dg(a).$$

Problem 2

Proof. Since f and g are differentiable at a, by definition there exist linear maps $Df(a), Dg(a) \colon \mathbb{R}^n \to \mathbb{R}$ and functions $r_f, r_g \colon \mathbb{R}^n \to \mathbb{R}$ such that

$$f(a+h) = f(a) + Df(a)(h) + r_f(h),$$
$$g(a+h) = g(a) + Dg(a)(h) + r_g(h),$$

where

$$\lim_{h \to 0} \frac{r_f(h)}{\|h\|} = 0, \quad \lim_{h \to 0} \frac{r_g(h)}{\|h\|} = 0.$$

We compute

$$\begin{aligned} (f \pm g)(a+h) &= f(a+h) \pm g(a+h) \\ &= \big(f(a) + Df(a)(h) + r_f(h)\big) \\ &\quad \pm \big(g(a) + Dg(a)(h) + r_g(h)\big) \\ &= (f \pm g)(a) + \big(Df(a) \pm Dg(a)\big)(h) + \big(r_f(h) \pm r_g(h)\big). \end{aligned}$$

Since $Df(a) \pm Dg(a)$ is linear and

$$\frac{|r_f(h) \pm r_g(h)|}{\|h\|} \leqslant \frac{|r_f(h)|}{\|h\|} + \frac{|r_g(h)|}{\|h\|} \to 0 \quad \text{as } h \to 0,$$

the function $f \pm g$ is differentiable at a with

$$D(f \pm g)(a) = Df(a) \pm Dg(a).$$

$\square$

Lemma 14. Lemma Let $f \colon \mathbb{R}^n \to \mathbb{R}$ be differentiable at a. Define the remainder function r by $r(h) = f(a+h) - f(a) - Df(a)(h)$. Then r is differentiable at 0 with $Dr(0) = 0$, and consequently $\lim_{h \to 0} r(h) = 0$.

Proof. First, observe that $r(0) = f(a) - f(a) - Df(a)(0) = 0$. To show differentiability at 0 with derivative $L = 0$, we check the limit:

$$\lim_{h \to 0} \frac{|r(0+h) - r(0) - 0(h)|}{\|h\|} = \lim_{h \to 0} \frac{|r(h)|}{\|h\|}.$$

By the definition of the differentiability of f at a, this limit is 0. Thus, r is differentiable at 0. Since differentiability implies continuity (see Theorem 12), we have $\lim_{h \to 0} r(h) = r(0) = 0$. □

Theorem 15. *Let $f, g \colon \mathbb{R}^n \to \mathbb{R}$ be differentiable at a point $a \in \mathbb{R}^n$. Then the function $f \cdot g$ is differentiable at a. Moreover,*

$$D(f \cdot g)(a) = g(a)Df(a) + f(a)Dg(a).$$

Problem 2

Proof. Since f and g are differentiable at a, there exist linear maps $Df(a), Dg(a) \colon \mathbb{R}^n \to \mathbb{R}$ and remainder functions $r_f, r_g \colon \mathbb{R}^n \to \mathbb{R}$ such that for h in a neighborhood of 0:

$$f(a+h) = f(a) + Df(a)(h) + r_f(h),$$

$$g(a+h) = g(a) + Dg(a)(h) + r_g(h),$$

where the remainder terms satisfy

$$\lim_{h \to 0} \frac{r_f(h)}{\|h\|} = 0 \quad \text{and} \quad \lim_{h \to 0} \frac{r_g(h)}{\|h\|} = 0.$$

We propose that the derivative of $f \cdot g$ at a is the linear map given by

$$L(h) = g(a)Df(a)(h) + f(a)Dg(a)(h).$$

To prove this, we examine the difference quotient. We expand the product $(f \cdot g)(a+h)$:

$$\begin{aligned}(f \cdot g)(a+h) &= \big(f(a) + Df(a)(h) + r_f(h)\big)\big(g(a) + Dg(a)(h) + r_g(h)\big) \\ &= f(a)g(a) + g(a)Df(a)(h) + f(a)Dg(a)(h) \\ &\quad + Df(a)(h)Dg(a)(h) + g(a)r_f(h) + f(a)r_g(h) \\ &\quad + Df(a)(h)r_g(h) + Dg(a)(h)r_f(h) + r_f(h)r_g(h).\end{aligned}$$

We can rewrite this as:

$$(f \cdot g)(a+h) = (f \cdot g)(a) + L(h) + R(h),$$

where $R(h)$ collects all higher-order terms:

$$\begin{aligned} R(h) = {} & Df(a)(h)Dg(a)(h) + g(a)r_f(h) + f(a)r_g(h) \\ & + Df(a)(h)r_g(h) + Dg(a)(h)r_f(h) + r_f(h)r_g(h). \end{aligned}$$

To verify differentiability, we must show that $\lim\limits_{h \to 0} \dfrac{|R(h)|}{\|h\|} = 0$.
By Theorem 1, there exists a constant $C > 0$ such that for all $h \in \mathbb{R}^n$:

$$|Df(a)(h)| \leqslant C\|h\| \quad \text{and} \quad |Dg(a)(h)| \leqslant C\|h\|.$$

Using the triangle inequality and the bounds above, for $h \neq 0$:

$$\begin{aligned} \frac{|R(h)|}{\|h\|} & \leqslant \frac{|Df(a)(h)||Dg(a)(h)|}{\|h\|} + |g(a)|\frac{|r_f(h)|}{\|h\|} + |f(a)|\frac{|r_g(h)|}{\|h\|} \\ & \quad + \frac{|Df(a)(h)||r_g(h)|}{\|h\|} + \frac{|Dg(a)(h)||r_f(h)|}{\|h\|} + \frac{|r_f(h)||r_g(h)|}{\|h\|} \\ & \leqslant \frac{C\|h\| \cdot C\|h\|}{\|h\|} + |g(a)|\frac{|r_f(h)|}{\|h\|} + |f(a)|\frac{|r_g(h)|}{\|h\|} \\ & \quad + \frac{C\|h\||r_g(h)|}{\|h\|} + \frac{C\|h\||r_f(h)|}{\|h\|} + \frac{|r_f(h)||r_g(h)|}{\|h\|} \\ & = C^2\|h\| + |g(a)|\frac{|r_f(h)|}{\|h\|} + |f(a)|\frac{|r_g(h)|}{\|h\|} \\ & \quad + C|r_g(h)| + C|r_f(h)| + \frac{|r_f(h)|}{\|h\|}|r_g(h)|. \end{aligned}$$

We now analyze the limit as $h \to 0$:

- The term $C^2\|h\| \to 0$.
- Since $\frac{|r_f(h)|}{\|h\|} \to 0$ and $\frac{|r_g(h)|}{\|h\|} \to 0$, the terms multiplied by constants $|g(a)|$ and $|f(a)|$ vanish.
- By Lemma 14, $\lim_{h \to 0} r_f(h) = 0$ and $\lim_{h \to 0} r_g(h) = 0$. Thus, $C|r_g(h)| \to 0$ and $C|r_f(h)| \to 0$.

- The final term approaches $0 \cdot 0 = 0$.

Since every term approaches 0, we conclude:

$$\lim_{h \to 0} \frac{|R(h)|}{\|h\|} = 0.$$

Thus, $f \cdot g$ is differentiable at a, and

$$D(f \cdot g)(a) = g(a)Df(a) + f(a)Dg(a). \qquad \square$$

Theorem 16. *Let $g_1, g_2 \colon \mathbb{R}^2 \to \mathbb{R}$ be continuous functions. Define*

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

Then:

(a) $\dfrac{\partial f}{\partial y}(x, y) = g_2(x, y).$

(b) A function f satisfying $\dfrac{\partial f}{\partial x}(x, y) = g_1(x, y)$ is given by

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

Problem 3

Proof. **(a)** Let

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

We compute the partial derivative with respect to y by treating x as a constant.

Observe that the term $\int_0^x g_1(t, 0)\, dt$ depends only on x. Therefore, its partial derivative with respect to y is 0.

For the term $\int_0^y g_2(x, t)\, dt$, since g_2 is continuous, the Fundamental Theorem of Calculus implies:

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

Combining these, we have:

$$\begin{aligned}\frac{\partial f}{\partial y}(x,y) &= \frac{\partial}{\partial y}\left(\int_0^x g_1(t,0)\,dt\right) + \frac{\partial}{\partial y}\left(\int_0^y g_2(x,t)\,dt\right)\\ &= 0 + g_2(x,y)\\ &= g_2(x,y).\end{aligned}$$

(b) Let

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

We compute the partial derivative with respect to x by treating y as a constant.

The term $\int_0^y g_2(0,t)\,dt$ depends only on y, so its partial derivative with respect to x is 0.

For the remaining term, since the integrand $g_1(t,y)$ is continuous in t, the Fundamental Theorem of Calculus gives:

$$\frac{\partial}{\partial x}\int_0^x g_1(t,y)\,dt = g_1(x,y).$$

Therefore:

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \frac{\partial}{\partial x}\left(\int_0^x g_1(t,y)\,dt\right) + \frac{\partial}{\partial x}\left(\int_0^y g_2(0,t)\,dt\right)\\ &= g_1(x,y) + 0\\ &= g_1(x,y).\end{aligned}$$

□

Theorem 17. *There exists a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ such that*

$$\frac{\partial f}{\partial x}(x,y) = x \quad \textit{and} \quad \frac{\partial f}{\partial y}(x,y) = y \quad \textit{for all } (x,y) \in \mathbb{R}^2.$$

Moreover, every such function has the form

$$f(x,y) = \frac{x^2+y^2}{2} + C,$$

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

Problem 4

Proof. Assume that $f\colon \mathbb{R}^2 \to \mathbb{R}$ satisfies

$$\frac{\partial f}{\partial x}(x,y) = x \quad \text{for all } (x,y) \in \mathbb{R}^2.$$

Fix $y \in \mathbb{R}$. Since $x \mapsto \frac{\partial f}{\partial x}(x,y)$ is continuous, the Fundamental Theorem of Calculus implies that

$$f(x,y) - f(0,y) = \int_0^x \frac{\partial f}{\partial x}(t,y)\,dt = \int_0^x t\,dt = \frac{x^2}{2}.$$

Thus, $f(x,y) = \frac{x^2}{2} + g(y)$, where $g\colon \mathbb{R} \to \mathbb{R}$ is a function depending only on y.

Now differentiate this expression with respect to y:

$$\frac{\partial f}{\partial y}(x,y) = g'(y).$$

By assumption,

$$g'(y) = y \quad \text{for all } y \in \mathbb{R}.$$

Integrating with respect to y yields

$$g(y) = \int_0^y t\,dt + C = \frac{y^2}{2} + C,$$

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

Substituting this expression for $g(y)$ back into the formula for f, we obtain

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

Finally, a direct computation shows that

$$\frac{\partial f}{\partial x}(x,y) = x \quad \text{and} \quad \frac{\partial f}{\partial y}(x,y) = y,$$

so the function satisfies the required conditions. $\square$

Theorem 18. *Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be a function such that*

$$\frac{\partial f}{\partial y}(x, y) = 0 \quad \textit{for all } (x, y) \in \mathbb{R}^2.$$

Then there exists a function $g\colon \mathbb{R} \to \mathbb{R}$ such that

$$f(x, y) = g(x) \quad \textit{for all } (x, y) \in \mathbb{R}^2.$$

In particular, f is independent of y.

Problem 5

Proof. Fix $x \in \mathbb{R}$ and consider the function

$$\varphi_x\colon \mathbb{R} \to \mathbb{R}, \qquad \varphi_x(y) = f(x, y).$$

Since $\frac{\partial f}{\partial y}(x, y)$ exists for all (x, y) and equals zero, the function φ_x is differentiable and satisfies

$$\varphi_x'(y) = 0 \quad \text{for all } y \in \mathbb{R}.$$

By the Mean Value Theorem, for any $y_1, y_2 \in \mathbb{R}$ there exists ξ between y_1 and y_2 such that

$$\varphi_x(y_2) - \varphi_x(y_1) = \varphi_x'(\xi)(y_2 - y_1).$$

Since $\varphi_x'(\xi) = 0$, it follows that

$$\varphi_x(y_2) = \varphi_x(y_1).$$

Thus φ_x is constant on $\mathbb{R}$.

Therefore, for each fixed x, there exists a real number $g(x)$ such that

$$f(x, y) = g(x) \quad \text{for all } y \in \mathbb{R}.$$

This defines a function $g\colon \mathbb{R} \to \mathbb{R}$ satisfying

$$f(x, y) = g(x) \quad \text{for all } (x, y) \in \mathbb{R}^2.$$

Hence, f is independent of y. □

Lemma 19. Let $U \subseteq \mathbb{R}^n$ be an open, convex set. If $f\colon U \to \mathbb{R}$ is a differentiable function such that $\nabla f(x) = 0$ for all $x \in U$, then f is constant on U.

Proof. Let $a, b \in U$ be arbitrary points. Since U is convex, the line segment connecting a and b lies entirely within U. Define the path $\gamma\colon [0,1] \to U$ by

$$\gamma(t) = (1-t)a + tb.$$

Consider the function $g\colon [0,1] \to \mathbb{R}$ defined by $g(t) = f(\gamma(t))$. By the Chain Rule, g is differentiable on $(0,1)$ and its derivative is

$$g'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t) = \nabla f(\gamma(t)) \cdot (b-a).$$

By hypothesis, $\nabla f(x) = 0$ for all $x \in U$. Thus,

$$g'(t) = 0 \cdot (b-a) = 0$$

for all $t \in (0,1)$.

By the Mean Value Theorem applied to g, we have

$$g(1) - g(0) = g'(c)(1-0) = 0$$

for some $c \in (0,1)$. Since $g(0) = f(a)$ and $g(1) = f(b)$, it follows that $f(a) = f(b)$. Since a and b were arbitrary, f is constant on U. □

Theorem 20. *Let*

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

(a) There exists a function $f\colon A \to \mathbb{R}$ such that

$$\frac{\partial f}{\partial y} = 0 \quad \text{on } A,$$

but f is not independent of y.

(b) If $f\colon A \to \mathbb{R}$ satisfies

$$\frac{\partial f}{\partial x} = 0 \quad \text{and} \quad \frac{\partial f}{\partial y} = 0 \quad \text{on } A,$$

then f is constant on A.

Problem 6

Proof. **(a)** Consider the function defined by:

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

For any fixed x, the function $f(x,y)$ is locally constant with respect to y: If $x > 0$, $f(x,y)$ is the constant x for all $y > 0$, and the constant 0 for all $y < 0$. In both open intervals, the derivative is 0. If $x \leqslant 0$, $f(x,y)$ is identically 0. Thus, $\frac{\partial f}{\partial y} = 0$ everywhere on A.

It is enough to check that f is continuous on the y-axis ($x = 0$), where the two definitions meet. Approaching any point $(0, y_0)$ from the right with $y > 0$: $\lim_{x \to 0^+} x = 0$. Approaching from any other direction: The value is already 0. Since the limits match, f is continuous on A.

Choose any $x > 0$, say $x = 1$.

$$f(1,1) = 1 \quad \neq \quad f(1,-1) = 0.$$

Therefore, f is not independent of y.

(b) Assume that

$$\frac{\partial f}{\partial x} = 0 \quad \text{and} \quad \frac{\partial f}{\partial y} = 0 \quad \text{on } A.$$

We will show that $f(x,y)$ is constant for all $(x,y) \in A$.

Let $A_1 := \{(x,y) \in \mathbb{R}^2 : y > 0\}$, $A_2 := \{(x,y) \in \mathbb{R}^2 : x < 0\}$, and $A_3 := \{(x,y) \in \mathbb{R}^2 : y < 0\}$. Then each A_i is an open, convex subset of $\mathbb{R}^2$. By Lemma 19, $f|A_i$ is constant for each i. Since $A_1 \cap A_2 \neq \varnothing$, $A_2 \cap A_3 \neq \varnothing$, and $A = A_1 \cup A_2 \cup A_3$, the map f is constant on A. □

Theorem 21. *Consider the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined by*

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

Then

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1, \qquad \frac{\partial^2 f}{\partial y \partial x}(0,0) = -1.$$

Problem 7

Proof. We first compute the first-order partial derivatives at $(0,0)$:

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

$$\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0}\frac{0-0}{k} = 0.$$

Now, we compute $\dfrac{\partial^2 f}{\partial x \partial y}(0,0)$

For fixed $x \neq 0$,

$$\begin{aligned}\frac{\partial f}{\partial y}(x,0) &= \lim_{k\to 0}\frac{f(x,k)-f(x,0)}{k}\\ &= \lim_{k\to 0}\frac{xk(x^2-k^2)/(x^2+k^2)-0}{k}\\ &= \lim_{k\to 0}\frac{x(x^2-k^2)}{x^2+k^2}\\ &= x.\end{aligned}$$

Then

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{h\to 0}\frac{\frac{\partial f}{\partial y}(h,0)-\frac{\partial f}{\partial y}(0,0)}{h} = \lim_{h\to 0}\frac{h-0}{h} = 1.$$

Finally, we compute $\dfrac{\partial^2 f}{\partial y \partial x}(0,0)$

For fixed $y \neq 0$,

$$\begin{aligned}\frac{\partial f}{\partial x}(0,y) &= \lim_{h\to 0}\frac{f(h,y)-f(0,y)}{h}\\ &= \lim_{h\to 0}\frac{hy(h^2-y^2)/(h^2+y^2)-0}{h}\\ &= \lim_{h\to 0}\frac{y(h^2-y^2)}{h^2+y^2}\\ &= -y.\end{aligned}$$

Then

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{k\to 0}\frac{\frac{\partial f}{\partial x}(0,k)-\frac{\partial f}{\partial x}(0,0)}{k} = \lim_{k\to 0}\frac{-k-0}{k} = -1.$$

Therefore,

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1, \qquad \frac{\partial^2 f}{\partial y \partial x}(0,0) = -1. \qquad \square$$

Remark 22. Schwarz's Theorem states that if mixed partial derivatives are continuous, they are equal regardless of the order. However, if mixed partial derivatives are not continuous, the order of partial differentiation matters. For example, Theorem 21 presents a function whose first partial derivatives f_x and f_y exist and are continuous everywhere, yet the mixed second partial derivatives at the origin are unequal ($f_{xy} \neq f_{yx}$).

Lemma 23. Let $m \in \mathbb{N} \cup \{0\}$. Then $\lim_{x\to 0} |x|^{-m} e^{-1/x^2} = 0$.

Proof. For $t \geqslant 0$ the exponential series gives

$$e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!} \geqslant \frac{t^{k+1}}{(k+1)!} \qquad (k \in \mathbb{N} \cup \{0\}).$$

Hence for $t > 0$

$$\frac{t^k}{e^t} \leqslant \frac{(k+1)!}{t} \xrightarrow[t\to\infty]{} 0,$$

so $\lim\limits_{t\to\infty} \dfrac{t^k}{e^t} = 0$.

Now let $m \geqslant 0$ be an integer and put $t = 1/x^2$ for $x \neq 0$. Then for $t \geqslant 1$, we have

$$\frac{e^{-1/x^2}}{|x|^m} = t^{m/2} e^{-t} \leqslant t^{\lceil m/2 \rceil} e^{-t} \xrightarrow[t\to\infty]{} 0,$$

which shows e^{-1/x^2} tends to 0 faster than any power of $|x|$ as $x \to 0$. $\square$

Theorem 24. *Define* $f : \mathbb{R} \to \mathbb{R}$ *by*

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then

Problem 8

(a) *f is continuous for all $x \in \mathbb{R}$.*

(b) *For every $n \geqslant 1$, the derivative $f^{(n)}$ exists and is continuous on $\mathbb{R}$, and $f^{(n)}(0) = 0$.*

Problem 8

Proof of (a). If $x \neq 0$, then f is the composition of the continuous functions $\mathbb{R}\backslash\{0\} \ni x \mapsto -1/x^2 \in \mathbb{R}\backslash\{0\}$ and $\mathbb{R}\backslash\{0\} \ni t \mapsto e^t \in \mathbb{R}\backslash\{0\}$, so f is continuous at every nonzero point. It remains to check continuity at 0. By Lemma 23, $\lim_{x\to 0} e^{-1/x^2} = 0$. Hence f is continuous at 0. Combining this with continuity away from 0 gives continuity on $\mathbb{R}$. □

Lemma 25. Let $f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ be a polynomial of degree m. Then

$$|f(x)| \leqslant |x|^m \left(|a_m| + |a_{m-1}| + \cdots + |a_0|\right)$$

for $|x| \geqslant 1$.

Proof. Let x be a real number such that $|x| \geqslant 1$. Then

$$\begin{aligned} |f(x)| &= \left|a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0\right| \\ &= |x^m| \left|a_m + a_{m-1}\frac{1}{x} + \cdots + a_1\frac{1}{x^{m-1}} + a_0\frac{1}{x^m}\right| \\ &\leqslant |x|^m \left(|a_m| + |a_{m-1}|\frac{1}{|x|} + \cdots + |a_1|\frac{1}{|x|^{m-1}} + |a_0|\frac{1}{|x|^m}\right) \\ &\leqslant |x|^m \left(|a_m| + |a_{m-1}| + \cdots + |a_0|\right). \end{aligned}$$

□

Proof of (b). We first prove by induction that for each $n \geqslant 1$ there exists a polynomial P_n (with real coefficients) such that for every $x \neq 0$

$$f^{(n)}(x) = P_n\big(1/x\big)\, e^{-1/x^2}. \tag{1}$$

For $n = 0$ take $P_0 \equiv 1$. Suppose (1) holds for some n. Differentiate (for $x \neq 0$):

$$f^{(n+1)}(x) = \big(P_n(1/x)\big)'\, e^{-1/x^2} + P_n(1/x)\,\big(e^{-1/x^2}\big)'.$$

Since $\left(e^{-1/x^2}\right)' = \frac{2}{x^3}e^{-1/x^2}$ and $\left(P_n(1/x)\right)'$ is again a rational function which can be written as a polynomial in $1/x$ (times a power of x^{-1}) , we see that $f^{(n+1)}(x)$ can be written in the form

$$f^{(n+1)}(x) = P_{n+1}(1/x)\, e^{-1/x^2}$$

for some polynomial P_{n+1}. This completes the induction.

Now fix $n \geqslant 0$. From (1) we have for $x \neq 0$

$$\left|f^{(n)}(x)\right| = \left|P_n(1/x)\right| e^{-1/x^2}.$$

The polynomial $\left|P_n(1/x)\right|$ grows at most like a fixed power of $|x|^{-1}$; hence, by Lemma 25, there exist constants $C > 0$ and $m \geqslant 0$ such that

$$\left|f^{(n)}(x)\right| \leqslant C\,|x|^{-m}e^{-1/x^2} \quad \text{for } |x| \leqslant 1.$$

As in part (a), with $t = 1/x^2$ we get

$$|x|^{-m}e^{-1/x^2} = t^{m/2}e^{-t} \to 0 \quad \text{as } x \to 0.$$

Thus $\lim_{x\to 0} f^{(n)}(x) = 0$. Define $f^{(n)}(0) := 0$. The preceding limit shows that this value agrees with the limit of $f^{(n)}(x)$ as $x \to 0$, so $f^{(n)}$ is continuous at 0. Together with smoothness on $\mathbb{R}\backslash\{0\}$, this proves $f^{(n)}$ exists and is continuous on all of $\mathbb{R}$, and $f^{(n)}(0) = 0$.

Finally, to see explicitly that the derivatives at 0 computed via the difference quotient agree with 0, one can check by induction that

$$\frac{d^n f}{dx^n}(0) = \lim_{x\to 0} \frac{f^{(n-1)}(x) - f^{(n-1)}(0)}{x - 0} = \lim_{x\to 0} \frac{f^{(n-1)}(x)}{x} = 0,$$

using the fact already established that $f^{(n-1)}(x)$ tends to 0 faster than any power of x. This gives another direct verification that all derivatives at 0 are 0. $\square$

Remark 26. In complex analysis, holomorphicity implies analyticity; if a function is differentiable once, it is infinitely differentiable and equal to its Taylor series. Theorem 24 shows that real calculus is different. It provides the standard example of a function that is smooth (C^∞) but not analytic (C^ω). Although f is infinitely differentiable, it is flat at the origin: all its derivatives are zero. Consequently, its Taylor series vanishes and does not represent the function. This property allows for the construction of smooth bump functions with compact support. These are essential in differential topology for building partitions of unity, which serve as the standard tool to glue local constructions into global ones on a manifold.

Remark 27. The function $g\colon \mathbb{R}^2\backslash\big([0,\infty)\times\{0\}\big)\to\mathbb{R}$ defined by

$$g(x,y)=\begin{cases} e^{-1/x^2} & \text{if } x>0 \text{ and } y>0,\\ 0 & \text{otherwise.}\end{cases}$$

gives a C^∞ example of the case (a) of Theorem 20.

Homework 3

Theorem 28. *Let $A \subseteq \mathbb{R}^n$ be an open set, and let $f\colon A \to \mathbb{R}^n$ be a continuously differentiable, one–one function. Assume that* $\det Df(x) \neq 0$ *for all $x \in A$. Then $f(A)$ is an open subset of $\mathbb{R}^n$, and the inverse function $f^{-1}\colon f(A) \to A$ is differentiable. Moreover, for every $y \in f(A)$,*

$$D(f^{-1})(y) = \left(Df(f^{-1}(y))\right)^{-1}.$$

Problem 1

Proof. Let $a \in A$ be arbitrary. Since A is open and f is continuously differentiable on A with $\det Df(a) \neq 0$, the Inverse Function Theorem applies at the point a. Hence, there exist an open set $V_a \subseteq A$ containing a and an open set $W_a \subseteq \mathbb{R}^n$ containing $f(a)$ such that the restriction $f|V_a \to W_a$ has a continuous inverse $(f|V_a)^{-1}\colon W_a \to V_a$, which is differentiable and satisfies

$$D\left((f|V_a)^{-1}\right)(w) = \left(Df\left((f|V_a)^{-1}(w)\right)\right)^{-1} \quad \text{for all } w \in W_a.$$

We first show that $f(A)$ is open. Let $y \in f(A)$. Then $y = f(a)$ for some $a \in A$. By the above, there exists an open neighborhood W_a of y such that $W_a = f(V_a) \subseteq f(A)$. Therefore, every point of $f(A)$ has an open neighborhood contained in $f(A)$, and hence $f(A)$ is open.

Since f is one–one, the inverse map $f^{-1}\colon f(A) \to A$ is well-defined. Let $y \in f(A)$ and write $y = f(a)$ for some $a \in A$. On the open neighborhood W_a of y, the inverse of f coincides with the differentiable local inverse $(f|_{V_a})^{-1}$. Consequently, f^{-1} is differentiable at y.

Since $y \in f(A)$ was arbitrary, it follows that f^{-1} is differentiable on all of $f(A)$. Moreover, for every $y \in f(A)$,

$$D(f^{-1})(y) = \left(Df\left(f^{-1}(y)\right)\right)^{-1},$$

as asserted. $\square$

Theorem 29. *For every $n \geqslant 2$, there is no C^1 injective map from an open convex subset of $\mathbb{R}^n$ into $\mathbb{R}^{n-1}$.*

Problem 2

Proof. We argue by induction on n.

Base case ($n = 2$). We show that there is no continuously differentiable, injective map from an open convex subset of $\mathbb{R}^2$ into $\mathbb{R}$.

Let U be an open convex subset of $\mathbb{R}^2$, and let $f\colon U \to \mathbb{R}$ be a C^1 map. We show that f is not injective.

Case 1: $Df(x) = 0$ for all $x \in U$. Let a and b be any two points of U. Since U is convex, the segment connecting them lies in U. Define $\gamma\colon [0,1] \to \mathbb{R}$ by $\gamma(t) := f((1-t)a + tb)$. By the Mean Value Theorem, there exists $s \in (0,1)$ such that:

$$\gamma(1) - \gamma(0) = \gamma'(s)(1-0) = Df((1-s)a + sb) \cdot (b-a).$$

Since $Df(x) = 0$ for all x, this term is 0. Thus $\gamma(1) = \gamma(0)$, which implies $f(b) = f(a)$. Therefore, f is constant, hence not injective.

Case 2: $Df(p) \neq 0$ for some $p \in U$. Write $p = (p_1, p_2) \in U$. Without loss of generality, we may assume $\frac{\partial f}{\partial x_2}(p) \neq 0$. Define $G\colon U \to \mathbb{R}^2$ by

$$G(x_1, x_2) := (x_1, f(x_1, x_2)). \tag{2}$$

The Jacobian matrix of G at p is

$$DG(p) = \begin{pmatrix} 1 & 0 \\ \frac{\partial f}{\partial x_1}(p) & \frac{\partial f}{\partial x_2}(p) \end{pmatrix},$$

which has a nonzero determinant. By the Inverse Function Theorem, there exist open neighborhoods $V \subseteq U$ of p and W of $G(p)$ such that $G|_V\colon V \to W$ is a diffeomorphism. Since W is open, there exist open intervals I and J such that $I \times J \subseteq W$, with $p_1 \in I$ and $f(p) \in J$.

Consider any $p_1' \in I \setminus \{p_1\}$. Let $a' = (a_1', a_2')$ be the unique point in V such that $G(a') = (p_1', f(p))$. Then:

$$(p_1', f(p)) = G(a') \overset{(2)}{=} (a_1', f(a')).$$

This implies $a_1' = p_1'$ and $f(a') = f(p)$. Since $p_1' \neq p_1$, we have $a' \neq p$. Thus, f maps two distinct points (p and a') to the same value, so f is not injective.

Inductive hypothesis. Assume that for some $n \geqslant 2$, there is no C^1 injective map from an open convex subset of $\mathbb{R}^n$ into $\mathbb{R}^{n-1}$.

Inductive step. Let $U \subset \mathbb{R}^{n+1}$ be open and convex, and let

$$f = (f_1, \dots, f_n) \colon U \to \mathbb{R}^n$$

be a C^1 map. We show that f is not injective.

Case 1: $Df(x) = 0$ for all $x \in U$. Consider any two points a and b of U. Let $\pi_\ell \colon \mathbb{R}^{n-1} \to \mathbb{R}$ be the projection map onto the ℓ-th coordinate. Then

$$D(\pi_\ell \circ f)(x) = D\pi_\ell(f(x)) \circ Df(x) = \pi_\ell \circ Df(x) = 0.$$

Define $\gamma \colon [0,1] \ni t \longmapsto \pi_\ell \circ f((1-t)a + tb) \in \mathbb{R}$. Then there exists $s \in (0,1)$ such that

$$\gamma(0) - \gamma(1) = D(\pi_\ell \circ f)((1-s)a + sb) \cdot (b-a) = 0 \cdot (b-a) = 0.$$

Thus, $\pi_\ell \circ f(b) = \pi_\ell \circ f(a)$. Since this holds for all ℓ, $f(a) = f(b)$. Thus f is constant, hence not injective.

Case 2: $Df(p) \neq 0$ for some $p \in U$. Then there exist indices $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, n+1\}$ such that

$$\frac{\partial f_i}{\partial x_j}(p) \neq 0.$$

Define $G \colon U \to \mathbb{R}^{n+1}$ by replacing the j-th coordinate with f_i:

$$G(x_1, \dots, x_{n+1}) := (x_1, \dots, x_{j-1}, f_i(x), x_{j+1}, \dots, x_{n+1}).$$

The Jacobian matrix $DG(p)$ is given by

$$\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
\frac{\partial f_i}{\partial x_1}(p) & \frac{\partial f_i}{\partial x_2}(p) & \cdots & \frac{\partial f_i}{\partial x_{j-1}}(p) & \frac{\partial f_i}{\partial x_j}(p) & \frac{\partial f_i}{\partial x_{j+1}}(p) & \cdots & \frac{\partial f_i}{\partial x_{n+1}}(p) \\
0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1
\end{pmatrix}$$

Its determinant is:

$$\det DG(p) = \frac{\partial f_i}{\partial x_j}(p) \neq 0.$$

By the Inverse Function Theorem, there exist open neighborhoods U_0 of p and V_0 of $G(p)$ such that $G\colon U_0 \to V_0$ is a diffeomorphism. We may assume V_0 is an open ball, since open balls form a basis for the topology of $\mathbb{R}^{n+1}$. (This ensures V_0 is convex).

Define

$$\widetilde{f} := f \circ G^{-1}\colon V_0 \to \mathbb{R}^n.$$

<u>Claim.</u> For all $y \in V_0$, one has $\widetilde{f}_i(y) = y_j$.
Indeed, if $y \in V_0$, then $y = G(x)$ for a unique $x \in U_0$. By definition of G, the j-th coordinate of y is $y_j = f_i(x)$. Hence

$$\widetilde{f}_i(y) = f_i\left(G^{-1}(y)\right) = f_i(x) = y_j.$$

Set $c := f_i(p)$. Let $H_c := \{y \in \mathbb{R}^{n+1} : y_j = c\}$ be the affine hyperplane, and define the slice $S_c := V_0 \cap H_c$.

Define the projection $\varphi\colon H_c \to \mathbb{R}^n$ by dropping the j-th coordinate:

$$\varphi(x_1, \ldots, x_{n+1}) = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{n+1}).$$

Define $h\colon S_c \to \mathbb{R}^{n-1}$ by deleting the i-th component of $\widetilde{f}$:

$$h(y) = \left(\widetilde{f}_1(y), \ldots, \widetilde{f}_{i-1}(y), \widetilde{f}_{i+1}(y), \ldots, \widetilde{f}_n(y)\right).$$

The map $\psi := h \circ \varphi^{-1}$ maps the open convex set $\varphi(S_c) \subset \mathbb{R}^n$ into $\mathbb{R}^{n-1}$. Since $\widetilde{f}$ is C^1, ψ is C^1.

By the inductive hypothesis, ψ is not injective. Thus, h is not injective on S_c. There exist distinct $a, b \in S_c$ such that $h(a) = h(b)$. Because $a, b \in S_c$, we have $a_j = b_j = c$. From the Claim above,

$$\widetilde{f}_i(a) = a_j = c = b_j = \widetilde{f}_i(b),$$

Combined with $h(a) = h(b)$ (equality of all other components), this implies $\widetilde{f}(a) = \widetilde{f}(b)$.

Since $\widetilde{f} = f \circ G^{-1}$ and G^{-1} is injective, $f(G^{-1}(a)) = f(G^{-1}(b))$ with $G^{-1}(a) \neq G^{-1}(b)$, contradicting the injectivity of f. $\square$

Theorem 30. *Let $f\colon \mathbb{R} \to \mathbb{R}$ be a C^1 function such that $f'(a) \neq 0$ for every $a \in \mathbb{R}$. Then f is one–one.*

Problem 3 (a)

Proof. Since f' has the Darboux property, the condition $f'(a) \neq 0$ for all $a \in \mathbb{R}$ implies that f' does not change sign on $\mathbb{R}$. Thus either $f'(x) > 0$ for all x or $f'(x) < 0$ for all x.

Let $x_1 < x_2$. By the Mean Value Theorem, there exists $c \in (x_1, x_2)$ such that

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1).$$

Since $x_2 - x_1 > 0$ and $f'(c) \neq 0$, we have $f(x_2) \neq f(x_1)$. Therefore, f is injective. □

Example 31. The condition $f'(a) \neq 0$ for all $a \in \mathbb{R}$ is not necessary for injectivity.
Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = x^3.$$

Then $f'(x) = 3x^2$, and hence $f'(0) = 0$. Nevertheless, f is strictly increasing on $\mathbb{R}$, and therefore one–one.

Problem 3 (b)

Theorem 32. *There exists a C^1 map $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ such that $\det Df(x,y) \neq 0$ for all $(x,y) \in \mathbb{R}^2$, but f is not one–one.*

Problem 3 (c)

Proof. Define $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ by

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

The Jacobian matrix of f is

$$Df(x,y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}.$$

Its determinant is

$$\det Df(x,y) = e^{2x}(\cos^2 y + \sin^2 y) = e^{2x},$$

which is strictly positive for all $(x, y) \in \mathbb{R}^2$.

However,

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

for all $(x, y) \in \mathbb{R}^2$, showing that f is not injective. □

Theorem 33 (Implicit Function Theorem). *Let $F\colon \mathbb{R}^{n+m} \to \mathbb{R}^m$ be a C^1 map. Write points as $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. Suppose there exists (a, b) such that*

$$F(a,b) = 0 \quad \textit{and} \quad D_y F(a,b) \textit{ is invertible.}$$

Then there exist neighborhoods U of a and V of b, and a unique C^1 map $\varphi\colon U \to V$ such that

$$F(x, \varphi(x)) = 0 \quad \textit{for all } x \in U.$$

Moreover, the derivative of φ is given by

$$D\varphi(x) = -\big(D_y F(x, \varphi(x))\big)^{-1} D_x F(x, \varphi(x)) \quad \textit{for all } x \in U.$$

Theorem 34. *Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be define by $f(x,y) = x^2 - y^2 - 1$. The zero set $\mathcal{Z} = \{(x,y) \in \mathbb{R}^2 : f(x,y) = 0\}$ is a hyperbola. The Implicit Function Theorem applies at every point of $\mathcal{Z}$.*

Problem 4

Proof. The equation $f(x, y) = 0$ is equivalent to $x^2 - y^2 = 1$, which is the standard hyperbola.

We compute the partial derivatives:

$$\frac{\partial f}{\partial x}(x,y) = 2x, \qquad \frac{\partial f}{\partial y}(x,y) = -2y.$$

Let $(a, b) \in \mathcal{Z}$. Then $a^2 - b^2 = 1$.

Case 1: $b \neq 0$.

Then

$$\frac{\partial f}{\partial y}(a,b) = -2b \neq 0.$$

By the Implicit Function Theorem, there exists an open interval I containing a and a C^1 function $\varphi\colon I \to \mathbb{R}$ such that $f(x, \varphi(x)) = 0$ for all $x \in I$. Thus, near (a,b), the hyperbola is the graph of a function $y = \varphi(x)$.

Solving explicitly, $y^2 = x^2 - 1$, so locally

$$y = \pm\sqrt{x^2 - 1},$$

with the sign determined by whether $b > 0$ or $b < 0$.

Case 2: $b = 0$.

Since $(a,b) \in \mathcal{Z}$, we have $a^2 = 1$, hence $a = \pm 1$. Then

$$\frac{\partial f}{\partial y}(a,0) = 0, \quad \text{but} \quad \frac{\partial f}{\partial x}(a,0) = 2a \neq 0.$$

Therefore, the Implicit Function Theorem applies with respect to x.

There exist an open interval J containing 0 and a C^1 function $\psi\colon J \to \mathbb{R}$ such that $f(\psi(y), y) = 0$ for all $y \in J$.

Solving explicitly, $x^2 = 1 + y^2$, so locally

$$x = \pm\sqrt{1 + y^2},$$

with the sign determined by whether $a = 1$ or $a = -1$.

In all cases, at least one partial derivative of f is nonzero at each point of the hyperbola. Hence the Implicit Function Theorem applies at every point of $\mathcal{Z}$. □

Theorem 35. *Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = x^2 - y^2$. Then:*

1. *The zero set $\mathcal{Z} = \{(x,y) \in \mathbb{R}^2 : f(x,y) = 0\}$ is the union of two lines.*
2. *The Implicit Function Theorem applies at every point of $\mathcal{Z}$ except the origin.*

Problem 5

> 3. *Near each such point, $\mathcal{Z}$ can be written locally as the graph of a C^1 function either of y in terms of x or of x in terms of y.*

Problem 5

Proof. The zero set of f is $\mathcal{Z} = \{(x, y) \in \mathbb{R}^2 : x^2 - y^2 = 0\}$. Factoring, we obtain $x^2 - y^2 = (x - y)(x + y)$, and hence

$$\mathcal{Z} = \{(x, y) : y = x\} \ \cup \ \{(x, y) : y = -x\}.$$

Thus $\mathcal{Z}$ is the union of two straight lines through the origin.

We compute the partial derivatives of f:

$$\frac{\partial f}{\partial x}(x, y) = 2x, \qquad \frac{\partial f}{\partial y}(x, y) = -2y.$$

The Implicit Function Theorem applies at a point $(a, b) \in \mathcal{Z}$ provided

$$\left(\frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b)\right) \neq (0, 0).$$

This occurs if and only if $(a, b) \neq (0, 0)$. Therefore, the Implicit Function Theorem applies at every point of $\mathcal{Z}$ except the origin.

<u>Solving for y as a function of x.</u> Let $(a, b) \in \mathcal{Z}$ with $b \neq 0$. Then

$$\frac{\partial f}{\partial y}(a, b) = -2b \neq 0,$$

so by the Implicit Function Theorem there exists a neighborhood of a and a unique C^1 function $y = \varphi(x)$ such that

$$f(x, \varphi(x)) = 0 \quad \text{and} \quad \varphi(a) = b.$$

From the equation $x^2 - y^2 = 0$, we obtain explicitly

$$y = \pm x.$$

Near points on the line $y = x$ with $x \neq 0$, the local solution is $y = x$, and near points on the line $y = -x$ with $x \neq 0$, the local solution is $y = -x$.

Solving for x as a function of y. Let $(a, b) \in \mathcal{Z}$ with $a \neq 0$. Then

$$\frac{\partial f}{\partial x}(a, b) = 2a \neq 0,$$

so the Implicit Function Theorem guarantees the existence of a C^1 function $x = \psi(y)$ defined near b such that

$$f(\psi(y), y) = 0 \quad \text{and} \quad \psi(b) = a.$$

Solving $x^2 - y^2 = 0$ gives

$$x = \pm y,$$

which yields the two local branches of $\mathcal{Z}$.

Failure at the origin. At $(0, 0)$,

$$\frac{\partial f}{\partial x}(0, 0) = 0, \qquad \frac{\partial f}{\partial y}(0, 0) = 0,$$

so the Implicit Function Theorem does not apply. Geometrically, this reflects the fact that $\mathcal{Z}$ is not locally the graph of a single function near the origin, but rather the transverse intersection of two smooth curves. $\square$

Remark 36. In Theorem 34 and Theorem 35, the level sets are defined by equations of the form $F(x, y) = 0$, where $F \colon \mathbb{R}^2 \to \mathbb{R}$ is a C^1 function. At points where $\frac{\partial F}{\partial y} \neq 0$, the Implicit Function Theorem yields a function $y = \varphi(x)$ whose derivative is given by

$$F_x(x, \varphi(x)) \;+\; \varphi'(x) F_y(x, \varphi(x)) \;=\; 0.$$

The formula follows by differentiating the identity $F(x, \varphi(x)) \equiv 0$ with respect to x and applying the chain rule. Similarly, at points where $\frac{\partial F}{\partial x} \neq 0$, one may solve for $x = \psi(y)$, with

$$F_y(\psi(y), y)) \;+\; \psi'(y) F_x(\psi(y), y) \;=\; 0.$$

Example 37 (Theorem 34). Let $F(x,y) = x^2 - y^2 - 1$. Then

$$F_x(x,y) = 2x, \qquad F_y(x,y) = -2y.$$

At points where $y \neq 0$, the Implicit Function Theorem applies and

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = -\frac{2x}{-2y} = \frac{x}{y}.$$

At points where $x \neq 0$,

$$\frac{dx}{dy} = -\frac{F_y}{F_x} = -\frac{-2y}{2x} = \frac{y}{x}.$$

Problem 6

Example 38 (Theorem 35). Let $F(x,y) = x^2 - y^2$. Then

$$F_x(x,y) = 2x, \qquad F_y(x,y) = -2y.$$

At points where $y \neq 0$,

$$\frac{dy}{dx} = -\frac{F_x}{F_y} = \frac{x}{y}.$$

On the branch $y = x$ (with $x \neq 0$), this gives

$$\frac{dy}{dx} = 1.$$

On the branch $y = -x$ (with $x \neq 0$), this gives

$$\frac{dy}{dx} = -1.$$

At points where $x \neq 0$,

$$\frac{dx}{dy} = -\frac{F_y}{F_x} = \frac{y}{x}.$$

On the branch $x = y$ (with $y \neq 0$),

$$\frac{dx}{dy} = 1,$$

Problem 7

and on the branch $x = -y$,

$$\frac{dx}{dy} = -1.$$

At points where both partial derivatives vanish (e.g. $(0,0)$), the Implicit Function Theorem does not apply, and no derivative formula is guaranteed.

Problem 7

Theorem 39. *Let $f\colon \mathbb{R}^3 \to \mathbb{R}$ be defined by*

$$f(x,y,z) = x^2 + y^2 + z^2 - 2xz - 4.$$

Then the zero set

$$\mathcal{Z} = \{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) = 0\}$$

can be locally expressed as the graph of a C^1 function of any two variables, whenever the corresponding partial derivative of f does not vanish.

Problem 8

Proof. We begin by rewriting the defining equation of $\mathcal{Z}$. A direct computation gives

$$x^2 + y^2 + z^2 - 2xz - 4 = (x-z)^2 + y^2 - 4.$$

Hence the equation $f(x,y,z) = 0$ is equivalent to

$$(x-z)^2 + y^2 = 4. \tag{3}$$

Solving for x in terms of (y,z). Assume (y,z) satisfies $|y| < 2$. From (3) we obtain

$$(x-z)^2 = 4 - y^2.$$

Taking square roots yields

$$x - z = \pm\sqrt{4-y^2}.$$

Thus, locally,

$$x = z \pm \sqrt{4 - y^2}.$$

Since the partial derivative

$$\frac{\partial f}{\partial x}(x, y, z) = 2(x - z)$$

is nonzero whenever $x \neq z$, the Implicit Function Theorem guarantees that each choice of sign determines a C^1 function $x = \varphi(y, z)$ in a neighborhood of any point with $x \neq z$.

Step 2: Solving for z in terms of (x, y). Similarly, equation (3) implies

$$(z - x)^2 = 4 - y^2,$$

and hence

$$z = x \mp \sqrt{4 - y^2}.$$

Since

$$\frac{\partial f}{\partial z}(x, y, z) = 2(z - x),$$

the Implicit Function Theorem applies whenever $z \neq x$, producing a C^1 function $z = \psi(x, y)$ locally.

Solving for y in terms of (x, z). Rewriting (3) as

$$y^2 = 4 - (x - z)^2,$$

we obtain

$$y = \pm\sqrt{4 - (x - z)^2}.$$

The partial derivative

$$\frac{\partial f}{\partial y}(x, y, z) = 2y$$

is nonzero whenever $y \neq 0$. Hence, by the Implicit Function Theorem, y may be locally expressed as a C^1 function of (x, z) at all points with $y \neq 0$. $\square$