# MA662 course notes

Krit Boonsiriseth

Spring 2019
(Updated May 22, 2019)

**Abstract**

Course notes for Hotchkiss class MA662 (Multivariable Calculus). Proofs in here are not guaranteed to be rigorous. The sections are split by class tests. The textbook used is Hubbard and Hubbard's *Vector Calculus, Linear Algebra, and Differential Forms: A Unified Approach*, 5th ed.

*Multivariable calculus is the same as single variable calculus done several times.*
— Dr. Weiss, January 14, 2019

# Contents

**1 Topology** **3**
1.1 Topology of $\mathbb{R}$ . . . 3
1.2 Open and closed sets . . . 4
1.3 Sequences and limits . . . 6
1.4 Continuity . . . 8
1.5 Compact sets . . . 10

**2 Derivatives** **11**
2.1 Review of derivatives in $\mathbb{R}$ . . . 11
2.2 Derivatives: full, partial, directional . . . 12
2.3 Jacobian matrix . . . 13
2.4 Differentiability . . . 14
2.5 Computing derivatives . . . 16
2.6 Newton's method . . . 18
2.7 Inverse and implicit function theorems . . . 18

**3 Manifolds** **20**
3.1 Definition and parametrization . . . 20
3.2 Tangent spaces . . . 22

**4 Optimizations** **25**
4.1 Critical points and Hessian matrix . . . 25
4.2 Lagrange Multipliers . . . 26

**5 Riemann integrals** **28**
5.1 Introduction . . . . . 28
5.2 Riemann integrals . . . . . 30
5.3 Change of coordinates . . . . . 31

**6 Forms and Vector calculus** **33**
6.1 Forms . . . . . 33
6.2 Integrals over parametrized manifolds . . . . . 35
6.3 Orientation . . . . . 37
6.4 Boundaries on manifolds . . . . . 41
6.5 Exterior Derivatives . . . . . 44
6.6 Stokes' Theorem . . . . . 45
6.7 Vector calculus: work, flux, and mass forms . . . . . 46
6.8 Vector calculus: grad, curl, div . . . . . 48
6.9 Review: exercises . . . . . 49
6.10 Review: a Princeton final . . . . . 51

# 1 Topology

## 1.1 Topology of $\mathbb{R}$

**Definition 1.1** (Upper and lower bounds).
Let $X \subseteq \mathbb{R}$. Then $u \in \mathbb{R}$ is called an *upper bound* of $X$ if $x \leqslant u$ for all $x \in X$, and $\ell \in \mathbb{R}$ is called a *lower bound* if $x \geqslant \ell$ for all $x \in X$.

It is an axiomatic property of $\mathbb{R}$ that each subset of $\mathbb{R}$ has a least upper bound and, likewise, each subset that is bounded below has a greatest lower bound.

**Definition 1.2** (Supremum and infimum).
Let $X \subseteq \mathbb{R}$ be bounded. Then the *supremum* $y = \sup X$ is the least upper bound of $X$, that is, for any upper bound $y'$ of $X$, $y' \geqslant y$. Likewise, the *infimum* $z = \inf X$ is the greatest lower bound of $X$, that is, for any lower bound $z'$ of $X$, $z' \leqslant z$.

**Definition 1.3** (Maximum and minimum).
For an $X \subseteq \mathbb{R}$, if $\sup X \in X$, then we call it the *maximum* of $X$ (denoted $\max X$) and if $\inf X \in X$, we call it the *minimum* of $X$ (denoted $\min X$).

**Example 1.4.** For the open interval $A = (0, 1)$, $\sup A = 1, \inf A = 0$, and $\max A, \min A$ do not exist. For the closed interval $A = [0, 1]$, $\sup A = \max A = 1$ and $\inf A = \min A = 0$.

**Proposition 1.5.** If $X \subseteq \mathbb{R}$ is bounded above then $y = \sup X$ iff

(i) $y$ is an upper bound of $X$, and

(ii) for all $\epsilon > 0$, there exists $x \in X$ such that $x > y - \epsilon$.

*Proof.* ($\Rightarrow$) Suppose $y = \sup X$. (i) follows by definition. If (ii) is false then there is an $\epsilon > 0$ such that for all $x \in X$, $x \leqslant y - \epsilon$, so $y - \epsilon < y$ is also an upper bound of $X$, which contradicts $y = \sup X$.

($\Leftarrow$) By (i) $y$ is an upper bound of $X$. Suppose there is a smaller upper bound $y' < y$ of $X$. Consider $\epsilon = y - y'$. (ii) implies the existence of $x \in X$ such that $x > y - \epsilon = y'$, so $y'$ is not an upper bound of $X$, a contradiction. Hence $y = \sup X$. ∎

**Proposition 1.6.** Let $X$ be bounded below. Show that $\inf X = -\sup(-X)$, where $-X = \{-x \mid x \in X\}$.

*Proof.* As $X$ is bounded below, $-X$ is bounded above, so $\sup -X$ exist. Call it $y$. By definition, for any $x \in -X$, $y \geqslant x$, and for any upper bound $y'$ of $-X$, $y \leqslant y'$. Therefore, for any $x \in X$, $-y \leqslant x$, so $-y$ is a lower bound of $X$.

Suppose that $z$ is a lower bound of $X$. Then, $z \leqslant x$ for all $x \in X$, so $-z \geqslant x$ for all $x \in -X$, so $-z$ is an upper bound of $-X$, and hence $y \leqslant -z$, in turn implying $-y \geqslant z$. Therefore, $-y$ is the infimum of $X$, so $-\sup(-X) = \inf X$. ∎

**Proposition 1.7.** If $A, B$ are bounded subsets of $\mathbb{R}$, then $A \cup B$ is bounded and

$$\sup A \cup B = \sup\{\sup A, \sup B\}.$$

*Proof.* Let $s = \sup A, t = \sup B$, and WLOG, suppose that $s \geqslant t$. Since $s \geqslant a$ for any $a \in A$ and $s \geqslant t \geqslant b$ for any $b \in B$, $s$ is an upper bound of $A \cup B$ so $A \cup B$ is bounded above. (Similarly, it is also bounded below, and thus bounded.) Now let $\epsilon > 0$. From Prop 1.5, there exists $a \in A$ such that $a > s - \epsilon$. Therefore, for all $\epsilon > 0$, there exists $a \in A \cup B$ such that $a > s - \epsilon$, therefore $s = \sup A \cup B$. ∎

## 1.2 Open and closed sets

**Definition 1.8** (Neighborhoods).
Let $x \in \mathbb{R}^n$ be a point, and $\epsilon > 0$. Then the *$\epsilon$-neighborhood* of $x$ is defined by

$$\mathcal{B}_\epsilon(x) = \{y \in \mathbb{R}^n \mid |x - y| < \epsilon\}.$$

**Definition 1.9** (Interiors and boundaries).
Let $X \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^n$. Then $x$ is called

- an *interior point* of $X$ if there exists an $\epsilon > 0$ such that $\mathcal{B}_\epsilon(x) \subseteq X$.
- a *boundary point* of $X$ if for all $\epsilon > 0$, $\mathcal{B}_\epsilon(x) \cap X \neq \emptyset$ and $\mathcal{B}_\epsilon(x) \cap X^c \neq \emptyset$.
- an *exterior point* of $X$ if it is an interior point of $X^c$.

The *interior* $\mathring{X}$ or $X^\circ$ of $X$ is the set of all interior points of $X$, and the *boundary* $\partial X$ of $X$ is the set of all boundary points of $X$.

**Definition 1.10** (Open and closed).
A set $X \subseteq \mathbb{R}^n$ is called

- *open* if it only consists of interior points, that is, $\mathring{X} = X$
- *closed* if its complement is open.

As a result, a set $X$ is open if it contains none of its boundary points, and closed if it contains all of its boundary points.

Also, this is not a dichotomy: $\emptyset$ and $\mathbb{R}^n$ are *clopen*, that is, both open and closed in $\mathbb{R}^n$. Lots of sets are neither open nor closed.

**Exercise 1.11** (1.5.1 in book)**.** For each of the following subsets, state whether it is open or closed (or both or neither), and say why.

a. $\{x \in \mathbb{R} \mid 0 < x \leqslant 1\}$ as a subset of $\mathbb{R}$.
b. $\{(x, y) \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} < 1\}$ as a subset of $\mathbb{R}^2$.
c. the interval $(0, 1]$ as a subset of $\mathbb{R}$
d. $\{(x, y) \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} \leqslant 1\}$ as a subset of $\mathbb{R}^2$.
e. $\{x \in \mathbb{R} \mid 0 \leqslant x \leqslant 1\}$ as a subset of $\mathbb{R}$.
f. $\{(x, y, z) \in \mathbb{R}^3 \mid \sqrt{x^2 + y^2 + z^2} = 1 \text{ and } x, y, z \neq 0\}$ as a subset of $\mathbb{R}^3$
g. the empty set as a subset of $\mathbb{R}$

*Solution.* a. neither: the boundary points are $\{0, 1\}$, but $0$ is not in the set while $1$ is in the set.
b. open: it does not contain any of the boundary points $\{(x, y) \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} = 1\}$
c. neither: see a.
d. closed: it contains all of its boundary points $\{(x, y) \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} = 1\}$
e. closed: the boundary points $\{0, 1\}$ are both in the set.
f. neither: it does not contain the boundary point $\mathbf{0}$ but it contains the boundary points $\{(x, y, z) \in \mathbb{R}^3 \mid \sqrt{x^2 + y^2 + z^2} = 1\}$.
g. clopen: the empty set is both open and closed.

**Exercise 1.12** (1.5.2 in book)**.** For each of the following subsets, state whether it is open or closed (or both or neither), and say why.

a. $(x, y)$-plane in $\mathbb{R}^3$

b. $\mathbb{R} \subset \mathbb{C}$

c. the line $x = 5$ in the $(x, y)$-plane

d. $(0, 1) \subset \mathbb{C}$

e. $\mathbb{R}^n \subset \mathbb{R}^n$

f. the unit sphere in $\mathbb{R}^3$

*Solution.* a. closed: it is its own boundary

b. closed: also its own boundary

c. closed: see above

d. neither: it contains 0.5 but not 0, and both are boundary points.

e. clopen

f. closed: the shell is its own boundary

**Exercise 1.13** (1.5.5, book p.102). For each of the following subsets of $\mathbb{R}$ and $\mathbb{R}^2$, state whether it is open or closed (or both or neither), and prove it.

a. $\{(x, y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 2\}$

b. $\{(x, y) \in \mathbb{R}^2 \mid xy \neq 0\}$

c. $\{(x, y) \in \mathbb{R}^2 \mid y = 0\}$

d. $\mathbb{Q} \in \mathbb{R}$

*Solution.* a. open. Let $\mathbf{p} \in A = \{(x, y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 2\}$, so $1 < |\mathbf{p}| < \sqrt{2}$. Choose $\epsilon = \min\{|\mathbf{p}| - 1, \sqrt{2} - |\mathbf{p}|\}/2$. By the triangle inequality, for any point $\mathbf{q} \in \mathcal{B}_\epsilon(\mathbf{p})$, $1 < |\mathbf{p}| - \epsilon < |\mathbf{q}| < |\mathbf{p}| + \epsilon < \sqrt{2}$, so $\mathbf{q} \in A$, hence $\mathcal{B}_\epsilon(\mathbf{p}) \subset A$. Therefore all points $\mathbf{p} \in A$ are interior points, so $A$ does not contain any of its boundary, so $\mathbf{p}$ is open.

b. open. Let $\mathbf{p} = (a, b) \in X = \{(x, y) \in \mathbb{R}^2 \mid xy \neq 0\}$, so $|a|, |b| > 0$. Choose $\epsilon = \min\{|a|, |b|\}/2$. For any point $\mathbf{q} = (u, v) \in \mathcal{B}_\epsilon(p)$, $|u| > |a| - \epsilon > 0$ and $|v| > |b| - \epsilon > 0$, so $\mathbf{q} \in X$, therefore $\mathcal{B}_\epsilon(\mathbf{p}) \subset X$. Hence all points $\mathbf{p} \in X$ are interior points, so $X$ is open.

c. closed. Consider its complement $C = \{(x, y) \in \mathbb{R}^2 \mid y \neq 0\}$. For any $\mathbf{p} = (a, b) \in C$, consider $\epsilon = |b|/2$. For any point $\mathbf{q} = (u, v) \in \mathcal{B}_\epsilon(\mathbf{p})$, $|v| > |b| - \epsilon > 0$ so $\mathbf{q} \in C$, so $\mathcal{B}_\epsilon(\mathbf{p}) \subset C$, so $C$ is open. Therefore $C^c = \{(x, y) \in \mathbb{R}^2 \mid y = 0\}$ is closed.

d. neither. Since both the rationals and irrationals are dense in $\mathbb{R}$, for any $x \in \mathbb{R}$ and $\epsilon > 0$, $\mathcal{B}_\epsilon(x)$ contains both rationals and irrationals. Therefore $\mathbb{R}$ is the boundary of $\mathbb{Q} \in \mathbb{R}$, so $\mathbb{Q}$ contains some, but not all, of its boundary, and hence $\mathbb{Q}$ is *ajar*[1].

**Exercise 1.14** (1.5.3 in book). Prove the following statements for open subsets of $\mathbb{R}^n$:

a. Any union (finite, countable, uncountable) of open sets is open

b. A finite intersection of open sets is open

c. An infinite intersection of open sets is not necessarily open

*Solution.* a. Let our union be $U = \bigcup S_x$, and let $u \in U$. Therefore $u$ is in one of the open sets $S_x$, and so $u$ is an interior point of $S_x$. Hence there is an $\epsilon > 0$ such that $\mathcal{B}_\epsilon(u) \subset S_x \subset U$, so $u$ is an interior point fo $U$. Therefore, $U$ is open.

b. Let our intersection be $Z = \bigcap_{i=1}^n S_i$. Suppose $z \in Z$, so $z$ is in $S_i$ for all $i = 1, \ldots, n$. Since $S_i$'s are all open, there are $\epsilon_i > 0$ such that $\mathcal{B}_{\epsilon_i}(z) \subset S_i$ for each $i$. Take $\epsilon = \min \epsilon_i$. It follows that $\mathcal{B}_\epsilon(z) \subset \mathcal{B}_{\epsilon_i}(z) \subset S_i$ for all $i$, so $\mathcal{B}_\epsilon(z) \subset Z$, so $z$ is an interior point. Therefore, $Z$ is open.

[1]This is not a real mathematical term

c. The sets $S_n = (-\frac{1}{n}, \frac{1}{n})$ for $n = 1, 2, \ldots$ are all open, but their intersection is $\{0\}$, which is not open because it contains 0, which is its own boundary.

**Theorem 1.15.** *The* closure $\overline{X}$ *of* $X$*, defined as the union of* $X$ *and* $\partial X$*, is the smallest closed set that contains* $X$*, in the sense that no proper subset of* $\overline{X}$ *is both closed and contains* $X$*.*

*Proof.* If $X$ is closed, we are done. Otherwise, assume that $Y \subset \mathbb{R}^n$ is closed with

$$X \subsetneq Y \subseteq \overline{X}.$$

We claim that $Y = \overline{X}$. Suppose $Y \neq \overline{X}$, so there is a point $\mathbf{x} \in \overline{X} \setminus Y$. As $\mathbf{x} \in \overline{X}$, for any $\epsilon > 0$, $\mathcal{B}_\epsilon(\mathbf{x})$ contains points in $X$, and therefore, in $Y$. As $\mathbf{x} \notin Y$, $\mathcal{B}_\epsilon(\mathbf{x})$ also contain a point $\mathbf{x}$ not in $Y$. Therefore, $\mathbf{x}$ is in the boundary of $Y$ but not in $Y$ itself, so $Y$ is not closed, which is a contradiction.[2] ■

**Exercise 1.16.** Let $A, B$ be sets in $\mathbb{R}$. Show that $A^\circ \subseteq A$, $(A^\circ)^\circ = A^\circ$, and that $(A \cap B)^\circ = A^\circ \cap B^\circ$. Show also that $A^\circ \cup B^\circ \subseteq (A \cup B)^\circ$, and give an example that the inclusion may be proper.

*Solution.*

- For any $a \in A^\circ$, there is an $\epsilon > 0$ such that $\mathcal{B}_\epsilon(a) \subset A$, so $a \in \mathcal{B}_\epsilon(a) \subset A$. Therefore, $A^\circ \subseteq A$.
- We already know that $(A^\circ)^\circ \subseteq A^\circ$, so it suffices to show that $A^\circ \subseteq (A^\circ)^\circ$. Let $a \in A^\circ$, so there is an $\epsilon > 0$ such that $\mathcal{B}_\epsilon(a) \subseteq A$. Consider a point $p \in \mathcal{B}_{\epsilon/2}(a)$. By the triangle inequality, $\mathcal{B}_{\epsilon/2}(p) \subseteq \mathcal{B}_\epsilon(a) \subseteq A$, $p \in A^\circ$. Hence, $\mathcal{B}_{\epsilon/2}(a) \subset A^\circ$, so $a \in (A^\circ)^\circ$. Therefore, $A^\circ \subseteq (A^\circ)^\circ$.
- If $c \in (A \cap B)^\circ$, then there is an $\epsilon > 0$ such that $\mathcal{B}_\epsilon(c) \in A \cap B$. Therefore $\mathcal{B}_\epsilon(c) \in A$, implying $c \in A^\circ$, and $\mathcal{B}_\epsilon(c) \in B$, implying $c \in B^\circ$. On the other hand, if $c \in A^\circ$ and $c \in B^\circ$, there exists $\epsilon_a, \epsilon_b > 0$ such that $\mathcal{B}_{\epsilon_a}(c) \subseteq A$ and $\mathcal{B}_{\epsilon_b}(c) \subseteq B$. Take $\epsilon = \min\{\epsilon_a, \epsilon_b\}$ to get $\mathcal{B}_\epsilon(c) \subseteq \mathcal{B}_{\epsilon_a}(c) \subseteq A$ and $\mathcal{B}_\epsilon(c) \subseteq \mathcal{B}_{\epsilon_b}(c) \subseteq B$, so $\mathcal{B}_\epsilon(c) \subseteq A \cap B$, which means $c \in (A \cap B)^\circ$.
- Let $c \in A^\circ \cup B^\circ$, so $c \in A^\circ$ or $c \in B^\circ$. WLOG suppose the former is true. Then, there is an $\epsilon > 0$ such that $\mathcal{B}_\epsilon(c) \subseteq A \cup B$, so $c \in (A \cup B)^\circ$. However, if we take, for example, $A = [-1, 0]$ and $B = [0, 1]$, it follows that $A^\circ \cup B^\circ = (-1, 0) \cup (0, 1) \neq (-1, 1) = (A \cup B)^\circ$. ■

## 1.3 Sequences and limits

The definition for convergence in $\mathbb{R}^n$ is the same as in $\mathbb{R}$:

**Definition 1.17.**
A sequence $\{\mathbf{a}_i\}$ of points in $\mathbb{R}^n$ converges to $\mathbf{a} \in \mathbb{R}^n$ if for all $\epsilon > 0$ there exists an $M$ such that for all $m > M$, $|\mathbf{a}_m - \mathbf{a}| < \epsilon$ (or equivalently, $\mathbf{a}_m \in \mathcal{B}_\epsilon(\mathbf{a})$.)

**Proposition 1.18.** A sequence $m \mapsto \mathbf{a}_m$ with $\mathbf{a}_m \in \mathbb{R}^n$ converges iff its individual components all converge.

[2] Alternate proof: $\mathbf{x} \in \mathbb{R}^n - Y$ which is open so there exists an $\epsilon > 0$ such that $\mathcal{B}_\epsilon(\mathbf{x}) \subseteq R^n - Y$ so $\mathcal{B}_\epsilon(\mathbf{x}) \subseteq \mathbb{R}^n - X$ which contradicts $\mathbf{x} \in \overline{X}$.

*Proof.* ($\Rightarrow$) Write $\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}^T$, $\mathbf{a}_m = \begin{pmatrix} a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}^T$. For any $k$, given an $\epsilon > 0$, there is an $M$ such that for all $m > M$, $|a_{m,k} - a_k| \leqslant |\mathbf{a}_m - \mathbf{a}| < \epsilon$, so $\{a_{m,k}\}$ converges to $a_k$.

($\Leftarrow$) Define $a_{m,k}$ as above; let $\{a_{m,k}\}$ converge to $a_k$; $\mathbf{a} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}^T$. Fix $\epsilon > 0$. For any $k$, there is an $M_k$ such that for all $m > M_k$, $|a_{m,k} - a_k| \leqslant \epsilon/\sqrt{n}$. Choose $M = \max M_k$. This gives

$$|\mathbf{a}_m - \mathbf{a}| = \sqrt{\sum_{k=1}^{n} (a_{m,k} - a_k)^2} < \sqrt{\frac{n\epsilon}{n}} = \epsilon$$

so $\{\mathbf{a_m}\}$ converges to $\mathbf{a}$. ∎

**Proposition 1.19.** If a sequence $i \mapsto \mathbf{a}_i$ of points in $\mathbb{R}^n$ converges to both $\mathbf{a}$ *and* $\mathbf{b}$, then $\mathbf{a} = \mathbf{b}$.

*Proof.* For any $\epsilon > 0$, there is an $M_a$ such that for all $m > M_a$, $|\mathbf{a}_m - \mathbf{a}| < \epsilon/2$ and also an $M_b$ such that for all $m > M_b$, $|\mathbf{a}_m - \mathbf{b}| < \epsilon/2$. Therefore there is an $M(\epsilon) := \max\{M_a, M_b\} + 1$ such that

$$|\mathbf{a}_{M(\epsilon)} - \mathbf{a}| < \frac{\epsilon}{2} \text{ and } |\mathbf{a}_{M(\epsilon)} - \mathbf{b}| < \frac{\epsilon}{2},$$

therefore, by the triangle inequality, $|\mathbf{a} - \mathbf{b}| < \epsilon$. As this holds for all $\epsilon > 0$, $|\mathbf{a} - \mathbf{b}| = 0$, and hence $\mathbf{a} = \mathbf{b}$. ∎

Next is an important fact linking limits to closed sets.

**Proposition 1.20** (Closed sets contain all limit points).

a. Let $i \to \mathbf{x}_i$ be a sequence in a closed set $C \subset \mathbb{R}^n$ converging to $\mathbf{x} \in \mathbb{R}^n$. Then $\mathbf{x} \in C$.

b. Conversely, if every convergent sequence in a set $C \subset \mathbb{R}^n$ converges to a point in $C$, then $C$ is closed.

*Proof.*

a. If $\mathbf{x} \notin C$, then $\mathbf{x} \in C^c$ which is open, so there exists $r > 0$ such that $\mathcal{B}_r(\mathbf{x}_0) \subset C^c$. Then for all $m$ we have $|\mathbf{x}_m - \mathbf{x}| \geqslant r$. But by the definition of convergence, for any $\epsilon > 0$ we have $|\mathbf{x}_m - \mathbf{x}| < \epsilon$ for $m$ large enough. This is a contradiction when $\epsilon = r$.

b. If $C$ is not closed then there is a boundary point $\mathbf{x}$ of $C$ that is not in $C$. Since for all $\epsilon > 0$, $\mathcal{B}_\epsilon(\mathbf{x})$ contains a point $\mathbf{x}(\epsilon)$ in $C$, we can find a sequence, say,

$$\mathbf{x}(1), \mathbf{x}(1/2), \mathbf{x}(1/4), \cdots$$

that converges to $\mathbf{x} \notin C$. ∎

**Definition 1.21.**

If a sequence $\{\mathbf{a}_i\}$ converges to $\mathbf{a} \in \mathbb{R}$ then we say $\mathbf{a}$ is the *limit* of $\{\mathbf{a}_i\}$, denoted by

$$\mathbf{a} = \lim_{n \to \infty} \mathbf{a}_n.$$

**Definition 1.22.**

Let $X$ be a subset of $\mathbb{R}^n$ and $\mathbf{x}_0$ a point in $\overline{X}$. A function $f : X \to \mathbb{R}^m$ has the *limit* $\mathbf{a}$ at $\mathbf{x}_0$, written

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \mathbf{a},$$

if for all $\epsilon > 0$ there exists $\delta > 0$ such that for all $\mathbf{x} \in X$,

$$|\mathbf{x} - \mathbf{x}_0| < \delta \implies |f(\mathbf{x}) - \mathbf{a}| < \epsilon.$$

Limits behave as you'd expect on $\mathbb{R}^n$ (can be added, multiplied by a scalar limit, preserve dot products) with one caveat: the limit has to be the same when approach it from *all directions*, so instead of left and right limit, there are infinitely many paths.

Similar to convergent sequences, we can break limits into components.

**Proposition 1.23.** The limit of a function at a point exists iff the limit of all components of the function at that point exists.

*Proof.* Left as an exercise for the reader.

**Exercise 1.24** (1.5.14 in book)**.** State whether the following limits exist, and prove it.[3]

a. $\displaystyle\lim_{(x,y)\to(1,2)} \frac{x^2}{x+y}$

b. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sqrt{|x|}y}{x^2+y^2}$

c. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sqrt{|xy|}}{\sqrt{x^2+y^2}}$

d. $\displaystyle\lim_{(x,y)\to(1,2)} x^2+y^3-3$

*Solution.* a. The limit exists, and is $\frac{1^2}{1+2} = \frac{1}{3}$. The proof is straightforward.

b. For each $t > 0$ and for $(x,y) = (t,t)$,

$$\frac{\sqrt{|x|}y}{x^2+y^2} = \frac{t^{1.5}}{2t^2} = \frac{1}{2\sqrt{t}}.$$

As $t \to 0$, this goes to $+\infty$, so the limit does not exist.

c. Fix a $s > 0$. For each $t > 0$ and for $(x,y) = (t, st)$,

$$\frac{\sqrt{|xy|}}{\sqrt{x^2+y^2}} = \frac{t\sqrt{s}}{t\sqrt{1+s^2}} = \frac{1}{\sqrt{1+s^2}}.$$

Therefore, as we approach $(0,0)$ along lines $y = sx$ for different values of $s$, say $s = 0$ and $1$, $\frac{\sqrt{|xy|}}{\sqrt{x^2+y^2}}$ goes to different values, so the limit does not exist.

d. The limit exists, and is $1^2 + 2^3 - 3 = 6$. The proof is straightforward.

## 1.4 Continuity

**Definition 1.25.**
Let $X \subset \mathbb{R}^n$. A mapping $f : X \to \mathbb{R}^m$ is *continuous at* $\mathbf{x}_0 \in X$ if

$$\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0);$$

$f$ is *continuous on* $X$ if it is continuous at every point of $X$. Equivalently, $f : X \to \mathbb{R}^m$ is continuous at $\mathbf{x}_0 \in X$ iff for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$|\mathbf{x} - \mathbf{x}_0| < \delta \implies |f(\mathbf{x}) - f(\mathbf{x}_0)| < \epsilon.$$

Approaching $\mathbf{x}_0$ continuously from all directions is pretty hard—thankfully, just considering sequences is enough.

[3] The points $\binom{x}{y}$ are written $(x,y)$ to preserve space.

**Proposition 1.26.** Let $X \subset \mathbb{R}^n$ and $\mathbf{x} \in X$. mapping $f : X \to \mathbb{R}^m$ is continuous at $\mathbf{x}$ iff for every sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ approaching $\mathbf{x}$,

$$\lim_{n\to\infty} f(\mathbf{x}_n) = f(\mathbf{x})$$

Continuity has some nice properties: it is preserved under addition, scalar multiplication, division (not by zero), multiplication.

As a result, polynomials are continuous on all of $\mathbb{R}^n$, and rational functions are continuous on the subset of $\mathbb{R}^n$ where the quotient does not vanish.

**Exercise 1.27** (1.5.21 in book)**.** For the following functions, can you choose a value for $f$ at $(0,0)$ to make the function continuous at the origin?

a. $f(x,y) = \dfrac{1}{x^2+y^2+1}$

b. $f(x,y) = \dfrac{\sqrt{x^2+y^2}}{|x|+|y|^{1/3}}$

c. $f(x,y) = (x^2+y^2)\ln(x^2+2y^2)$

d. $f(x,y) = (x^2+y^2)\ln|x+y|$

*Solution.* a. $f$ is a rational function that does not vanish, so it is already continuous and we can choose $f(0,0) = \frac{1}{0^2+0^2+1} = 1$.

b. Approaching $(0,0)$ on the line $x=0$ gives $f(0,y) = |y|^{2/3}$ which goes to $+\infty$ as $y \to 0$ so a limit does not exist, and we cannot make $f$ continuous.

c. For any $\mathbf{p} = (x,y) = (r,\theta)$ with $r < 1$,

$$r^2 \ln r^2 = (x^2+y^2)\ln(x^2+y^2) < f(x,y) \leqslant 0$$

so $0 = \lim_{r\to0} -r^2\ln r^2 \leqslant \lim_{r\to0} f(\mathbf{p}) \leqslant 0$, squeezing $\lim_{r\to0} f(\mathbf{p})$ to 0, so we define $f(0,0)=0$.

d. The function goes to $-\infty$ among the line $y=-x$, so the limit does not exist.

**Exercise 1.28** (1.5.16 in book)**.** a. Let $D^* \in \mathbb{R}^2$ be the region $0 < x^2+y^2 < 1$, and let $f : D^* \to \mathbb{R}$ be a function. What does the following assertion mean?

$$\lim_{(x,y)\to(0,0)} f(x,y) = a$$

b. For the following two functions, defined on $\mathbb{R}^2 - \{0\}$, either show that the limit exists at $\mathbf{0}$ and find it, or show that it does not exist:

$$f(x,y) = \frac{\sin(x+y)}{\sqrt{x^2+y^2}} \qquad g(x,y) = (|x|+|y|)\ln(x^2+y^4)$$

*Solution.* a. As we wander closer to hole in the disc, $f$ approaches $a$.

b. For $f$, the limit does not exist: if we approach $(0,0)$ on the line $x=y$ from $x>0$,

$$\lim_{x\to0^+} f(x,x) = \lim_{x\to0} \frac{\sin 2x}{\sqrt{2}x} = \lim_{x\to0} \frac{2\cos 2x}{\sqrt{2}} = \sqrt{2},$$

however, if we approach $(0,0)$ on the line $x=-y$,

$$\lim_{x\to0} f(x,-x) = \lim_{x\to0} \frac{0}{\sqrt{2}x} = 0.$$

For $g$, the limit exists and is zero. For any $\mathbf{p} = (x, y)$, define $s = |x| + |y|$. Then, for $s$ (and thus $|x|, |y|$) small enough,

$$1 > x^2 + y^4 > x^4 + y^4 \geqslant \frac{s^4}{8},^4$$

so $0 > \ln(x^2 + y^4) > 4 \ln s - \ln 8$, so

$$s(4 \ln s - \ln 8) < g(x, y) < 0.$$

Taking $s \to 0$ gives $\lim_{(x,y)\to(0,0)} = 0$.

**Digression 1.29** ($L^p$ norms). The *p-norm* or *$L^p$-norm* of $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots v_n \end{bmatrix}$ is defined by

$$\|\mathbf{v}\|_p = (|v_1|^p + |v_2|^p + \cdots + |v_n|^p)^{\frac{1}{p}}$$

When $p = \infty$ this becomes

$$\|\mathbf{v}\|_\infty = \max\{|v_1|, |v_2|, \ldots, |v_n|\}.$$

The vector space of all real sequences such that $\sum |x_i|^p$ converges is called the *$L^p$ space*, and this is studied in functional analysis.

## 1.5 Compact sets

**Definition 1.30** (Bounded sets).
A subset $X \subset \mathbb{R}^n$ is *bounded* if it is contained in a ball in $\mathbb{R}^n$ centered at the origin:

$$X \subset B_R(\mathbf{0}) \text{ for some } R < \infty.$$

**Definition 1.31** (Compact sets).
A nonempty subset $C \subset \mathbb{R}^n$ is *compact* if it is closed and bounded

Compactness is a powerful definition, as it is enough to imply the existence of a convergent subsequence for any sequence in the compact set:

**Theorem 1.32** (Bolzano-Weierstrass theorem). *If a compact set $C \subset \mathbb{R}^n$ contains a sequence $i \mapsto \mathbf{x}_i$, then that sequence has a convergent subsequence $j \mapsto \mathbf{x}_{i(j)}$ whose limit is in $C$.*

*Proof.* Put $C$ in a bounded box, say $B_1 : \{|v_1|, |v_2|, \ldots, |v_n| \leqslant N\}$. Choose $a_1$ (from the sequence) anywhere in $B_1$. Chop the box into 4 smaller boxes–one of these, call $B_2$ must have infinitely many items of the sequence. Choose $a_2$ from the sequence anywhere in $B_2$. Repeat ad infinitum. ■

**Definition 1.33** (Bounds of a function).
A number $M$ is the *supremum*, denoted by $\sup_{\mathbf{c}\in C} f(\mathbf{c})$ of a function[5] $f : C \to \mathbb{R}$ if $M$ is the least upper bound of the values of $f$. If there is an $\mathbf{a} \in C$ such that $f(\mathbf{a}) = M$ then $M$ is also the *maximum* of $f$.

Similarly, a number $M$ is the *infimum*, denoted by $\inf_{\mathbf{c}\in C} f(\mathbf{c})$ of a function $f : C \to \mathbb{R}$ if $M$ is the greatest lower bound of the values of $f$. If there is a $\mathbf{b} \in C$ such that $f(\mathbf{b}) = M$ then $M$ is also the *minimum* of $f$.

[4] Power mean inequality: for $x_i > 0, p > q > 0$, $\left(\sum x_i^p\right)^{1/p} > \left(\sum x_i^q\right)^{1/q}$

[5] Here, $C$ can be any subset of $\mathbb{R}^n$—not necessarily compact.

**Theorem 1.34.** *A continuous function on a compact domain has both maximums and minimums. More formally, let $C \subset \mathbb{R}^n$ be a compact subset, and let $f : C \to R$ be a continuous function. Then there exists a point $\mathbf{a} \in C$ such that $f(\mathbf{a}) \geqslant f(\mathbf{c})$ for all $\mathbf{c} \in C$ as well as a point $\mathbf{b} \in C$ such that $f(\mathbf{b}) \leqslant f(\mathbf{c})$ for all $\mathbf{c} \in C$.*

*Proof.* Here is a sketch: if $f$ is unbounded then there exists a sequence $n \mapsto \mathbf{x}_n$ such that $f(\mathbf{x}_N) > N$ for all positive integers $N$. By Bolzano-Weierstrass, there is a subsequence of $j \mapsto \mathbf{x}_{n_j}$ that converges to a point in $C$, but this implies $f$ is not continuous at $C$. Therefore $f$ is bounded, and so $f$ has a supremum $M$. Therefore there is a sequence $n \mapsto \mathbf{y}_n$ such that

$$\lim_{n\to\infty} f(\mathbf{y}_n) = M.$$

Take a subsequence $j \mapsto \mathbf{y}_{n_j}$ that converges to a point $\mathbf{y}$, and it follows that

$$f(\mathbf{y}) = \lim_{j\to\infty} f(\mathbf{y}_{n_j}) = M. \quad \blacksquare$$

# 2 Derivatives

## 2.1 Review of derivatives in $\mathbb{R}$

**Exercise 2.1** (1.7.3 in book). Find $f'(x)$ for the following functions $f$:

a. $f(x) = \sin^3(x^2 + \cos x)$

b. $f(x) = \cos^2\left((x + \sin x)^2\right)$

c. $f(x) = (\cos x)^4 \sin x$

d. $f(x) = (x + \sin^4 x)^3$

e. $f(x) = \frac{\sin x^2 \sin^3 x}{2+\sin x}$

f. $f(x) = \sin\left(\frac{x^3}{\sin x^2}\right)$

*Answer:* a. $3\sin^2(x^2 + \cos x) \cdot \cos(x^2 + \cos x) \cdot (2x - \sin x)$

**Exercise 2.2** (1.7.4 in book). Using the definition, check whether the following functions are differentiable at 0.

a. $f(x) = |x|^{3/2}$

b. $f(x) = \begin{cases} x \ln|x| & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

c. $f(x) = \begin{cases} x/\ln|x| & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

In each case, if $f$ is differentiable at 0, is $f(0 + h) - f(0) - f'(0)h$ comparable to $h^2$?

*Solution.* a. For $x > 0$, $\frac{d}{dx}x^{3/2} = \frac{3}{2}x^{-1/2}$. For $x < 0$, $\frac{d}{dx}(-x)^{3/2} = -\frac{3}{2}x^{-1/2}$. These both approach 0 as $x \to 0$, so $f$ is differentiable. However, the error is $O(h^{3/2})$.

b. We have $\frac{h\ln|h|-0}{h-0} = \ln|h|$ which goes to $-\infty$ as $h \to 0$, so $f$ is not differentiable.

c. We have $\frac{\frac{h}{\ln|h|}-0}{h-0} = \frac{1}{\ln|h|}$ which indeed goes to 0 as $h \to 0$, so $f$ is differentiable. Size of error is left to the reader.

**Theorem 2.3** (Mean value theorem). *If $f : [a, b] \to \mathbb{R}$ is continuous, and $f$ is differentiable on $(a, b)$, then there exists $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

This theorem don't really fit anywhere, but I'll put it here.

**Theorem 2.4** (Fundamental theorem of algebra). *Let $P$ be a single-variable polynomial with complex coefficients and positive degree. Then $P$ has a root.*

## 2.2 Derivatives: full, partial, directional

The idea behind derivatives in $\mathbb{R}^n$ is like that of derivatives in $\mathbb{R}$: *local linearization.* Given $\mathbf{f}: \mathbb{R}^n \to \mathbb{R}^m$, the *derivative of $f$ at* $\mathbf{x}$ is linear approximation $\mathbf{Df}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$ of $\mathbf{f}$ at $\mathbf{x}$ with error vanishing faster than $\mathbf{h}$, that is,

$$(\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})) - \mathbf{Df}(\mathbf{x})(\mathbf{h}) \in o(\mathbf{h}),$$

or, in more familiar terms,

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{(\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})) - [\mathbf{Df}(\mathbf{x})](\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0}.$$

If $[\mathbf{Df}(\mathbf{x})]$ exists, we can find it as follows: for each standard basis vector $\mathbf{e}_i$,

$$\begin{aligned} 0 &= \lim_{h\to 0} \tfrac{1}{|h\mathbf{e}_i|} \left(\mathbf{f}(\mathbf{x}+h\mathbf{e}_i) - \mathbf{f}(\mathbf{x}) - [\mathbf{Df}(\mathbf{x})](h\mathbf{e}_i)\right) \\ &= \lim_{h\to 0} \tfrac{1}{|h|} \left(\mathbf{f}(\mathbf{x}+h\mathbf{e}_i) - \mathbf{f}(\mathbf{x}) - h[\mathbf{Df}(\mathbf{x})](\mathbf{e}_i)\right) \\ &= \tfrac{\mathbf{f}(\mathbf{x}+h\mathbf{e}_i) - f(x)}{h} - [\mathbf{Df}(\mathbf{x})]\mathbf{e}_i \end{aligned}$$

so if $[\mathbf{Df}(\mathbf{x})]$ exists,

$$[\mathbf{Df}(\mathbf{x})]\mathbf{e}_i = \lim_{h\to 0} \frac{\mathbf{f}(\mathbf{x}+h\mathbf{e}_i) - \mathbf{f}(x)}{h}.$$

This can be calculated by considering $\mathbf{x}_i$ (that is, the $i^{\text{th}}$ component of $\mathbf{x}$) the only variable and holding all other components constant. In calculating this, only a part of $\mathbf{x}$ is changed, and so the above limit is called the *partial derivative* of $\mathbf{f}$ at $\mathbf{x}$ with respect to $\mathbf{x}_i$.

There are a variety of notations for partial derivatives:

- $D_i\mathbf{f}(x)$
- $D_x\mathbf{f}(\mathbf{x}), D_y\mathbf{f}(\mathbf{x}), D_z f(\mathbf{x})$
- $\dfrac{\partial \mathbf{f}}{\partial x_i}$
- $f_x, f_{x_1}, \ldots$

Note that computing partial derivatives is just like computing derivatives in $\mathbb{R}$.

**Example 2.5.** For $\mathbf{f}(x,y) = \sin(x^2+y^3)$, $D_x\mathbf{f}(x,y) = \cos(x^2+y^3)\cdot 2x$ and $D_y\mathbf{f}(x,y) = \cos(x^2+y^3)\cdot 3y^2$, so

$$\left[\mathbf{Df}\begin{pmatrix} x \\ y \end{pmatrix}\right] = \cos(x^2+y^3)\cdot \begin{bmatrix} 2x & 3y^2 \end{bmatrix}.$$

**Exercise 2.6** (1.7.5 in book). Find the partial derivatives of the following functions $f: \mathbb{R}^2 \to \mathbb{R}$:

a. $f\begin{pmatrix} x \\ y \end{pmatrix} = \sqrt{x^2+y}$

b. $f\begin{pmatrix} x \\ y \end{pmatrix} = x^2y + y^4$

c. $f\begin{pmatrix} x \\ y \end{pmatrix} = \cos xy + y\cos y$

d. $f\begin{pmatrix} x \\ y \end{pmatrix} = \dfrac{xy^2}{\sqrt{x+y^2}}$

Partial derivatives concerns situations where we approach a point on line parallel to an axis, and there is nothing stopping us from approaching a point on other lines. This is the idea of directional derivatives.

**Definition 2.7.**
The *directional derivative* of $f$ at $\mathbf{x}$ in direction $\mathbf{v}$ is the rate of change of $f$ as we step into direction $\mathbf{v}$, which is

$$\lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{v}) - f(\mathbf{x})}{h}.$$

**Proposition 2.8** (1.7.14 in book)**.** If $U \subset \mathbb{R}^n$ is open, and $f : U \to \mathbb{R}^m$ is differentiable at $\mathbf{a} \in U$, then all directional derivatives of $f$ at $\mathbf{a}$ exist, and the directional derivative in the direction $\mathbf{v}$ is given by

$$[\mathbf{D}f(\mathbf{x})]\mathbf{v} = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{v}) - f(\mathbf{x})}{h}.$$

*Proof.* Define the *remainder* $r(\mathbf{h})$ as the error of the linear approximation:

$$r(\mathbf{h}) = (f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a})) - [\mathbf{D}f(\mathbf{a})]\mathbf{h}.$$

As $f$ is differentiable at $\mathbf{a}$, $\lim_{\mathbf{h} \to \mathbf{0}} r(\mathbf{h})/\|\mathbf{h}\| = \mathbf{0}$.

Substituting $h\mathbf{v}$ for $\mathbf{h}$ then dividing everything by $h$ gives

$$|\mathbf{v}| \frac{r(h\mathbf{v})}{\|h\mathbf{v}\|} = \frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h}.$$

When taking $h \to 0$, LHS goes to $0$, so the limit of RHS exists, and is $0$. ■

## 2.3 Jacobian matrix

The *Jacobian matrix* of $f$ at $\mathbf{x}$, or simply *Jacobian*, denoted $[\mathbf{Jf}(\mathbf{x})]$, is the matrix listing out all the partial derivatives of $\mathbf{f}$ at $\mathbf{x}$, that is, if we write

$$\mathbf{f}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

then

$$\left[\mathbf{Jf}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right] = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \frac{\partial \mathbf{f}}{\partial x_2} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

Here we might notice that $[\mathbf{Jf}(\mathbf{x})]\mathbf{e}_i$ is precisely $\frac{\partial \mathbf{f}}{\mathbf{x}_i}$, so *if* $[\mathbf{Df}(\mathbf{x})]$ *exists* then $[\mathbf{Df}(\mathbf{x})] = [\mathbf{J}(\mathbf{x})]$. In words, if the derivative exists, then the Jacobian matrix is also the matrix of the derivative. Just to make sure we don't forget:

**Warning 2.9.**
The Jacobian is only the matrix of the derivative if the function is actually differentiable!

The natural question to ask after this is thus: when is a function differentiable?

## 2.4 Differentiability

**Definition 2.10** ($C^p$ function).
A $C^p$ *function* on $U \subset \mathbb{R}^n$ is a function that is $p$ times continuously differentiable: all of its partial derivatives up to order $p$ exist and are continuous on $U$.

**Theorem 2.11** (Criterion for differentiability). *If $U$ is an open subset of $\mathbb{R}^n$, and $f : U \to \mathbb{R}^m$ is a $C^1$ mapping, then $f$ is differentiable on $U$, and its derivative is given by the Jacobian matrix.*

**Exercise 2.12** (1.9.1 in book). Show that the function

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \frac{x^4+y^4}{x^2+y^2} & \text{if } \binom{x}{y} \neq \binom{0}{0} \\ 0 & \text{if } \binom{x}{y} = \binom{0}{0} \end{cases}$$

is differentiable at every point in $\mathbb{R}^2$.

*Solution.* Since $f$ is symmetric in $x$ and $y$, it suffices to consider partial derivatives of $f$ w.r.t. $x$. For any point $\binom{x}{y} \neq \binom{0}{0}$, we have

$$D_x f\begin{pmatrix} x \\ y \end{pmatrix} = \frac{4x^3(x^2+y^2) - 2x(x^4+y^4)}{(x^2+y^2)^2}$$

which exists.

For $\binom{0}{0}$ we have

$$D_x f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \lim_{h\to 0} \frac{f\binom{h}{0} - f\binom{0}{0}}{h} = \lim_{h\to 0} \frac{h^4}{h \cdot h^2} = 0.$$

It is left to show that

$$\lim_{\binom{x}{y}\to\binom{0}{0}} D_x f\begin{pmatrix} x \\ y \end{pmatrix} = D_x f\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

but that is true since

$$\left| \frac{4x^3(x^2+y^2) - 2x(x^4+y^4)}{(x^2+y^2)^2} \right| \leqslant \left| \frac{4x^3}{x^2+y^2} \right| + \left| \frac{2x(x^4+y^4)}{(x^2+y^2)^2} \right| \leqslant |4x| + |2x|$$

which goes to 0 as $x, y \to 0$.

Differentiability is a *very strong* condition, and even the existence of directional derivatives is not enough:

**Exercise 2.13** (1.9.2a in book). Show that for

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \frac{3x^2y-y^3}{x^2+y^2} & \text{if } \binom{x}{y} \neq \binom{0}{0} \\ 0 & \text{if } \binom{x}{y} = \binom{0}{0}, \end{cases}$$

all directional derivatives exist, but that $f$ is not differentiable at the origin.

*Solution.* We first calculate the Jacobian. We have

$$D_x f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \lim_{h\to 0} \frac{f\binom{h}{0}}{h} = \lim_{h\to 0} \frac{3h^2}{h^3} = 0$$

and

$$D_y f\begin{pmatrix}0\\0\end{pmatrix} = \lim_{h\to 0} \frac{f\binom{0}{h}}{h} = \lim_{h\to 0} \frac{-h^3}{h^3} = -1$$

so

$$\left[Jf\begin{pmatrix}0\\0\end{pmatrix}\right] = \begin{bmatrix}0 & -1\end{bmatrix}$$

For a direction $\mathbf{v} = \begin{bmatrix}v_1\\v_2\end{bmatrix}$, $\lim_{h\to 0} \frac{f(h\mathbf{v})}{h} = \lim_{h\to 0} \frac{3v_1^2v_2 - v_2^3}{v_1^2 + v_2^2}$ which clearly exists, so all directional derivatives of $f$ exist.

If $f$ is differentiable, the directional derivative has to be $[Jf(\mathbf{0}]\mathbf{v} = -v_2$ but for $\mathbf{v} = \begin{bmatrix}1\\1\end{bmatrix}$ this gives $-1 = 1$, which is a contradiction.

**Exercise 2.14** (1.9.2b in book). Show that for

$$g\begin{pmatrix}x\\y\end{pmatrix} = \begin{cases}\frac{x^2y}{x^4+y^2} & \text{if } \binom{x}{y} \neq \binom{0}{0}\\ 0 & \text{if } \binom{x}{y} = \binom{0}{0},\end{cases}$$

all directional derivatives exist at every point, but that $g$ is not continuous

*Solution.* As $g$ is a rational function that does not vanish at any point that is not $\binom{0}{0}$, directional derivatives of $g$ exist at any $\binom{x}{y} \neq \binom{0}{0}$.

At $\binom{0}{0}$, so the directional derivatives exist.

However, approaching $\binom{x}{y}$ on the line $y = x^2$ gives

$$\lim_{x\to 0} g\begin{pmatrix}x\\x^2\end{pmatrix} = \lim_{x\to 0} \frac{x^4}{2x^4} = \frac{1}{2} \neq 0$$

so $f$ is not continuous.

**Exercise 2.15** (1.9.2c in book). Show that for

$$h\begin{pmatrix}x\\y\end{pmatrix} = \begin{cases}\frac{x^2y}{x^6+y^2} & \text{if } \binom{x}{y} \neq \binom{0}{0}\\ 0 & \text{if } \binom{x}{y} = \binom{0}{0},\end{cases}$$

all directional derivatives exist at every point, but that $h$ is unbounded in a neighborhood of 0.

*Solution.* As $h$ is a rational function that does not vanish at any point that is not $\binom{0}{0}$, directional derivatives of $h$ exist at any $\binom{x}{y} \neq \binom{0}{0}$.

At $\binom{0}{0}$, so the directional derivatives exist.

However, approaching $\binom{x}{y}$ on the line $y = x^3$ gives

$$g\begin{pmatrix}x\\x^3\end{pmatrix} = \frac{x^5}{2x^6} = \frac{1}{2x}$$

which grows unbounded as $x \to 0$.

**Exercise 2.16** (1.34 in book). Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by the formula

$$f\begin{pmatrix}x\\y\end{pmatrix} = \begin{cases}\dfrac{xy}{x^2+y^2} & \text{if } \begin{pmatrix}x\\y\end{pmatrix} \neq \begin{pmatrix}0\\0\end{pmatrix}\\ 0 & \text{if } \begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}.\end{cases}$$

a. Show that both partial derivatives exist everywhere.

b. Where is $f$ differentiable?

*Solution.* At any point $\mathbf{p}$ except the origin, $f$ is a rational function that does not vanish, so $f$ has partial derivatives and is differentiable at $\mathbf{p}$. At the origin,

a. $D_x f(\mathbf{0}) = \lim_{h \to 0} \frac{f(h\mathbf{e}_1) - f(\mathbf{0})}{h} = 0$ so $D_x f(\mathbf{0})$ exists. Similarly, as $f$ is symmetric, $D_y f(\mathbf{0})$ also exists as well.

b. Approaching the origin on the line $x = y$ gives

$$\lim_{x \to 0} f\begin{pmatrix} x \\ x \end{pmatrix} = \lim_{x \to 0} \frac{x^2}{2x^2} = \frac{1}{2} \neq 0$$

so $f$ is not continuous, and is not differentiable.

## 2.5 Computing derivatives

Most rules in $\mathbb{R}$ transfer to $\mathbb{R}^m$:

- A constant function is differentiable, and its derivative is the zero matrix $[\mathbf{0}]$.
- A linear function $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is differentiable, and is its own derivative at all points.
- Sum rule: Derivatives are additive.
- Product rule: if $f : U \to \mathbb{R}$ and $\mathbf{g} : U \to \mathbb{R}^m$ are differentiable at $\mathbf{a}$, then so is $f\mathbf{g}$, and the derivative is given by

$$[\mathbf{D}(f\mathbf{g})(\mathbf{a})]\mathbf{v} = f(\mathbf{a})[\mathbf{D}\mathbf{g}(\mathbf{a})]\mathbf{v} + ([\mathbf{D}f(\mathbf{a})]\mathbf{v})\mathbf{g}(\mathbf{a}).$$

The division rule works as well, but it's pretty hard to type out in LaTeX.

For a *differentiable* function where partial derivatives are easy to calculate, also recall the Jacobian: if

$$f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

then

$$\left[\mathbf{Df}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right] = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

**Theorem 2.17** (Chain rule). *Let $U \subset \mathbb{R}^n, V \subset \mathbb{R}^m$ be open sets, let $\mathbf{g} : U \to V$ and $\mathbf{f} : V \to \mathbb{R}^p$ be mappings, and let $\mathbf{a}$ be a point of $U$. If $g$ is differentiable at $\mathbf{a}$ and $\mathbf{f}$ is differentiable at $\mathbf{g}(\mathbf{a})$, then the composition $\mathbf{f} \circ \mathbf{g}$ is differentiable at $\mathbf{a}$, and its derivative is given by*

$$[\mathbf{D}(\mathbf{f} \circ \mathbf{g})(a)] = [\mathbf{Df}(\mathbf{g}(a))] \circ [\mathbf{Dg}a].$$

**Warning 2.18.**
$(\mathbf{f} \circ \mathbf{g})(\mathbf{a}) = \mathbf{f}(\mathbf{g}(\mathbf{a}))$ does not mean $[\mathbf{D}(\mathbf{f} \circ \mathbf{g})(\mathbf{a})]) = [\mathbf{Df}(\mathbf{g}(\mathbf{a}))]$: the former is the derivative of $\mathbf{f} \circ \mathbf{g}$ at $\mathbf{a}$ but the latter is the derivative of $\mathbf{f}$ at point $\mathbf{g}(\mathbf{a})$.

**Example 2.19** (1.8.4 in book). Define $\mathbf{g} : \mathbb{R} \to \mathbb{R}^3$ and $f : \mathbb{R}^3 \to \mathbb{R}$ by

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2, \text{ and } \mathbf{g}(t) = \begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}.$$

Then, $[\mathbf{D}(f \circ \mathbf{g})(t)]$ can be found by first evaluating $[\mathbf{Dg}(t)] = \begin{bmatrix} 1 \\ 2t \\ 3t^2 \end{bmatrix}$ and $[\mathbf{D}f(\mathbf{g}(t))] = \begin{bmatrix} 2t & 2t^2 & 2t^3 \end{bmatrix}$.

Therefore, we get

$$[\mathbf{D}(f \circ \mathbf{g})(t)] = \begin{bmatrix} 2t & 2t^2 & 2t^3 \end{bmatrix} \begin{bmatrix} 1 \\ 2t \\ 3t^2 \end{bmatrix} = 2t + 4t^3 + 6t^5.$$

Finding $f \circ \mathbf{g}$ directly gives $f \circ \mathbf{g}(t) = t^2 + t^4 + t^6$ so $(f \circ \mathbf{g})'(t) = 2t + 4t^3 + 6t^5$ as it should.

**Exercise 2.20** (1.8.9 in book). Let $\varphi : \mathbb{R} \to \mathbb{R}$ be any differentiable function. Show that the function

$$f\left(y\varphi(x^2 - y^2)\right)$$

satisfies the equation

$$\frac{1}{x} D_1 f \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{y} D_2 f \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{y^2} f \begin{pmatrix} x \\ y \end{pmatrix}$$

*Solution.* As $f\binom{x}{y} = y\varphi(x^2 - y^2)$,

$$\begin{aligned} \frac{1}{x} D_1 f \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{y} D_2 f \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{x} D_x \left(y\varphi(x^2 - y^2)\right) + \frac{1}{y} D_y \left(y\varphi(x^2 - y^2)\right) \\ &= \frac{1}{x}\left(2xy\varphi'(x^2 - y^2)\right) + \frac{1}{y}\left(-2y^2\varphi'(x^2 - y^2) + \varphi'(x^2 - y^2)\right) \\ &= \frac{1}{y}\varphi(x^2 - y^2) \\ &= \frac{1}{y^2} f \begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned}$$

**Exercise 2.21** (1.8.11 in book). Show that if $f\begin{pmatrix} x \\ y \end{pmatrix} = \phi\left(\frac{x+y}{x-y}\right)$ for some differentiable function $\phi : \mathbb{R} \to \mathbb{R}$, then

$$xD_x f + yD_y f = 0.$$

*Solution.* Let $g : \mathbb{R}^2 \to \mathbb{R}$ be defined by $g\binom{x}{y} = \frac{x+y}{x-y}$. Then, $f = \phi \circ g$, so by the chain rule

$$\left[\mathbf{D}f\begin{pmatrix} x \\ y \end{pmatrix}\right] = \left[\mathbf{D}\phi\left(g\begin{pmatrix} x \\ y \end{pmatrix}\right)\right] \circ \left[\mathbf{D}g\begin{pmatrix} x \\ y \end{pmatrix}\right] = \phi'\left(\frac{x+y}{x-y}\right) \cdot \frac{2}{(x-y)^2}\begin{bmatrix} -y & x \end{bmatrix}$$

Therefore,

$$xD_x f + yD_y f = \phi'\left(\frac{x+y}{x-y}\right)(-2xy + 2xy) = 0.$$

**Example 2.22.** Further down the road, we will need to transform coordinates, so as an example, if we write $x = r\cos\theta, y = r\sin\theta$ then we can consider the function $\mathbf{g}$ sending $\binom{r}{\theta} \to \binom{x}{y}$. This gives

$$\left[\mathbf{Dg}\begin{pmatrix} r \\ \theta \end{pmatrix}\right] = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

Therefore, for a function $f : \mathbb{R}^2 \to \mathbb{R}$

$$\left[\mathbf{D}f\begin{pmatrix} x(r,\theta) \\ y(r,\theta) \end{pmatrix}\right] = \left[\mathbf{D}(f\circ\mathbf{g})\begin{pmatrix} r \\ \theta \end{pmatrix}\right] = \left[\mathbf{D}f\begin{pmatrix} x \\ y \end{pmatrix}\right] \circ \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}.$$

**Theorem 2.23** (Multivariable mean value theorem). *Let $U \subset \mathbb{R}^n$ be open, let $f : U \to \mathbb{R}$ be differentiable, and let the segment $[\mathbf{a},\mathbf{b}]$ joining $\mathbf{a} \to \mathbf{b}$ be contained in $U$. Then there exists $\mathbf{c}_0 \in [\mathbf{a},\mathbf{b}]$ such that*

$$f(\mathbf{b}) - f(\mathbf{a}) = [\mathbf{D}f(\mathbf{c}_0)](\mathbf{b}-\mathbf{a}).$$

## 2.6 Newton's method

Newton's method is a way to approximately solve a nonlinear equation by making repeated guesses and linear approximations. Essentially, to find a solution to $f(\mathbf{a}) = 0$, we start with a guess $\mathbf{a}_0$. In each step, given guess $\mathbf{a}_n$, we make a new guess

$$\mathbf{a}_{n+1} = \mathbf{a}_n - [\mathbf{Df}(\mathbf{a}_n)]^{-1} f(\mathbf{a}_n).$$

Mathematica snippet:

```
f[{x_, y_}] = {Cos[x] + y - 1.1, x + Cos[x + y] - 0.9};
g[{x_, y_}] = {x, y} - Inverse[D[f[{x, y}], {{x, y}}]].f[{x, y}]
```

## 2.7 Inverse and implicit function theorems

Given a function from $\mathbb{R}^n \to \mathbb{R}^m$, is there a neighborhood $U$ in $\mathbb{R}^m$ with a function $g : U \to \mathbb{R}^n$ such that $\mathbf{f}\circ\mathbf{g} = \mathbf{g}\circ\mathbf{f} = \mathrm{id}$? Answer: if the derivative is invertible.

**Theorem 2.24** (Inverse function theorem). *If $\mathbf{f}$ is continuously differentiable, and its derivative is invertible at some point $\mathbf{x}_0$, then $\mathbf{f}$ is locally invertible, with differentiable inverse, in some neighborhood of the point $\mathbf{f}(\mathbf{x}_0)$.*

Given an equation $F(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) : \mathbb{R}^n \to \mathbb{R}^m$. Is there a neighborhood $U$ in $\mathbb{R}^n$ so that some of the $\mathbf{x}_i$ are functions of the others?

**Theorem 2.25** (Implicit function theorem, short form). *Let $U \subset \mathbb{R}^n$ be open and $\mathbf{c}$ a point in $U$. Let $\mathbf{F} : U \to \mathbb{R}^{n-k}$ be a $C^1$ mapping such that $\mathbf{F}(\mathbf{c}) = \mathbf{0}$ and $[\mathbf{DF}(\mathbf{c})]$ is onto. Then the system of linear equations $[\mathbf{DF}(\mathbf{c})](\mathbf{x}) = \mathbf{0}$ has $n-k$ pivotal variables and $k$ non-pivotal variables, and there exists a neighborhood of $\mathbf{c}$ in which $\mathbf{F} = \mathbf{0}$ implicitly defines the $n-k$ pivotal variables as a function $\mathbf{g}$ of the $k$ non-pivotal variables.*

**Example 2.26.** If we take $\mathbf{F}\binom{x}{y} = x^2 + y^2 - 1$, then $[\mathbf{DF}\binom{x}{y}] = \begin{bmatrix} 2x & 2y \end{bmatrix}$. Take a point, say, $\mathbf{c} = \binom{1}{0}$ which satisfy $\mathbf{F}(\mathbf{c}) = \mathbf{0}$. Then, $[\mathbf{DF}(\mathbf{c})] = \begin{bmatrix} 2 & 0 \end{bmatrix}$. As $x$ is pivotal and $y$ cannot be pivotal,

$x$ is a function of $y$ in a neighborhood of **c**. In this case it happens that $y$ cannot be a function of $x$ in any neighborhood of **c** too, but this is not always true.

On the other hand, if we take $\mathbf{c} = \binom{1/\sqrt{2}}{1/\sqrt{2}}$ then $[\mathbf{DF}(\mathbf{c})] = \begin{bmatrix}\sqrt{2} & \sqrt{2}\end{bmatrix}$ where both $x$ and $y$ can be pivotal, so $x$ and $y$ are functions of each other in some neighborhood of **c**.

**Exercise 2.27** (2.10.1 in book)**.** Does the inverse function theorem guarantee that the following functions are locally invertible with differentiable inverse?

a. $\mathbf{F}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x^2y\\-2x\\y^2\end{pmatrix}$ at $\begin{pmatrix}1\\1\end{pmatrix}$

b. $\mathbf{F}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x^2y\\-2x\end{pmatrix}$ at $\begin{pmatrix}1\\1\end{pmatrix}$

*Solution.* a. First of all, **F** consists entirely of polynomials, so **F** is continuously differentiable. At point $\binom{1}{1}$ the derivative of **F** is

$$\left[\mathbf{DF}\begin{pmatrix}1\\1\end{pmatrix}\right] = \begin{bmatrix}2 & 1\\-2 & 0\\0 & 2\end{bmatrix}$$

which is not even a square matrix, and hence not invertible. This applies to any functions from $\mathbb{R}^n \to \mathbb{R}^m$ with $n \neq m$.

b. We have

$$\left[\mathbf{DF}\begin{pmatrix}1\\1\end{pmatrix}\right] = \begin{bmatrix}2 & 1\\-2 & 0\end{bmatrix}$$

which is invertible.

**Exercise 2.28** (2.10.5 in book)**.** Apply the implicit function theorem to $y^2 + y + 3x + 1 = 0$, and determine where $y$ can be defined implicitly as a function of $x$.

*Solution.*

$$\left[\mathbf{DF}\begin{pmatrix}x\\y\end{pmatrix}\right] = \begin{bmatrix}3 & 2y+1\end{bmatrix}.$$

Clearly, for any point $\binom{x}{y}$ with $y \neq -1/2$, $y$ is pivotal, so **F** defines $y$ implicitly as a function of **F** in some neighborhood of each point **c** with $y_{\mathbf{c}} \neq -1/2$.

**Exercise 2.29** (2.10.9 in book)**.** Does the system of equations

$$x + y + \sin(xy) = a, \qquad \sin(x^2 + y) = 2a$$

have a solution for sufficiently small (but nonzero) $a$?

*Solution.* Consider $\mathbf{F}\begin{pmatrix}x\\y\\a\end{pmatrix} = \begin{pmatrix}x + y + \sin(xy) - a\\ \sin(x^2+y) - 2a\end{pmatrix}$. We can calculate

$$\left[\mathbf{DF}\begin{pmatrix}x\\y\\a\end{pmatrix}\right] = \begin{bmatrix}1 + y\cos(xy) & 1 + x\cos(xy) & -1\\ 2x\cos(x^2+y) & \cos(x^2+y) & -2\end{bmatrix},$$

so

$$\left[\mathbf{DF}\begin{pmatrix}0\\0\\0\end{pmatrix}\right] = \begin{bmatrix}1 & 1 & -1\\0 & 1 & -2\end{bmatrix}.$$

As $x$ and $y$ can both be made pivotal and $a$ non-pivotal, it follows that $x$ and $y$ are functions of $a$ in some neighborhood of $\mathbf{0}$.

**Exercise 2.30** (2.10.15a in book). Show that the mapping $\mathbf{F}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}e^x + e^y\\e^x + e^{-y}\end{pmatrix}$ is locally invertible at every point $\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2$.

*Solution.* First, as each component of $\mathbf{F}$ is formed by continuously differentiable fuctions, $\mathbf{F}$ is also continuously differentiable. At each point $\begin{pmatrix}x\\y\end{pmatrix}$, we have

$$\left[\mathbf{DF}\begin{pmatrix}x\\y\end{pmatrix}\right] = \begin{bmatrix}e^x & e^y\\e^x & -e^{-y}\end{bmatrix}$$

This gives

$$\det\left[\mathbf{DF}\begin{pmatrix}x\\y\end{pmatrix}\right] = -e^{x-y} - e^{x+y} < 0$$

so $[\mathbf{DF}(\mathbf{p})]$ is always invertible, so $\mathbf{F}$ is locally invertible at every point $\mathbf{p} \in \mathbb{R}^2$.

**Exercise 2.31** (2.31 in book).

a. True or false? The equation $\sin(xyz) = z$ expresses $x$ implicitly as a differentiable function of $y$ and $z$ near the point $(x, y, z) = (\pi/2, 1, 1)$.

b. True or false? The equation $\sin(xyz) = z$ expresses $z$ implicitly as a differentiable function of $x$ and $y$ near the point $(x, y, z) = (\pi/2, 1, 1)$.

*Solution.* Let $F : \mathbb{R}^3 \to \mathbb{R}$ be defined by $F(x, y, z) = \sin(xyz) - z$. We first calculate

$$[\mathbf{D}F(x, y, z)] = \begin{bmatrix}yz\cos(xyz) & xz\cos(xyz) & xy\cos(xyz) - 1\end{bmatrix}.$$

Therefore,

$$[\mathbf{D}F(\pi/2, 1, 1)] = \begin{bmatrix}0 & 0 & -1\end{bmatrix}$$

so $z$ is the only pivotal variable, so $z$ can be expressed as a differentiable function of $x$ and $y$, but $x$ cannot be expressed as a differentiable function of $y$ and $z$.

# 3 Manifolds

## 3.1 Definition and parametrization

- Idea: in BC Calc, the main object of study was *functions*. This is too restrictive as many objects that are smooth (have a best linear approximation at each point) are not graphs of functions globally. Example: circles, spirals, etc.

  Since the derivative only tells us about the local properties of a set of points, it suffices to ask that the set is a graph of a differentiable function in some neighborhood of every point.

**Definition 3.1.**
A subset $M \subset \mathbb{R}^n$ is a *smooth $k$-dimensional manifold* if locally it is the graph of a $C^1$ mapping $\mathbb{R}^k \to \mathbb{R}^{n-k}$. In other words, for any point $\mathbf{m} \in M$, there is a mapping $f : \mathbb{R}^k \to \mathbb{R}^{n-k}$ which coincides with $M$ on a neighborhood of the projection of $\mathbf{m}$ onto $\mathbb{R}^k$.

There are two important ways to define a manifold:

i) By equation, for example, $x^2 + y^2 - 1 = 0$.

ii) By parametrization: $j(t) = \binom{\cos t}{\sin t}$ for $t \in (0, 2\pi)$.

When does an equation define a smooth manifold, and when does a parametrization define a smooth manifold?

**Theorem 3.2.** *Let $U \subset \mathbb{R}^n$ be open, and let $F : U \to \mathbb{R}^{n-k}$ be a $C^1$ mapping. Let $M$ be a subset of $\mathbb{R}^n$ such that*

$$M \cap U = \{\mathbf{z} \in U \mid F(\mathbf{z}) = 0\}.$$

*If $[\mathbf{D}F(\mathbf{z})]$ is onto for every $\mathbf{z} \in M \cap U$, then $M \cap U$ is a smooth $k$-dimensional manifold embedded in $\mathbb{R}^n$. If every $\mathbf{z} \in M$ is in such a $U$, then $M$ is a $k$-dimensional manifold.*

*Conversely, if $M$ is a smooth $k$-dimensional manifold embedded in $\mathbb{R}^n$, then every point $\mathbf{z} \in M$ has a neighborhood $U \subset \mathbb{R}^n$ such that there exists a $C^1$ mapping $F : U \to \mathbb{R}^{n-k}$ with $[\mathbf{D}F(\mathbf{z})]$ onto and $M \cap U = \{\mathbf{y} \mid F(\mathbf{y}) = 0\}$.*

**Definition 3.3** (Parametrization of a manifold).
A *parametrization* of a $k$-dimensional manifold $M \subset \mathbb{R}^n$ is a mapping $\gamma : U \subset \mathbb{R}^k \to M$ satisfying the following conditions:

i. $U$ is open

ii. $\gamma$ is a $C^1$ bijection

iii. $[\mathbf{D}\gamma(\mathbf{u})]$ is injective for every $\mathbf{u} \in U$.

**Exercise 3.4** (3.1.11 in book). a. Find a parametrization for the union $X$ of the lines through the origin and a point of the parametrized curve $t \mapsto \begin{pmatrix} t \\ t^2 \\ t^2 \end{pmatrix}$.

b. Find an equation for the closure $\overline{X}$ of $X$. Is $\overline{X}$ exactly $X$?

c. Show that $\overline{X} - \{\mathbf{0}\}$ is a smooth surface.

d. Show that the map $\begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} r(1+\sin\theta) \\ r\cos\theta \\ r(1-\sin\theta) \end{pmatrix}$ is another parametrization of $\overline{X}$. [6]

e. Relate $\overline{X}$ to the set of noninvertible symmetric $2 \times 2$ matrices.

*Solution.* a. $\begin{pmatrix} t \\ u \end{pmatrix} \to \begin{pmatrix} ut \\ ut^2 \\ ut^3 \end{pmatrix}$

b. Equation of $\overline{X}$: $xz - y^2 = 0$. Here, $\overline{X} \neq X$ because $\begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$ is in $\overline{X}$ but not $X$.

[6] It seems that *parametrization of $\overline{X}$* actually means a parametrization of an open set with closure $\overline{X}$.

c. For $F\begin{pmatrix}x\\y\\z\end{pmatrix} = xz - y^2$, $\left[\mathbf{D}F\begin{pmatrix}x\\y\\z\end{pmatrix}\right] = \begin{bmatrix}z & -2y & x\end{bmatrix}$. Unless $\begin{pmatrix}x\\y\\z\end{pmatrix} = \mathbf{0}$, this is injective, so $\overline{X} - \{\mathbf{0}\}$ is a smooth surface.

d. It is easy to see that for any point $\begin{pmatrix}r(1+\sin\theta)\\r\cos\theta\\r(1-\sin\theta)\end{pmatrix}$. Consider a point $\begin{pmatrix}x\\y\\z\end{pmatrix}$ such that $xz - y^2 = 0$. If $x + z = 0$ it follows that $x = y = z = 0$ and this is given by $r = 0$. Else choose $\theta$ so that $\cos\theta = \frac{2y}{x+z}$, and $r = \frac{x+z}{2}$.

e. $xz - y^2 = \det\begin{vmatrix}x & y\\y & z\end{vmatrix}$.

## 3.2 Tangent spaces

As we saw previously, a $k$-manifold in $\mathbb{R}^n$ can be thought of as the solution to an equation $\mathbf{F}(\mathbf{x}) = 0$ or the image of a parametrization $\gamma : \mathbb{R}^k \to \mathbb{R}^n$. Since both $\mathbf{F}$ and $\gamma$ are assumed to be $C^1$, their derivatives at a point $\mathbf{x}$ give a local linear approximation of the manifold, and this is called the *tangent space to the manifold at point* $\mathbf{x}$.

The linear equivalent of $\{x \mid \mathbf{F}(\mathbf{x}) = \mathbf{0}\}$ is $\ker[\mathbf{DF}(\mathbf{x})]$, and the linear equivalent of $\operatorname{im}\gamma(\mathbf{u})$ is $\operatorname{im}[\mathbf{D}\gamma(\mathbf{u})]$.

**Example 3.5** (Unit circle equation). Our favorite example, the unit circle, can be written as the equation

$$F\begin{pmatrix}x\\y\end{pmatrix} = x^2 + y^2 - 1 \text{ which gives } \left[\mathbf{D}F\begin{pmatrix}x\\y\end{pmatrix}\right] = \begin{bmatrix}2x & 2y\end{bmatrix}.$$

Therefore, the tangent space is given by

$$\ker\begin{bmatrix}2x & 2y\end{bmatrix} = \left\{\begin{bmatrix}\dot{x}\\\dot{y}\end{bmatrix} \in \mathbb{R}^2 \,\middle|\, \begin{bmatrix}2x & 2y\end{bmatrix}\cdot\begin{bmatrix}\dot{x}\\\dot{y}\end{bmatrix} = 0\right\}.$$

Here the tangent space is based on coordinates $(\dot{x}, \dot{y})$ centered at the point $(x, y)$.

At $(x, y) = (1, 0)$, for example, our kernel equation reduces to $2\dot{x} = 0$, so the tangent space is $\dot{x} = 0$. This is the same as the tangent line at $(1, 0)$, which is $x - 1 = 0$, but with coordinates shifted so that $(1, 0)$ becomes the origin.

At $(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, the kernel equation is $\sqrt{2}\dot{x} + \sqrt{2}\dot{y} = 0$, so the tangent space is $\dot{x} + \dot{y} = 0$. Again, note that here, $(\dot{x}, \dot{y}) = \mathbf{0}$ when $(x, y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

In general, for a point $(x, y)$ on the circle, the tangent space is given by $x\dot{x} + y\dot{y} = 0$.

**Example 3.6** (Unit circle parametrization). The unit circle also admits a parametrization

$$\gamma(t) = \begin{pmatrix}\cos t\\\sin t\end{pmatrix},\ t \in (0, 2\pi) \text{ or } t \in (-\pi, \pi), \text{ which gives } [\mathbf{D}\gamma(t)] = \begin{bmatrix}-\sin t\\\cos t\end{bmatrix}.$$

Therefore, each $t \in (0, 2\pi)$ gives a point on the circle with the tangent space spanned by $\begin{bmatrix}-\sin t\\\cos t\end{bmatrix}$. Just like before, the tangent space is centered at $\gamma(t)$.

At $t = 0$, for example, this shows that the tangent space is spanned by $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as we expect, and at $t = \frac{pi}{4}$, the tangent space is spanned by $\begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$, which is also as expected.

**Exercise 3.7** (3.2.4 in book)**.** For each of the following functions $f$ and points $\begin{pmatrix} a \\ b \end{pmatrix}$, state whether there is a tangent plane to the graph of $f$ and point $\begin{pmatrix} a \\ b \\ f\binom{a}{b} \end{pmatrix}$. If there is such a tangent plane, find its equation, and compute the intersection of the tangent plane with the graph.[7]

*Solution.* Using parametrizations:

a. The parametrization $\gamma : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ x^2 - y^2 \end{pmatrix}$ has derivative $[\mathbf{D}\gamma(t)] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2x & -2y \end{bmatrix}$ which is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & -2 \end{bmatrix}$ at $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. The columns are linearly independent, so there is a tangent space centered at $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ spanned by the columns. This gives the tangent plane $\begin{pmatrix} 1+s \\ 1+t \\ 2s - 2t \end{pmatrix}$.

b. The parametrization $\gamma : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ \sqrt{x^2 + y^2} \end{pmatrix}$ has derivative $[\mathbf{D}\gamma(t)] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \end{bmatrix}$, which does not exist at the origin, so there are no tangent planes at the origin.

c. At $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, the derivative is given by $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$. Therefore the tangent plane is $\begin{pmatrix} 1+s \\ -1+t \\ \frac{1}{\sqrt{2}}(2-s-t) \end{pmatrix}$.

d. The parametrization $\gamma : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ \cos(x^2 + y) \end{pmatrix}$ has derivative

$$[\mathbf{D}\gamma(t)] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -2x\sin(x^2+y) & -\sin(x^2+y) \end{bmatrix},$$

which reduces to $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ at the origin, so the tangent plane is $\begin{pmatrix} s \\ t \\ 1 \end{pmatrix}$.

Using equations:

[7] We won't do this last part.

a. The equation $F\begin{pmatrix}x\\y\\z\end{pmatrix}: z - x^2 + y^2 = 0$ has derivative $\left[\mathbf{D}F\begin{pmatrix}x\\y\\z\end{pmatrix}\right] = \begin{bmatrix}-2x & 2y & 1\end{bmatrix}$. The tangent space at $\begin{pmatrix}1\\1\end{pmatrix}$ is given by $\ker\left[\mathbf{D}F\begin{pmatrix}1\\1\\0\end{pmatrix}\right] = \ker\begin{bmatrix}-2 & 2 & 1\end{bmatrix}$ so the tangent plane is given by $-2(x-1)+2(y-1)+z=0$, which is $-2x+2y+z=0$.

b. The equation $F\begin{pmatrix}x\\y\\z\end{pmatrix}: z^2 - x^2 - y^2 = 0$ has derivative $\left[\mathbf{D}F\begin{pmatrix}x\\y\\z\end{pmatrix}\right] = \begin{bmatrix}-2x & -2y & 2z\end{bmatrix}$. At the origin, $[\mathbf{D}F(\mathbf{0})]$ is not onto, so there are no tangent spaces.

c. At $\begin{pmatrix}1\\-1\end{pmatrix}$, the tangent space is $\ker\left[\mathbf{D}F\begin{pmatrix}1\\-1\\\sqrt{2}\end{pmatrix}\right] = \ker\begin{bmatrix}-2 & 2 & 2\sqrt{2}\end{bmatrix}$ so the tangent plane is given by $-2(x-1)+2(y+1)+2\sqrt{2}(z-\sqrt{2})=0$, which is $-x+y+\sqrt{2}z=-2$.

d. The equation $F\begin{pmatrix}x\\y\\z\end{pmatrix}: z - \cos(x^2+y) = 0$ has derivative

$$\left[\mathbf{D}F\begin{pmatrix}x\\y\\z\end{pmatrix}\right] = \begin{bmatrix}-2x\sin(x^2+y) & -\sin(x^2+y) & -1\end{bmatrix}.$$

At $\begin{pmatrix}0\\0\end{pmatrix}$, this tangent space is given by $\ker\left[\mathbf{D}F\begin{pmatrix}0\\0\\0\end{pmatrix}\right]$ ...

**Exercise 3.8** (3.2.6 in book). a. Show that the subset $X \subset \mathbb{R}^4$ where $x_1^2+x_2^2-x_3^2-x_4^2=0$ and $x_1+2x_2+3x_3+4x_4=4$ is a manifold in $\mathbb{R}^4$ in a neighborhood of the point $\mathbf{p}=(1,0,1,0)$.

b. What is the tangent space to $X$ at $\mathbf{p}$?

c. What pair of variables do the equations above *not* express as functions of the other two?

d. Is the entire set $X$ a manifold?

*Solution.* a. Let $\mathbf{F}(x_1,x_2,x_3,x_4) = \begin{pmatrix}x_1^2+x_2^2-x_3^2-x_4^2\\x_1+2x_2+3x_3+4x_4\end{pmatrix}$. It follows that

$$[\mathbf{DF}(x_1,x_2,x_3,x_4)] = \begin{bmatrix}2x_1 & 2x_2 & -2x_3 & -2x_4\\1 & 2 & 3 & 4\end{bmatrix}$$

so at $\mathbf{p}$,

$$[\mathbf{DF}(p)] = \begin{bmatrix}2 & 0 & -2 & 0\\1 & 2 & 3 & 4\end{bmatrix}.$$

As the first two columns are linearly independent, $[\mathbf{DF}(\mathbf{p})]$ has full rank, so $\mathbf{F}=\mathbf{0}$ describes a smooth manifold in a neighborhood of $\mathbf{p}$.

b. The tangent space to $X$ at $\mathbf{p}$ is ker $[\mathbf{DF}(p)]$. Row-reducing gives

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 2 \end{bmatrix}$$

so ker $[\mathbf{DF}(\mathbf{p})]$ is given by $\dot{x_1} + \dot{x_3} = 0$ and $\dot{x_2} + 2\dot{x_3} + 2\dot{x_4} = 0$ which is

$$\text{span}\left(\begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}\right).$$

c. In $[\mathbf{DF}(\mathbf{p})]\,\mathbf{x} = \mathbf{0}$, $x_2$ and $x_4$ cannot be pivotal at the same time.

d. For $X$ not to be a manifold in a neighborhood of a point $\mathbf{q} = (y_1, y_2, y_3, y_4)$, it follows that $[\mathbf{DF}(\mathbf{q})]$ has rank at most one. As no column of $[\mathbf{DF}(\mathbf{q})]$ can be zero, all columns of $[\mathbf{DF}(\mathbf{q})]$ must be scalar multiples of each other, and this gives $\mathbf{q} = k(1, 2, -3, -4)$ for some scalar $k$. However, the two equations for $X$ reduce to

$$-20k^2 = 0 \text{ and } -4k = 4$$

which cannot simultaneously hold. Therefore, for any point $\mathbf{q} \in X$, $X$ is a manifold in a neighborhood of $\mathbf{q}$, so $X$ is a manifold.

# 4 Optimizations

## 4.1 Critical points and Hessian matrix

Recall the Taylor polynomial at $x = 0$:

$$f(x) = f(0) + f'(0)x + f''(0)\frac{x^2}{2!} + f'''(0)\frac{x^3}{3!} + \cdots.$$

In general, the Taylor polynomial at $x = a$ is

$$f(x) = f(a) + f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2!} + f'''(a)\frac{(x-a)^3}{3!} + \cdots.$$

If $f'(a) = 0$, the behavior of $f$ near $a$ is determined by the quadratic term $f''(a)$.

For a function $f : \mathbb{R}^2 \to \mathbb{R}$, if $f$ is at least $C^2$, we can write, at $\mathbf{0}$:

$$f(\mathbf{p}) = f(\mathbf{0}) + D_x f(\mathbf{0})x + D_y(\mathbf{0})y + \frac{1}{2!}\left(D_{xx}f(\mathbf{0})x^2 + D_{xy}f(\mathbf{0})xy + D_{yx}f(\mathbf{0})yx + D_{yy}f(\mathbf{0})y^2\right) + \cdots$$

therefore

$$f(\mathbf{p}) \approx f(\mathbf{0}) + [\mathbf{D}f(\mathbf{0})]\mathbf{p} + \frac{1}{2}\mathbf{p}^T[\mathbf{H}f(\mathbf{0})]\mathbf{p}$$

where $[\mathbf{H}f(\mathbf{0})]$ is the *Hessian matrix*:

$$[\mathbf{H}f(\mathbf{0})] = \begin{bmatrix} D_{xx}f(\mathbf{0}) & D_{xy}f(\mathbf{0}) \\ D_{yx}f(\mathbf{0}) & D_{yy}f(\mathbf{0}) \end{bmatrix}.$$

It turns out that in higher-order partials, the order of differentiation does not matter:

$$D_{xy}f = D_{yx}f$$

whenever they exist, so the Hessian is symmetric.

If $[\mathbf{D}f(\mathbf{p})] = \begin{bmatrix} 0 & 0 \end{bmatrix}$, $\mathbf{p}$ is a critical point of $f$, and $[\mathbf{H}f(\mathbf{p})]$ determines the behavior of $f$ near $\mathbf{p}$. If $[\mathbf{H}f(\mathbf{p})]$ is positive definite[8], $\mathbf{p}$ is a local minimum; if it is negative definite, $\mathbf{p}$ is a local maximum. If it is semidefinite, then $\mathbf{p}$ is a local maximum or minimum, but they may not be unique. If it is neither, $\mathbf{p}$ is a saddle point.

**Exercise 4.1** (3.6.1 in book). a. Show that $f(x, y, z) = x^2 + xy + z^2 - \cos y$ has a critical point at the origin.

b. What kind of critical point does it have?

*Solution.* a. We have $[\mathbf{D}f(x, y, z)] = \begin{bmatrix} 2x + y & x + \sin y & 2z \end{bmatrix}$ so $[\mathbf{D}f(\mathbf{0}] = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$.

b. The Hessian matrix at a point $(x, y, z)$ is

$$[\mathbf{H}f(x, y, z)] = \begin{bmatrix} 2 & 1 & 0 \\ 1 & \cos y & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

At the origin it is

$$[\mathbf{H}f(\mathbf{0})] = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

which has eigenvalues $\left\{ 2, \frac{3 - \sqrt{5}}{2}, \frac{3 + \sqrt{5}}{2} \right\}$ all positive, so $[\mathbf{H}f(\mathbf{0})]$ is positive definite, so $\mathbf{0}$ is a local minimum of $f$.

## 4.2 Lagrange Multipliers

Main idea: find points on manifold such that the derivative is zero in all directions tangent to the manifold $\to$ this will give a critical point on the manifold.

**Theorem 4.2** (Lagrange Multipliers, 3.7.5 in book). *Let $U \subset \mathbb{R}^n$ be open, and let $\mathbf{F} : U \to \mathbb{R}^m$ be a $C^1$ mapping defining a manifold $X$, with $[\mathbf{DF}(\mathbf{x})]$ onto for every $\mathbf{x} \in X$. Let $f : U \to \mathbb{R}$ be a $C^1$ mapping. Then $\mathbf{a}$ is a critical point of $f$ restricted to $X$ if and only if there exist numbers $\lambda_1, \ldots, \lambda_m$, called* Lagrange multipliers *such that*

$$[\mathbf{D}f(\mathbf{a})] = \lambda_1[\mathbf{D}F_1(\mathbf{a})] + \cdots + \lambda_m[\mathbf{D}F_m(\mathbf{a})].$$

**Example 4.3** (3.7.6 in book). Suppose we want to maximize $f(x, y) = x+y$ on the ellipse $x^2+2y^2 = 1$. We have

$$F(x, y) = x^2 + 2y^2 - 1 \text{ and } [\mathbf{D}F(x, y)] = [2x, 4y]$$

while $[\mathbf{D}f(x, y)] = [1, 1]$. At a critical point, there will be $\lambda$ such that

$$[1, 1] = \lambda[2x, 4y].$$

[8] Recall: if both eigenvalues are positive, the quadratic form is positive definite; if one is positive and one is zero, it is positive semidefinite; if one is positive and one is negative, it is indefinite, etc.

Plugging this into the $F$ gives

$$\lambda = \pm\sqrt{\frac{3}{8}}, \text{ so } (x,y) = \pm\left(\sqrt{\frac{2}{3}}, \sqrt{\frac{1}{6}}\right)$$

At these critical points,

$$x + y = \pm\sqrt{\frac{3}{2}}$$

so the maximum of $f$ on $F$ is $\sqrt{\frac{3}{2}}$.

**Exercise 4.4** (3.7.6 in book). Find all the critical points of the function

$$f(x,y,z) = 2xy + 2yz - 2x^2 - 2y^2 - 2z^2$$

on the unit sphere in $\mathbb{R}^3$.

*Solution.* The constraint function $F$ is $F(x,y,z) = x^2+y^2+z^2-1$ so $[\mathbf{D}F(x,y,z)] = [2x, 2y, 2z]$. On the other hand, $[\mathbf{D}f(x,y,z)] = [2y-4x, 2x+2z-4y, 2y-4z]$. At a critical point, there must be a $\lambda$ such that

$$[2x, 2y, 2z] = \lambda[2y-4x, 2x+2z-4y, 2y-4z].$$

Let $\mathbf{v} = (x,y,z)$, and rewrite the equation as

$$\begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \mathbf{v} = \lambda \mathbf{v}$$

so $\mathbf{v}$ must be an eigenvector of the above matrix. We find the (unscaled) eigenvectors to be

$$\begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}.$$

After scaling to fit the constraint $F$, the eigenvectors, and thus the six critical points are

$$\pm\begin{pmatrix} \frac{1}{2} \\ \pm\frac{1}{\sqrt{2}} \\ \frac{1}{2} \end{pmatrix}, \pm\begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

At these critical points, the values of $f$ are $-2 \pm \sqrt{2}, -2$.

**Exercise 4.5** (3.7.8 in book). Find the maximum of the function $x^a e^{-x} y^b e^{-y}$ on the triangle $x \geqslant 0, y \geqslant 0, x + y \leqslant 1$, in terms of $a$ and $b$, for $a, b > 0$.

*Solution.* Let $f(x,y) = x^a e^{-x} y^b e^{-y}$. We have

$$[\mathbf{D}f(x,y)] = e^{-x-y} x^{a-1} y^{b-1} \begin{bmatrix} ay - xy & bx - xy \end{bmatrix}.$$

- Inside the triangle, the only critical point is when

  $$[\mathbf{D}f(x,y)] = \begin{bmatrix} 0 & 0 \end{bmatrix},$$

  and this happens when $(x,y) = (a,b)$ (if it lies inside the triangle).

- On the boundary $F_1 : x = 0$, which has $[\mathbf{D}F_1] = [1, 0]$, we want

$$e^{-x-y}x^{a-1}y^{b-1}\begin{bmatrix}ay - xy & bx - xy\end{bmatrix} = \lambda\begin{bmatrix}1 & 0\end{bmatrix}$$

  which happens when $(x, y) = (b, 0)$.
- On the boundary $F_2 : y = 0$, similarly, the only possible critical point is when $(x, y) = (a, 0)$.
- On the boundary $F_3 : x + y = 1$, which has $[\mathbf{D}F_3] = [1, 1]$, we want

$$e^{-x-y}x^{a-1}y^{b-1}\begin{bmatrix}ay - xy & bx - xy\end{bmatrix} = \lambda\begin{bmatrix}1 & 1\end{bmatrix}$$

  which gives $(x, y) = \left(\dfrac{a}{a+b}, \dfrac{b}{a+b}\right)$.
- There are also the three corner points $(0, 0), (1, 0), (0, 1)$.

Luckily, it is easy to see that any critical point $(x, y) = 0$ with $xy = 0$ gives $F(x, y) = 0$ which is definitely not the maximum. There are two remaining possible critical points: $(x, y) = (a, b), \left(\dfrac{a}{a+b}, \dfrac{b}{a+b}\right)$.

- If $a + b \geqslant 1$, $(a, b)$ is on/out of the boundaries of the triangle, so the maximum happens when $(x, y) = \left(\dfrac{a}{a+b}, \dfrac{b}{a+b}\right)$, and is $\dfrac{a^a b^b}{e(a+b)^{a+b}}$.
- If $a + b < 1$, there is an additional critical point $(x, y) = (a, b)$, where $f$ attains the value $e^{-(a+b)}a^a b^b$. We have

$$\frac{e^{-(a+b)}a^a b^b}{a^a b^b e^{-1}(a+b)^{-(a+b)}} = e^{1-s}s^s$$

  where $s = a + b$. It is not hard to show that $e^{1-s}s^s > 1$ when $s > 1$, so $f(a, b)$ is greater than $f\left(\frac{a}{a+b}, \frac{b}{a+b}\right)$. In this case, $f$ attains the maximum value $e^{-(a+b)}a^a b^b$ at $(x, y) = (a, b)$.

# 5 Riemann integrals

## 5.1 Introduction

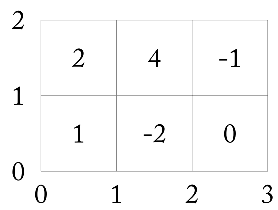
Goal: calculate a "total amount" from a "density function" defined over a region in $\mathbb{R}^n$.

In order for this to work, two stipulations will be made:

- the region is a bounded subset of $\mathbb{R}^n$, and
- the density function is a bounded function $f : \mathbb{R}^n \to \mathbb{R}$.

As in BC Calc, the idea is to refine a discrete problem by taking a limit.

**Example 5.1** (Rectangular regions). Take the region $U = [0, 3] \times [0, 2]$, and the density function $f : \mathbb{R}^2 \to R$ defined by



Then, the total amount is simply $2+4+(-1)+1+(-2)+0=4$. If we refine the grid so the number of grid squares go to infinity, we can consider this problem for a function $f(x,y)$ defined pointwise, for example $f(x,y)=2x+y^2$. If this is the density on $U$, what is the total amount now?

Let us, for example, hold the $x$-value constant at $x=x_0$. With the $x$-value held constant, we have now a function $h(y)=f(x_0,y)=2x_0+y^2$. We can then calculate the "total amount" on the vertical segment from $(x_0,0)$ to $(x_0,2)$ as

$$\int_0^2 2x_0+y^2\,dy = 4x_0+\frac{8}{3}.$$

Finally, the "total amount" over $U$ is obtained by integrating over the remaining variable: $x$. This is

$$\int_0^3 4x_0+\frac{8}{3}\,dx_0 = 26$$

Intuitively, the result should be the same if we go horizontally first:

$$\int_0^2\left(\int_0^3 2x+y^2\,dx\right)\,dy = \int_0^2\left(x^2+xy^2|_0^3\right)\,dy = \int_0^2 9+3y^2\,dy = 26$$

We can generalize that idea: intuitively[9],

$$\int_{[a,b]\times[c,d]} f(x,y)\left|d^2(x,y)\right| = \int_a^b\int_c^d f(x,y)\,dy\,dx = \int_c^d\int_a^b f(x,y)\,dx\,dy.$$

It turns out that this, as well as higher dimensional analogues, is true. This is called Fubini's theorem.

**Example 5.2.** We can use the above ideas without much complication to consider other types of regions. Let's consider the triangle $R$ with vertices $(0,0),(3,0),(0,2)$, which is bounded by $x=0, y=0, \frac{x}{3}+\frac{y}{2}=1$. Here we have

$$\int_R 2x+y^2\,d^2(x,y) = \int_0^3\left(\int_0^{2-\frac{2x}{3}} 2x+y^2\,dy\right)\,dx = \int_0^2\left(\int_0^{3-\frac{3x}{2}} 2x+y^2\,dx\right)\,dy = 8.$$

**Exercise 5.3** (Q31 in problem set). Evaluate

$$\int_0^1\int_{e^y}^e \frac{x}{\ln x}\,dx\,dy$$

*Solution.* Swapping the integrals give

$$\begin{aligned}\int_0^1\int_{e^y}^e \frac{x}{\ln x}\,dx\,dy &= \int_1^e\int_0^{\ln x}\frac{x}{\ln x}\,dy\,dx\\ &= \int_1^e x\,dx\\ &= \frac{e^2-1}{2}.\end{aligned}$$

[9]Here the absolute value sign denotes the fact that we are not considering orientation.

## 5.2 Riemann integrals

We will begin with several useful definitions.

**Definition 5.4** (Support).
The *support* $\text{Supp}(f)$ of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the closure of the set

$$\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \neq 0\}.$$

**Definition 5.5** ($M_A(f)$ and $m_A(f)$).
For an $f : \mathbb{R}^n \to \mathbb{R}$ and an $A \subset \mathbb{R}^n$, then $M_A(f)$ is the supremum of $f$ over $A$, and $m_A(f)$ is the infimum of $f$ over $A$.

**Definition 5.6** (Oscillation).
The *oscillation* of $f$ over $A$ is the difference between its supremum and infimum over $A$.

Riemann integrals in $\mathbb{R}^n$ can now be defined as follows:

- Partition $\mathbb{R}^n$ into cubes of side length 1 (where a cube is a region of the form $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$, and its volume is defined as $\prod_{i=1}^n |b_i - a_i|$.)
- A refinement of this partition is obtained by subdividing each cube into $2^n$ subcubes of side length $\frac{1}{2}$, using the midpoints of each side. (It turns out that with different partitions, the integrals, if exists, are all the same value).
- We can then continue refining this partition to get cubes or arbitrarily small side length and volume.

The $N^{\text{th}}$ *upper* and *lower sums* of $f$ over $A$, denoted $U_N(f)$ and $L_N(f)$, are now defined as the sum of (the supremums of $f$ over each cube in $A$ times the volume of the cube), and (the infimums ...)

It turns out that as $N$ increases, the upper sums are nonincreasing, and the lower sums are non-decreasing. This gives rise to the definitions of upper and lower integrals:

**Definition 5.7** (Upper and lower integrals; 4.1.10 in book).
We call

$$U(f) = \lim_{N \to \infty} U_N(f) \quad \text{and} \quad L(f) = \lim_{N \to \infty} L_N(f)$$

the *upper and lower integrals* of $f$.

If these are equal, then it makes sense to call the value the integral of $f$:

**Definition 5.8** (Integral; 4.1.12 in book).
A function $f : \mathbb{R}^n \to \mathbb{R}$, bounded with bounded support, is *integrable* if its upper and lower integrals are equal. Its integral is then

$$\int_{\mathbb{R}^n} f|d^n\mathbf{x}| = U(f) = L(f).$$

**Exercise 5.9** (4.1.10 in book). a. What are the upper and lower sums $U_1(f)$ and $L_1(f)$ for the function

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} x^2 + y^2 & \text{if } x, y \in (0, 1) \\ 0 & \text{otherwise,} \end{cases}$$

i.e. the upper and lower sums for the partition of $\mathbb{R}^2$ into squares of side length $1/2$?

b. Compute the integral of $f$ and show that it is between the upper and lower sum.

*Solution.* a. The lower sum $L_1(f)$ is

$$\frac{1}{4}\left(0+\frac{1}{4}+\frac{1}{4}+\frac{1}{2}\right)=\frac{1}{4}.$$

(Note that the infimum only considers the support of the function!) The upper sum $U_1(f)$ is

$$\frac{1}{4}\left(\frac{1}{2}+\frac{5}{4}+\frac{5}{4}+2\right)=\frac{5}{4}.$$

b. The integral is

$$\int_0^1\int_0^1 x^2+y^2\,dy\,dx=\int_0^1 x^2+\frac{1}{3}\,dx=\frac{2}{3}$$

which is between the lower sum and the upper sum.

**Proposition 5.10** (Properties of the Riemann integral).

1. The set of Riemann integrable functions is closed under addition and scalar multiplications (and so it is a vector space).
2. The Riemann integral commutes with addition and scalar multiplication (and so it is a linear transformation on the aforementioned vector space)
3. If $f(\mathbf{x})\leqslant g(\mathbf{x})$ for all $\mathbf{x}$ then

$$\int_{\mathbb{R}^n} f|d^n\mathbf{x}|\leqslant\int_{\mathbb{R}^n} g|d^n\mathbf{x}|.$$

**Definition 5.11** (Volume in $\mathbb{R}^n$).
When $\mathbf{1}_A$, the indicator function of $A$, is integrable, then the *$n$-dimensional volume* of $A$ is

$$\mathrm{vol}_n A=\int_{\mathbb{R}^n}\mathbf{1}_A|d^n\mathbf{x}|.$$

The set $A$ is then said to be *Riemann measurable*.

As expected, the volume is preserved under translations, and the volume of the union of two disjoint sets is the sum of the individual volumes.

Also, sets of volume 0 are of an interest because we can ignore them in integrals. A formal definition is given in the book[10], but roughly speaking, a set $X$ has volume 0 if it is possible to cover $X$ with cubes of arbitrarily small total volume.

## 5.3 Change of coordinates

Suppose we have an integral over some region $R$ which is ugly in the normal coordinate space ("$X$-space") but easy to describe over a "$U$-space" (instead of $X$-space). Let's say the relation between the $U$-space and the $X$-space is described by $\mathbf{x}=\gamma(\mathbf{u})$.

Zoom into a single point $(u,v)$ in the $U$-space. The box defined by $(u,v)$, $(u+\Delta u,v)$, $(u,v+\Delta v)$ maps to a distorted box defined by $\gamma(u,v),\gamma(u+\Delta u,v),\gamma(u,v+\Delta v)$.

The length of the spanning vectors of the distorted box are

$$\mathbf{a}=\gamma(u+\Delta u,v)-\gamma(u,v)=D_u\gamma(u,v)\Delta u+O(\Delta u^2)$$

$$\mathbf{b}=\gamma(u,v+\Delta v)-\gamma(u,v)=D_v\gamma(u,v)\Delta v+O(\Delta v^2)$$

[10] Proposition 4.1.23

so as $\Delta u, \Delta v \to 0$, the volume of the distorted box is

$$\text{vol}_2\mathcal{P} \sim |\det \mathbf{D}\gamma(u,v)|\Delta u \Delta v$$

so the *correction factor* for distortion under $\gamma$ is

$$|\det \mathbf{D}\gamma(u,v)|.$$

**Definition 5.12** (Polar coordinates).
Polar coordinates describe a point in a plane by its distance $r$ from the origin and its angle.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} =: \gamma \begin{pmatrix} r \\ \theta \end{pmatrix}.$$

We have

$$\mathbf{D}\gamma \begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}$$

so $\left| \det \mathbf{D}\gamma \begin{pmatrix} r \\ \theta \end{pmatrix} \right| = r.$

**Example 5.13.** We will try to find the area of the region $R$ bounded by $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Define $x = ar\cos\theta, y = br\sin\theta$. Then our boundaries are $r \in [0,1]$ and $\theta \in [0, 2\pi]$ and our correction factor is

$$\det \begin{vmatrix} a\cos\theta & -ar\sin\theta \\ b\sin\theta & br\cos\theta \end{vmatrix} = abr,$$

hence

$$\int_R 1|d^2(x,y)| = \int_0^1 \int_0^{2\pi} 1 \cdot abr \, d\theta \, dr = \int_0^1 2\pi abr \, dr = ab\pi.$$

**Exercise 5.14.** Compute $\displaystyle\int_{D_R} (x^2+y^2)\, dx\, dy$, where $D_R = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \middle| x^2 + y^2 \leqslant R^2 \right\}$.

*Solution.* Transforming to polar coordinates, we have

$$\int_{D_R} (x^2+y^2)\, dx\, dy = \int_0^R \int_0^{2\pi} r^2 \cdot r \, d\theta \, dr = \int_0^R 2\pi r^3 \, dr = \frac{1}{2}\pi R^4.$$

**Definition 5.15** (Spherical coordinates).
Spherical coordinates describe a point in space by its distance $\rho$ from the origin, its longitude $\theta$, and its latitude $\phi$:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \rho\cos\theta\cos\phi \\ \rho\sin\theta\cos\phi \\ \rho\sin\phi \end{pmatrix} =: S \begin{pmatrix} \rho \\ \theta \\ \phi \end{pmatrix}$$

We have $\left| \det \mathbf{D}S \begin{pmatrix} \rho \\ \theta \\ \phi \end{pmatrix} \right| = \rho^2 \cos\phi.$

**Definition 5.16** (Cylindrical coordinates).
Cylindrical coordinates describe a point in space by its $z$-coordinate and its polar coordinate on the $xy$-plane:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ z \end{pmatrix} =: C\begin{pmatrix} r \\ \theta \\ z \end{pmatrix}.$$

We have $\left|\det \mathbf{D}C\begin{pmatrix} r \\ \theta \\ z \end{pmatrix}\right| = r$.

# 6 Forms and Vector calculus

Quick note on this section: vector calculus is traditionally the study of vector fields in $\mathbb{R}^3$. Exterior calculus / forms, which comes later, generalizes vector calculus to all dimensions $\mathbb{R}^n$. However, for this note I think it makes more sense to go over the language of exterior calculus / forms first.

Previously we have done more and more general integrals:

↑ *Less general*

- Riemann integral in $\mathbb{R}^n$, over a box
- Riemann integral in $\mathbb{R}^n$, over "well-behaved" subsets of $\mathbb{R}^n$
- Riemann integral in $\mathbb{R}^n$, w.r.t. various coordinate systems

$$x = \Phi(u) \Rightarrow \int f(x)\,dx = \int f(\Phi(u))|\det \mathbf{D}\Phi(u)||d^n u|$$

- Next step: allow $\Phi$ to be any parametrization $\Phi : U \to R^n$ where $U \subseteq \mathbb{R}^k$ with $k < n$.

↓ *More general*

How do we accomplish this next step? We need to find an appropriate correction factor. The Jacobian $\mathbf{D}\Phi$ is not a square matrix anymore, so we can't simply use the determinant.

## 6.1 Forms

To find the correct correction factor, we need to generalize the concept of a determinant, and this comes in form of $k$-forms, first discussed in the last days of MA661.

**Definition 6.1.**
A *$k$-form* on $\mathbb{R}^n$ is an anti-symmetric multilinear function $(\mathbb{R}^n)^k$ taking in $k$ vectors and returning a number. The set of $k$-forms is donated as $A_c^k(\mathbb{R}^n)$.

**Example 6.2.** The 2-form $dx_1 \wedge dx_2$ takes in two vectors and outputs the determinant of the square matrix formed by the first and second entries of the vectors.

$$dx_1 \wedge dx_2 \left( \begin{bmatrix} 1 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 1 \\ 2 \end{bmatrix} \right) = \begin{vmatrix} 1 & 3 \\ 2 & -2 \end{vmatrix} = -8$$

**Exercise 6.3** (6.1.3 in book)**.** Compute the following numbers:

(a) $dx_1 \wedge dx_4 \left( \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ -1 \\ 2 \end{bmatrix} \right) = \begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = 0.$

(d) $dx_1 \wedge dx_2 \wedge dx_2$ (something) $= 0$ because $dx_2$ is repeated.

**Definition 6.4.**
$dx_1 \wedge dx_2$ is an example of *elementary $k$-forms*: those of the form

$$dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_k}, \quad i_1 < i_2 < \cdots < i_k.$$

**Example 6.5.** There are $2^4$ elementary $k$-forms in $\mathbb{R}^4$, corresponding to subsets of $\{1, 2, 3, 4\}$:

$$1, dx_1, \cdots, dx_4, dx_1 \wedge dx_2, \cdots, dx_3 \wedge dx_4, dx_1 \wedge dx_2 \wedge dx_3, \cdots, dx_2 \wedge dx_3 \wedge dx_4, dx_1 \wedge dx_2 \wedge dx_3 \wedge dx_4.$$

Since $k$-forms can be added together or multiplied by a scalar, they form a vector space. As the last example illustrates, the vector space of $k$-forms in $\mathbb{R}^n$ has dimension $\binom{n}{k}$.

Let $x$ be a point, and $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_k}$ be vectors in $\mathbb{R}^n$. Then, $P_x(\mathbf{v_1}, \ldots, \mathbf{v_k})$ is the parallelogram spanned by $(\mathbf{v_1}, \ldots, \mathbf{v_k})$ attached to point $x$.

**Definition 6.6.**
A *$k$-form field* in $\mathbb{R}^n$ is just a field of $k$-forms attached to each point of $\mathbb{R}^n$. In other words, it is like a $k$-form, but the scalars depend on $(x_1, \ldots, x_n)$. $k$-form fields take in anchored parallelograms and return a number.

**Example 6.7.** $\phi = 3dx_1 \wedge dx_3$ is a 2-form; $\omega = e^{x+y} dx \wedge dy$ is a 2-form field.

**Example 6.8.** $\cos(xz) dx \wedge dy$ is a 2-form field on $\mathbb{R}^3$. As an example of evaluation,

$$\cos(xz) dx \wedge dy \left( P_{\begin{pmatrix} 1 \\ 2 \\ \pi \end{pmatrix}} \left( \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \right) \right) = \cos(1 \cdot \pi) \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = -2.$$

We used the *wedge* $\wedge$ symbol in our notation for $k$-forms. This represents a wedge product, which has the following formal definition, which I guess comes from the need to make sure that $\wedge$ preserves antisymmetry as well as multilinearity.

**Definition 6.9.**
The *wedge product* of the forms $\phi \in A_c^k(\mathbb{R}^n)$ and $\omega \in A_c^\ell(\mathbb{R}^n)$ is the element $\phi \wedge \omega \in A_c^{k+\ell}(\mathbb{R}^n)$ defined by

$$(\phi \wedge \omega)(\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_{k+l}}) = \sum_{\substack{\sigma \text{ permutes } \{1,2,\ldots,k+\ell\} \\ \sigma(1) < \cdots < \sigma(k) \\ \sigma(k+1) < \cdots < \sigma(k+\ell)}} \operatorname{sgn}(\sigma) \phi(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(k)}) \omega(\mathbf{v}_{\sigma(k+1)}, \ldots, \mathbf{v}_{\sigma(k+\ell)}).$$

**Example 6.10.** The wedge product of $\phi \in A_c^2(\mathbb{R}^n)$ and $\omega \in A_c(\mathbb{R}^n)$ is

$$\phi \wedge \omega(\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}) = \phi(\mathbf{v_1}, \mathbf{v_2})\omega(\mathbf{v_3}) - \phi(\mathbf{v_1}, \mathbf{v_3})\omega(\mathbf{v_2}) + \phi(\mathbf{v_2}, \mathbf{v_3})\omega(\mathbf{v_1}).$$

**Exercise 6.11** (6.1.11 in book). Let $\phi$ and $\psi$ be 2-forms. Write out $\phi \wedge \psi(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)$.

*Answer:*

$$\begin{aligned}\phi \wedge \psi(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4) &= \phi(\mathbf{v}_1, \mathbf{v}_2)\psi(\mathbf{v}_3, \mathbf{v}_4) - \phi(\mathbf{v}_1, \mathbf{v}_3)\psi(\mathbf{v}_2, \mathbf{v}_4) + \phi(\mathbf{v}_1, \mathbf{v}_4)\psi(\mathbf{v}_2, \mathbf{v}_3) \\ &\quad + \phi(\mathbf{v}_2, \mathbf{v}_3)\psi(\mathbf{v}_1, \mathbf{v}_4) - \phi(\mathbf{v}_2, \mathbf{v}_4)\psi(\mathbf{v}_1, \mathbf{v}_3) + \phi(\mathbf{v}_3, \mathbf{v}_4)\psi(\mathbf{v}_1, \mathbf{v}_1)\end{aligned}$$

## 6.2 Integrals over parametrized manifolds

Using the language of forms, we can now describe integrals over parametrized manifolds.

**Definition 6.12** (Integrating a $k$-form field, 6.2.1 in book).
Let $U \subset \mathbb{R}^k$ be a bounded open set with boundary of volume 0. Let $V \subset \mathbb{R}^n$ be open, and let $[\gamma(U)]$ be a parametrized domain in $V$. Let $\varphi$ be a $k$-form field on $V$. Then the *integral of $\varphi$ over* $[\gamma(U)]$ is

$$\int_{[\gamma(U)]} \varphi := \int_U \varphi \left( P_{\gamma(U)} \left( \mathbf{D}_1\gamma(\mathbf{u}), \ldots, \mathbf{D}_k\gamma(\mathbf{u}) \right) \right)$$

**Example 6.13** (6.2.4 in book). We will integrate $dx \wedge dy + y\, dx \wedge dz$ over the parametrized domain

$$\gamma\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} s+t \\ s^2 \\ t^2 \end{pmatrix}, S = \left\{ \begin{pmatrix} s \\ t \end{pmatrix} \middle| 0 \leqslant s \leqslant 1, 0 \leqslant t \leqslant 1 \right\}.$$

We find

$$\int_{[\gamma(S)]} dx \wedge dy + y\, dx \wedge dz = \int_0^1 \int_0^1 (dx \wedge dy + y\, dx \wedge dz) \left( P_{\begin{pmatrix} s+t \\ s^2 \\ t^2 \end{pmatrix}} \left( \begin{bmatrix} 1 \\ 2s \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2t \end{bmatrix} \right) \right) ds\, dt$$

$$= \int_0^1 \int_0^1 \left( \det \begin{bmatrix} 1 & 1 \\ 2s & 0 \end{bmatrix} + s^2 \det \begin{bmatrix} 1 & 1 \\ 0 & 2t \end{bmatrix} \right) ds\, dt = \int_0^1 \int_0^1 (-2s + 2s^2 t)\, ds\, dt = -\frac{2}{3}.$$

**Exercise 6.14** (6.2.1 in book). Set up each of the following integrals of form fields over parametrized domains as an ordinary multiple integral.

a. $\int_{[\gamma(I)]} x\, dy + y\, dz$, where $I = [-1, 1]$ and $\gamma(t) = (\sin t, \cos t, t)$.

b. $\int_{[\gamma(U)]} x_1\, dx_2 \wedge dx_3 + x_2\, dx_3 \wedge dx_4$, where $U = \left\{ \binom{u}{v} \middle| 0 \leqslant u, v; u + v \leqslant 2 \right\}$ and $\gamma\binom{u}{v} = (uv, u^2 + v^2, u - v, \ln(u + v + 1))$.

*Solution.* a.

$$\int_{[\gamma(I)]} x\, dy + y\, dz = \int_0^1 (x\, dy + y\, dz) \left( P_{\begin{pmatrix} \sin t \\ \cos t \\ t \end{pmatrix}} \left( \begin{bmatrix} \cos t \\ -\sin t \\ 1 \end{bmatrix} \right) \right) |dt|$$

$$= \int_0^1 (-\sin^2 t + \cos t) |dt|$$

b.

$$\int_{[\gamma(U)]} x_1\, dx_2 \wedge dx_3 + x_2\, dx_3 \wedge dx_4$$

$$= \int_U (x_1\, dx_2 \wedge dx_3 + x_2\, dx_3 \wedge dx_4) \left( P_{\begin{pmatrix} uv \\ u^2+v^2 \\ u-v \\ \ln(u+v+1) \end{pmatrix}} \left( \begin{bmatrix} v \\ 2u \\ 1 \\ \frac{1}{u+v+1} \end{bmatrix}, \begin{bmatrix} u \\ 2v \\ -1 \\ \frac{1}{u+v+1} \end{bmatrix} \right) \right) |du\, dv|$$

$$= \int_0^2 \int_0^{2-u} uv(-2u-2v) + (u^2+v^2)\frac{2}{u+v+1}\, dv\, du$$

**Exercise 6.15** (6.2.2 in book). Repeat 6.2.1 for $\int_{[\gamma(U)]} x\, dy \wedge dz$ where $U = [-1,1] \times [-1,1]$ and $\gamma\binom{u}{v} = (u^2, u+v, v^3)$

*Solution.*

$$\int_{[\gamma(U)]} x\, dy \wedge dz = \int_{-1}^1 \int_{-1}^1 (x\, dy \wedge dz) \left( P_{\begin{pmatrix} u^2 \\ u+v \\ v^3 \end{pmatrix}} \left( \begin{bmatrix} 2u \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 3v^2 \end{bmatrix} \right) \right) |du\, dv|$$

$$= \int_{-1}^1 \int_{-1}^1 3u^2v^2 |du\, dv|$$

**Exercise 6.16** (6.2.4 in book). Let $z_1 = x_1 + iy_1, z_2 = x_2 + iy_2$ be coordinates in $\mathbb{C}^2$. Let $S \subset \mathbb{C}$ be the square $\{z = x + iy \mid |x| \leqslant 1, |y| \leqslant 1\}$, and define $\gamma : S \to \mathbb{C}^2$ by

$$\gamma : z \mapsto \begin{pmatrix} e^z \\ e^{-z} \end{pmatrix}, z = x + iy, |x| \leqslant 1, |y| \leqslant 1.$$

What is $\int_{[\gamma(S)]} dx_1 \wedge dy_1 + dy_1 \wedge dx_2 + dx_2 \wedge dy_2$?

*Solution.* First we note that $e^z = e^x(\cos y + i \sin y)$ and $e^{-z} = e^{-x}(\cos y - i \sin y)$ so $\gamma$ sends $\binom{x}{y}$ to $\begin{pmatrix} e^x \cos y \\ e^x \sin y \\ e^{-x} \cos y \\ -e^{-x} \sin y \end{pmatrix}$. Therefore,

$$\int_{[\gamma(S)]} dx_1 \wedge dy_1 + dy_1 \wedge dx_2 + dx_2 \wedge dy_2$$

$$= \int_{-1}^1 \int_{-1}^1 dx_1 \wedge dy_1 + dy_1 \wedge dx_2 + dx_2 \wedge dy_2 \left( P_{\begin{pmatrix} e^x \cos y \\ e^x \sin y \\ e^{-x} \cos y \\ -e^{-x} \sin y \end{pmatrix}} \left( \begin{bmatrix} e^x \cos y \\ e^x \sin y \\ -e^{-x} \cos y \\ e^{-x} \sin y \end{bmatrix}, \begin{bmatrix} -e^x \sin y \\ e^x \cos y \\ -e^{-x} \sin y \\ -e^{-x} \cos y \end{bmatrix} \right) \right) |dx\, dy|$$

$$= \int_{-1}^{1}\int_{-1}^{1} \begin{vmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{vmatrix} + \begin{vmatrix} e^x \sin y & e^x \cos y \\ -e^{-x} \cos y & -e^{-x} \sin y \end{vmatrix} + \begin{vmatrix} -e^{-x} \cos y & -e^{-x} \sin y \\ e^{-x} \sin y & -e^{-x} \cos y \end{vmatrix} |dx\, dy|$$

$$= \int_{-1}^{1}\int_{-1}^{1} e^{2x} - \sin^2 y + \cos^2 y + e^{-2x} |dx\, dy|$$

$$= \int_{-1}^{1}\int_{-1}^{1} e^{2x} + e^{-2x} + \cos 2y\, |dx\, dy|$$

$$= \int_{-1}^{1} 2e^{2x} + 2e^{-2x}\, dx + \int_{-1}^{1} 2\cos 2y\, dy$$

$$= \left(e^{2x} - e^{-2x}\right)_{x=-1}^{1} + \left(\sin 2y\right)_{y=-1}^{1} = 2e^2 - 2e^{-2} + 2\sin 2$$

**Exercise 6.17** (6.2.5 in book). Let $S : \{z \in \mathbb{C} \mid |z| < 1\}$ be the interior of the unit circle in the complex plane. Define $\gamma : S \to \mathbb{C}^3$ by $\gamma : z \mapsto \begin{pmatrix} z \\ z^2 \\ z^3 \end{pmatrix}$. What is $\int_{[\gamma(S)]} dx_1 \wedge dy_1 + dx_2 \wedge dy_2 + dx_3 \wedge dy_3$?

## 6.3 Orientation

We have just calculated integrals without caring for the sign, and everything seemed fine. Why, then, should we care about signs? The answer is that the integrals we have calculated implicitly depended on the parametrization, and especially the boundaries of the domain.

In order to integrate forms over manifolds and *get a consistent result no matter the parametrization*, we need to define orientation. Now what is an orientation?

An orientation of a manifold is simply a way to distinguish between the "two sides" of a manifold: inside or outside of a sphere, the $+z$ or $-z$ side of the $xy$-plane, or (insert more examples).

Obviously this relies on the manifold actually having two sides. For manifolds with one side, such as the Mobius strip, we just ... don't care about them. They are not orientable.

Some technical definitions follow. Honestly, feel free to skip this part.

**Definition 6.18** (Orientation of vector spaces).
Let $V$ be a finite-dimensional real vector space, and let $\mathcal{B}_V$ be the set of bases of $V$. An *orientation* of $V$ is a map $\Omega : \mathcal{B}_V \to \{+1, -1\}$ such that if $B$ and $B'$ are two bases with change of basis matrix $[\Phi_{B' \to B}]$, then

$$\Omega(B') = \operatorname{sgn}(\det[\Phi_{B' \to B}])\, \Omega(B).$$

A basis $B \in \mathcal{B}_V$ is called *direct* if $\Omega(B) = +1$; it is called *indirect* otherwise.

Conventionally, in $\mathbb{R}^n$, the orientation should be defined so that the standard basis is direct. As a result, $\Phi(B) = \operatorname{sgn}(\det[B])$. This is called the *standard orientation*

Reflection changes orientation, but rotation and translation does not. (Think of a mirror).

Orienting manifolds is a more generalized version of orienting vector spaces: we need to select an orientation for every tangent space, and it must *vary continuously*.

**Definition 6.19** (Orientation of manifolds).
Let $M$ be a manifold. For a point $\mathbf{x} \in M$, let $\mathcal{B}_{\mathbf{x}}(M)$ denote the set of bases of the tangent space $T_{\mathbf{x}}M$ to $M$ at $\mathbf{x}$. Let $\mathcal{B}(M)$ be the set of all "vectors" of the form $(\mathbf{x}, \mathbf{v}_1, \ldots, \mathbf{v}_n)$ where $(\mathbf{v}_1, \ldots, \mathbf{v}_n) \in \mathcal{B}_{\mathbf{x}}(M)$.

Then, an *orientation* of a $k$-dimensional manifold $M \subset \mathbb{R}^n$ is a continuous map[11] from $\mathcal{B}(M) \to \{+1, -1\}$ such that the restrictions to $\mathcal{B}_{\mathbf{x}}(M)$ determines an orientation of $T_{\mathbf{x}}(M)$.

That doesn't really help in understanding how to actually orient a manifold; here are some specific examples.

**Proposition 6.20** (6.3.4 in book).

1. **Points.** An orientation for a 0-dimensional manifold in $\mathbb{R}^n$, i.e. a discrete set of points, is simply an assignment of either $+1$ or $-1$ to each of the points.
2. **Open subsets of $\mathbb{R}^n$.** An open subset in $\mathbb{R}^n$ simply carries the standard orientation of $\mathbb{R}^n$.
3. **Curves.** Let $C \subset \mathbb{R}^n$ be a smooth curve. A non-vanishing tangent vector field $\mathbf{t}$ that varies continuously with $\mathbf{x}$ defines an orientation

$$\Omega_{\mathbf{x}}^{\mathbf{t}}(\mathbf{v}) := \operatorname{sgn}(\mathbf{t}(\mathbf{x}) \cdot \mathbf{v}).$$

4. **Surfaces in $\mathbb{R}^3$.** For each point $\mathbf{x}$ on the surface, add a transverse vector $\mathbf{n}(\mathbf{x})$ that is continuous as $\mathbf{x}$ varies and is not in the tangent space. Then we can define an orientation $\Omega^{\mathbf{n}}$ of $S$ by

$$\Omega^{\mathbf{n}}(\mathbf{v}_1, \mathbf{v}_2) := \operatorname{sgn}(\det[\mathbf{n}(\mathbf{x}), \mathbf{v}_1, \mathbf{v}_2]).$$

**Exercise 6.21.** Find a vector field that orients the curve given by $x + x^2 + y^2 = 2$.

*Solution.* (Using parametrization) Rewrite the equation as $\left(x + \frac{1}{2}\right)^2 + y^2 = \frac{9}{4}$, so we have the parametrization

$$\Gamma : t \mapsto \begin{pmatrix} -\frac{1}{2} + \frac{3}{2}\cos t \\ \frac{3}{2}\sin t \end{pmatrix}, t \in [0, 2\pi]$$

The tangent vectors are given by the Jacobian

$$\mathbf{D}\Gamma : \begin{bmatrix} -\frac{3}{2}\sin t \\ \frac{3}{2}\cos t \end{bmatrix}$$

and since we are always going counterclockwise, these vectors point in the right direction. Therefore the vector field $F(t) := \begin{bmatrix} -\sin t \\ \cos t \end{bmatrix}$ works.

(Using equation) The kernel of the Jacobian gives

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 2y \\ -(2x+1) \end{bmatrix}$$

as another possible vector field.

**Example 6.22** (6.3.12 in book). Consider the manifold $M \subset \mathbb{R}^4$ of equation $x_1^2 + x_2^2 + x_3^2 - x_4 = 0$. Find a basis for the tangent space to $M$ at the point $\mathbf{p} = (1, 0, 0, 1)$ that is direct for the orientation ... ?

First we find

$$\mathbf{D}T(\mathbf{p}) = \begin{bmatrix} 2x_1 & 2x_2 & 2x_3 & -1 \end{bmatrix}_{\mathbf{p}} = \begin{bmatrix} 2 & 0 & 0 & -1 \end{bmatrix}$$

[11] Huh. Continuous map to a discrete set?!?

The tangent space $T_{\mathbf{p}}M$ is then given by the kernel of $\mathbf{D}T(\mathbf{p})$, which is exactly set of vectors $\mathbf{x}$ where $\begin{bmatrix} 2 \\ 0 \\ 0 \\ -1 \end{bmatrix} \cdot \mathbf{x} = 0$. As a result, $\begin{bmatrix} 2 \\ 0 \\ 0 \\ -1 \end{bmatrix}$ is a transverse vector and can be used for orientation: a basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ of $T_{\mathbf{p}}M$ has the same sign as $\det \left( \begin{bmatrix} 2 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \right)$.

Consider the basis $\left( \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix} \right)$ of $T_{\mathbf{p}}M$. We have

$$\det \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 2 \end{bmatrix} = 5 > 0$$

so this basis works.

**Example 6.23** (Using parametrization). The manifold $M$ from the previous question can also be described by the parametrization

$$\gamma \begin{pmatrix} s \\ t \\ u \end{pmatrix} := \begin{pmatrix} s \\ t \\ u \\ s^2 + t^2 + u^2 \end{pmatrix}$$

so

$$\mathbf{D}\gamma \begin{pmatrix} s \\ t \\ u \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2s & 2t & 2u \end{bmatrix}.$$

At $\begin{pmatrix} s \\ t \\ u \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, this generates the basis $\left( \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix} \right)$. The rest follows the previous example.

**Exercise 6.24** (6.4.1 in book). If the cone $M$ of equation $f \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x^2 + y^2 - z^2 = 0$ is oriented by $\nabla f$[12], does the parametrization $\gamma : \begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ r \end{pmatrix}$ preserve orientation?

[12] which is defined as simply $\mathbf{D}f^T$

*Solution.* At a point $\mathbf{p} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, we find $\nabla f(\mathbf{p}) = \begin{bmatrix} 2x \\ 2y \\ -2z \end{bmatrix}$. We also find $\mathbf{D}\gamma \begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \\ 1 & 0 \end{bmatrix}$.

Assuming $M$ is orientable, it suffices to check one point on $M$, so we will choose $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, represented by $\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. At this point, we have

$$\det \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = -4 < 0,$$

so $\gamma$ is *not* orientation preserving. If we want to find an orientation preserving parametrization we can, say, switch the first and second rows of $\gamma$.

**Exercise 6.25** (6.4.4 in book). What is the integral $\int_S x_3\, dx_1 \wedge dx_2 \wedge dx_4$, where $S$ is the part of the three-dimensional manifold of equation

$$x_4 = x_1 x_2 x_3 \quad \text{where } 0 \leqslant x_1, x_2, x_3 \leqslant 1,$$

oriented by $\Omega = \operatorname{sgn} dx_1 \wedge dx_2 \wedge dx_3$?

*Solution.* Consider the standard parametrization $\gamma : \begin{pmatrix} u \\ v \\ w \end{pmatrix} \mapsto \begin{pmatrix} u \\ v \\ w \\ uvw \end{pmatrix}$. We compute

$$\mathbf{D}\gamma \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ vw & wu & uv \end{bmatrix},$$

so

$$\Omega \left( \begin{bmatrix} 1 \\ 0 \\ 0 \\ vw \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ wu \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ uv \end{bmatrix} \right) = \operatorname{sgn} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = +1$$

which means $\gamma$ preserves the correct orientation. Now

$$\int_S x_3\, dx_1 \wedge dx_2 \wedge dx_4 = \int_0^1 \int_0^1 \int_0^1 w \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ vw & wu & uv \end{vmatrix} dw\, dv\, du$$

$$= \int_0^1 \int_0^1 \int_0^1 uvw\, dw\, dv\, du = \frac{1}{8}.$$

**Exercise 6.26** (6.4.5 in book). Let $z_1 = x_1 + iy_1, z_2 = x_2 + iy_2$ be coordinates in $\mathbb{C}^2$. Compute the integral of $dx_1 \wedge dy_1 + dy_1 \wedge dx_2$ over the part of the locus of equation $z_2 = z_1^k$ where $|z_1| < 1$, oriented by $\Omega = \operatorname{sgn} dx_1 \wedge dx_2$.

*Solution.* Let $S$ denote the region in question. Consider the parametrization $\gamma : \begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ r^k \cos k\theta \\ r^k \sin k\theta \end{pmatrix}$.

We compute $\mathbf{D}\gamma \begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \\ kr^{k-1}\cos k\theta & -kr^k \sin k\theta \\ kr^{k-1}\sin k\theta & kr^k\cos k\theta \end{bmatrix}$, so

$$\Omega\left(\mathbf{D}\gamma\begin{pmatrix} r \\ \theta \end{pmatrix}\right) = r > 0$$

hence $\gamma$ is orientation preserving. Therefore

$$\begin{aligned}
\int_S dx_1 \wedge dy_1 + dy_1 \wedge dx_2 &= \int_0^1 \int_0^{2\pi} \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} + \begin{vmatrix} \sin\theta & r\cos\theta \\ kr^{k-1}\cos k\theta & -kr^k\sin k\theta \end{vmatrix} \, d\theta \, dr \\
&= \int_0^1 \int_0^{2\pi} r - kr^k(\sin\theta\sin k\theta + \cos\theta\cos k\theta) \, d\theta \, dr \\
&= \int_0^1 \int_0^{2\pi} r - kr^k \cos(k-1)\theta \, d\theta \, dr \\
&= \int_0^1 2\pi r \, dr - \int_0^1 \frac{kr^k}{k-1}\sin(2\pi(k-1)) \, dr \\
&= \pi - \frac{k}{k^2-1}\sin(2\pi k) = \pi \text{ for } k \in \mathbb{Z}.
\end{aligned}$$

### Digression: Complex functions

Assume $z, a, c \in \mathbb{C}, r \in \mathbb{R}$.

- $z \mapsto z + a$ translates the complex plane by the vector $a$.
- $z \mapsto cz$ is a rotation by $\arg c$ and scaled by $|c|$.
- $z \mapsto \overline{z}$ is a reflection over the real axis.
- $z \mapsto e^z$ scales $1$ by $e^{\Re(z)}$ and then rotates by $\Im(z)$.

From $e^{iz} = \cos z + i\sin z$ and $e^{-iz} = \cos z - i \sin z$, we have

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} \text{ and } \sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

Therefore, $\cos iz = \cosh z$ and $\sin iz = -i\sinh z$.

## 6.4 Boundaries on manifolds

I guess mathematicians are naturally curious, and hence for some reason they decided to see if there is any relation between an integral on a boundary of a (piece of) manifold, and an integral on the piece of manifold itself. Turns out there is!

**Definition 6.27** (Boundary of a subset of a manifold, 6.6.1 in book).
Let $M \subset \mathbb{R}^n$ be a $k$-dimensional manifold, and $X \subset M$ a subset. The *boundary of $X$ in $M$*, written $\partial_M X$, is the set of points $\mathbf{x} \in M$ such that every neighborhood of $\mathbf{x}$ contains points of $X$ and points of $M - X$.

As sets of points with volume 0 do not matter in an integral, it suffices to consider smooth points.

**Definition 6.28.**
A *smooth point* of a boundary of a manifold is a point on a section of the boundary that can be described by a $C^1$ function.

**Exercise 6.29** (6.6.1 in book). Verify that the following inequalities describe a piece-with-boundary of $\mathbb{R}^3$

a. $xyz \leqslant 1, \quad x^2+y^2+z^2 \leqslant 4$

b. $xyz \leqslant 1, \quad x^2+y^2+z^2 \leqslant 4, \quad x+y+z \geqslant 0$

Boundary orientation:

- Let $X$ be a piece-with-boundary of manifold $M$.
- Let $\partial_M X$ be the boundary of $X$ in $M$ and $\mathbf{p} \in \partial_M X$.
- An orientation of $T_{\mathbf{p}}\partial_M X$ is given by an outward-pointing vector $\mathbf{v}_{\text{out}}$ at $\mathbf{p}$: if $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1})$ is the basis of the boundary tangent space, then $(\mathbf{v}_{\text{out}}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{k-1})$ should be a direct basis of $T_p\partial M$.

**Exercise 6.30** (6.6.5 in book). Consider the region $X = P \cap B \in \mathbb{R}^3$, where $P$ is the plane of equation $x+y+z=0$ and $B$ is the ball $x^2+y^2+z^2 \leqslant 1$. Orient $P$ by the normal $\mathbf{N} = \begin{bmatrix}1\\1\\1\end{bmatrix}$ and orient the sphere $x^2+y^2+z^2=1$ by the outward-pointing normal.

a. Which of sgn $dx \wedge dy$, sgn $dx \wedge dz$, sgn $dy \wedge dz$ give the same orientation of $P$ as $\mathbf{N}$?

b. Show that $X$ is a piece-with-boundary of $P$ and that the mapping

$$\gamma : t \mapsto \begin{pmatrix} \frac{\cos t}{\sqrt{2}} - \frac{\sin t}{\sqrt{6}} \\ -\frac{\cos t}{\sqrt{2}} - \frac{\sin t}{\sqrt{6}} \\ \frac{2\sin t}{\sqrt{6}} \end{pmatrix}$$

for $0 \leqslant t \leqslant 2\pi$, is a parametrization of $\partial X$.

c. Is the parametrization in part b. compatible with the boundary orientation of $\partial X$?

d. Do any of sgn $dx$, sgn $dy$, sgn $dz$ define the orientation of $\partial X$ at every point?

e. Do any of sgn $x\,dy - y\,dx$, sgn $x\,dz - z\,dx$, sgn $y\,dz - z\,dy$ define the orientation of $\partial X$ at every point?

*Solution.* a. Consider the basis $B_P = \left( \begin{bmatrix}1\\-1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\-1\end{bmatrix} \right)$. As $\begin{vmatrix}1 & 1 & 1\\1 & -1 & 0\\1 & 0 & -1\end{vmatrix} = 3 > 0$, this basis gives the direct orientation. We can check that

$$(dx \wedge dy)B_P = 1, (dx \wedge dz)B_P = -1, (dy \wedge dz)B_P = 1,$$

so only $dx \wedge dy$ and $dy \wedge dz$ give the same orientation as $\mathbf{N}$.

b. Intuitively, $X$ is a closed disk in $P$. (We'll skip the rigorous part for now.) Its boundary is given by the solution to $x+y+z=0$ and $x^2+y^2+z^2=1$, so a point $(x,y,z)$ in the boundary satisfies

$$(x-y)^2+3z^2 = 2(x^2+y^2+z^2)-(x+y+z)(x+y-z)=2,$$

so we can write $z=\sqrt{\frac{2}{3}}\sin t, x-y=\sqrt{2}\cos t$ for some $t$. Using this and $x+y+z=0$ gives the desired parametrization, which is easy to verify.

c. Pick a point $\mathbf{p}=\gamma(0)=\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},0\right)$ on $\partial X$. We have $\mathbf{n_p}=(1,-1,0)$, so the tangent vector with the correct orientation is $\mathbf{w}=(1,1,-2)$ (found by setting $(dx\wedge dy)(\mathbf{n_p},\mathbf{w})>0$, or by $\mathbf{w}=\mathbf{N}\times\mathbf{n}$) Now $\mathbf{D}\gamma(0)=-\frac{1}{\sqrt{6}}\mathbf{w}$, so $\gamma$ gives the wrong orientation.

d. We guess the answer is no

e. We find $\mathbf{D}\gamma=\begin{bmatrix}-\frac{\sin t}{\sqrt{2}}-\frac{\cos t}{\sqrt{6}}\\ \frac{\sin t}{\sqrt{2}}-\frac{\cos t}{\sqrt{6}}\\ 2\frac{\cos t}{\sqrt{6}}\end{bmatrix}$ so

$$x\,dy-y\,dx=\left(\frac{\cos t}{\sqrt{2}}-\frac{\sin t}{\sqrt{6}}\right)\left(\frac{\sin t}{\sqrt{2}}-\frac{\cos t}{\sqrt{6}}\right)-\left(\frac{-\cos t}{\sqrt{2}}-\frac{\sin t}{\sqrt{6}}\right)\left(-\frac{\sin t}{\sqrt{2}}-\frac{\cos t}{\sqrt{6}}\right)=-\frac{1}{\sqrt{3}},$$

$$x\,dz-z\,dx=\frac{1}{\sqrt{3}},\text{ and } y\,dz-z\,dy=-\frac{1}{\sqrt{3}}$$

so all three define the orientation of $\partial X$ at every point

**Exercise 6.31.** a. Let $M\subset\mathbb{R}^4$ be a manifold defined by the equation $x_4=x_1^2+x_2^2+x_3^2$ and oriented by sgn $dx_1\wedge dx_2\wedge dx_3$. Consider the subset $X\subset M$ where $x_4\leqslant 1$. Show that it is a piece-with-boundary

b. Let $\mathbf{x}$ be a point of $\partial_M X$. Find a basis for the tangent space $T_{\mathbf{x}}\partial_M X$ that is direct for the boundary orientation.

*Solution.* a. Informally, the boundary of $X$ is given by the equations $x_4=1=x_1^2+x_2^2+x_3^2$, which is a $C_1$ function, and so $X$ is a piece-with-boundary.

b. The manifold is given the parametrization $\gamma:\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\mapsto\begin{pmatrix}x_1\\x_2\\x_3\\x_1^2+x_2^2+x_3^2\end{pmatrix}$ so

$$\mathbf{D}\gamma=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\\2x_1&2x_2&2x_3\end{bmatrix}$$

so sgn $dx_1\wedge dx_2\wedge dx_3=+1$ always. To orient $\partial_M X$ at a point $\mathbf{x}=\begin{bmatrix}x_1\\x_2\\x_3\\1\end{bmatrix}$ we pick $\mathbf{v}_{\text{out}}=$

$\begin{bmatrix}x_1\\x_2\\x_3\\2\end{bmatrix}$. Now, for $\mathbf{v}_1,\mathbf{v}_2$, as we can swap variables or vectors and switch signs, WLOG $x_1>0$.

Pick $\mathbf{v}_1 = \begin{bmatrix} x_3 \\ 0 \\ -x_1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} x_2 \\ -x_1 \\ 0 \\ 0 \end{bmatrix}$. This gives $dx_1 \wedge dx_2 \wedge dx_3 = x_1$, so sgn $dx_1 \wedge dx_2 \wedge dx_3 = +1$.

## 6.5 Exterior Derivatives

Exterior derivatives are generalizations of (normal) derivatives to apply to forms; their main purpose seems to be to allow the Fundamental Theorem (see next section) to be stated elegantly.

**Definition 6.32** (Exterior derivative).
Let $U \in \mathbb{R}^n$ be an open subset. The exterior derivative $\mathbf{d} : A^k(U) \to A^{k+1}(U)$ is defined by the formula

$$\mathbf{d}\varphi(P_{\mathbf{x}}(\mathbf{v}_1, \dots, \mathbf{v}_{k+1})) := \lim_{h \to 0} \frac{1}{h^{k+1}} \int_{\partial P_{\mathbf{x}}(h\mathbf{v}_1, \dots, h\mathbf{v}_{k+1})} \varphi.$$

**Exercise 6.33.** Let $\varphi = -y\, dx + x\, dy$. Calculate $\mathbf{d}\varphi$ directly from the definition.

*Solution.* Recall

$$\mathbf{d}\varphi = \lim_{h \to 0} \int_{\partial P_{\binom{x}{y}}(h\mathbf{u}, h\mathbf{v})} \varphi$$

Consider the parallelogram formed by $h\mathbf{u}$ and $h\mathbf{v}$; let $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$. There are four sections: $+h\mathbf{u}$, $+h\mathbf{v}$, $-h\mathbf{u}$, and $-h\mathbf{v}$. [I should probably put a pic here] Now we find the integral of each part:

- the $+h\mathbf{u}$ part $P_1$ is parametrized by $\gamma(t) = \binom{x}{y} + th\mathbf{u}, t \in [0, 1]$, so $\mathbf{D}\gamma(t) = h\mathbf{u}$. Therefore

$$\int_{P_1} -y\, dx + x\, dy = \int_0^1 -(y + thu_2)hu_1 + (x + thu_1)hu_2\, dt = -yhu_1 + xhu_2.$$

- the $+h\mathbf{v}$ part $P_2$ is parametrized by $\gamma(t) = \binom{x}{y} + h\mathbf{u} + th\mathbf{v}, t \in [0, 1]$, so $\mathbf{D}\gamma(t) = h\mathbf{v}$. Hence

$$\int_{P_2} -y\, dx + x\, dy = \int_0^1 -(y + hu_2 + thv_2)hv_1 + (x + hu_1 + thv_1)hv_2\, dt$$
$$= -yhv_1 + xhv_2 - h^2 u_2 v_1 + h^2 u_1 v_2.$$

- the $-h\mathbf{u}$ part $P_3$ is parametrized by $\gamma(t) = \binom{x}{y} + h\mathbf{v} + th\mathbf{u}, t \in [1, 0]$, so $\mathbf{D}\gamma(t) = h\mathbf{u}$. So

$$\int_{P_3} -y\, dx + x\, dy = \int_1^0 -(y + hv_2 + thu_2)hu_1 + (x + hv_1 + thu_1)hu_2 = yhu_1 - xhu_2 + h^2 u_1 v_2 - h^2 u_2 v_1$$

- the $-h\mathbf{v}$ part $P_4$ is parametrized by $\gamma(t) = \binom{x}{y} + th\mathbf{v}, t \in [1, 0]$, so $\mathbf{D}\gamma(t) = h\mathbf{v}$. Therefore

$$\int_{P_4} -y\, dx + x\, dy = \int_1^0 -(y + thv_2)hv_1 + (x + thv_1)hv_2 = yhv_1 - xhv_2.$$

Summing everything up, we have

$$\int_{\partial P_{\binom{x}{y}}(h\mathbf{u}, h\mathbf{v})} \varphi = 2h^2(u_1 v_2 - u_2 v_1).$$

In practice it is much easier to compute the exterior derivate using rules:

**Theorem 6.34** (Computing the exterior derivative; 6.7.4 in book). • *The exterior derivative of a function $f$ is given by*

$$\mathbf{d}f = [\mathbf{D}f] = \sum_{i=1}^{n}(D_i f)dx_i$$

• *If $f : U \to \mathbb{R}$ is a $C^2$ function, then*

$$\mathbf{d}\left(f\, dx_i \wedge \cdots \wedge dx_j\right) = \mathbf{d}f \wedge dx_i \wedge \wedge \cdots dx_j$$

## 6.6 Stokes' Theorem

We are ready to present *the Fundamental Theorem* (of multivariable calculus), otherwise known as (Generalized) Stokes' Theorem.

**Theorem 6.35** (Stokes' Theorem). *Let $R$ be a piece-with-boundary of a $k$-dimensional oriented smooth manifold $M$ in $\mathbb{R}^n$. Give the boundary $\partial R$ of $R$ the boundary orientation, and let $\varphi$ be a $C^2$ $(k-1)$-form defined on an open set containing $R$. Then*

$$\int_R \mathbf{d}\varphi = \int_{\partial R} \varphi$$

**Example 6.36.** Given $\varphi = -y\, dx + x\, dy$, we have $\mathbf{d}\varphi = 2\, dx \wedge dy$ so

$$\int_{\partial R} \varphi = \int_R \mathbf{d}\varphi = 2\text{vol}(R)$$

The unit circle, for example, is parametrized by $\gamma(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$, $t \in [0, 2\pi]$, so $\mathbf{D}\gamma(t) = \begin{bmatrix} -R\sin t \\ R\cos t \end{bmatrix}$, hence

$$\text{vol(unit circle)} = \frac{1}{2}\int_{\text{unit circle}} -y\, dx + x\, dy = \frac{1}{2}\int_0^{2\pi} \sin^2 t + \cos^2 t = \frac{1}{2}\cdot 2\pi = \pi.$$

Stokes' theorem is a generalization of the Fundamental Theorem of (single-variable) calculus:

$$f(b) - f(a) = \int_{\{-a,+b\}} d = \int_{[a,b]} f = \int_a^b f'(x)\, dx.$$

Quick review of integrals:

↑ *Less general*

- Riemann sums in $\mathbb{R}^n$; bounded functions with bounded support; density → total amount

⋮ *What if we don't have a cartesian coordinate system?*

- Change-of-variables, correction factors to account for distortions: $|\det \mathbf{D}\Phi|$

⋮ *What if I have a region in $\mathbb{R}^n$ parametrized in $\mathbb{R}^k$; $k < n$?*

- Integrate $k$-forms over parametrized domains in $\mathbb{R}^n$

⋮ *Under what changes of parameters is the integral invariant?*

- Integrals depend on *orientation*: the order of the basis vectors in $T_pM$.

⋮ *Is there a connection between $\int_X \varphi$ and some integral on the boundary $\partial\varphi$?*

- Exterior derivatives and Stokes' theorem.

↓ *More general*

## 6.7 Vector calculus: work, flux, and mass forms

These two subsections serve to introduce the terminology of vector calculus in case we should ever need them.

Recap: 0-forms in $\mathbb{R}^n$ assign a value to each point in $\mathbb{R}^n$, and are simply functions $f : \mathbb{R}^n \to \mathbb{R}$.

A *form* looks like this: $3\,dx \wedge dy + 4\,dy \wedge dz$. They are evaluated on a tuple of vectors.

A *form field* looks like this: $e^{x+z}\,dx \wedge dy + 4\sin(x/y)\,dy$ and resolves to different forms at different points. They are evaluated on *parallelograms*, that is, a point and a tuple of vectors.

Every 1-form is the *work form* of a vector field.

**Definition 6.37** (Work form, 6.5.1 in book).
The *work form* $W_{\mathbf{F}}$ of a vector field $\mathbf{F}$ is $\mathbf{F}(\mathbf{x}) \cdot \mathbf{Dx} = F_1\,dx + F_2\,dy + \cdots$

**Definition 6.38** (Work, 6.5.5 in book).
The *work* of a vector field $\mathbf{F}$ along an oriented curve $C$ is $\int_C W_{\mathbf{F}}$.

**Example 6.39.** We will find work of $\mathbf{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} y \\ -x \\ 0 \end{bmatrix}$ done over the helix $\gamma(t) = \begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix}, t \in [0, 4\pi]$ and oriented by the tangent vector field $\mathbf{t} = \begin{bmatrix} -\sin t \\ \cos t \\ 1 \end{bmatrix}$.

First we check orientation: since $\mathbf{t} = \gamma'(t)$, $\mathbf{t} \cdot \gamma'(t) > 0$ so this parametrization preserves orientation. Therefore,

$$\int_C W_{\mathbf{F}} = \int_{t=0}^{4\pi} (y\,dx - x\,dy)\left( P_{\begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix}}\left(\begin{bmatrix} -\sin t \\ \cos t \\ 1 \end{bmatrix}\right)\right) dt$$

$$= \int_0^{4\pi} \sin t(-\sin t) - \cos t \cos t\,dt = \int_0^{4\pi} -1\,dt = -4\pi.$$

**Exercise 6.40** (6.5.18 in book). Find the work of $\mathbf{F}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} x^2 \\ y^2 \\ z^2 \end{bmatrix}$ over the arc $C$ of helix parametrized by $\gamma : t \mapsto \begin{pmatrix} \cos t \\ \sin t \\ at \end{pmatrix}$ for $0 \leqslant t \leqslant \alpha$, and oriented so that $\gamma$ is orientation preserving.

**Example 6.41.** We have $W_{\mathbf{F}} = x^2\,dx + y^2\,dy + z^2\,dz$ so

$$\int_C W_{\mathbf{F}} = \int_0^{\alpha} (x^2\,dx + y^2\,dy + z^2\,dz)\left( P_{\begin{pmatrix}\cos t\\ \sin t\\ at\end{pmatrix}}\left(\begin{bmatrix}-\sin t\\ \cos t\\ a\end{bmatrix}\right)\right) dt$$

$$= \int_0^{\alpha} -\cos^2 t\sin t + \sin^2 t\cos t + a^3t^2\,dt$$

$$= \frac{1}{3}\left(\cos^3 t + \sin^3 t + a^3t^3\right)\Big|_0^{\alpha}$$

$$= \frac{1}{3}\left(\cos^3\alpha + \sin^3\alpha + a^3\alpha^3 - 1\right)$$

**Definition 6.42** (Flux form).
Given a vector field $\mathbf{F}$ on $\mathbb{R}^3$, the *flux form* $\Phi_{\mathbf{F}}$ is the 2-form field

$$\Phi_{\mathbf{F}}\left(P_{\mathbf{x}}(\mathbf{v},\mathbf{w})\right) := \det[\mathbf{F}(\mathbf{x}),\mathbf{v},\mathbf{w}]$$

**Definition 6.43** (Flux).
The *flux* of a vector field $\mathbf{F}$ over an oriented surface $S$ is $\int_S \Phi_{\mathbf{F}}$.

**Definition 6.44** (Mass form).
Let $U$ be a subset of $\mathbb{R}^3$ and $f : U \to \mathbb{R}$ a function. The *mass form* $M_f$ is the 3-form defined by

$$M_f\left(P_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)\right) := f(\mathbf{x})\det[\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3]$$

Summary: let $f$ be a function on $\mathbb{R}^3$ and $\mathbf{F} = \begin{bmatrix}F_1\\ F_2\\ F_3\end{bmatrix}$ a vector field. Then

$$W_{\mathbf{F}} = F_1\,dx + F_2\,dy + F_3\,dz$$

$$\Phi_{\mathbf{F}} = F_1\,dy\wedge dz - F_2\,dx\wedge dz + F_3\,dx\wedge dy$$

$$M_f = f\,dx\wedge dy\wedge dz$$

Also see table 6.5.6

**Exercise 6.45** (6.5.1 in book).

*Answer:* (Work form: a, j, l) (Work: b, i) (Flux form: d, k, h) (Flux: c, e, f) (Mass form: g)

**Exercise 6.46** (6.5.3 in book). a. $\to \Phi_{\mathbf{F}}\left(P_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2)\right)$

b. $\to W_{\mathbf{F}}$

c. $\to M_f\left(P_{\mathbf{x}}(\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3)\right)$

d. ? to be clear, either $\mathbf{v}_1\cdot(\mathbf{v}_2\times\mathbf{v}_3)$, $(\mathbf{v}_1\times\mathbf{v}_2)\cdot\mathbf{v}_3$, or $(\mathbf{v}_1\cdot\mathbf{v}_2)\,\mathbf{v}_3$, probably the former ($=\det[\mathbf{v}_1,\mathbf{v}_2,\mathbf{v}_3]$)

e. ✓

f. $\to \Phi_{\mathbf{F}} = F_1\ dy \wedge dz - F_2\ dx \wedge dz + F_3\ dx \wedge dy$

g. $\to W_{\mathbf{F}}\left(P_{\mathbf{x}}(\mathbf{v})\right)$

h. $\to M_f$

i. ✓

**Exercise 6.47** (6.5.4 in book). Show that $\Phi_{\mathbf{F}\times\mathbf{G}} = W_{\mathbf{F}} \wedge W_{\mathbf{G}}$

*Solution.* Let $\mathbf{F} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix}$, $\mathbf{G} = \begin{bmatrix} G_1 \\ G_2 \\ G_3 \end{bmatrix}$. Then

$$\mathbf{F} \times \mathbf{G} = \begin{bmatrix} F_2G_3 - F_3G_2 \\ F_3G_1 - F_1G_3 \\ F_1G_2 - F_2G_1 \end{bmatrix}$$

Therefore

$$\Phi_{\mathbf{F}\times\mathbf{G}} = \sum_{cyc}(F_2G_3 - F_3G_2)\ dy \wedge dz.$$

As $W_{\mathbf{F}} = F_1\ dx + F_2\ dy + F_3\ dz$ and $W_{\mathbf{G}} = G_1\ dx + G_2\ dy + G_3\ dz$,

$$W_{\mathbf{F}} \wedge W_{\mathbf{G}} = \sum_{cyc} F_1G_1(dx \wedge dx) + F_1G_2(dx \wedge dy) + F_1G_3(dx \wedge dz)$$

$$= \sum_{cyc}(F_2G_3 - F_3G_2)\ dy \wedge dz = \Phi_{\mathbf{F}\times\mathbf{G}}.$$

## 6.8 Vector calculus: grad, curl, div

grad, curl, and div are the traditional vector calculus operators.

- For 0-forms we have

$$\mathbf{d}f = D_1 f\ dx_1 + \cdots + D_n f\ dx_n =: W_{\text{grad } f} = W_{\nabla f}.$$

- For 1-forms we have

$$\mathbf{d}W_{\mathbf{F}} = \mathbf{d}\left(\sum_{i=1}^{n} F_i\ dx_i\right) = \sum_{i=1}^{n} \mathbf{d}F_i\ dx_i = \sum_{i=1}^{n} \mathbf{d}F_i \wedge dx_i$$

$$= \sum_{i,j=1}^{n} F_{i,j} dx_j \wedge dx_i = \sum_{1\leqslant i<j\leqslant n} (F_{j,i} - F_{i,j}) dx_i \wedge dx_j.$$

In $\mathbb{R}^3$, this gives the flux form of a vector field we define as the *curl*.

$$\mathbf{d}W_{\mathbf{F}} =: \Phi_{\text{curl}\,\mathbf{F}} = \Phi_{\nabla\times\mathbf{F}}$$

- For 2-forms in $\mathbb{R}^3$, we have

$$\mathbf{d}\Phi_{\mathbf{F}} = \mathbf{d}(F_1\ dy \wedge dz + F_2\ dz \wedge dx + F_3\ dx \wedge dy) = \mathbf{d}F_1 \wedge dy \wedge dz + \mathbf{d}F_2 \wedge dz \wedge dx + \mathbf{d}F_3 \wedge dx \wedge dy$$

$$= (F_{1,x} + F_{2,y} + F_{3,z}) dx \wedge dy \wedge dz =: M_{\text{div}\,\mathbf{F}} = M_{\nabla\cdot\mathbf{F}}$$

Physical interpretations:

- As $[\mathbf{D}f(\mathbf{x})]\mathbf{v} = \nabla f \cdot \mathbf{v} = |\nabla f||\mathbf{v}| \cos \angle(\nabla f, \mathbf{v})$, $\operatorname{grad} f$ tells us the direction of the steepest incline.
- for curl, see fig 6.8.2

**Exercise 6.48** (6.8.1 in book). Notation review

*Solution.* a. i,ii: vector fields, iii, iv, v: numbers, vi: function

b.
- $\operatorname{grad} f = \nabla f$
- $\operatorname{curl} \mathbf{F} = \nabla \times \mathbf{F}$
- $\operatorname{div} \mathbf{F} = \nabla \cdot \mathbf{F}$
- $\mathbf{d}f = D_1 f\, dx_1 + D_2 f\, dx_2 + D_3 f\, dx_3 = W_{\nabla f} = W_{\operatorname{grad} f}$
- $\mathbf{d}W_{\mathbf{F}} = \Phi_{\operatorname{curl} \mathbf{F}} = \Phi_{\nabla \times \mathbf{F}}$
- $\mathbf{d}\Phi_{\mathbf{F}} = M_{\operatorname{div} \mathbf{F}} = M_{\nabla \cdot \mathbf{F}}$

c. i $\to$ $\operatorname{grad} f$, ii $\to$ $\operatorname{curl} \mathbf{F}$, iii $\to$ $\Phi_{\mathbf{F}}(\mathbf{v}_1, \mathbf{v}_2)$, iv $\to$ $W_{\mathbf{F}}$, v ✓,vi ✓

## 6.9 Review: exercises

**Exercise 6.49** (6.10.1 in book). Let $U$ be a compact piece-with-boundary of $\mathbb{R}^3$. Show that

$$\operatorname{vol}_3 U = \int_{\partial U} \frac{1}{3} (x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy)$$

*Solution.* Let $\varphi = \frac{1}{3} (x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy)$. We have

$$\mathbf{d}\varphi = \frac{1}{3} (dx \wedge dy \wedge dz + dy \wedge dz \wedge dx + dz \wedge dx \wedge dy) = dx \wedge dy \wedge dz.$$

By Stokes' theorem,

$$\int_{\partial U} \varphi = \int_U \mathbf{d}\varphi = \int_U dx \wedge dy \wedge dz = \operatorname{vol}_3 U$$

**Exercise 6.50** (6.10.2 in book). Let $C$ be the part of the cone of equation $z = a - \sqrt{x^2 + y^2}$ where $z \geqslant 0$, oriented by the upward-pointing normal. What is the integral

$$\int_C x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy?$$

*Solution.* Let $\varphi = x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy$, so $\mathbf{d}\varphi = 3dx \wedge dy \wedge dz$. Let $B$ be the base plate $z = 0$; $\sqrt{x^2 + y^2} \leqslant a$—$B$ and $C$ together form the boundary of the solid cone $\mathcal{C}$. By Stokes' theorem,

$$\int_C \varphi = 3 \operatorname{vol}_3 \mathcal{C} - \int_B \varphi,$$

but as $z = 0$ for all points in $B$, $\int_B \varphi = 0$, so

$$\int_C \varphi = 3 \operatorname{vol}_3 \mathcal{C} = 3 \cdot \frac{1}{3} \cdot \pi \cdot a^2 \cdot a = \pi a^3.$$

**Exercise 6.51** (6.10.3 in book). Compute the integral of $x_1\, dx_2 \wedge dx_3 \wedge dx_4$ over the part of the 3-dimensional manifold of equation $x_1 + x_2 + x_3 + x_4 = a$ where $x_1, x_2, x_3, x_4 \geqslant 0$, oriented so that the projection to $(x_1, x_2, x_3)$-coordinate 3-space is orientation-preserving.

*Solution.* Let $\varphi = x_1\, dx_2 \wedge dx_3 \wedge dx_4$, and let $S$ be the simplex bounded by the plane $P : x_1 + x_2 + x_3 + x_4 = a$ and the four hyperplanes $x_i = 0$. Using Stokes' theorem,

$$\operatorname{vol}_4 S = \int_S \mathbf{d}\varphi = \int_{\partial S} \varphi = \int_{S\cap P} \varphi + \sum_{i=1}^{4} \int_{S\cap\{x_i=0\}} \varphi.$$

However, as $\varphi$ contains all of $x_1, \ldots, x_4$, $\int_{S\cap\{x_i=0\}} \varphi$ vanishes for all $i$, so

$$\int_{S\cap P} \mathbf{d}\varphi = \operatorname{vol}_4 S = \frac{1}{4!}a^4.$$

**Exercise 6.52** (6.11.1 in book)**.** Let $S$ be the torus obtained by rotating around the $z$-axis the circle of equation $(x-2)^2 + z^2 = 1$. Orient $S$ by the outward-pointing normal. Compute

$$\int_S \Phi_{\mathbf{F}}, \text{ where } \mathbf{F} = \begin{bmatrix} x + \cos(yz) \\ y + e^{x+z} \\ z - x^2y^2 \end{bmatrix}.$$

*Solution.* We have $\Phi_{\mathbf{F}} = F_1\, dy \wedge dz + F_2\, dz \wedge dx + F_3\, dx \wedge dy$, so $\mathbf{d}\Phi = 3\, dx \wedge dy \wedge dz$, hence

$$\int_S \Phi_F = \int_{\mathcal{S}} \mathbf{d}\Phi_F = \operatorname{vol}_3 \mathcal{S} = 4\pi^2$$

where $\mathcal{S}$ is the solid torus bounded by $S$.

**Exercise 6.53** (6.11.2 in book)**.** Suppose $U \subset \mathbb{R}^3$ is open, $\mathbf{F}$ is a $C^1$ vector field on $U$, and $\mathbf{a}$ is a point of $U$. Let $S_r(\mathbf{a})$ be the sphere of radius $r$ centered at $\mathbf{a}$, oriented by the outward-pointing normal. Compute $\lim_{r\to 0} \frac{1}{r^3} \int_{S_r(\mathbf{a})} \Phi_{\mathbf{F}}$.

**Exercise 6.54** (6.10.4 in book)**.** a. Show that the 2-form on $\mathbb{R}^3 - \{\mathbf{0}\}$ given by

$$\varphi = \frac{x\, dy \wedge dz + y\, dz \wedge dx + z\, dx \wedge dy}{(x^2+y^2+z^2)^{3/2}}$$

satisfies $\mathbf{d}\varphi = 0$.

b. Compute $\int_S \varphi$, where $S$ is the sphere of radius $R \neq \sqrt{3}$ centered at $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, oriented by the outward-pointing normal. (The result depends on $R$, of course)

*Solution.* a. Define $r = \sqrt{x^2+y^2+z^2}$, so

$$\mathbf{d}\varphi = \mathbf{d}\left(\frac{x}{r^3} dy \wedge dz + \frac{y}{r^3} dz \wedge dx + \frac{z}{r^3} dz \wedge dx\right) = \mathbf{d}\Phi_{\mathbf{F}}$$

for $\mathbf{F} = \frac{1}{r^3}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$, hence

$$\mathbf{d}\varphi = M_{\operatorname{div}\mathbf{F}} = \left(D_1\left(\frac{x}{r^3}\right) + D_2\left(\frac{y}{r^3}\right) + D_3\left(\frac{z}{r^3}\right)\right) dx \wedge dy \wedge dz.$$

Note $D_1(r) = \frac{x}{r}$ so $D_1\left(\frac{x}{r^3}\right) = \frac{r^3 - 3rx^2}{r^6}$, so

$$D_1\left(\frac{x}{r^3}\right) + D_2\left(\frac{y}{r^3}\right) + D_3\left(\frac{z}{r^3}\right) = \frac{1}{r^6}\left(r^3 - 3rx^2 + r^3 - 3ry^2 + r^3 - 3rz^2\right)$$

$$= \frac{3}{r^6}\left(r^3 - r(x^2+y^2+z^2)\right) = 0.$$

b. If $R < \sqrt{3}$ then the ball $B$ bounded by $R$ does not contain the origin, so $\varphi$ is $C^1$ over $B$ and hence
$$\int_S \varphi = \int_B \mathbf{d}\varphi = \int_B 0 = 0.$$
If $R > \sqrt{3}$ then as $\mathbf{d}\varphi = 0$ at every point except the origin, we can deform $S$ to the unit sphere $C$ without changing the flux, though note that the Stokes' Theorem is not applicable here because $\varphi$ is not continuous at the origin. Back to the problem, consider the parametrization
$$\gamma : \begin{pmatrix} \theta \\ \phi \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta\cos\phi \\ \sin\theta\cos\phi \\ \sin\phi \end{pmatrix}$$
of the unit sphere. We find
$$\mathbf{D}\gamma \begin{pmatrix} \theta \\ \phi \end{pmatrix} = \begin{bmatrix} -\sin\theta\cos\phi & -\cos\theta\sin\phi \\ \cos\theta\cos\phi & -\sin\theta\sin\phi \\ 0 & \cos\phi \end{bmatrix}.$$
At a point $\begin{pmatrix} \cos\theta\cos\phi \\ \sin\theta\cos\phi \\ \sin\phi \end{pmatrix}$, the outward-pointing normal is given by the same vector as the point, so to check orientation, we verify that
$$\begin{vmatrix} \cos\theta\cos\phi & -\sin\theta\cos\phi & -\cos\theta\sin\phi \\ \sin\theta\cos\phi & \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \sin\phi & 0 & \cos\phi \end{vmatrix} = \sin\phi(\sin\phi\cos\phi) + \cos\phi(\cos\phi^2) = \cos\phi > 0$$
as $-\frac{\pi}{2} < \phi < \frac{\pi}{2}$ Therefore
$$\begin{aligned}\int_C \varphi &= \int_C x\ dy \wedge dz + y\ dz \wedge dx + z\ dx \wedge dy \\ &= \int_{-\pi/2}^{\pi/2}\int_0^{2\pi} (\cos\theta\cos\phi)(\cos\theta\cos^2\phi) + (\sin\theta\cos\phi)(\sin\theta\cos^2\phi) + \sin\phi(\sin\phi\cos\phi)\ d\theta\ d\phi \\ &= \int_{-\pi/2}^{\pi/2}\int_0^{2\pi} \cos^3\phi(\sin^2\theta + \cos^2\theta) + \cos\phi\sin^2\phi\ d\theta\ d\phi \\ &= \int_{-\pi/2}^{\pi/2}\int_0^{2\pi} \cos\phi\ d\theta\ d\phi \\ &= 4\pi.\end{aligned}$$

## 6.10 Review: a Princeton final

**Problem 1.** Let $M \subset \mathbb{R}^3$ be given by $M = \{(x, y, z) \mid x^4 + 3y^4 + 2z^4 = 36\}$ and consider the function $f : M \to \mathbb{R}$ given by $f(x, y, z) = x + 3y + 16$

a. Prove that $f$ attains a maximum value on $M$

b. Find this value and the point on $M$ where it is attained.

*Solution.* a. A continuous function on a compact set attains a maximum value.

b. Lagrange multipliers

**Problem 3.** Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^3 + x$, and let

$$R = \{(x, y) \in \mathbb{R}^2 \mid \big(f^{-1}(x)\big)^2 + \big(f^{-1}(y)\big)^2 \leqslant 1\}.$$

Find the area of $R$.

*Solution.* As $f$ has an inverse, $R$ is simply the image of the unit sphere when $f$ is applied; the correction factor is simply $\mathbf{D}f$. Alternatively, parametrize the boundary and use Stokes' theorem.

**Problem 4.** Let $M$ be the three-dimensional submanifold of $\mathbb{R}^4$ defined by

$$M = \{(x, y, z, w) \mid x^2 + y^2 + z^2 \leqslant 1, w = x - \frac{1}{3}y^3 + \frac{1}{5}z^5\},$$

oriented so that the component of its normal vector in the $w$ direction is positive. Compute

$$\int_M x^2 z^2 \, dy \wedge dz \wedge dw + z^2 \, dx \wedge dz \wedge dw + dx \wedge dy \wedge dw.$$

**Problem 5.** Let $f : \mathbb{R}^2 \to \mathbb{R}^3$ and $g : \mathbb{R}^3 \to \mathbb{R}^5$ be given by

$$f(x_1, x_2) = (x_1^3 + x_1, x_1 + x_2, x_2^2 + 1)$$
$$g(y_1, y_2, y_3) = (y_1^2, y_1 y_2, y_2^2, y_2 y_3, y_3^2)$$

and let $h = g \circ f$. Find a basis for the tangent space to $\operatorname{im} h$ at $h(1, 1)$

*Solution.* Use the chain rule: $\mathbf{D}(g \circ f)(\mathbf{x}) = \mathbf{D}(g)(f(\mathbf{x})) \circ \mathbf{D}f(\mathbf{x})$.

**Problem 6.** The function $\mathbb{R}^4 \to \mathbb{R}$ is given by $f(x_1, x_2, x_3, x_4) = x_1 x_3^2 + 2x_2^2 x_4^2$.

a. If I'm standing at the point $(1, 1, 2, 2)$, in which direction should I move to make $F$ increase as quickly as possible?

b. Same question, but now supposing that I am only allowed to move on the hypersurface

$$\{(x_1, x_2, x_3, x_4) \mid x_1^4 - x_2^4 + 2x_2 x_3 - 3x_1 x_4 + x_4^2$$

*Solution.* a. Just find $\nabla f$.

b. Project $\nabla f$ onto the tangent space of $f$. An easy way is to find a vector $\mathbf{v}_\perp$ perpendicular to the tangent space and find the $\mathbf{v}_\perp$-component $\mathbf{v}_\perp \nabla f$ of $\nabla f$; the projection would simply be $\nabla f - \mathbf{v}_\perp \nabla f$.

**Problem 7.** See Exercise 6.44 on page 49.