HOMEWORK 1 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 7.1.

1. For the vectors $x = (1, 0, 2)$ and $y = (1, -3, 1)$ in $\mathbb{R}^3$, find

(1) $2x + y$;
(2) $x \cdot y$;
(3) $\|x\|$ and $\|y\|$;
(4) the cosine of the angle between x and y;
(5) the distance from x to y.

Solution: (a) $2x + y = 2(1, 0, 2) + (1, -3, 1) = (2, 0, 4) + (1, -3, 1) = (3, -3, 5)$.
(b) $x \cdot y = (1)(1) + (0)(-3) + (2)(1) = 1 + 0 + 2 = 3$.
(c) $\|x\| = \sqrt{1^2 + 0^2 + 2^2} = \sqrt{5}$.
$\|y\| = \sqrt{1^2 + (-3)^2 + 1^2} = \sqrt{11}$.
(d) $\cos\theta = \frac{x \cdot y}{\|x\|\|y\|} = \frac{3}{\sqrt{5}\sqrt{11}} = \frac{3}{\sqrt{55}}$.
(e) $\|x - y\| = \|(0, 3, 1)\| = \sqrt{0 + 9 + 1} = \sqrt{10}$. □

4. (just part b) Let $u, v, w \in \mathbb{R}^d$ and $a \in \mathbb{R}$. We want to show that $(u + v) \cdot w = u \cdot w + v \cdot w$. I'll use the notation $u = (u_1, \ldots, u_d)$, etc.

Proof.

$$\begin{aligned}(u + v) \cdot w &= (u_1 + v_1, \ldots, u_d + v_d) \cdot (w_1, \ldots, w_d) \\ &= \sum_{i=1}^{d} (u_i + v_i)w_i \\ &= \sum_{i=1}^{d} u_i w_i + v_i w_i \text{ (by the distributive property in } \mathbb{R}) \\ &= \sum_{i=1}^{d} u_i w_i + \sum_{i=1}^{d} v_i w_i \text{ (using commutativity of + in } \mathbb{R}) \\ &= u \cdot w + v \cdot w.\end{aligned}$$

(I'm ok with you not specifically saying that you used the distributive and commutative properties here). □

5. Prove that $|\|x\| - \|y\|| \leq \|x - y\|$ holds for any pair of vectors x, y in a normed vector space.

Proof. We have

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

Date: September 6, 2019.

where the last inequality is from the usual triangle inequality, Theorem 7.1.10(a). Now subtracting $\|y\|$ from the far-left side and far-right side above yields $\|x\| - \|y\| \leq \|x - y\|$. Similarly (you can just say by switching the x's and y's),

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

and so $\|y\| - \|x\| \leq \|y - x\|$, and we know $\|y - x\| = \|(-1)(x - y)\| = |-1|\|x - y\| = \|x - y\|$. Equivalently (multiplying by (-1)), $\|x\| - \|y\| \geq -\|x - y\|$. Combining these, we have

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

or equivalently, $|\|x\| - \|y\|| \leq \|x - y\|$, as desired. □

7. For a norm on a vector space X, defined by an inner product $\cdot$ as in Definition 7.1.7 (i.e., $\|x\| := \sqrt{x \cdot x}$), prove the parallelogram law

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

holds for all $x, y \in X$.

Proof. (Since we've done Example 7.1.5, I'll assume we know how to expand $(x + y) \cdot (x + y)$ and $(x - y) \cdot (x - y)$). We have

$$\begin{aligned}
\|x + y\|^2 + \|x - y\|^2 &= (x + y) \cdot (x + y) + (x - y) \cdot (x - y) \\
&= (\|x\|^2 + 2x \cdot y + \|y\|^2) + (\|x\|^2 - 2x \cdot y + \|y\|^2) \\
&= 2\|x\|^2 + 2\|y\|^2
\end{aligned}$$

as desired. □

11. Prove that the sup norm $\|f\|_\infty := \sup_{x \in I} |f(x)|$ really is a norm on $C(I)$ (the vector space of continuous real-valued functions on the nonempty closed bounded interval I).

Proof. We need to prove that $\|\cdot\|_\infty$ satisfies properties (a), (b), and (c) as listed in Theorem 7.1.10. Let $f, g \in C(I)$ and $a \in \mathbb{R}$.

(a)

$$\begin{aligned}
\|f + g\|_\infty &= \sup_I |f + g| \\
&\leq \sup_I (|f| + |g|) \text{ (by the triangle inequality for } \mathbb{R}) \\
&\leq \sup_I |f| + \sup_I |g| \text{ (by Theorem 1.5.10(c))} \\
&= \|f\|_\infty + \|g\|_\infty.
\end{aligned}$$

(b)

$$\begin{aligned}
\|af\|_\infty &= \sup_I |af(x)| \\
&= \sup_I |a||f(x)| \\
&= |a| \sup_I |f(x)| \text{ (by Theorem 1.5.10(a))} \\
&= |a|\|f\|_\infty.
\end{aligned}$$

(c) Suppose $\|f\|_\infty = 0$. This means that $\sup_I |f(x)| = 0$. This means in particular that 0 is an upper bound for $\{|f(x)| : x \in I\}$, so $|f(x)| \leq 0$ for all $x \in I$. On the other hand, we of course have $|f(x)| \geq 0$ for all x in I. Hence, $|f(x)| = 0$ for all $x \in I$, i.e., $f = 0$ (meaning the 0-function, which is the additive identity in $C(I)$), as desired. □

12. Let $\{x_k\}$ and $\{y_k\}$ be sequence of real numbers such that $\sum_{k=1}^\infty x_k^2 < \infty$ and $\sum_{k=1}^\infty y_k^2 < \infty$. Show that $\sum_{k=1}^\infty |x_k y_k| < \infty$.

Proof. Let $s_n := \sum_{k=1}^n |x_k y_k|$ denote the n-th partial sum of $\sum_{k=1}^\infty |x_k y_k|$. Similarly, let $s_n^x := \sum_{k=1}^n x_k^2$ and $s_n^y := \sum_{k=1}^n y_k^2$. Note that we can view s_n as a dot product in $\mathbb{R}^n$:

$$s_n = (|x_1|, \ldots, |x_n|) \cdot (|y_1|, \ldots, |y_n|).$$

Let $|\vec{x}|_n$ denote the vector $(|x_1|, \ldots, |x_n|)$ and let $|\vec{y}|_n$ denote $(|y_1|, \ldots, |y_n|)$. Note that $s_n = |\vec{x}|_n \cdot |\vec{y}|_n$, $\sqrt{s_n^x} = \||\vec{x}|_n\|$, and $\sqrt{s_n^y} = \||\vec{y}|_n\|$. By the Cauchy-Schwarz inequality, we have

$$s_n \leq \||\vec{x}|_n\|\||\vec{y}|_n\| = \sqrt{s_n^x}\sqrt{s_n^y}.$$

Since the sequences of partial sums $\{s_n^x\}$ and $\{s_n^y\}$ are convergent, so is the sequence $\{\sqrt{s_n^x}\sqrt{s_n^y}\}$ (because $\sqrt{\cdot}$ is a continuous function and products of bounded sequences are bounded). Hence, $\{s_n\}$ is bounded above by a convergent, hence bounded, sequence. Since $\{s_n\}$ is also monotone increasing, this implies that s_n converges. □

Section 7.2

1. Show from the definition of the limit of a sequence in $\mathbb{R}^d$ that

$$\lim \left(\frac{n}{1+n}, \frac{1-n}{n} \right) = (1, -1).$$

Proof. We have

$$\begin{aligned} \left\| (1,-1) - \left(\frac{n}{1+n}, \frac{1-n}{n} \right) \right\| &= \left\| \left(\frac{1}{1+n}, \frac{-1}{n} \right) \right\| \\ &= \sqrt{(\frac{1}{1+n})^2 + (\frac{1}{n})^2} \\ &= \sqrt{\frac{n^2 + (1+n)^2}{n^2(n+1)^2}} \\ &= \frac{\sqrt{2n^2+2n+1}}{n(n+1)} \\ &\leq \frac{\sqrt{2n^2}}{n^2+n} \leq \frac{n\sqrt{2}}{n^2} = \frac{\sqrt{2}}{n}. \end{aligned}$$

So given $\epsilon > 0$, we let $N = \frac{\sqrt{2}}{\epsilon}$, and then $n > N$ implies

$$\left\| (1,-1) - \left(\frac{n}{1+n}, \frac{1-n}{n} \right) \right\| \leq \frac{\sqrt{2}}{(\sqrt{2}/\epsilon)} = \epsilon,$$

as desired. □

2. Let $x_n = \left(\frac{n^2+n-1}{3n^2+2}, \frac{n-1}{n+1}\right)$. We claim x_n converges to $(\frac{1}{3}, 1)$. By Theorem 7.1.13, it suffices to check this for each component. That is, we just have to show that $\frac{n^2+n-1}{3n^2+2} \to \frac{1}{3}$ and $\frac{n-1}{n+1} \to 1$, and we see that these hold because they are limits of rational functions in n where the numerator and denominator have the same degree, and so we can just look at the leading coefficients (you can feel free to make this more precise however you want).

□

3. Let $x_n = (1 + (-1)^n, 1/n, 1 + 1/n)$. Then x_n cannot converge because the first components $\{1+(-1)^n\}$ do not converge (you can prove this if you want, but I'm ok with you just stating it since I'm sure you essentially did this last semester).

□

6. Let $\{x_n\}$ and $\{y_n\}$ be sequences in $\mathbb{R}^d$. Suppose $\lim x_n = 0$ and $\{y_n\}$ is bounded. Prove that $\lim x_n \cdot y_n = 0$.

Proof. Since $\{y_n\}$ is bounded, there is a number $M \in \mathbb{R}_{\geq 0}$ such that $\|y_n\| \leq M$ for all n. Using Cauchy-Schwarz, we have

$$|x_n \cdot y_n| \leq \|x_n\|\|y_n\| \leq M\|x_n\|$$

for all n. Since $x_n \to 0$, we have that $\|x_n\| \to 0$ by Theorem 7.2.11, and so $M\|x_n\| \to 0$. It now follows from Theorem 2.3.1 that $x_n \cdot y_n \to 0$.

Note: There are of course many other ways you could do this proof. □

9. Let $x_n = (\sin n, \cos n, 1 + (-1)^n)$. Does $\{x_n\}$ have a convergent subsequence?

Solution: For each n, we have

$$\begin{aligned}\|x_n\| &= \sqrt{\sin^2 n + \cos^2 n + (1 + (-1)^n)^2} \\ &= \sqrt{1 + (1 + (-1)^n)^2} \\ &\leq \sqrt{1 + 2^2} = \sqrt{5}.\end{aligned}$$

Hence, $\{x_n\}$ is bounded, so by the Bolzano-Weierstrass Theorem, it does have a convergent subsequence. □

12. For $x, y \in \mathbb{R}$, set $\delta(x, y) = 0$ if $x = y$ and $\delta(x, y) = 1$ if $x \neq y$. Show that δ is a metric on $\mathbb{R}$.

Proof. Let $x, y, z \in \mathbb{R}$. We will show that δ satisfies conditions (a), (b), and (c) from Definition 7.2.1.

(a) If $x = y$, then $y = x$, and so $\delta(x, y) = 0 = \delta(y, x)$. Similarly, if $x \neq y$, then $y \neq x$, so then $\delta(x, y) = 1 = \delta(y, x)$. In any case, $\delta(x, y) = \delta(y, x)$.

(b) By definition, $\delta(x, y) = 0$ if $x = y$, and conversely, if $x \neq y$, then $\delta(x, y) = 1 \neq 0$ (this is the contrapositive of the converse). So $\delta(x, y) = 0$ if and only if $x = y$.

(c) If $x = z$, then $\delta(x, z) = 0$, and this is less than or equal to $\delta(x, y) + \delta(y, z)$. Now suppose $x \neq z$, so $\delta(x, z) = 1$. If $x = y$ and $y = z$, then $x = z$, a contradiction, so either $x \neq y$ or $y \neq z$ (or both). So then $\delta(x, y) = 1$ or $\delta(y, z) = 1$, or both. In any case, $\delta(x, y) + \delta(y, z) \geq 1 = \delta(x, z)$, as desired. □

13. Which sequences converge in the metric space of the previous exercise?

Claim: A sequence $\{a_n\}$ in $\mathbb{R}$, with the metric δ of the previous exercise, converges if and only if there exists some $N \in \mathbb{R}$ such that $\{a_n\}$ is constant for $n > N$ (i.e., $a_n = a_m$ for all $n, m > N$).

Proof. Let $\{a_n\}$ be such a sequence. Let $a \in \mathbb{R}$ be the constant which a_n equals whenever $n > N$. We claim that $\lim a_n = a$. Indeed, for $\epsilon > 0$, we have by assumption an N such that $\delta(a, a_n) = 0 < \epsilon$ whenever $n > N$.

Now, let $\{a_n\}$ be a sequence which does not satisfy the condition from the claim. That is, for any $N \in \mathbb{R}$, there exist $n, m > N$ such that $a_n \neq a_m$. Let $\epsilon = \frac{1}{2}$. Then, as we have just said, for any $N \in \mathbb{R}$, there exist $n, m > N$ such that $a_n \neq a_m$. So for any $a \in \mathbb{R}$, at least one of a_n and a_m does not equal a, and so $\delta(a, a_n) = 1 > \epsilon$ or $\delta(a, a_m) = 1 > \epsilon$. So a_n does not converge to any $a \in \mathbb{R}$.

□

HOMEWORK 2 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 7.3.

1. Prove that $H := \{(x,y) \in \mathbb{R}^2 : y > 0\}$ is an open subset of $\mathbb{R}^2$.

Proof. Let $\vec{p} = (x_0, y_0)$ be a point in H. We claim that $B_{y_0}(p) \subset H$. Let $\vec{q} = (x_1, y_1) \in B_{y_0}(p)$. Then

$$y_0 > \|\vec{q} - \vec{p}\| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \geq \sqrt{(y_1 - y_0)^2} = |y_1 - y_0|,$$

So $-y_0 < y_1 - y_0 < y_0$, hence $0 < y_1 < 2y_0$. In particular, since $y_1 > 0$, we have $\vec{q} \in H$, as desired. □

2. Prove that every finite subset of $\mathbb{R}^d$ is closed.

Proof. Let $S = \{\vec{x}_1, \ldots, \vec{x}_n\}$ be a finite set of points in $\mathbb{R}^d$. Let $\vec{y} \in \mathbb{R}^d \setminus S$. For each $i = 1, \ldots, n$, let $r_i := \|\vec{y} - \vec{x}_i\|$. For each i here, $r_i > 0$ since $\vec{y} \neq \vec{x}_i$ (because $\vec{y}$ is in the complement of S). Let $r := \min_{i=1}^n \{r_i\} > 0$. Then for each $i = 1, \ldots, n$, we have $\|\vec{y} - \vec{x}_i\| = r_i \geq r$, hence $\vec{x}_i \notin B_r(\vec{y})$. So $B_r(\vec{y}) \subset \mathbb{R}^d \setminus S$. Thus, $\mathbb{R}^d \setminus S$ is open, so S is closed, as desired. □

Alternative proof: You could of course just prove that a single point is closed and then use the fact that finite unions of closed sets are closed. To prove that a single point is closed, you could work from the definition like above, or you could use Theorem 7.3.10: If S is a single point p, then any sequence in S must be the constant sequence $x_k = p$ for all k, hence the limit must be $p \in S$. So S contains the limit of any convergent sequence in S and therefore must be closed. As yet another approach, you could say that a point x is equal to the intersection of all closed balls centered at x, and arbitrary intersections of closed sets are closed, so x is closed. □

3. Find the interior, closure, and boundary of the set $\{(x,y) \in \mathbb{R}^2 : 0 \leq x < 2, 0 \leq y < 1\}$. **As I posted on Canvas, no justification is needed for this problem.**

Solution:

Interior: $\{(x,y) \in \mathbb{R}^2 : 0 < x < 2, 0 < y < 1\}$.

Closure: $\{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 2, 0 \leq y \leq 1\}$.

Boundary: $\{(x,y) \in \mathbb{R}^2 : x = 0 \text{ or } 2, 0 \leq y \leq 1\} \cup \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 2, y = 0 \text{ or } 1\}$.

There are of course other acceptable ways to write this. □

4. Find the interior, closure, and boundary of the set $\{(x,y) \in \mathbb{R}^2 : \|(x,y)\| < 1\} \cup \{(x,y) \in \mathbb{R}^2 : y = 0, -2 < x < 2\}$. **As I posted on Canvas, no justification is needed for this problem.**

Solution:

Interior: $\{(x,y) \in \mathbb{R}^2 : \|(x,y)\| < 1\}$.

Closure: $\{(x,y) \in \mathbb{R}^2 : \|(x,y)\| \leq 1\} \cup \{(x,y) \in \mathbb{R}^2 : y = 0, -2 \leq x \leq 2\}$.

Boundary: $\{(x,y) \in \mathbb{R}^2 : \|(x,y)\| = 1\} \cup \{(x,y) \in \mathbb{R}^2 : y = 0, 1 \leq |x| \geq 2\}$.

There are of course other acceptable ways to write this. □

Date: September 6, 2019.

6. Let A be an open set and B a closed set. If $B \subset A$, prove that $A \setminus B$ is open. If $A \subset B$, prove that $B \setminus A$ is closed.

I'll use the superscript ^{C} to denote taking a complement.

Proof. Since B is closed, B^C is open. If $B \subset A$, then $A \setminus B = A \cap B^C$, and this is open because it is an intersection of two open sets.

Similarly, since A is open, A^C is closed. If $A \subset B$, then $B \setminus A = B \cap A^C$, and this is closed since it is an intersection of two closed sets. □

7. Prove Theorem 7.3.7 (just parts (b) and (c) since I did part (a) in class). That is, for $E \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$, prove that:

(b) $x \in \overline{E}$ if and only if every neighborhood of x contains a point of E;

(c) $x \in \partial E$ if and only if every neighborhood of x contains points of E and points of E^C.

Proof. (b) Suppose $x \in \overline{E}$. Let U be a neighborhood of x, i.e., an open set containing x. Then U^C is a closed set which does not contain x, hence does not contain $\overline{E}$. Since $\overline{E}$ is the intersection of all closed sets containing E, this implies that U^C is not a closed set containing E. But U^C is closed, so it must be that it does not contain E. Thus, some points of E must be contained in U. So every neighborhood of x contains points of E, as desired.

Now suppose $x \notin \overline{E}$. Then since $\overline{E}$ is closed and $E \subset \overline{E}$, $\overline{E}^C$ is a neighborhood of x which does not contain a point of E. This proves the (contrapositive of the) converse.

(c) Suppose $x \in \partial E = \overline{E} \setminus E^\circ$. Since $x \in \overline{E}$, (b) tells us that every neighborhood of x contains a point of E. On the other hand, taking the contrapositive of the forward direction of (a), $x \notin E^\circ$ means that no neighborhood of x is contained in E, i.e., every neighborhood of x contains a point of E^C, as desired.

Conversely, if every neighborhood of x contains a point of E, then $x \in \overline{E}$ by (b). If, furthermore, every neighborhood of x contains a point of E^C, then there is no neighborhood of x contained in E, and so (a) says that $x \notin E^\circ$. So $x \in \overline{E} \setminus E^\circ = \partial E$, as desired. □

8. It is possible that $\overline{E}^\circ$ is strictly larger than E°. For example, let $E = \mathbb{R}^d \setminus \{0\}$. Then $\overline{E} = \mathbb{R}^d$, and so $\overline{E}^\circ = \mathbb{R}^d$, but $E^\circ = \mathbb{R}^d \setminus \{0\}$.

9. Let A and B be subsets of $\mathbb{R}^d$. Show that $\overline{A \cup B} = \overline{A} \cup \overline{B}$. Is the analogous statement true for $A \cap B$? Justify your answer.

Proof. $\overline{A \cup B}$ is a closed set containing $A \cup B$. In particular, it is a closed set containing A and therefore contains $\overline{A}$. Similarly, $\overline{A \cup B}$ is a closed set containing B and therefore contains $\overline{B}$. So $\overline{A \cup B} \supset \overline{A} \cup \overline{B}$.

For the reverse containment, let $x \in \overline{A \cup B}$. Using Theorem 7.3.10 to characterize closures, x is a limit of some sequence $\{x_n\}$ of points in $A \cup B$. Note that $\{x_n\}$ must have a subsequence $\{x_{n_k}\}$ which is contained entirely in A or entirely in B. Since $x = \lim x_{n_k}$, we have $x \in \overline{A}$ or $x \in \overline{B}$ using Theorem 7.3.10 again. So $x \in \overline{A} \cup \overline{B}$, as desired.

The analogous statement for $A \cap B$ is false. This would be the statement that $\overline{A \cap B}$ equals $\overline{A} \cap \overline{B}$. But consider $A := \{x \in \mathbb{R}^d : \|x\| < 1\}$ and $B := \{x \in \mathbb{R}^d : \|x\| > 1\}$. Then $\overline{A \cap B} = \emptyset$, but $\overline{A} \cap \overline{B} = \{x \in \mathbb{R}^d : \|x\| = 1\}$. □

10. Let A and B be subsets of $\mathbb{R}^d$. Show that $(A \cap B)^\circ = A^\circ \cap B^\circ$. Is the analogous statement true for $A \cup B$? Justify your answer.

Proof. $A^\circ \cap B^\circ$ is an open set because it's an intersection of two open sets. Since $A^\circ \subset A$, $A^\circ \cap B^\circ \subset A$ too. Similarly, $A^\circ \cap B^\circ \subset B$. Hence, $A^\circ \cap B^\circ$ is an open set contained in both A and B, so it is contained in $(A \cap B)^\circ$.

For the reverse containment, $(A \cap B)^\circ$ is an open set contained in both A and B and is therefore contained in A° and B°, as desired.

It is not true that $(A \cup B)^\circ = A^\circ \cup B^\circ$. As a counterexample, consider $A = \mathbb{R}^d \setminus \{0\}$, $B = \{0\}$. Then $(A \cup B)^\circ = (\mathbb{R}^d)^\circ = \mathbb{R}^d$, but $A^\circ = \mathbb{R}^d \setminus \{0\}$ and $B^\circ = \emptyset$, so $A^\circ \cup B^\circ = \mathbb{R}^d \setminus \{0\}$. □

11. Let $\{x_n\}$ be a convergent sequence in $\mathbb{R}^d$ with limit x. Set $A = \{x_1, x_2, x_3, \dots\} \cup \{x\}$. Show that A is a closed set.

Proof. Let $\{y_n\}$ be a convergent sequence of points in A. We want to show that $\lim y_n \in A$. Suppose that $\{y_n\}$ consists of only finitely many distinct points. Then $\{y_n\}$ is a sequence of points in some finite subset $B \subset A$. By Problem 2, B is closed, so $\lim y_n \in B \subset A$, as desired.

Now suppose y_n has infinitely many distinct points. By restricting to a subsequence, we can assume all terms of the sequence $\{y_n\}$ are distinct. Then for any $N \in \mathbb{R}$, there are only finitely many numbers $k \in \mathbb{N}$ with $k \leq N$, so there are at most finitely many terms of $\{y_n\}$ with $y_n \in \{x_{k_n} : k_n < N\} \cup \{x\}$. That is, there is some number N' such that $n > N'$ implies that $y_n = x_{k_n}$ for some $k_n > N$. For any $\epsilon > 0$, there is some $N \in \mathbb{R}$ such that $\|x - x_{k_n}\| < \epsilon$ whenever $k_n > N$, hence whenever n is greater than the corresponding N'. So then $y_n \to x \in A$, as desired.

(I suspect there should be a nicer proof here, but I haven't found it yet). □

14. Find the interior and closure of the set $\mathbb{Q}$ of rationals in $\mathbb{R}$. (I say on Canvas to justify your answer.)

Claim: $\mathbb{Q}^\circ = \emptyset$, and $\overline{\mathbb{Q}} = \mathbb{R}$.

Proof. Let $x \in \mathbb{R}$, and let U be an arbitrary open neighborhood of x. Then for some $r > 0$, $B_r(x) \subset U$, i.e., $(x - r, x + r) \subset U$. Since $(x - r, x + r)$ contains irrational numbers, U is not contained in $\mathbb{Q}$, so by Theorem 7.3.7(a), $x \notin \mathbb{Q}^\circ$. On the other hand, since $(x - r, x + r)$ also contains rational numbers, U contains points in $\mathbb{Q}$, so by Theorem 7.3.7(b), $x \in \overline{Q}$. □

15. If $E \subset \mathbb{R}^d$, show $(\overline{E})^C = (E^C)^\circ$.

Proof. Let $\mathcal{V}$ denote the set of all closed sets $V \subset \mathbb{R}^d$ such that $V \supset E$. Let $\mathcal{U}$ denote the set of all open sets $U \subset \mathbb{R}^d$ such that $U \subset E^C$. We have that V is closed if and only if V^C is open, and $V \supset E$ if and only if $V^C \subset E^C$. So $\mathcal{U} = \{V^C : V \in \mathcal{V}\}$. Now,

$$\begin{aligned}
(\overline{E})^C &= \left(\bigcap_{V \in \mathcal{V}} V \right)^C \\
&= \bigcup_{V \in \mathcal{V}} V^C \text{ (using the infinite version of DeMorgan's Law, Exercise 1.1.7)} \\
&= \bigcup_{U \in \mathcal{U}} U = (E^C)^\circ.
\end{aligned}$$

□

HOMEWORK 3 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 7.4.

2. Let K be a compact subset of $\mathbb{R}^d$ and let $A_1 \supset A_2 \supset \cdots \supset A_j \supset \cdots$ be a nested downward sequence of closed subsets of $\mathbb{R}^d$. Show that if $A_k \cap K \neq \emptyset$ for each k, then $(\bigcap_k A_k) \cap K \neq \emptyset$.

Proof. **Method 1:** For each $k \in \mathbb{N}$, define $B_k := A_k \cap K$. Each B_k is closed since A_k and K are both closed. Also, each B_k is bounded since it is a subset of K, which is bounded. We are given that each B_k is nonempty. Thus, $B_1 \supset B_2 \supset \cdots$ is a nested downward sequence of nonempty closed bounded subsets of $\mathbb{R}^d$, so by Theorem 7.4.6, there exists a point $x \in \bigcap_k B_k$. Of course, since $B_k \subset A_k$ and $B_k \subset K$ for each k, we have $\bigcap_k B_k \subset \bigcap_k A_k$, and $\bigcap_k B_k \subset K$. Hence, $x \in (\bigcap_k A_k) \cap K$, as desired.

Method 2: Suppose $(\bigcap_k A_k) \cap K$ is empty. Then $K \subset (\bigcap_k A_k)^C = \bigcup_k A_k^C$ (by DeMorgan's Law). Since each A_k is closed, A_k^C is open, so this means that $\{A_k : k \in \mathbb{N}\}$ is an open cover for K. Since K is compact, there is a finite subcover, hence a finite collection $k_1 < k_2 < \ldots < k_n$ of natural numbers such that $\bigcup_{i=1}^n A_{k_i}^C \supset K$. Because the A_k's are nested, this union is in fact equal to the single set $A_{k_n}^C$, so $K \subset A_{k_n}^C$. But we are given that $K \cap A_{k_n} \neq \emptyset$, so this is impossible. This contradiction means that $(\bigcap_k A_k) \cap K$ must be nonempty, as desired. □

4. Prove that if K is a compact subset of $\mathbb{R}^d$, then K contains a point of maximal norm. That is, there is a point $x_1 \in K$ such that $\|x\| \leq \|x_1\|$ for all $x \in K$.

Hint: Set $m = \sup\{\|x\| : x \in K\}$ and consider the open balls $B_{m-1/n}(0)$.

Proof. As the hint suggests, take $m = \sup\{\|x\| : x \in K\}$, and consider the collection of open sets $\mathcal{U} := \{B_{m-1/n}(0) : n \in \mathbb{N}\}$. Suppose there is no $x_1 \in K$ with $\|x_1\| = m$. Then $\|x\| < m$ for all $x \in K$, and so for each $x \in K$ there is some $n \in \mathbb{N}$ such that $\|x\| < m - 1/n$. Hence, $K \subset \bigcup_{n\in\mathbb{N}} B_{m-1/n}(0)$. So then $\mathcal{U}$ is an open cover for K, and since K is compact, there is a finite subcover. Because these balls are nested (cf. Exercise 7.4.1 which I went over in class, or the argument in Method 2 of Exercise 2 above), there must in fact be some ball $B_{m-1/n}(0)$ from $\mathcal{U}$ which itself contains K, that is, $K \subset B_{m-1/n}(0)$ for some $n \in \mathbb{N}$. But then $m - 1/n$ is an upper bound for $\{\|x\| : x \in K\}$ which is smaller than the least upper bound m, a contradiction. So there must in fact be some $x_1 \in K$ with $\|x_1\| = m$, as desired (m being an upper bound for $\{\|x\| : x \in K\}$ means that $\|x\| \leq m$ for all $x \in K$). □

5. Prove that if K is a compact subset of $\mathbb{R}^d$ and y is a point of $\mathbb{R}^d$ which is not in K, then there is a closest point to y in K. That is, there is an $x_0 \in K$ such that $\|x_0 - y\| \leq \|x - y\|$ for all $x \in K$.

Proof. Let $m = \inf\{\|x - y\| : x \in K\|\}$. We want to show that there exists some $x_0 \in K$ such that $\|x_0 - y\| = m$ (because m being a lower bound means that $m \leq \|x - y\|$ for all $x \in K$, so this would prove the claim).

Date: September 6, 2019.

Suppose there is no such x_0. Let $\mathcal{U} := \{\overline{B_r}(y)^C : r > m\}$ be the set of complements of closed balls of radii greater than m. By our assumption, for every $x \in K$, $\|x - y\| > m$, hence there is some $r > m$ with $r < \|x - y\|$, and then $x \in \overline{B_r}(y)^C$. That is, $\mathcal{U}$ forms an open cover for K. So then since K is compact, there is a finite subcover, and then because this open cover is nested, there is some $\overline{B_{r_0}}(y)^C$ with $r_0 > m$ which contains K. But then this r_0 is a lower bound for $\{\|x - y\| : x \in K\|\}$ which is greater than the greatest lower bound, and this gives the desired contradiction. □

7. Prove that if K_1, K_2 is a disjoint pair of compact sets, then there exists a disjoint pair of open sets V_1, V_2 such that $K_1 \subset V_1$ and $K_2 \subset V_2$. Hint: Use Theorem 7.4.10.

Proof. K_2^C is an open set containing K_1, so by Theorem 7.4.10, there is some open set V_1 such that $K_1 \subset V_1 \subset \overline{V}_1 \subset K_2^C$. Then $V_2 := \overline{V}_1^C$ is an open set which contains K_2 and is disjoint from the open set $V_1 \supset K_1$, as desired. □

9. Show that it is true that the union of any finite collection of compact subsets of $\mathbb{R}^d$ is compact, but it is not true that the union of an infinite collection of compact subsets is necessarily compact. Show the latter statement by finding an example of an infinite union of compact sets which is not compact.

Solution: For the **first statement**, let $K_1, \ldots, K_n$ be a finite collection of compact subsets of $\mathbb{R}^d$. Then each K_i is closed and bounded (by Heine-Borel). Hence, the union $K := \bigcup_{i=1}^n K_i$ is closed (finite unions of closed sets are closed). Also, K_i being bounded means that there is some $M_i \in \mathbb{R}$ such that $\|x\| \leq M_i$ for all $x \in K_i$. So then K is bounded because $\|x\| \leq \max\{M_1, \ldots, M_n\}$ for all $x \in K$. So K is closed and bounded, hence compact by Heine-Borel again.

Now for the **second statement** (there are of course many possible answers here): Points are compact since they are closed and bounded (in fact, points in any topological space are compact—any open cover obviously has a single-set subcover). So take any non-compact subset S of $\mathbb{R}^d$ with infinitely many points (for example, take $S = \mathbb{R}^d$ itself). Then even though S is not compact, it equals the infinite union of its points, which are compact. □

10. Prove that if A and B are compact subsets of a metric space, then $A \cup B$ and $A \cap B$ are also compact.

Proof. Let $\mathcal{U}$ be an open cover for $A \cup B$. Then $\mathcal{U}$ is also an open cover for A and an open cover for B. Since A and B are compact, there exist finite subcovers $\mathcal{U}_A$ for A and $\mathcal{U}_B$ for B. Then $\mathcal{U}_A \cup \mathcal{U}_B$ is still finite and covers both A and B, hence is a finite subcover for $A \cup B$. This proves that $A \cup B$ is compact.

Now let $\mathcal{V}$ be an open cover for $A \cap B$. The proof of Theorem 7.4.5 works for arbitrary metric spaces (the posting on Canvas mentions this), so A and B being compact implies that they are both closed, hence $A \cap B$ is closed. Then $(A \cap B)^C$ is open, and

$$(A \cap B)^C \cup \bigcup_{V \in \mathcal{V}} V \supset (A \cap B)^C \cup A \cap B,$$

and this is the entire space! Hence, $\widetilde{\mathcal{V}} := \{(A \cap B)^C\} \cup \mathcal{V}$ is in particular an open cover for, say, A, and so there is a finite subcover for A. Since $A \cap B \subset A$, this is also a finite subcover of $\widetilde{\mathcal{V}}$ for $A \cap B$. Since $A \cap B$ has no points in common with $(A \cap B)^C$, we can remove $(A \cap B)^C$ from this finite subcover and get a new finite subcover of $A \cap B$, this time consisting only of sets in $\mathcal{V}$, as desired. □

12. Prove that if X is a compact metric space, then every closed subset of X is also compact.

Proof. Let $A \subset X$ be closed, and let $\mathcal{U}$ be an open cover for A. Then $\widetilde{\mathcal{U}} := \mathcal{U} \cup \{A^C\}$ is an open cover for X, and since X is compact it has a finite subcover $\mathcal{U}'$ for X. This subcover $\mathcal{U}'$ may include $\{A^C\}$, but since $A \cap A^C = \emptyset$, removing $\{A^C\}$ from $\mathcal{U}'$ yields an open cover for A which is now a finite subcover of $\mathcal{U}$ (rather than of $\widetilde{\mathcal{U}}$). Hence, A is compact, as desired. □

15. Consider the metric space of Exercise 7.2.12 ($\mathbb{R}$ with the metric $\delta(x, y) = 0$ if $x = y$, and $\delta(x, y) = 1$ if $x \neq y$). Show that it is complete and bounded but not compact. Thus, the analogue of the Heine-Borel Theorem does not hold in general metric spaces.

Proof. $\mathbb{R}$ is bounded under this metric since $\mathbb{R} \subset B_2(0)$ ($\delta(0, y)$ is 0 or 1, hence < 2, for every $y \in \mathbb{R}$).

We now show that $\mathbb{R}$ with this metric is complete. Let $\{x_n\}$ be a Cauchy sequence. So, taking $\epsilon = 1/2$, we have that there exists an N such that $\delta(x_n, x_m) < 1/2$ for all $n, m > N$. But with this metric, $\delta(x_n, x_m) < 1/2$ implies that $x_n = x_m$. So there is an N such that x_n is always equal to the same number x when $n > N$, and then $\lim x_n = x$. Hence, Cauchy implies convergent, so this space is complete.

Under this metric, points are open sets, since $B_{1/2}(x) = x$ for any $x \in \mathbb{R}$. (This metric makes sense on any set X, and the corresponding metric topology is the discrete topology, meaning that every subset of X is an open set). So $\mathcal{U} := \{\{x\} : x \in \mathbb{R}\}$ is an infinite open cover for $\mathbb{R}$, and it clearly has no proper subcover, let alone an finite subcover. So $\mathbb{R}$ with this metric is not compact. □

HOMEWORK 4 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 7.5.

1. Tell whether or not the set A is connected. If A is not connected, describe its connected components. Justify your answers.

$$A = \{(x, y) \in \mathbb{R}^2 : \|(x, y)\| < 1\} \cup \{(x, y) \in \mathbb{R}^2 : 1 \leq x \leq 2, y = 0\}.$$

Claim: A is connected.

Proof. We know that the set $B := \{(x, y) \in \mathbb{R}^2 : \|(x, y)\| < 1\}$ is connected because every open or closed ball in $\mathbb{R}^d$ is connected (Theorem 7.5.7), and this set B is the open ball $B_1(0)$. We also know that the set $C := \{(x, y) \in \mathbb{R}^2 : 1 \leq x \leq 2, y = 0\}$ is connected, because we know that line-segments in $\mathbb{R}^d$ are connected.

Now, the line segment $\{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, y = 0\}$ is contained in A, and it includes points in B and the point $(1, 0) \in C$. So the connected component of A containing $(1, 0)$ must include both B and C, hence all of A. So A is connected. □

2. Tell whether or not the set A is connected. If A is not connected, describe its connected components. Justify your answers.

$$A = \{(x, y) \in \mathbb{R}^2 : \|(x, y)\| < 1\} \cup \{(x, y) \in \mathbb{R}^2 : 1 < x \leq 2, y = 0\}.$$

Claim: A is not connected. The connected components of A are $B := \{(x, y) \in \mathbb{R}^2 : \|(x, y)\| < 1\}$ and $C := \{(x, y) \in \mathbb{R}^2 : 1 < x \leq 2, y = 0\}$.

Proof. B is connected because it is the open ball $B_1(0)$, and open balls are connected (Theorem 7.5.7). C is connected because it is a line segment in $\mathbb{R}^2$, and line segments (with or without their endpoints) are connected. So since $A = B \cup C$, we just have to show that A is not itself connected, and then we know that B and C must be its connected components.

Let $U := B = B_1(0)$, $V := \{(x, y) \in \mathbb{R}^2 : x > 1\}$. Clearly, U and V are open (U is an open ball, and V being open is similar to Exercise 7.3.1—I could alternatively use $V = B_1((2, 0))$ if you prefer that). We claim that U and V seperate A.

- $U = B$ and $C \subset V$, so $A \subset U \cup V$.
- $U \cap V = \emptyset$, so $(U \cap B) \cap (V \cap C) = \emptyset$.
- $U \cap A = B \neq \emptyset$, and $V \cap A = C \neq \emptyset$.

So U and V do indeed separate A, as desired. □

5. What are the connected components of the complement of the set of integers in $\mathbb{R}$?

Claim: The connected components of $\mathbb{R} \setminus \mathbb{Z}$ are the open intervals $\{(n, n + 1)\}$ for $n \in \mathbb{Z}$.

Date: September 6, 2019.

Proof. Each open interval $(n, n+1)$ is clearly a nonempty connected open subset of $\mathbb{R} \setminus \mathbb{Z}$ (we know that intervals are the connected subsets of $\mathbb{R}$). Furthermore, these intervals are pairwise disjoint, and $\mathbb{R} \setminus \mathbb{Z}$ is the disjoint union of all these,

$$\mathbb{R} \setminus \mathbb{Z} = \bigsqcup_{n \in \mathbb{Z}} (n, n+1),$$

so the claim now follows immediately from Theorem 7.5.11.

Alternatively (without using Theorem 7.5.11, but using what we said above about $(n, n+1)$ being connected): Any connected subset I of $\mathbb{R} \setminus \mathbb{Z}$ larger than $(n, n+1)$ must contain n or $n+1$ because connected subsets of $\mathbb{R} \setminus \mathbb{Z}$ are intervals (so we can pick $x \in (n, n+1)$ and $y \notin (n, n+1)$ and then say that I contains every point between x and y). But n and $n+1$ are integers, so they are not in $\mathbb{R} \setminus \mathbb{Z}$, so there must not be any connected subsets of $\mathbb{R} \setminus \mathbb{Z}$ larger than $(n, n+1)$. So these are the maximal connected subsets, i.e., the connected components.

Another alternative (sketch, using what we said before about $(n, n+1)$ being connected): For each $n \in \mathbb{Z}$, you can show that the open sets $U := (n, n+1)$ and $V := (-\infty, n) \cup (n+1, \infty)$ separate $(n, n+1)$ from the rest of $\mathbb{R} \setminus \mathbb{Z}$, so $(n, n+1)$ must be a connected component. □

7. Which subsets of $\mathbb{R}$ are both compact and connected? Justify your answer.

Claim: The compact connected subsets of $\mathbb{R}$ are exactly the intervals of the form $[a, b]$, $a \leq b$, along with the empty set (which I guess could be viewed as $[a, b]$ with $a > b$). By the interval $[a, a]$, I mean the single point $\{a\}$.

Proof. We know that compact in $\mathbb{R}$ is equivalent to closed and bounded (Heine-Borel Theorem), and being nonempty and connected in $\mathbb{R}$ is equivalent to being an interval (Theorem 7.5.4), so the nonempty compact connected subsets of $\mathbb{R}$ must be exactly the nonempty closed bounded intervals in $\mathbb{R}$. These are precisely the intervals described in the claim.

For the empty set, note that in any topological space, the empty set is compact: for any open cover of $\emptyset$ (note: every collection of open sets is an open cover for $\emptyset$), the empty sub-collection is a finite subcover of $\emptyset$. In terms of Heine-Borel (special for $\mathbb{R}^d$): the empty set is closed (in any topological space) and bounded (this is vacuously true), so it is compact. Also, the empty set is connected: If you try to separate it with some U and V, you have $U \cap \emptyset = \emptyset$, contradicting one of the conditions for separating sets. □

9. Prove that if E is an open connected subset of $\mathbb{R}^d$, then each pair of points in E can be connected by a piecewise linear path in E. Hint: Fix a point $x_0 \in E$ and consider two sets: (1) the set U of all points in E that can be connected to x_0 by a piecewise linear path in E and (2) the set V of points in E that cannot be connected to x_0 by a piecewise linear path in E.

Proof. Our goal is to show that $V = \emptyset$. To do this, we will show that if $V \neq \emptyset$, then U and V separate E, contradicting that E is connected.

Since U is defined to be a subset of E, and V is essentially defined as $E \setminus U$, we have $E = U \sqcup V$. This immediately implies conditions (a) and (b) from Definition 7.5.1). Condition (c) follows because $x_0 \in U \cap E$, so $U \cap E \neq \emptyset$, and we have assumed that $V \neq \emptyset$, and $V \subset E$, so then $V \cap E \neq \emptyset$. So now we just have to check that U and V are open.

For any $x \in U$, since E is open, there is some $\delta > 0$ such that $B_\delta(x) \subset E$. Since any point in $B_\delta(x)$ can be connected by a line segment to x (a segment of a radius in fact), $x \in U$ implies that all of $B_\delta(x)$ is in U. So U must be open.

Similarly, for any $x \in V$, we have some $\delta > 0$ such that $B_\delta(x) \subset E$. If some $y \in B_\delta(x)$ were in U, then since y and x can be connected by a line segment, this would mean $x \in U$, hence $x \notin V$, a contradiction. So $B_\delta(x)$ is indeed contained in V, hence V is open, as desired. □

10. Prove that the closure of a connected set is connected.

Proof. Let E be a set. Suppose $\overline{E}$ is not connected. Then there exist open sets U, V separating $\overline{E}$. Then we have

- $E \subset \overline{E} \subset U \cup V$,
- $(E \cap U) \cap (E \cap V) \subset (\overline{E} \cap U) \cap (\overline{E} \cap V) = \emptyset$, and
- $\overline{E} \cap U \neq \emptyset$ and $\overline{E} \cap V \neq \emptyset$. Hence, by Theorem 7.3.7(b), $E \cap U \neq \emptyset$, and $E \cap V \neq \emptyset$.

But then U and V separate E. So $\overline{E}$ not connected implies E not connected, hence E connected implies $\overline{E}$ connected. □

15'. Find a nonempty, totally disconnected subset of $\mathbb{R}$ which has no isolated points.

Note: The original problem also specified compact (and forgot to specify nonempty). This makes things quite hard, so I modified it as above.

Solution: Consider the subset $\mathbb{Q} \subset \mathbb{R}$. For any $x \in \mathbb{Q}$, any neighborhood $U \ni x$ contains an interval $(x - \epsilon, x + \epsilon)$ for some $\epsilon > 0$, and this interval contains other rational numbers (Exercise 1.4.7). Since every neighborhood of x contains other points of $\mathbb{Q}$, x is not an isolated point.

To show totally disconnected, consider any distinct points $x, y \in \mathbb{Q}$. There is an irrational number z between x and y (Exercise 1.4.9), so the sets $(-\infty, z)$ and (z, ∞) separate $\mathbb{Q}$ with x and y in different connected components. (Alternatively: If x and y are in the same connected component, then, since any nonempty connected subset of $\mathbb{R}$ is an interval, everything between x and y must be in this connected component too, contradicting that there are irrational numbers between x and y). So no two distinct points share a connected component, hence every connected component must be a single point. □

Note: The above example is not compact since it is neither closed nor bounded. Intersecting with $[0, 1]$ would give a bounded example, but getting something closed is quite a bit harder, which is why I modified the problem.

There does exist a nonempty, compact, totally disconnected subset of $\mathbb{R}$ with no isolated points. The standard example is the Cantor set. See `https://en.wikipedia.org/wiki/Cantor_set` for its construction. Briefly, the idea is that you start with a closed bounded interval, then remove the open interval consisting of the middle third of this interval, then remove the open middle thirds of each of the new intervals, and repeat forever.

HOMEWORK 5 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 8.1.

1. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

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

Is this function continuous at $(0,0)$? Justify your answer.

Solution: Yes, f is continuous at $(0,0)$. One can show that $|xy| \leq \frac{1}{2}(x^2+y^2)$ for all $x, y \in \mathbb{R}$ (I did this in class using Cauchy-Schwartz applied to $|(x,y)\cdot(y,x)|$, and your book does it a different way in Example 8.1.4, so it's ok if you use this without proof). So then

$$\left|\frac{xy^2}{x^2+y^2} - 0\right| = \frac{|xy||y|}{x^2+y^2} \leq \frac{\frac{1}{2}(x^2+y^2)|y|}{x^2+y^2} = \frac{1}{2}|y|. \tag{1}$$

Set $\delta = 2\epsilon$. Then $\|(x,y)-(0,0)\| < \delta$ means that $\sqrt{x^2+y^2} < 2\epsilon$. Since $x^2 \geq 0$, this implies that $\sqrt{y^2} < 2\epsilon$. Since $\sqrt{y^2} = |y|$, this implies that $\frac{1}{2}|y| < \epsilon$. So then $\|(x,y)-(0,0)\| < \delta$ implies that $\frac{1}{2}|y| < \epsilon$, hence $|f(x,y) - f(0,0)| < \epsilon$ by (1), as desired.

3. Does the function $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$ defined by

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

have a limit as (x,y) approaches $(0,0)$? Justify your answer.

Solution: No, $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist. Let us restrict to the line $y=0$. Then the limit becomes

$$\lim_{x\to 0} f(x,0) = \lim_{x\to 0} \frac{x}{\sqrt{x^2}} = \lim_{x\to 0} \frac{x}{|x|},$$

and this does not exist: $\frac{x}{|x|} = 1$ for $x > 0$ and -1 for $x < 0$, so the right- and left-hand limits are 1 and -1, respectively. In particular, they are not equal, so the limit cannot exist.

If we approach $(0,0)$ along the line $y = ax$ for $a \in \mathbb{R}$, we get

$$\lim_{x\to 0} f(x,ax) = \lim_{x\to 0} \frac{x}{\sqrt{x^2+a^2x^2}} = \lim_{x\to 0} \frac{x}{|x|\sqrt{1+a^2}},$$

and even this one-dimensional limit does not exist: The right-hand limit is $\lim_{x\to 0^+} \frac{x}{|x|\sqrt{1+a^2}} = \frac{1}{\sqrt{1+a^2}}$, and the does not equal the left-hand limit, which is $\lim_{x\to 0^-} \frac{x}{|x|\sqrt{1+a^2}} = -\frac{1}{\sqrt{1+a^2}}$. Since the limit does not exist along a line $y = ax$ through $(0,0)$, it does not exist in $\mathbb{R}^2$.

Note: You could **alternatively** consider the one-sided limit along a line $y = ax$, and get something like $\frac{1}{\sqrt{1+a^2}}$ for the limit (or negative this if you have $x \to 0$ from the left hand side). The fact that you get different values for different a then implies that the limit cannot exist.

Date: September 6, 2019.

As another **alternative**, you could work directly from the definition. If you take $\epsilon = \frac{1}{2}$, then for any $\delta > 0$, the points $(x, y) = (\pm\frac{\delta}{2}, 0)$ satisfies $\|(x, y)\| < \delta$, but $f(\frac{\delta}{2}, 0) = 1$ while $f(-\frac{\delta}{2}, 0)$ equals -1, and there is no L within ϵ of both of these.

5. For the function $f : \mathbb{R}^2 \setminus \{(0, 0)\} \to \mathbb{R}$ defined by $f(x, y) = \frac{x^2 y}{x^4+y^2}$, show that f has limit 0 as $(x, y) \to (0, 0)$ along any straight line through the origin, but that it does not have a limit as $(x, y) \to (0, 0) \in \mathbb{R}^2$.

Note: Your book had all of $\mathbb{R}^2$ as the domain of f, but that can't be right since f is not defined at $(0, 0)$.

Solution: Along the line $x = 0$, we have $f(x, y) = f(0, y) = 0$ for all $y \neq 0$, so the limit is 0. Along the line $y = ax$ for $a \in \mathbb{R}$, we have

$$\lim_{x\to 0} f(x, ax) = \lim_{x\to 0} \frac{ax^3}{x^4 + a^2x^2} = \lim_{x\to 0} \frac{ax}{x^2 + a^2} = 0.$$

This proves that the limit is 0 along any straight line through the origin.

Now consider the limit along the parabola $y = x^2$ through the origin. We have

$$\lim_{x\to 0} f(x, x^2) = \lim_{x\to 0} \frac{x^4}{x^4 + x^4} = \lim_{x\to 0} \frac{x^4}{2x^4} = \frac{1}{2} \neq 0.$$

Since the limit along $y = x^2$ is different than along $y = x$, the limit in $\mathbb{R}^2$ must not exist.

9. Prove that a is a limit point of a set $D \subset \mathbb{R}^p$ if and only if there is a sequence of points in D but not equal to a which converges to a.

Proof. Suppose a is a limit point of D. Then for each $n \in \mathbb{N}$, $B_{1/n}(a)$ contains points of $D \setminus \{a\}$. Pick one such point x_n for each n. Then the sequence $\{x_n\}$ is a sequence of points in $D \setminus \{a\}$. Since $\|x_n - a\| < \frac{1}{n}$ and $\lim_{n\to\infty} \frac{1}{n} = 0$, we have $\lim_{n\to\infty} x_n = a$, as desired.

Conversely, suppose there exists a sequence of points $\{x_n\}$ in $D \setminus \{a\}$ which converges to a. Let U be a neighborhood of a. Then since U is open, there exists an $r > 0$ such that $B_r(a) \subset U$. Since $x_n \to a$, there exists some N such that $n > N$ implies $x_n \in B_r(a)$, hence $x_n \in U$. Since U was an arbitrary neighborhood and this x_n is in $D \setminus \{a\}$, this means that a is a limit point of D, as desired. □

Section 8.2.

1. If $A = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y \leq 1\}$, which of the following sets cannot be the image of the set A under a continuous function $F : A \to \mathbb{R}^2$? Justify your answers.

(a) $\overline{B}_2(0, 0)$.
(b) $B_1(0)$. **Note:** By 0 here, I think your book meant the 0-vector $(0, 0)$ in $\mathbb{R}^2$. Since part (a) used $(0, 0)$ for this, I think some people misinterpreted this as $B_1(0)$ as $(-1, 1)$ in $\mathbb{R}^1$, rather than the unit disk in $\mathbb{R}^2$. People should not lose points for this.
(c) $\{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, 0 \leq y\}$.
(d) $\overline{B}_1(0, 0) \cup \overline{B}_1(3, 0)$.
(e) $\{(t, t) \in \mathbb{R}^2 : t \in \mathbb{R}; 0 \leq t \leq 1\}$.

Solution: I claim that (b), (c), and (d) cannot be the image of A under a continuous function $F : A \to \mathbb{R}^2$.

Note that A is closed and bounded, hence compact. So by Theorem 8.2.3, $F(A)$ must be compact too, hence closed and bounded. However, (b) is not closed, and (c) is not bounded, so that proves that these two cannot be $F(A)$ for any continuous F.

For (d), note that A is connected, so $F(A)$ is connected by Theorem 8.2.7. But the set in (d) is not connected, so it cannot be $F(A)$ for any continuous F.

Note: I said that you do not have to justify that (a) and (e) can be the image of a continuous function F, but let me justify that here anyway. For (a), $\overline{B}_2(0,0)$ is the image of $F : A \to \mathbb{R}^2$ defined by $F(x,y) := (2x\cos(2\pi y), 2x\cos(2\pi y))$ (in terms of polar coordinates, $2x$ is the radius r and $2\pi y$ is the angle θ). For (e), the set there is the image of $F(x,y) := (x,x)$.

3. If K is a compact, connected subset of $\mathbb{R}^p$, and $f : K \to \mathbb{R}$ is a continuous function, what can you say about $f(K)$?

Solution: By Theorem 8.2.3, $F(K)$ is compact, hence closed and bounded by the Heine-Borel theorem. By Theorem 8.2.7, $f(K)$ is connected, hence an interval in $\mathbb{R}$. So $F(K)$ is a closed bounded interval in $\mathbb{R}$, i.e., a set of the form $[a,b]$ for $a,b \in \mathbb{R}$ (if you just said this last bit with no justification, that's probably ok).

Note: If $a = b$, this set is a point. If $a > b$, you can interpret this as the empty set (or you can mention the empty set as a possibility separately—you get $f(K) = \emptyset$ if and only if $K = \emptyset$).

5. The image of a compact set under a continuous function is compact, hence closed, by Theorem 8.2.3. Is the image of a closed set under a continuous function necessarily closed? Prove that it is or give an example where it is not.

Solution: The image of a closed set under a continuous function is not necessarily closed. Here are some possible counterexamples (the first two I actually talked about in class when going over Exercise 8.2.4

- $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \arctan(x)$, in which case the domain is all of $\mathbb{R}$, which is closed, but the image is $(-\pi/2, \pi/2)$, which is not closed.
- $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{1}{1+x^2}$. Then the domain is again all of $\mathbb{R}$, which is closed, but the image is $(0,1]$, which is not closed.
- The Gaussian $f(x) = e^{-x^2}$ has domain $\mathbb{R}$ but image $(0,1]$.
- $f : \mathbb{R}_{\geq 1} \to \mathbb{R}$, $f(x) = \frac{1}{x}$ has domain $[1,\infty)$, which is closed, but image $(0,1]$, which is not closed.

6. Is the image of an open set under a continuous function necessarily an open set? Prove that it is or give an example where it is not.

Solution: The image of an open set under a continuous function does not need to be open. For example, consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 0$ for all $x \in \mathbb{R}$. Then the image of the open set $\mathbb{R}$ is just the single point $\{0\}$, which is not open.

More generally, for any nonempty open set U and f any constant map $f : U \to \mathbb{R}^q$, the image of U will be a point, which is not open.

8. Prove that if $f : T \to \mathbb{R}$ is a continuous real-valued function on the unit circle $T = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$, then there is a pair of diametrically opposed points (x,y) and $(-x,-y)$ on T at which f has the same value.

Proof. As in the hint I posted, consider $g(x, y) := f(x, y) - f(-x, -y)$. We want to show that there is a point $(x, y) \in T$ where $g(x, y) = 0$, since this is equivalent to $f(x, y) = f(-x, -y)$.

Pick any point $(x_0, y_0) \in T$. If $g(x_0, y_0) = 0$, we are done, so let's suppose $g(x_0, y_0) \neq 0$. Note that $g(x, y) = -g(-x, -y)$, so if $g(x_0, y_0) > 0$, then $g(-x_0, -y_0) < 0$, and similarly, if $g(x_0, y_0) < 0$, then $g(-x_0, -y_0) > 0$. Either way, there is a point where $g > 0$ and a point where $g < 0$. Then since g is continuous,[1] and T is connected,[2] $g(T)$ is an interval (cf. Corollary 8.2.8), and so $g(T)$ must contain 0, as desired. □

Note: By composing with the parameterization of T I mentioned, you can actually turn this into a single variable Intermediate Value Theorem problem.

[1] g is continuous because f is continuous, and $-x$ and $-y$ are continuous, and compositions of continuous functions are continuous, and sums and scalar multiples of continuous functions are continuous. I'm ok with you just saying that g is continuous though.

[2] As I posted, T is the image of the continuous function $\gamma(t) = (\cos(t), \sin(t))$ with connected domain $[0, 2\pi]$, so T must itself be connected. I'm ok with you just saying T is connected though.

HOMEWORK 6 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 8.2.

10. Is the function $f : \mathbb{R}^2 \setminus \{(2,0)\} \to \mathbb{R}$ defined by

$$f(x,y) = \frac{1}{(x-2)^2 + y^2}$$

uniformly continuous on $B_1(0,0)$? Is it uniformly continuous on $B_2(0,0)$? Justify your answers.

Solution: f is continuous on $B_2(0,0)$, so in particular, it is continuous on $\overline{B}_1(0,0)$. By Theorem 8.2.12, since $\overline{B}_1(0,0)$ is compact, it follows that f is uniformly continuous on $\overline{B}_1(0,0)$, hence also on $B_1(0,0)$. [You could also use Theorem 8.2.13, which says that f admitting a continuous extension from $B_1(0,0)$ to $\overline{B}_1(0,0)$ implies that f is uniformly continuous on $B_1(0,0)$.]

On the other hand, $\lim_{(x,y)\to(2,0)} f(x,y) = \infty$, so there is no continuous extension of f to a domain that includes the point $(2,0)$. Since $(2,0) \in \overline{B}_2(0,0)$, f does not admit a continuous extension from $B_2(0,0)$ to the closure $\overline{B}_2(0,0)$, so by Theorem 8.2.13, f is not uniformly continuous on $B_2(0,0)$. □

Section 8.3

1. Show that the sequence $\{\gamma_n(t)\}$, where

$$\gamma_n(t) = \left(\frac{1}{1+nt}, \frac{t}{n}\right),$$

does not converge uniformly on $[0,1]$.

Proof. For $t \neq 0$, $\{\gamma_n(t)\}$ converges to $(0,0)$, because (you don't really have to justify this), for any fixed t, the numerators of the entries of $\gamma_n(t)$ are constant while the denominators go to infinity. However, $\gamma_n(0) = (1,0)$ for all n, so $\gamma_n(0)$ converges to $(1,0)$. Hence, $\gamma_n(t)$ converges pointwise to the function

$$\gamma(t) := \begin{cases} (1,0) \text{ if } t = 0 \\ (0,0) \text{ if } t \neq 0. \end{cases}$$

So if γ_n converges uniformly, it must be to this $\gamma(t)$. But each γ_n is continuous, so by Theorem 8.3.4, if the convergence were uniform, γ would have to be continuous, and this is not the case. So the convergence is not uniform. □

2. Show that the sequence $\{\lambda_n(t)\}$, where

$$\lambda_n(t) = \left(\frac{t}{1+nt}, \frac{t}{n}\right),$$

does converge uniformly on $[0,1]$.

Date: September 6, 2019.

Proof. For each n, if $t=0$, we have $|\frac{t}{1+nt}|=0$, and if $t\neq 0$, we have $|\frac{t}{1+nt}|<\frac{t}{nt}=\frac{1}{n}$. In any case, $|\frac{t}{1+nt}|<\frac{1}{n}$. So for all n and t, we have

$$\|\lambda_n(t)-(0,0)\|=\sqrt{\left(\frac{t}{1+nt}\right)^2+\left(\frac{t}{n}\right)^2}<\sqrt{\left(\frac{1}{n}\right)^2+\left(\frac{1}{n}\right)^2}=\sqrt{\frac{2}{n^2}}=\frac{\sqrt{2}}{n}.$$

Since $\frac{\sqrt{2}}{n}\to 0$, λ_n converges uniformly to the 0-function $(0,0)$ by Theorem 8.3.2. [Alternatively, given $\epsilon>0$, take $N=\frac{\sqrt{2}}{\epsilon}$, and then you can get uniform convergence to $(0,0)$ directly from the definition (8.3.1)]. □

3. Does the sequence $\{F_k(x,y):=(k^{-1}\sin kx, k^{-1}\cos ky)\}$ converge pointwise on $\mathbb{R}^2$? Does it converge uniformly on $\mathbb{R}^2$? Justify your answer.

Solution: The sequence converges pointwise and uniformly to $(0,0)$. We have

$$\|F_k(x,y)-(0,0)\|=\sqrt{(k^{-1}\sin kx)^2+(k^{-1}\cos ky)^2)}\leq\sqrt{k^{-2}+k^{-2}}=\frac{\sqrt{2}}{k},$$

and since $\frac{\sqrt{2}}{k}\to 0$, F_k converges uniformly (hence pointwise) to $(0,0)$ by Theorem 8.3.2.

5. Find $\|F\|_D$ if $D=\{(x,y)\in\mathbb{R}^2 : x^2+y^2\leq 1\}$ and $F:\mathbb{R}^2\to\mathbb{R}^2$ is defined by $F(x,y)=(x+1,y+1)$.

Solution: I've realized that this is really a Calc III problem. From what we've covered, you should know that since D is compact and F is continuous, $\|F\|_D$ will be the maximum (not just supremum) value of F on D. To find this maximum value, you can either just use your geometric intuiton, or use methods from Calc III, or some combination. I'll go over both approaches, and will ask the grader not to grade this one for correctness since, as I mentioned, it's not really focused on stuff we've covered in this class.

Geometric intuition approach: The image $F(D)$ is the disk $\overline{B}_1(1,1)$, and $\|F\|_D$ will be the maximal norm of points in this disk. It's not hard to guess that this maximal norm is attained at the point $(1+\frac{1}{\sqrt{2}},1+\frac{1}{\sqrt{2}})$ (i.e., at $F(1/\sqrt{2},1/\sqrt{2})$), and this norm is $\sqrt{(1+1/\sqrt{2})^2+(1+1/\sqrt{2})^2}=1+\sqrt{2}$.

More precise Calc III approach: You want to find where $\|F(x,y)\|$ is maximized on D, or equivalently, where $\|F(x,y)\|^2$ is maximized on D. We have

$$\|F(x,y)\|^2=(x+1)^2+(y+1)^2=x^2+y^2+2x+2y+2.$$

Taking the partial derivatives, we get $F_x(x,y)=2x+2$ and $F_y(x,y)=2y+2$, and these are never 0 on D. So the extrema of F must occur on the boundary of D, i.e., where $x^2+y^2=1$. Substituting this into x^2+y^2 above reduces $\|F(x,y)\|^2$ to simply $2x+2y+3$, and now we want to maximize this on the unit circle. You can do this using your geometric intuition, or by drawing level sets, or you can solve $x^2+y^2=1$ to get $y=\pm\sqrt{1-x^2}$, substitute, and use Calc I methods to find the max. Alternatively, let's use the method of Lagrange multipliers.

We want to maximize $f(x,y)=2x+2y+3$ subject to the constraint $g(x,y)=x^2+y^2=1$. Taking the gradients, we have $\nabla f(x,y)=(2,2)$, and $\nabla g(x,y)=(2x,2y)$. We solve $\nabla g=\lambda\nabla f$ to get $x=y=\lambda$. Substituting into the constraint $x^2+y^2=1$ yields $2x^2=1$, hence $x=\pm\frac{1}{\sqrt{2}}$ and $y=x$. The point where x and y equal $-1/\sqrt{2}$ will yield the minimum for $\|F(x,y)\|$, and the max occurs at $(1/\sqrt{2},1/\sqrt{2})$. Plugging this in to $\|F(x,y)\|$ yields $1+\sqrt{2}$. □

7. Prove that if $\{F_n\}$ is a sequence of bounded functions from a set $D\subset\mathbb{R}^p$ into $\mathbb{R}^q$, and if $\{F_n\}$ converges uniformly to F on D, then F is also bounded.

Proof. By the definition of uniform convergence, given any $\epsilon > 0$, there is some N such that $n > N$ implies $\|F(x) - F_n(x)\| < \epsilon$ for all $x \in D$. Let's pick one choice of ϵ, say $\epsilon = 1$, and then pick one choice of $n > N$. So we have $\|F(x) - F_n(x)\| < 1$ for all $x \in D$. Since $F_n(x)$ is bounded, there is some number M such that $\|F_n(x)\| < M$ for all $x \in D$. So using the triangle inequality, we have

$$\|F(x)\| = \|F(x) - F_n(x) + F_n(x)\| \leq \|F(x) - F_n(x)\| + \|F_n(x)\| < 1 + M$$

for all $x \in D$. So then $\|F(x)\|$ is bounded by $1 + M$. □

HOMEWORK 7 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 8.4:

For Exercises 8.4.1–8.4.5, we have the following matrices:

$$A = \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix},\ B = \begin{pmatrix} 2 & 5 \\ -2 & 2 \end{pmatrix},\ C = \begin{pmatrix} 1 & -1 \\ 4 & -6 \\ -1 & 2 \end{pmatrix},\ D = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 3 \end{pmatrix}.$$

1. Find $2A + B$, $A - B$, and BA.

Solution:

$$2A + B = 2\begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 5 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 6 & -2 \\ 4 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 5 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 8 & 3 \\ 2 & 4 \end{pmatrix}.$$

$$A - B = \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 5 \\ -2 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -6 \\ 4 & -1 \end{pmatrix}.$$

$$BA = \begin{pmatrix} 2 & 5 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 16 & 3 \\ -2 & 4 \end{pmatrix}.$$

2. Find $\det A$, $\det B$, A^{-1}, and B^{-1}.

Solution:

$\det A = \det \begin{pmatrix} 3 & -1 \\ 2 & 1 \end{pmatrix} = 3 - (-2) = 5.$

$\det B = \det \begin{pmatrix} 2 & 5 \\ -2 & 2 \end{pmatrix} = 4 - (-10) = 14.$

$A^{-1} = \frac{1}{5}\begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{-2}{5} & \frac{3}{5} \end{pmatrix}.$

$B^{-1} = \frac{1}{14}\begin{pmatrix} 2 & -5 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} \frac{2}{14} & \frac{-5}{14} \\ \frac{2}{14} & \frac{2}{14} \end{pmatrix}.$

3. Find CD and DC.

Solution:

Date: September 6, 2019.

$$CD = \begin{pmatrix} 1 & -1 \\ 4 & -6 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & -1 & -2 \\ 14 & -6 & -14 \\ -4 & 2 & 5 \end{pmatrix}.$$

$$DC = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 4 & -6 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

5. Find $\det CD$.

Solutions: Using the computation of CD in Exercise 8.4.3 above:

$$\det CD = \det \begin{pmatrix} 3 & -1 & -2 \\ 14 & -6 & -14 \\ -4 & 2 & 5 \end{pmatrix} = (-90 - 56 - 56) - (-48 - 70 - 84) = (-202) - (-202) = 0.$$

Alternatively: Since C is a 3×2 matrix, its rank is at most 2, i.e., the image of C has dimension at most 2. So CD is a 3×3 matrix of rank at most 2 (because its image is contained in the image of C), meaning that CD must be singular, i.e., $\det CD$ must be 0.

7. Is the function $F : \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = (x + y, x - y)$ a linear transfromation? If so, what is its matrix?

Solution: Yes, F is linear, because $F(x,y) = \begin{pmatrix} x+y \\ x-y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$. This also shows that $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is the matrix for F.

Alternatively, you could just say that F is linear because each term of each component has degree 1. Then since $F(1, 0) = (1, 1)$ and $F(0, 1) = (1, -1)$, the matrix must be as above [using Equation 8.4.3 in your book, i.e., that the j-th column of the matrix for F is $F(e_j)$].

9. What is the matrix for the linear transfromation of $\mathbb{R}^2$ which reflects each point through the diagonal line $y = x$? (This transformation interchanges the x and y coordinates of each point).

Solution: This transformation takes $e_1 = (1, 0)$ to $(0, 1)$, and it takes $e_2 = (0, 1)$ to $(1, 0)$. These form the columns of the matrix for the transformation, so the matrix is $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Alternatively, you could show that this is the right matrix by noting that $\begin{pmatrix} y \\ x \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

14. Prove that if K and L are linear transformations from $\mathbb{R}^p \to \mathbb{R}^q$, then $\|K + L\| \le \|K\| + \|L\|$.

Proof. **Method 1:** We have

$$\begin{aligned}\|K+L\| &:= \sup\left\{\frac{\|(K+L)(x)\|}{\|x\|} : x\in\mathbb{R}^p\setminus\{0\}\right\}\\ &= \sup\left\{\frac{\|K(x)+L(x)\|}{\|x\|} : x\in\mathbb{R}^p\setminus\{0\}\right\}\\ &\le \sup\left\{\frac{\|K(x)\|}{\|x\|}+\frac{\|L(x)\|}{\|x\|} : x\in\mathbb{R}^p\setminus\{0\}\right\}.\end{aligned}$$

In this last step, we used the triangle inequality to say that $\|K(x)+L(x)\|\le\|K(x)\|+\|L(x)\|$ ($K(x)$ and $L(x)$ live in $\mathbb{R}^q$, which we're assuming is a normed vector space, and norms, by definition, satisfy the triangle inequality). Also, this is subtle (and I'm fine with you not mentioning it), but we're using what I called Lemma 1 in the announcements I posted in order to say that the inequality of the sup's follows from $\frac{\|K(x)+L(x)\|}{\|x\|}\le\frac{\|K(x)\|}{\|x\|}+\frac{\|L(x)\|}{\|x\|}$.

Finally (using Theorem 1.5.10(c)), we can say that this last line above is

$$\le \sup\left\{\frac{\|K(x)\|}{\|x\|} : x\in\mathbb{R}^p\setminus\{0\}\right\}+\sup\left\{\frac{\|L(x)\|}{\|x\|} : x\in\mathbb{R}^p\setminus\{0\}\right\},$$

and this equals $\|K\|+\|L\|$. □

Method 2: Alternatively, think of $\|L\|$ as the smallest number such that $\|L(x)\|\le\|L\|\|x\|$ for all $x\in\mathbb{R}^p$, and similarly for $\|K\|$ and $\|L+K\|$. Then for each x, we have

$$\|(L+K)(x)\| = \|L(x)+K(x)\|\le\|L(x)\|+\|K(x)\|\le\|L\|\|x\|+\|K\|\|x\| = (\|L\|+\|K\|)\|x\|,$$

and so $\|L+K\|\le\|L\|+\|K\|$, as desired. □

15. Prove that if $K:\mathbb{R}^p\to\mathbb{R}^q$ and $L:\mathbb{R}^q\to\mathbb{R}^r$ are linear transformations, then $\|L\circ K\|\le\|L\|\|K\|$.

Proof. By Lemma 2 in the announcement I posted, the right-hand side can be written as

$$\begin{aligned}\|L\|\|K\| &= \left(\sup\left\{\frac{\|L(y)\|}{\|y\|} : x\in\mathbb{R}^q\setminus\{0\}\right\}\right)\left(\sup\left\{\frac{\|K(x)\|}{\|x\|} : x\in\mathbb{R}^p\setminus\{0\}\right\}\right)\\ &\ge \sup\left\{\frac{\|L(y)\|}{\|y\|}\cdot\frac{\|K(x)\|}{\|x\|} : x\in\mathbb{R}^p\setminus\{0\}, y\in\mathbb{R}^q\setminus\{0\}\right\}.\end{aligned}$$

Now, restricting to the subset where $y=K(x)$ (and using Theorem 1.5.7(e)), we get that this last line above is

$$\begin{aligned}&\ge \sup\left\{\frac{\|L(K(x))\|}{\|K(x)\|}\cdot\frac{\|K(x)\|}{\|x\|} : x\in\mathbb{R}^p\setminus\{0\}\right\}\\ &= \sup\left\{\frac{\|L(K(x))\|}{\|x\|} : x\in\mathbb{R}^p\setminus\{0\}\right\},\end{aligned}$$

and this equals $\|L\circ K\|$, this proving the desired inequality. □

Section 8.5:

1. Do the vectors $(1,2,1)$, $(2,0,1)$, and $(1,-1,1)$ form a basis for $\mathbb{R}^3$? Justify your answer.

Solution: $\det\begin{pmatrix}1 & 2 & 1\\ 2 & 0 & -1\\ 1 & 1 & 1\end{pmatrix} = (0-2+2)-(0-4-1) = 5 \neq 0$, so the vectors do form a basis.

Note: Arranging the vectors into a matrix and checking whether or not the determinant is 0 is a standard way to check if the vectors form a basis. See Example 8.5.1 for why this works.

2. Do the vectors $(1,2,1)$, $(2,0,1)$, and $(0,4,1)$ form a basis for $\mathbb{R}^3$? Justify your answer.

Solution: $\det\begin{pmatrix}1 & 2 & 0\\ 2 & 0 & 4\\ 1 & 1 & 1\end{pmatrix} = (0+8+0)-(0+4+4) = 8-8 = 0$, so the vectors do not form a basis (see the note from the previous problem).

3. What is the rank of the matrix $\begin{pmatrix}1 & 2 & 1\\ 2 & 3 & -1\\ 1 & 1 & -2\end{pmatrix}$?

Solutions:

Method 1: The determinant is $(-6-2+2)-(3-8-1) = -6-(-6) = 0$, so the rank is less than 3.

On the other hand, the determinant of the top-left 2×2 submatrix is $\det\begin{pmatrix}1 & 2\\ 2 & 3\end{pmatrix} = 3-4 = -1 \neq 0$. So by Theorem 8.5.5, the rank is 2.

Method 2: The row operation $R_2 \mapsto R_2 - R_1 - R_3$ (i.e., subtracting the first and third rows from the second row) yields the matrix $\begin{pmatrix}1 & 2 & 1\\ 0 & 0 & 0\\ 1 & 1 & -2\end{pmatrix}$. Since the first and third rows are not proportional to each other, they cannot be linearly dependent (this trick works when there are only two vectors), so the rank must be 2.

Method 3: Picking up with the matrix from Method 2, we can subtract the third row from the first row to get $\begin{pmatrix}0 & 1 & 3\\ 0 & 0 & 0\\ 1 & 1 & -2\end{pmatrix}$. From here, subtract the first row from the third, and the re-order the rows to get the reduced row-echelon form: $\begin{pmatrix}1 & 0 & -5\\ 0 & 1 & 3\\ 0 & 0 & 0\end{pmatrix}$. Since this has 2 nonzero rows, the rank is 2.

4. What is the rank of the matrix $\begin{pmatrix}1 & -2 & 3\\ -2 & 4 & -6\end{pmatrix}$?

Solution: The second row is (-2) times the first row, so the rows are linearly dependent and the rank is just 1.

(As in the previous problem, other methods are possible, but this is by far the easiest method here).

7. Find parametric equations for the line in $\mathbb{R}^3$ containing both $(1,1,1)$ and $(3,-1,3)$.

Solution: The vector $u = (3,-1,3) - (1,1,1) = (2,-2,2)$ is tangent to the line, and $(1,1,1)$ is on the line, so the line is given parametrically by

$$\gamma(t) = (1,1,1) + t(3,-1,3).$$

Writing this as your answer is ok. Here are some **other acceptable forms** (there are many others too):

$$\gamma(t) = (1+3t, 1-t, 1+3t).$$

Alternatively,

$$x(t) = 1+3t, y(t) = 1-t, z(t) = 1+3t.$$

Alternatively, changing your base-point and negating u,

$$\gamma(t) = (3,-1,3) + t(-2,2,-2),$$

or any of the analogous equivalent forms.

HOMEWORK 8 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 9.1:

1. If $f(x,y) = \sqrt{x^2+y^2}$, find $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$. Are there any points in the plane where they don't exist?

Solution: $\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2+y^2}}$, and $\frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2+y^2}}$. These exist everywhere except for $(0,0)$, because at this point the denominators equal 0.

Alternatively, to show why the partials do not exist at the origin, you might say $\frac{\partial f}{\partial x}(0,0) = \frac{d}{dt}f(t,0) = \frac{d}{dt}|t| = \lim_{h\to 0}\frac{|h|-0}{h}$, and this limit does not exist—the right-hand limit is 1, and the left-hand limit is -1. Similarly for $\frac{\partial f}{\partial y}$.

4. If $f(x,y) = e^{xy}\sin y$, find $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial^2 f}{\partial x\partial y}$, and $\frac{\partial^2 f}{\partial y\partial x}$.

Solution: $\frac{\partial f}{\partial x} = ye^{xy}\sin y$, and (using the product rule) $\frac{\partial f}{\partial y} = xe^{xy}\sin y + e^{xy}\cos y$.

$$\frac{\partial^2 f}{\partial x\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}\left(xe^{xy}\sin y + e^{xy}\cos y\right) = e^{xy}\sin y + xye^{xy}\sin y + ye^{xy}\cos y.$$

By Theorem 9.1.6, $\frac{\partial^2 f}{\partial y\partial x}$ equals the same thing. Or we can compute directly:

$$\frac{\partial^2 f}{\partial y\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(ye^{xy}\sin y\right) = e^{xy}\sin y + xye^{xy}\sin y + ye^{xy}\cos y.$$

6. Let $f:\mathbb{R}\to\mathbb{R}$ be differentiable on $\mathbb{R}$ and define a function $g:\mathbb{R}^2\to\mathbb{R}$ by $g(x,y) = f(x+y)$. Use (9.1.1) to show that $\frac{\partial g}{\partial x} = \frac{\partial g}{\partial y}$ on $\mathbb{R}^2$.

Solution: By (9.1.1), we have

$$\begin{aligned}\frac{\partial g}{\partial x}(x,y) &= \lim_{h\to 0}\frac{g((x,y)+he_1) - g(x,y)}{h} = \lim_{h\to 0}\frac{g(x+h,y)-g(x,y)}{h} \\ &= \lim_{h\to 0}\frac{f(x+h+y)-f(x+y)}{h}.\end{aligned}$$

Date: September 6, 2019.

On the other hand,

$$\frac{\partial g}{\partial y}(x,y) = \lim_{h\to 0}\frac{g((x,y)+he_2)-g(x,y)}{h} = \lim_{h\to 0}\frac{g(x,y+h)-g(x,y)}{h}$$
$$= \lim_{h\to 0}\frac{f(x+y+h)-f(x+y)}{h}.$$

These are clearly equal. □

10. If f is defined on $\mathbb{R}^2$ by

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

show that both $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist everywhere but they are not continuous at $(0,0)$. In fact f itself is not continuous at $(0,0)$ (Example 8.1.3).

Solution: At $(0,0)$, we have

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

and similarly,

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

That the derivative exists (and is continuous) everywhere else is easy, because away from the origin this is the derivative of a rational function at a point where the denominator is nonzero, and we know that such derivatives exist (Theorem 4.2.6(d)). So by Theorem 9.1.6, if the partials were continuous at $(0,0)$, then f would be differentiable here, hence continuous, which is not the case (as stated in the problem).

Alternatively (what I think the problem intended for you to do), we can check directly that the partial derivatives are not continuous at $(0,0)$. Using the quotient rule, we have for $(x,y)\neq(0,0)$ that

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

and

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

Taking the limit of $\frac{\partial f}{\partial x}(x,y)$ as we approach $(0,0)$ along the line $x=0$ yields

$$\lim_{y\to 0}\frac{\partial f}{\partial x}(0,y) = \lim_{y\to 0}\frac{y^3}{y^4} = \lim_{y\to 0}\frac{1}{y},$$

and this does not exist, so $\frac{\partial f}{\partial x}$ cannot be continuous at $(0,0)$. Similarly, taking the limit of $\frac{\partial f}{\partial y}(x,y)$ as we approach $(0,0)$ along the line $y=0$ yields

$$\lim_{x\to 0}\frac{\partial f}{\partial y}(x,0) = \lim_{x\to 0}\frac{x^3}{x^4} = \lim_{y\to 0}\frac{1}{x},$$

and this is again undefined, and so $\frac{\partial f}{\partial y}$ is also not continuous at $(0,0)$. □

Section 9.2:

1. If $L : \mathbb{R}^p \to \mathbb{R}^p$ is a linear function, show that $dL = L$. In other words, if L has matrix A, then A is the differential matrix of the linear function $L(x) = Ax$.

Proof. **Note:** *I'm hoping to emphasize the idea that it often helps to approach proofs by first writing what you know and what you want to show. I'll illustrate that in this solution, even though it's not really the shortest way to write it up.*

We are given that $L : \mathbb{R}^p \to \mathbb{R}^p$ is linear, meaning that $L(x+y) = L(x) + L(y)$, and $L(cx) = cL(x)$. We want to show that $dL = L$. By definition, $dL(x)$ is the linear function (if it exists) such that

$$\lim_{h \to 0} \frac{L(x+h) - L(x) - [dL(x)](h)}{\|h\|} = 0.$$

We know that L is linear, so to show that $dL = L$ (i.e., that $dL(x) = L$ for each x), we just need to show that

$$\lim_{h \to 0} \frac{L(x+h) - L(x) - L(h)}{\|h\|} = 0$$

for each x. By linearity of L, $L(x+h) = L(x) + L(h)$, so the above limit becomes

$$\lim_{h \to 0} \frac{L(x) + L(h) - L(x) - L(h)}{\|h\|} = \lim_{h \to 0} \frac{0}{\|h\|} = 0,$$

as desired. □

Method 2 (using Theorem 9.2.5 instead of just the definition): Suppose L has matrix $A = (a_{ij})_{ij}$, i.e., $a_{ij} = e_i \cdot L(e_j)$. Then by Theorem 9.2.5, the ij-entry of the matrix for $dL(x)$ is

$$e_i \cdot \frac{\partial L}{\partial x_j}(x) = e_i \cdot \lim_{h \to 0} \frac{L(x + he_j) - L(x)}{h} = e_i \cdot \lim_{h \to 0} \frac{hL(e_j)}{h} = e_i \cdot L(e_j) = a_{ij},$$

so the matrix for $dL(x)$ is A, as desired. □

2. Find the best affine approximation near $(0,0)$ to the function $F : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $F(x,y) = (xy - 2x + y + 1, x^2 + y^2 + x - 3y + 6)$.

Solution: We compute $F(0,0) = (1,6)$, and

$$dF(x,y) = \begin{pmatrix} y-2 & x+1 \\ 2x+1 & 2y-3 \end{pmatrix}. \tag{1}$$

So then $dF(0,0) = \begin{pmatrix} -2 & 1 \\ 1 & -3 \end{pmatrix}$. Now, the best affine approximation near $(0,0)$ is

$$\begin{aligned} T(x,y) = F(0,0) + dF(0,0)((x,y) - (0,0)) &= \begin{pmatrix} 1 \\ 6 \end{pmatrix} + \begin{pmatrix} -2 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \end{pmatrix} + \begin{pmatrix} -2x + y \\ x - 3y \end{pmatrix} \\ &= \begin{pmatrix} 1 - 2x + y \\ 6 + x - 3y \end{pmatrix}. \end{aligned}$$

If you like, you can rewrite this as $T(x,y) = (1 - 2x + y, 6 + x - 3y)$.

3. If F is the function of the previous exercise, find the best affine approximation to F near $(1,-1)$.

Solution: We compute $F(1,-1) = (-1-2-1+1, 1+1+1+3+6) = (-3, 12)$, and using (1), $dF(1,-1) = \begin{pmatrix} -3 & 2 \\ 3 & -5 \end{pmatrix}$. So then the best affine approximation near $(1,-1)$ is

$$\begin{aligned} T(x,y) = F(1,-1) + dF(1,-1)((x,y)-(1,-1)) &= \begin{pmatrix} -3 \\ 12 \end{pmatrix} + \begin{pmatrix} -3 & 2 \\ 3 & -5 \end{pmatrix} \begin{pmatrix} x-1 \\ y+1 \end{pmatrix} \\ &= \begin{pmatrix} -3 \\ 12 \end{pmatrix} + \begin{pmatrix} -3(x-1)+2(y+1) \\ 3(x-1)-5(y+1) \end{pmatrix} \\ &= \begin{pmatrix} 2-3x+2y \\ 4+3x-5y \end{pmatrix}. \end{aligned}$$

If you like, you may rewrite this as $T(x,y) = (2-3x+2y, 4+3x-5y)$.

6. Find the differential of the curve $\gamma(t) = (\sin(2\pi t), \cos(2\pi t), t^2)$. Then find the best affine approximation to the curve γ at the point $t = 1$.

Solution: $d\gamma(t) = (2\pi\cos(2\pi t), -2\pi\sin(2\pi t), 2t)$ (or if you want to write this as a matrix like we normally do in this class, it would be $\begin{pmatrix} 2\pi\cos(2\pi t) \\ -2\pi\sin(2\pi t) \\ 2t \end{pmatrix}$). Then $d\gamma(1) = (2\pi, 0, 2)$. Also, $\gamma(1) = (0,1,1)$. So the best affine approximation at $t = 1$ is

$$T(t) = \gamma(1) + d\gamma(1)(t) = (0,1,1) + (2\pi, 0, 2)t = (2\pi t, 1, 1+2t).$$

8. Prove that if U is a neighborhood of $0 \in \mathbb{R}^p$, and if $F : U \to \mathbb{R}^q$ is a function such that $F(0) = 0$, then F is differentiable at 0 with $dF = 0$ if and only if $\lim_{x\to 0} \|F(x)\|/\|x\| = 0$.

Proof. Suppose F is differentiable at 0 with $dF = 0$. This means that

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

Since $F(0) = 0$, and $0(h) = 0$, this reduces to $\lim_{h\to 0} \frac{F(h)}{\|h\|} = 0$. In general, the limit of a function is the 0-vector if and only if the limit of its norm is 0. (More generally, for any function G, $\lim_{x\to a} G(x) = L$ if and only if $\lim_{x\to a} \|G(x) - L\| = 0$. This becomes obvious if you write down the definition of each limit, as both definitions are essentially the same.) So it follows that $\lim_{h\to 0} \frac{\|F(h)\|}{\|h\|} = 0$, as desired.

Conversely, suppose $\lim_{h\to 0} \frac{\|F(h)\|}{\|h\|} = 0$. As mentioned above, this is equivalent to $\lim_{h\to 0} \frac{F(h)}{\|h\|} = 0$. In fact, the other steps above are reversible too: Since $F(0) = 0$ and $0(h) = 0$, we have

$$0 = \lim_{h\to 0} \frac{F(h)}{\|h\|} = \lim_{h\to 0} \frac{F(h) - F(0) - 0(h)}{\|h\|},$$

and so we see that $0 = dF(0)$, by the definition of $dF(0)$. □

Section 9.3:

12(b). Suppose F and G are functions defined on an open set $U \subset \mathbb{R}^p$, with values in $\mathbb{R}^q$. If F and G are differentiable at a point $x \in U$, then $F+G$ is differentiable at x, and $d(F+G)(x) = dF(x)+dG(x)$.

Proof. **Note:** *I'll again illustrate the strategy of writing out the definitions of what you know and what you want to show.*

We are given that F and G are differentiable at x, meaning that we have linear functions $dF(x)$ and $dG(x)$ such that

$$\lim_{h \to 0} \frac{F(x+h) - F(x) - dF(x)(h)}{\|h\|} = 0,$$

and similarly,

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

We want to show that $d(F+G)(x)$ exists and equals $dF(x) + dG(x)$. Since $dF(x)$ and $dG(x)$ are linear, their sum is linear, and so we just have to show that this sum satisfies the following:

$$\lim_{h \to 0} \frac{(F+G)(x+h) - (F+G)(x) - [dF(x) + dG(x)](h)}{\|h\|} = 0.$$

Using the definition of $F+G$ and then rearranging terms, we can rewrite this limit as

$$\begin{aligned} &\lim_{h \to 0} \frac{F(x+h) - F(x) - dF(x) + G(x+h) - G(x) - dG(x)(h)}{\|h\|} \\ &= \lim_{h \to 0} \frac{F(x+h) - F(x) - dF(x)}{\|h\|} + \lim_{h \to 0} \frac{G(x+h) - G(x) - dG(x)}{\|h\|} = 0 + 0 = 0, \end{aligned}$$

as desired. □

HOMEWORK 9 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 9.3

1. *If F is a function from an open subset $U \subset \mathbb{R}^p$ to $\mathbb{R}^q$ which is differentiable at a, and if B is an $r \times q$ matrix, then show that $d(BF)(a) = BdF(a)$. Here, $BF(x)$ is the matrix B applied to the vector $F(x)$ and $BdF(a)$ is the product of the matrix B and the matrix $dF(a)$.*

Solution:

Method 1 (Chain rule): B can be viewed as the linear function that maps a vector x in $\mathbb{R}^q$ to the vector Bx in $\mathbb{R}^r$, and then BF is the composition of this with the function F. So we can apply the chain rule to get

$$d(BF)(a) = dB(F(a))dF(a).$$

From Exercise 9.2.1, $dB = B$ (what was meant by this in 9.2.1 is that the differential $dB(x)$ is equal to B no matter what x is. I.e., if you view dB as a matrix-valued function, it's value is always the same matrix). So this becomes

$$d(BF)(a) = BdF(a),$$

as desired.

□

Method 2 (definition of the differential): B and $dF(a)$ are linear, so the product $BdF(a)$ is linear. So now by the definition of $d(BF)(a)$, what we want to show is that $BdF(a)$ satisfies the equation

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

We can factor B out of the numerator and even pull it out of the limit (because linear functions are continuous), and then this becomes

$$B\left(\lim_{h \to 0} \frac{F(a+h) - F(a) - dF(a)(h)}{\|h\|}\right) = 0.$$

This holds because the part inside the parentheses is 0 by the definition of $dF(a)$.

□

2. *If $f(x,y)$ is a differentiable function of $(x,y) \in \mathbb{R}^2$, and $g(t) = f(tx,ty)$ for all $t \in \mathbb{R}$, find $g'(1)$ in terms of the partial derivatives of f.*

Date: September 6, 2019.

Solution: The statement of the problem wasn't totally clear. I mentioned a couple equivalent interpretations. I'll work out both of these.

Viewing g as a function $\mathbb{R} \to \mathbb{R}$ with (x, y) fixed: We can view $g(t)$ as $f \circ H(t)$ where $H : \mathbb{R} \to \mathbb{R}^2$ is given by $H(t) := (tx, ty)$. Then by the Chain rule,

$$g'(1) = df(H(1))dH(1).$$

We have $H(1) = (x, y)$, $df(x, y) = (\frac{\partial f}{\partial x}(x, y)\ \frac{\partial f}{\partial y}(x, y))$, and $dH(t) = (x, y)$, so this becomes

$$g'(1) = \begin{pmatrix} \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y).$$

□

Viewing g as a function $g(x, y, t)$ from $\mathbb{R}^3$ to $\mathbb{R}$, and then $g'(1)$ means $\frac{\partial g}{\partial t}(x, y, 1)$: For this perspective, we take $H : \mathbb{R}^3 \to \mathbb{R}^2$, $H(x, y, t) = (xt, yt)$, so $g = f \circ H$ again and $dH(x, y, t) = \begin{pmatrix} t & 0 & x \\ 0 & t & y \end{pmatrix}$, and so we have

$$\begin{aligned} dg(x, y, 1) = df(H(x, y, 1))dH(x, y, 1) &= \begin{pmatrix} \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \end{pmatrix} \begin{pmatrix} 1 & 0 & x \\ 0 & 1 & y \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) & x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y) \end{pmatrix}. \end{aligned}$$

And so $\frac{\partial g}{\partial t}(x, y, 1)$ is given by the third entry, $x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y)$. □

Another approach using the "dependent variable notation" perspective. View $g(t)$ as $f(x(t), y(t))$ for $x(t) = tx$ and $y(t) = ty$ (this notation is a bit confusing because x and y are denoting both functions and constants). Then

$$\begin{aligned} \frac{\partial g}{\partial t}(1) &= \frac{\partial f}{\partial x}(x(1), y(1))\frac{\partial x}{\partial t}(1) + \frac{\partial f}{\partial y}(x(1), y(1))\frac{\partial y}{\partial t}(1) \\ &= x\frac{\partial f}{\partial x}(x, y) + y\frac{\partial f}{\partial y}(x, y). \end{aligned}$$

□

8. *If $F(x, y) = (f_1(x, y), f_2(x, y))$ is differentiable function from $\mathbb{R}^2 \to \mathbb{R}^2$ and if we define $G : \mathbb{R}^2 \to \mathbb{R}^2$ by $G(s, t) = F(st, s + t)$, find an expression for the differential matrix of G in terms of the partial derivatives of f_1 and f_2.*

Solution:

Method 1: Define a function $H : \mathbb{R}^2 \to \mathbb{R}^2$ by $H(s,t) = (st, s+t)$, so $G = F \circ H$. Then by the Chain rule,

$$\begin{aligned} dG(s,t) &= dF(H(s,t))dH(s,t) \\ &= \begin{pmatrix} \frac{\partial f_1}{\partial x}(st,s+t) & \frac{\partial f_1}{\partial y}(st,s+t) \\ \frac{\partial f_2}{\partial x}(st,s+t) & \frac{\partial f_2}{\partial y}(st,s+t) \end{pmatrix} \begin{pmatrix} t & s \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} t\frac{\partial f_1}{\partial x}(st,s+t) + \frac{\partial f_1}{\partial y}(st,s+t) & s\frac{\partial f_1}{\partial x}(st,s+t) + \frac{\partial f_1}{\partial y}(st,s+t) \\ t\frac{\partial f_2}{\partial x}(st,s+t) + \frac{\partial f_2}{\partial y}(st,s+t) & s\frac{\partial f_2}{\partial x}(st,s+t) + \frac{\partial f_2}{\partial y}(st,s+t) \end{pmatrix}. \end{aligned}$$

□

Method 2 (using dependant variable notation): View x and y as functions of s and t given by $x(s,t) = st$ and $y(s,t) = s+t$. Write $G(s,t)$ as $G(s,t) = (f_1(x(s,t)), f_2(x(s,t)))$. $dG(s,t)$ is given by

$$\begin{pmatrix} \frac{\partial f_1}{\partial s}(s,t) & \frac{\partial f_1}{\partial t}(s,t) \\ \frac{\partial f_2}{\partial s}(s,t) & \frac{\partial f_2}{\partial t}(s,t) \end{pmatrix},$$

so we just have to find the partial derivatives of f_1 and f_2 with respect to s and t. We have

$$\begin{aligned} \frac{\partial f_1}{\partial s}(s,t) &= \frac{\partial f_1}{\partial x}(x(s,t))\frac{\partial x}{\partial s}(s,t) + \frac{\partial f_1}{\partial y}(x(s,t))\frac{\partial y}{\partial s}(s,t) \\ &= t\frac{\partial f_1}{\partial x}(st,s+t) + \frac{\partial f_1}{\partial y}(st,s+t), \end{aligned}$$

and similarly for the other partials. □

9. *If (x,y,z) are the Cartesian coordinates of a point in $\mathbb{R}^3$ and the spherical coordinates of the same point are r, θ, ϕ, then $x = r\cos\theta\sin\phi$, $y = r\sin\theta\sin\phi$, and $z = r\cos\phi$. Let u be a variable which is a differentiable function of (x,y,z) on $\mathbb{R}^3$. Find a formula for the partial derivatives of u with respect to r, θ, ϕ in terms of its partial derivatives with respect to x, y, z.*

Solution: (I'm just going to do the dependant variable notation perspective, but this could also be done with matrices. I'm also fine with you leaving your answer as an equality of matrices here, and with you not working out the matrix multiplication for this problem).

From Equation (9.3.2) in your book (the dependant variable notation version of the Chain rule), we have

$$\begin{aligned} \frac{\partial u}{\partial r} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial r} \\ &= \frac{\partial u}{\partial x}\cos\theta\sin\theta + \frac{\partial u}{\partial y}\sin\theta\sin\phi + \frac{\partial u}{\partial z}\cos\phi. \end{aligned}$$

[Here, by $\frac{\partial u}{\partial x}$, we really mean the partial derivative at (x,y,z), i.e., $\frac{\partial u}{\partial x}(x,y,z)$, and similarly for $\frac{\partial u}{\partial y}$ and $\frac{\partial u}{\partial z}$. Similarly, by $\frac{\partial x}{\partial r}$, we meant $\frac{\partial x}{\partial r}(r,\theta,\phi)$, etc. I'm fine with you just making these notational assumptions here.].

$$\begin{aligned} \frac{\partial u}{\partial \theta} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \theta} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial \theta} \\ &= \frac{\partial u}{\partial x}(-r\sin\theta\sin\phi) + \frac{\partial u}{\partial y}(r\cos\theta\sin\phi). \end{aligned}$$

$$\begin{aligned}\frac{\partial u}{\partial \phi} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial \phi} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial \phi} \\ &= \frac{\partial u}{\partial x}(r\cos\theta\cos\phi) + \frac{\partial u}{\partial y}(r\sin\theta\cos\phi) + \frac{\partial u}{\partial z}(-r\sin\phi).\end{aligned}$$

□

11. *Show that if F is a differentiable function on an open set $U \subset \mathbb{R}^p$ with values in $\mathbb{R}^q$, then the real-valued function $\|F(x)\|^2$ on U has zero differential at x if and only if the vector $F(x)$ is orthogonal to each of the columns of $dF(x)$.*

Solution: Since $\|F(x)\|^2 = F(x) \cdot F(x)$, we have $d\|F(x)\|^2 = 0$ if and only if $d(F(x) \cdot F(x)) = 0$. By Theorem 9.3.6, $d(F(x) \cdot F(x)) = F(x)^T dF(x) + F(x)^T dF(x) = 2F(x)^T dF(x)$. [The book doesn't say they're taking the transpose, but they really should since we typically view $F(x)$ as a column vector, but here it must be a row vector].

So now we have that $d(F(x) \cdot F(x)) = 0$ if and only if $F(x)^T dF(x) = 0$. The entries of $F(x)^T dF(x)$ are just the dot products of $F(x)$ with the columns of $dF(x)$, and these dot products are 0 if and only if $F(x)$ is orthogonal to the columns of $dF(x)$, so both directions of the claim follow. □

Section 9.4

2. *For the function $f(x, y) = x^2 + y^3 + xy$, find the gradient at the point $(1, 1)$, the direction of greatest ascent of f at this point, and a direction in which the rate of increase of this function is 0 (the answers to the last two questions should be unit vectors).*

Solution: $\nabla f(x, y) = (2x + y, 3y^2 + x)$, so $\nabla f(1, 1) = (3, 4)$.

The direction of greatest increase at $(1, 1)$ is then the unit vector in this direction,

$$\frac{\nabla f(1,1)}{\|\nabla f(1,1)\|} = \frac{(3,4)}{\sqrt{3^2+4^2}} = (\frac{3}{5}, \frac{4}{5}).$$

Since $D_u f = df \cdot u$, or since df is orthogonal to the level sets of f, we see that the rate of increase is 0 in any direction orthogonal to ∇f. A unit vector orthogonal to $\nabla f(1, 1)$ is $(\frac{4}{5}, -\frac{3}{5})$. You could also use negative this, $(-\frac{4}{5}, \frac{3}{5})$. □

3. *Find a parametric equation for the tangent line to the curve $\gamma(t) = (t^3, 1/t, e^{2t-2})$ at the point where $t = 1$.*

Solution: We have $\gamma(1) = (1, 1, 1)$, and $\gamma'(t) = (3t^2, -t^{-2}, 2e^{2t-2})$, so $\gamma'(1) = (3, -1, 2)$. So the tangent line is parameterized by

$$\tau(t) = (1, 1, 1) + (t - 1)(3, -1, 2)$$

or equivalently,

$$\tau(t) = (3t - 2, 2 - t, 2t - 1).$$

[You may also break this up into parametric equations for the x,y, and z components separately if you like.] □

Alternatively: If you multiplied $\gamma'(t)$ above by t instead of $(t-1)$, that's actually ok! The resulting Φ is no longer the "best affine approximation to γ," but it does parameterize of the same line. Doing this gives:

$$\tau(t) = (1, 1, 1) + t(3, -1, 2),$$

or equivalently,

$$\tau(t) = (1 + 3t, 1 - t, 1 + 2t).$$

□

8. *Find the tangent space at $(2, 4, 1)$ for the parameterized surface in $\mathbb{R}^3$ parameterized by the function $G : U \to \mathbb{R}^3$, where $U = \{(u, v) : u > 0, v > 0\}$, and $G(u, v) = (uv, u^2, v^2)$.*

Solution: First, something I think several people were confused on, we need to find the point (u_0, v_0) in U which maps to $(2, 4, 1)$. We need $(u_0 v_0, u_0^2, v_0^2) = (2, 4, 1)$, and since being in U forces $u_0, v_0 > 0$, the only possibility is $u_0 = 2$, $v_0 = 1$. Now,

$$dG(u, v) = \begin{pmatrix} v & u \\ 2u & 0 \\ 0 & 2v \end{pmatrix},$$

so the tangent plane at $(2, 4, 1)$ is parametrerized by

$$\begin{aligned} \Phi(u, v) &= (2, 4, 1) + dG(2, 1)(u - 2, v - 1) \\ &= \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 4 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} u - 2 \\ v - 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix} + \begin{pmatrix} (u - 2) + 2(v - 1) \\ 4(u - 2) \\ 2(v - 1) \end{pmatrix} \\ &= \begin{pmatrix} u + 2v - 2 \\ 4u - 4 \\ 2v - 1 \end{pmatrix}. \end{aligned}$$

□

Alternatively: If you multiplied $dG(2, 1)$ above by (u, v) instead of $(u - 2, v - 1)$, that's actually ok! The resulting Φ is no longer the "best affine approximation to G," but it does parameterize the same plane as the Φ above. □

HOMEWORK 10 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 9.4

11. *Find the equation for the tangent plane to the cone* $z = x^2 + y^2$ *at the point* $(1, 2, 5)$.

Solution:

Method 1: Let $f(x, y, z) = z - x^2 - y^2$. The cone C we are interested in is just the level set $f(x, y, z) = 0$. We know from Example 9.4.13 that this can be represented as a smoothly parameterized 2-surface, and we just want to find the tangent space at $(1, 2, 5) \in C$. We have $df(x, y, z) = (-2x, -2y, 1)$, and so $df(1, 2, 5) = \begin{pmatrix} -2 & -4 & 1 \end{pmatrix}$. From Theorem 9.4.11 then, the tangent space is given by $df(1, 2, 5)[(x, y, z) - (1, 2, 5)] = 0$, i.e., by the solutions to

$$\begin{pmatrix} -2 & -4 & 1 \end{pmatrix} \begin{pmatrix} x-1 \\ y-2 \\ z-5 \end{pmatrix} = 0,$$

i.e.,

$$-2(x - 1) - 4(y - 2) + (z - 5) = 0.$$

□

Method 2, viewing this as a graph and using the extension of Example 9.4.13 I did in class: In class, I extended Example 9.4.13 to show that the tangent space to $z = g(x, y)$ at a point $(a, b, g(a, b))$ is given by

$$z = g(a, b) + \frac{\partial g}{\partial x}(a, b)(x - a) + \frac{\partial g}{\partial y}(a, b)(y - b).$$

This is the case where $g(x, y) = x^2 + y^2$, and $(a, b) = (1, 2)$. We have $g(a, b) = 5$, $\frac{\partial g}{\partial x} = 2x$ so $\frac{\partial g}{\partial x}(1, 2) = 2$, and $\frac{\partial g}{\partial y} = 2y$ so $\frac{\partial g}{\partial y}(1, 2) = 4$, so the tangent space is

$$z = 5 + 2(x - 1) + 4(y - 2).$$

□

14. *Find an equation for the tangent plane to the surface* $x^2 + y^2 - z^2 = 1$ *at each point* (a, b, c) *on the surface.*

Solution: This is exactly the same as Example 9.4.12, except that this level set does not contain the origin, so we don't have to worry about something going wrong there. Just like in that example,

Date: September 6, 2019.

we write $f(x,y,z) = x^2 + y^2 - z^2$, and then $df(a,b,c) = \begin{pmatrix} 2a & 2b & 2c \end{pmatrix}$, and this has rank 1 (since $(a,b,c) \neq (0,0,0)$ on the level set $f(x,y,z) = 1$.) Since $1 = 3-2$ (3 being the dimension of the domain of f and 2 being the dimension of the level set), we can apply Theorem 9.4.11 to say that the tangent space at (a,b,c) is given by $df(a,b,c)[(x,y,z) - (a,b,c)] = 0$. Doing the matrix multiplication makes this

$$2a(x-a) + 2b(y-b) + 2c(z-c) = 0,$$

just like in Example 9.4.12 (except that (a,b,c) now lives on a different surface than in that example).

□

Section 9.5

1. *Find the degree $n = 2$ Taylor formula for $f(x,y) = x^2 + xy$ at the point $a = (1,2)$.*

Solution: (Note that since f is a degree 2 polynomial, the degree 2 Taylor formula should have 0 for the remainder term R_2. We'll see that this does in fact hold).

We have $f(1,2) = 3$.

$df(x,y) = \begin{pmatrix} 2x+y & x \end{pmatrix}$, and so $df(1,2) = \begin{pmatrix} 4 & 1 \end{pmatrix}$.

$d^2 f(x,y) = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$.

Since the second-order partials are all constants, all the third order partials are 0, so $R_3 = 0$. So the $n = 2$ Taylor formula is:

$$\begin{aligned} f(x,y) &= f(a) + df(a)(x-a) + \frac{1}{2}d^2 f(a)(x-a)^2 + 0 \\ &= 3 + \begin{pmatrix} 4 & 1 \end{pmatrix}\begin{pmatrix} x-1 \\ y-2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x-1 & y-2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x-1 \\ y-2 \end{pmatrix} \\ &= 3 + 4(x-1) + (y-2) + \frac{1}{2}\begin{pmatrix} x-1 & y-2 \end{pmatrix}\begin{pmatrix} 2(x-1)+(y-2) \\ x-1 \end{pmatrix} \\ f(x,y) &= 3 + 4(x-1) + (y-2) + (x-1)^2 + (x-1)(y-2). \end{aligned}$$

□

2. *Find the degree $n = 2$ Taylor formula for $f(x,y) = e^{xy}$ at the point $a = (0,0)$.*

Solution:

$f(0,0) = 1$.

$df(x,y) = \begin{pmatrix} ye^{xy} & xe^{xy} \end{pmatrix}$, and so $df(0,0) = \begin{pmatrix} 0 & 0 \end{pmatrix}$.

$d^2 f(x,y) = \begin{pmatrix} y^2 e^{xy} & (1+xy)e^{xy} \\ (1+xy)e^{xy} & x^2 e^{xy} \end{pmatrix}$, so $d^2 f(0,0) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

So then Taylor's formula says:

$$\begin{aligned} f(x,y) &= f(a) + df(a)(x-a) + \frac{1}{2}d^2 f(a)(x-a)^2 + R_2 \\ &= 1 + 0 + \frac{1}{2}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + R_2 \\ &= 1 + xy + R_2. \end{aligned}$$

So now we just need to compute R_2.

For our third order partials, we have $\frac{\partial^3 f}{\partial x^3} = y^3 e^{xy}$, $\frac{\partial^3 f}{\partial x^2 y} = (2y + xy^2)e^{xy}$, $\frac{\partial^3 f}{\partial x^2 y} = (2x + x^2 y)e^{xy}$, and $\frac{\partial^3 f}{\partial y^3} = x^3 e^{xy}$.

Now, we want to evaluate $d^3 f$ at $\vec{c}$, where $\vec{c}$ is some point on the line segment from $(0,0)$ to (x,y). We can view $\vec{c}$ as (cx, cy), where c here is a scalar in $[0,1]$. R_3 is then equal to

$$\begin{aligned} R_3 &= \frac{1}{3!}\left[(cy)^3 e^{(cx)(cy)}(x^3) + 3(2(cy) + (cx)(cy)^2)e^{(cx)(cy)}(x^2 y) + 3(2(cx) + (cx)^2(cy))e^{(cx)(cy)}(xy^2) + (cx)^3 e^{(cx)(cy)}(y^3)\right] \\ &= \frac{1}{3!}\left[(cxy)^3 e^{c^2 xy} + 3[2cx^2y^2 + (cxy)^3]e^{c^2xy} + 3[2cx^2y^2 + (cxy)^3]e^{c^2xy} + (cxy)^3 + e^{c^2xy}\right] \\ &= \frac{1}{3!}\left[(8(cxy)^3 + 12cx^2y^2)e^{c^2xy}\right] \\ &= \left(\frac{4}{3}(cxy)^3 + 2cx^2y^2\right)e^{c^2xy}, \end{aligned}$$

for some $c \in [0,1]$. □

3. *Suppose $a \in \mathbb{R}^p$ and f is a real-valued function whose second-order partial derivatives all exist and are continuous on $B_r(a)$. Also, suppose that the operator norm $\|d^2 f(x)\|$ of the matrix $d^2 f(x)$ is bounded by M on $B_r(a)$. Prove that*

$$|f(x) - f(a) - df(a)(x-a)| \le \frac{M}{2}\|x-a\|^2$$

for all $x \in B_r(a)$.

Proof. By Taylor's formula with $n = 1$, we have $f(x) = f(a) + df(a)(x-a) + \frac{1}{2}d^2 f(c)(x-a)^2$ for some $c \in [a,x] \subset B_r(a)$, where $[a,x]$ denotes the line segment from a to x. Rearranging, this becomes

$$f(x) - f(a) - df(a)(x-a) = \frac{1}{2}d^2 f(c)(x-a)^2,$$

so it suffices to show that $|d^2 f(c)(x-a)^2| \le M\|x-a\|^2$ for all $c \in [a,x]$. Using the Cauchy-Schwarz inequality, we have

$$|d^2 f(c)(x-a)^2| = |(x-a)\cdot d^2 f(c)(x-a)| \le \|x-a\|\|d^2 f(c)(x-a)\|. \tag{1}$$

By the definition of the operator norm, we have

$$\|d^2 f(c)(x-a)\| \le \|d^2 f(c)\|\|x-a\|. \tag{2}$$

Combining (1) and (2) above and the fact that $\|d^2 f(c)\| \le M$ for all $c \in B_r(a)$, we get

$$|d^2 f(c)(x-a)^2| \le \|x-a\|\|d^2 f(c)(x-a)\| \le \|d^2 f(c)\|\|x-a\|^2 \le M\|x-a\|^2,$$

as desired. □

HOMEWORK 11 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 9.5

7. *Show that the following version of the Mean Value Theorem for vector-valued functions is true: if U is an open set in $\mathbb{R}^p$ containing the line segment joining a to b, and if $F : U \to \mathbb{R}^q$ is a differentiable function on U, then, for each vector $u \in \mathbb{R}^q$, there is a point c on the line segment joining a to b such that*

$$u \cdot (F(b) - F(a)) = u \cdot Df(c)(b - a). \tag{1}$$

Solution: Let $f : U \to \mathbb{R}$ denote the real-valued function $u \cdot F$, i.e., $f(x) := u \cdot F(x)$. By the Mean Value Theorem (Theorem 9.5.3), there exists a point c on the line segment from a to b such that

$$f(b) - f(a) = df(c)(b - a).$$

The left-hand side above is, by definition, $u \cdot F(b) - u \cdot F(a)$, which we can factor as $u \cdot (F(b) - F(a))$ to get the left-hand side of (1).

On the right-hand side, we have $df(c)(b - a) = d(u \cdot F)(b - a)$. By Theorem 9.3.6 (viewing u in $u \cdot F$ as the constant function $U \to \mathbb{R}^q$, $x \mapsto u$), we know that $d(u \cdot F)(b - a) = u \cdot dF(b - a)$ (since u is constant, $du = 0$), and this gives the right-hand side of (1), as desired.

Alternatively: To compute $d(u \cdot F)$, you can, for example, write $F(x) = (f_1(x), \ldots, f_q(x))$, $u = (u_1, \ldots, u_q)$, and then $(u \cdot F(x)) = (u_1 f_1(x), \ldots, u_q f_q(x))$. Since constants like the u_i's can be pulled in front of partial deriviatvies, $d(u_i f_i) = u_i df_i$ for each i, so then the differential of $d(u \cdot F)(x)$ is $\Big(u_1 df_1(x), \ldots, u_q df_q(x)\Big) = u \cdot dF(x)$.

Comments on the Mean Value Theorem for vector valued functions: Problem 6, which I didn't assign, asks you to prove that a more natural sounding generalization of the mean value theorem to vector valued functions, and they ask you to show that this actually doesn't work. Here's an intuitive explanation (which can probably be made into a more precise counterexample).

I like to view the Mean Value Theorem as saying that if you travel in a straight line and go 10 miles in 1 hour, then at some point in time, you must have been going exactly 10 miles per hour (intuitively, this is because if you were always going faster, you would have gone more than 10 miles, and if you were always going slower, you would have gone less than 10 miles). But suppose you weren't going in a straight line. Maybe your final point is 10 miles from your initial point, but you traveled along a semi-circle to get there. Then you actually travelled a distance of $5\pi > 10$ miles, so it's possible that you were going 5π miles per hour the whole time, and so you never necessarily went 10 miles per hour.

Date: September 6, 2019.

8. *Find all points of relative maximum and relative minimum and all saddle points for* $f(x,y) = 1 - 2x^2 - 2xy - y^2$.

Solution: We have $df(x,y) = \begin{pmatrix} -4x-2y & -2x-2y \end{pmatrix}$. This equalling $\begin{pmatrix} 0 & 0 \end{pmatrix}$ means that $-4x - 2y = 0$ and $-2x - 2y = 0$. The latter of these gives $x = y$, and then $-4x - 2y = 0$ becomes $-6x = 0$, hence $x = 0$ and $y = 0$. So $df(x,y)$ is only 0 at the origin $(0,0)$.

For the Hessian $d^2f(x,y)$, we get $d^2f(x,y) = \begin{pmatrix} -4 & -2 \\ -2 & -2 \end{pmatrix}$ (no matter what x and y are, so in particular at $(0,0)$). This has determinant $8 - 4 = 4 > 0$, so since the $(1,1)$-entry is $-4 < 0$, $d^2f(0,0)$ is negative definite, and so f **has a local maximum at** $(0,0)$**.**

9. *Find all points of relative maximum and relative minimum and all saddle points for* $f(x,y) = y^3 + y^2 + x^2 - 2xy - 3y$.

Solution: We have $df(x,y) = \begin{pmatrix} 2x-2y & 3y^2+2y-2x-3 \end{pmatrix}$. This equalling $\begin{pmatrix} 0 & 0 \end{pmatrix}$ means that $2x - 2y = 0$ and $3y^2 + 2y - 2x - 3 = 0$. $2x - 2y = 0$ implies $x = y$, and then the second equation becomes $3x^2 - 3 = 0$, i.e., $x^2 = 1$, so we have $x = y = \pm 1$. That is, local extrema can only occur at $(1,1)$ and $(-1,-1)$.

The Hessian $d^2f(x,y)$ is

$$\begin{pmatrix} 2 & -2 \\ -2 & 6y+2 \end{pmatrix},$$

and the determinant of this is $12y + 4 - 4 = 12y$. So at $(1,1)$, the determinant is $12 > 0$, and since the $(1,1)$-entry is $2 > 0$, f **has a local minimum at** $(1,1)$**.** At $(-1,-1)$, the determinant is $-12 < 0$, so f **has a saddle point at** $(-1,-1)$**.**

Section 9.6

2. *Show that the function* $F(x,y) = (x^2+y^2, xy)$ *has a smooth local inverse near points* (x,y) *where* $x \neq \pm y$. *Find the inverse function* F^{-1} *on the set* $\{(x,y) : -x < y < x\}$ *and identify its domain. Calculate the differential of this inverse function (1) directly and (2) by using the Inverse Function Theorem. Verify that the two methods give the same answer.*

Hint and comments I posted on Canvas: The inverse function is $G(u,v) = \left(\frac{\sqrt{u+2v}+\sqrt{u-2v}}{2}, \frac{\sqrt{u+2v}-\sqrt{u-2v}}{2}\right)$. The domain is $u > |2v|$. You should check that this really is the inverse, either by deriving it yourself or by just showing that $F \circ G$ and $G \circ F$ are both the identity function. Don't worry about checking the domain. The more important point is for you to check that computing the differential of G directly gives the same thing as what you get from applying the inverse function theorem.)

Solution: Here's how I came up with that inverse function. Let $u = x^2 + y^2$, $v = xy$. Notice that $(x+y)^2 = u + 2v$ and $(x-y)^2 = u - 2v$. So then $x + y = \sqrt{u+2v}$ (because $-x < y$ implies that $x + y > 0$, so we take the positive square root), and $x - y = \sqrt{u-2v}$ (again, we take the positive square root because $y < x$ implies $x - y > 0$). Adding/subtracting these two equations and dividing by 2 yields $x = \frac{\sqrt{u+2v}+\sqrt{u-2v}}{2}$ and $y = \frac{\sqrt{u+2v}-\sqrt{u-2v}}{2}$, as claimed for the inverse. I then picked that domain $u > |2v|$ simply because this is where G is defined, but technically we should check that G maps this to the set $\{(x, y) : -x < y < x\}$ (I'll skip this).

Now, computing the partial derivatives directly, we get

$$dG(u,v) = \begin{pmatrix} \frac{(u+2v)^{-1/2}+(u-2v)^{-1/2}}{4} & \frac{(u+2v)^{-1/2}-(u-2v)^{-1/2}}{2} \\ \frac{(u+2v)^{-1/2}-(u-2v)^{-1/2}}{4} & \frac{(u+2v)^{-1/2}+(u-2v)^{-1/2}}{2} \end{pmatrix}. \tag{2}$$

On the other hand, we compute

$$dF(x,y) = \begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix},$$

so

$$[dF(x,y)]^{-1} = \frac{1}{2x^2-2y^2}\begin{pmatrix} x & -2y \\ -y & 2x \end{pmatrix}.$$

We have $2x^2 - 2y^2 = 2(x + y)(x - y) = 2\sqrt{u+2v}\sqrt{u-2v}$. So substituting in the equations for x and y in terms of u and v yields

$$[dF(x,y)]^{-1} = \frac{1}{2\sqrt{u+2v}\sqrt{u-2v}}\begin{pmatrix} \frac{\sqrt{u+2v}+\sqrt{u-2v}}{2} & -\sqrt{u+2v}+\sqrt{u-2v} \\ \frac{-\sqrt{u+2v}+\sqrt{u-2v}}{2} & \sqrt{u+2v}+\sqrt{u-2v} \end{pmatrix}$$

and simplifying reduces this to the matrix from (2), as desired.

3. *Near which points of $\mathbb{R}^3$ does the spherical change of coordinates function $F(\rho,\theta,\phi) = (\rho\cos\theta\sin\phi, \rho\sin\theta\sin\phi, \rho\cos\phi)$ have a smooth local inverse? What is the differential of the local inverse at those points where it exists? To avoid tedious computation, express this in terms of (r,θ,ϕ) rather than in terms of the image variables $(x,y,z) = F(r,\theta,\phi)$.*

Late Edit added on Canvas: Don't bother computing the inverse matrix.

Solution: The differential of F is

$$dF(r,\theta,\phi) = \begin{pmatrix} \cos\theta\sin\phi & -\rho\sin\theta\sin\phi & \rho\cos\theta\cos\phi \\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi \\ \cos\phi & 0 & -\rho\sin\phi \end{pmatrix}. \tag{3}$$

The determinant is

$$\begin{aligned} &-\rho^2\cos^2\theta\sin^3\phi - \rho^2\sin^2\theta\sin\phi\cos^2\phi - \rho^2\cos^2\theta\cos^2\phi\sin\phi - \rho^2\sin^2\theta\sin^3\phi \\ &= -\rho^2\sin\phi[\cos^2\theta(\sin^2\phi+\cos^2\phi) + \sin^2\theta(\cos^2\phi+\sin^2\phi)] \\ &= -\rho^2\sin\phi. \end{aligned}$$

This equals 0 if and only if $\rho = 0$ or $\phi = k\pi$ for $k \in \mathbb{Z}$. So there is a smooth local inverse everywhere else. The differential for the inverse is then the inverse of the matrix in (3).

4. *Show that the system of equations $x = u^4 - u + uv + v^2$, $y = \cos u + \sin v$, can be solved for (u, v) as a smooth function F of (x, y) in some neighborhood of $(0, 0)$ in such a way that $(u, v) = (0, 0)$ when $(x, y) = (0, 1)$. What is the differential of the resulting function F at $(0, 1)$?*

Solution: Define $G : \mathbb{R}^2 \to \mathbb{R}^2$ by $G(u, v) = (u^4 - u + uv + v^2, \cos u + \sin v)$, so $(x, y) = G(u, v)$. Then

$$dG(u, v) = \begin{pmatrix} 4u^3 - 1 + v & u + 2v \\ -\sin u & \cos v \end{pmatrix},$$

and so

$$dG(0, 0) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

and $\det G(0, 0) = -1 \neq 0$. So by the Inverse Function Theorem, G has a smooth local inverse F at $(0, 0)$. Then $(u, v) = F(x, y)$ for each (x, y) in the domain of F, and since $G(0, 0) = (0, 1)$, $(0, 1)$ is in the domain of F, and $F(0, 1) = (0, 0)$. Finally,

$$dF(0, 1) = [dG(0, 0)]^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

9. *Show that if $F = (f_1, f_2) : \mathbb{R}^3 \to \mathbb{R}^2$ is a $\mathcal{C}^1$ function and $\vec{a}$ is a point of $\mathbb{R}^3$ at which dF has rank 2, then there is a $\mathcal{C}^1$ function $f_3 : \mathbb{R}^3 \to \mathbb{R}$ such that $\Phi = (f_1, f_2, f_3) : \mathbb{R}^3 \to \mathbb{R}^3$ has a $\mathcal{C}^1$ inverse function near $\vec{a}$.*

Hint I gave on Canvas: $dF(\vec{a})$ is a 2-by-3 matrix. $dF(\vec{a})$ having rank 2 means that the two rows are linearly independent. This implies that there exists a vector (infinitely many in fact) $u \in \mathbb{R}^3$ such that the rows of $df(\vec{a})$, together with u, form a basis for $\mathbb{R}^3$ (you should use this fact, don't worry about proving it). So now it suffices to find some f_3 such that $df_3(\vec{a})$ (the third row of $d\Phi(\vec{a})$) equals u.

Solution: Consider a vector u as in the hint (viewed as a row vector). Let $u = (a, b, c)$. Define $f(x, y, z) = ax + by + cz$. Then $df(x, y, z) = \begin{pmatrix} a & b & c \end{pmatrix}$, which is just u. So then, by the way u was chosen, $df_1(\vec{a})$, $df_2(\vec{a})$, and $df_3(\vec{a}) = u$ are linearly independent. These form the rows of $d\Phi(\vec{a})$, and so $d\Phi(\vec{a})$ is non-singular (being non-singular is equivalent to the rows being linearly independent). Thus, by the Inverse Function Theorem, Φ has a smooth local inverse near $\vec{a}$, as desired.

HOMEWORK 12 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

<u>Section 9.7</u>

1. *Are there any points on the graph of the equation $x^3 + 3xy^2 + 2y^3 = 1$ where it may not be possible to solve for y as a smooth function of x in some neighborhood of the point?*

Solution: Let $f(x, y) = x^3 + 3xy^2 + 2y^3 - 1$, so the graph of the equation is the solution set to $f(x, y) = 0$. By the implicit function theorem, it is possible to solve for y as a smooth function of x in a neighborhood of any point where $\frac{\partial f}{\partial y} \neq 0$. We have

$$\frac{\partial f}{\partial y} = 6xy + 6y^2 = 6y(x + y).$$

This equals 0 when $y = 0$, and also when $y = -x$. Plugging $y = 0$ into $f(x, y)$, we have

$$f(x, 0) = x^3 - 1,$$

which equals 0 at $x = 1$. Plugging $y = -x$ into $f(x, y)$, we get

$$f(x, -x) = x^3 + 3x^3 - 2x^3 - 1 = 2x^3 - 1,$$

and this equals 0 at $x = (\frac{1}{2})^{1/3}$. So the only points on the graph where we might not locally be able to express y as a smooth function of x are $(1, 0)$ and $((\frac{1}{2})^{1/3}, -(\frac{1}{2})^{1/3})$. □

2. *Can the equation $xz + yz + \sin(x + y + z) = 0$ be solved, in a neighborhood of $(0, 0, 0)$ for z as a smooth funciton $z = g(x, y)$ of (x, y), with $g(0, 0) = 0$?*

Solution: We have $\frac{\partial f}{\partial z}(x, y, z) = x+y+\cos(x+y+z)$, and so $\frac{\partial f}{\partial z}(0, 0, 0) = 1$. Since this is nonzero, the implicit function theorem tells us that there is indeed a neighborhood of $(0, 0, 0)$ in which z can be solved as a smooth function $z = g(x, y)$ of (x, y) with $g(0, 0) = 0$.

4. *Show that the system of equations*

$$u^2 + v^2 + 2u - xy + z = 0$$
$$u^3 + \sin v - xu + yv + z^2 = 0$$

has a solution for (u, v) as a smooth function of (x, y, z), in some neighborhood of $(0, 0, 0)$, with the property that $(u, v) = 0$ when $(x, y, z) = 0$.

Date: September 6, 2019.

Solution: Let $F(x,y,z,y,v) = (u^2+v^2+2u-xy+z, u^3+\sin v - xu+yv+z^2)$, so we are interested in where $F(x,y,z,u,v)=0$. One easily checks that $F(0,0,0,0,0)$ does indeed equal 0, so this point is in the solution set. We have

$$\frac{\partial(f_1,f_2)}{\partial(u,v)} = \begin{pmatrix} 2u+2 & 2v \\ 3u^2-x & \cos v + y \end{pmatrix}$$

and at $(0,0,0,0,0)$ this equals $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$. Since this is non-singular (it has determinant 2), the implicit function theorem tells us that there is a neighborhood of $(0,0,0,0,0)$ where it is possible to express (u,v) as a smooth function $G(x,y,z)$ of (x,y,z) in some neighborhood of $(0,0,0)$ such that $G(0,0,0) = (0,0)$, as desired. □

6. *For the equation $xy+yz+xz=1$, at which points on the solution set S is there a neighborhood in which S is a smooth 2-surface? At each such point (a,b,c), find an equation of the tangent plane.*

Solution: Let $f(x,y,z) = xy+yz+xz-1$, so S is the solution set of $f(x,y,z)=0$. We have

$$df(x,y,z) = \begin{pmatrix} y+z & x+z & x+y \end{pmatrix}.$$

By the corollary of the implicit function theorem (Corollary 9.7.3), S is locally a smooth 2-surface at each point on S where the rank of $df(x,y,z)$ is 1. The only other possibility is that the rank is 0 (because there is only 1 row, and the rank is the dimension of the space spanned by the rows, hence is at most 1). The rank being 0 would mean that $df(x,y,z)$ is the zero matrix, which would mean that $y=-z$, $x=-z$, and $x=-y$. The first two of these equations imply that $x=y$, and combined with $x=-y$, this implies that $y=-y$, hence $y=0$, and then $x=0$ and $z=0$ too. The point $(0,0,0)$ is not in S since $f(0,0,0) = -1 \neq 0$. So for every point $(a,b,c) \in S$, there is some neighborhood of (a,b,c) in which S is a smooth 2-surface.

Finally, the equation for the tangent plane at (a,b,c) is the set of (x,y,z) satisfying

$$df(a,b,c)((x,y,z)-(a,b,c)) = 0,$$

i.e.,

$$\begin{pmatrix} b+c & a+c & a+b \end{pmatrix} \begin{pmatrix} x-a \\ y-b \\ z-c \end{pmatrix}$$

=0. Multiplying the matrices, we get the following equation for the tangent plane to S at (a,b,c):

$$(b+c)(x-a)+(a+c)(y-b)+(a+b)(z-c) = 0.$$

□

7. *For the system of equations*

$$x^2+y^2-z^2 = 0,$$
$$x+y+z = 0,$$

at which points of the solution set S is there a neighborhood in which S is a smooth curve? At each such point, find an equation of the tangent line.

Solution: Let $F(x, y, z) = (x^2 + y^2 - z^2, x + y + z)$, so S is the solution set to $F(x, y, z) = 0$. We have

$$dF(x, y, z) = \begin{pmatrix} 2x & 2y & -2z \\ 1 & 1 & 1 \end{pmatrix}.$$

The corollary of the implicit function theorem (Corollary 9.7.3) tells us that S is locally a smooth curve near any point $(a, b, c) \in S$ for which $dF(a, b, c)$ has rank 2 (and possibly near other points too, but the problem intends for you to just look at the points where the implicit function theorem applies). Clearly, dF has rank at least 1 since the second row is nonzero. There are 2 rows, so the rank will equal 2 if and only if the first row is not a scalar multiple of the second (i.e., if and only if the two rows are linearly independent).

If $(2a, 2b, -2c)$ is a multiple λ of $(1, 1, 1)$, then $(2a, 2b, -2c) = (\lambda, \lambda, \lambda)$, and so $a = b = -c$. That is, $dF(a, b, c)$ has rank 2 unless $a = b = -c$, in which case it has rank 1. Plugging $x = y = -z$ in to f yields $F(x, x, -x) = (x^2, x)$, and this equals 0 only at $x = 0$, i.e., only at the point $(0, 0, 0)$. So $(0, 0, 0)$ is the only point in S for which there is not a neighborhood where S is a smooth curve.

Now consider some nonzero point $(a, b, c) \in S$. The tangent line to S at (a, b, c) is the set of (x, y, z) satisfying

$$dF(a, b, c)((x, y, z) - (a, b, c)) = 0,$$

i.e.,

$$\begin{pmatrix} 2a & 2b & -2c \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x - a \\ y - b \\ z - c \end{pmatrix} = 0.$$

Performing the matrix multiplication yields the following system of equations for your tangent line:

$$2a(x - a) + 2b(y - b) - 2c(z - c) = 0$$
$$(x - a) + (y - b) + (z - c) = 0.$$

(You could of course simplify this a bit, but I won't bother). □

HOMEWORK 13 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

Section 10.1

1. *Let $R = [0,1] \times [0,1]$ be the square with vertices at $(0,0)$, $(1,0)$, $(1,1)$, and $(0,1)$ and let P be the partition of R consisting of the partition $\{0, 1/4, 1/2, 3/4, 1\}$ in both factors of $[0,1] \times [0,1]$. Find $U(f,P)$ and $L(f,P)$ if $f(x,y) = xy$.*

Solution: The volume of each subrectangle R_j of R corresponding to the partition P is $(\frac{1}{4})(\frac{1}{4}) = \frac{1}{16}$. For each R_j, $\sup_{R_j} f$ is the value of $f(x,y)$ at the top right corner of R_j, while $\inf_{R_j} f$ is the value of f at the bottom left corner of R_j. We now compute the upper and lower sums (factoring out the $V(R_j)$'s since they each equal $\frac{1}{16}$, adding the contributions of the R_j's from the bottom left to the top right, and factoring an additional $\frac{1}{16}$ out of each $\sup_{R_j} f$ and $\inf_{R_j} f$) as:

$$\begin{aligned} L(f,P) &= \sum_{j=1}^{16} V(R_j) \inf_{R_j} f(x,y) \\ &= \left(\frac{1}{16}\right)\left(\frac{1}{16}\right)(0+0+0+0+0+1+2+3+0+2+4+6+0+3+6+9) = \frac{36}{16^2} = \frac{9}{64} \end{aligned}$$

(or 0.140625), and

$$\begin{aligned} U(f,P) &= \sum_{j=1}^{16} V(R_j) \sup_{R_j} f(x,y) \\ &= \left(\frac{1}{16}\right)\left(\frac{1}{16}\right)(1+2+3+4+2+4+6+8+3+6+9+12+4+8+12+16) = \frac{100}{16^2} = \frac{25}{64} \end{aligned}$$

(or 0.390625).

3. *Suppose f and g are functions defined on an aligned rectangle R. Suppose there is a positive constant K such that $|f(x) - f(y)| \leq K|g(x) - g(y)|$ for all $x, y \in R$. Prove that if g is integrable on R, then so is f.*

Date: September 6, 2019.

Proof. Since g is integrable on R, given $\epsilon > 0$, there exists a partition P such that $U(g,P)-L(g,P) \leq \frac{\epsilon}{K}$. Let $R_1,\dots,R_n$ be the subrectangles of R for this P. We have

$$\begin{aligned}
U(f,P)-L(f,P) &= \sum_{j=1}^{n}(\sup_{R_j} f - \inf_{R_j} f)V(R_j) \\
&= \sum_{j=1}^{n} \sup\{|f(x)-f(y)| : x,y \in R_j\}V(R_j) \\
&\leq \sum_{j=1}^{n} \sup\{K|g(x)-g(y)| : x,y \in R_j\}V(R_j) \\
&= K\sum_{j=1}^{n}(\sup_{R_j} g - \inf_{R_j} g)V(R_j) \\
&= K(U(g,P)-L(g,P)) \\
&< \epsilon.
\end{aligned}$$

Hence, f is integrable (by Theorem 10.1.7), as desired. [Note: The proof can be modified to use Theorem 10.1.8 instead.] □

4. *Use the result of the preceding exercise to prove that if f is an integrable function on an aligned rectangle R, then $|f|$ is also integrable on R.*

Proof. For each $x,y \in R$, the second form of the triangle inequality tells us that $||f(x)|-|f(y)|| \leq |f(x)-f(y)|$. So the previous problem applies with $K=1$, f there being $|f|$ here, and g there being f here. The result then is that f being integrable on R implies that $|f|$ is integrable on R, as desired. □

5. *Prove that if f is integrable on R, then f^2 is also integrable on R.*

Proof. Suppose f is integrable on R. By our definition of integrability, f must be bounded on R. Let $M>0$ be an upper bound for $|f(x)|$, $x \in R$. For each $x,y \in R$, we have

$$|f(x)^2-f(y)^2| = |f(x)-f(y)||f(x)+f(y)| \leq |f(x)-f(y)|(|f(x)|+|f(y)|) \leq 2M|f(x)-f(y)|.$$

So Exercise 10.1.3 above applies with f there being f^2 here, g there being f here, and K there being $2M$ here. The result then is that, since f is integrable, f^2 must be too, as desired. □

6. *Use the result of the preceeding exercise to prove that if f and g are integrable on R, then fg is also integrable on R.*

Proof. Assume f and g are integrable on R. Note that $(f+g)^2 = f^2+2fg+g^2$, so

$$fg = \frac{1}{2}\left[(f+g)^2 - f^2 - g^2\right]. \tag{1}$$

Theorem 10.1.10 tells us that $f+g$ is integrable on R, and then previous exercise ensures that $(f+g)^2$, f^2, and g^2 are each integrable on R. Applying Theorem 10.1.10 again implies that the right-hand side of (1) is integrable, so fg is integrable, as desired. □

Section 10.2

2. *Prove Theorem 10.2.7 – that is, show that if a subset A of $\mathbb{R}^d$ has outer volume zero, the it and each of its subsets is a Jordan region of volume* 0.

Proof. If A has outer volume zero, then by Theorem 10.2.3, so does every subset of A. So it suffices to prove that A must be a Jordan region of volume 0, since the same argument would then apply to each subset of A.

Since $0 \leq \chi_A(x)$ for all x, and since $\int_R 0 dV(x) = 0$, Theorem 10.1.11 ensures that $0 \leq \underline{V}(A)$ (in fact, by this argument, lower volumes are always non-negative). On the other hand, Theorem 10.1.5 implies that $\underline{V}(A) \leq \overline{V}(A) = 0$ (and by this argument, lower volumes are always $\leq$ the corresponding upper volumes). So we have $0 \leq \underline{V}(A) \leq 0$, hence $\underline{V}(A) = 0 = \overline{V}(A)$, meaning that A is a Jordan region of volume 0, as desired.

[Note: I'm ok with you not being as careful as me here with proving that lower volumes are always non-negative and less than or equal to the corresponding upper volumes.] □

4. *If E is the subset of the unit square $[0,1] \times [0,1]$ consisting of points with both coordinates rational numbers, find its inner volume $\underline{V}(E)$ and outer volume $\overline{V}(E)$. Is E a Jordan region.*

Solution: Note that the interior of E is $E^\circ = \emptyset$, and the closure of E is $\overline{E} = [0,1] \times [0,1]$. So by Theorem 10.2.5, we have

$$\underline{V}(E) = \underline{V}(E^\circ) = \underline{V}(\emptyset) = 0$$

and

$$\overline{V}(E) = \overline{V}(\overline{E}) = \overline{V}([0,1] \times [0,1]) = 1.$$

Since $0 \neq 1$, E is not a Jordan region. □

6. *Let U be an open subset of $\mathbb{R}^2$, and let $K \subset U$ be a compact set. Suppose $f : U \to R$ is a smooth function and $E = \{(x,y) \in K : f(x,y) = 0\}$. If df is never 0 on E, then show that E is a set of area 0 in $\mathbb{R}^2$.*

Proof. Let (a,b) be a point of E. Since $df(a,b) \neq 0$, either $\frac{\partial f}{\partial x}(a,b) \neq 0$ or $\frac{\partial f}{\partial y}(a,b) \neq 0$ (or both). If the latter, the Implicit function theorem says that there is some neighborhood $V \subset U$ of (a,b), a neighborhood A of a, and a smooth function $G : A \to \mathbb{R}$ such that $E \cap V$ is the graph of G, i.e., the set of points of the form $(x, G(x))$ for $x \in A$. If the former (i.e., if $\frac{\partial f}{\partial x}(a,b) \neq 0$), the Implicit function theorem says that there is some neighborhood $V \subset U$ of (a,b), a neighborhood A of b, and a smooth function $G : A \to \mathbb{R}$ such that $E \cap V$ is the graph of G, i.e., the set of points $(G(y), y)$ for $y \in A$. Either way, Example 10.2.11 implies that $E \cap V$ has volume 0.

Now, we have that for every $(a,b) \in E$, there is some neighborhood V such that $E \cap V$ has volume 0. These neighborhoods V thus form an open cover $\mathcal{V}$ for E. Since $f^{-1}(0)$ is closed (because it is the inverse image of a closed set) and K is compact (closed and bounded), we see that $E = f^{-1}(0) \cap K$ is compact (a finite intersection of closed sets is closed, and any intersection with a bounded set is bounded). Hence, $\mathcal{V}$ has a finite subcover. So E is a *finite* union of subsets of volume 0, hence E itself has volume 0 by Theorem 10.2.7 (or Exercise 10.2.5). □

[**Note:** The issue of compactness above is subtle, and probably a lot of people missed it.]

7. *Show that an ellipse in* $\mathbb{R}^2$ *is a set of area* 0 *in* $\mathbb{R}^2$ *and that the solid ellipse that it bounds is a Jordan region.*

Proof. An ellipse can be presented as a set of the form $\{(x,y) \in \mathbb{R}^2 : f(x,y) = 0\}$ for $f : \mathbb{R}^2 \to \mathbb{R}^2$ a smooth function of the form $f(x,y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$ for some constants A, B, C, D, E, F with $B^2 - 4AC < 0$. Also, ellipses are always compact, so we can instead write this as a set of the form $\{(x,y) \in K : f(x,y) = 0\}$ for some compact subset $K \subset \mathbb{R}^2$. Showing that *df* is never 0 on this set is, in general, a bit messier than I'd expected, but it could be done, and then the previous problem applies to say that the ellipse has zero volume (area).

[Alternatively, I'm fine with you just using geometric intuition to say that you can always split an ellipse up into, say, 4 pieces, and each can be expressed as a graph, so Example 10.2.11 applies to say it has 0 volume. I'm also ok with you restricting to, say, ellipses centered at the origin and with B above equal to 0 (i.e., not rotated).]

Finally, since the boundary of the solid ellipse is the ellipse which we just showed has volume 0, Theorem 10.2.9 implies that the solid ellipse is a Jordan region. □

Section 10.3

9. *Prove that if* f *is a bounded function on a set* A *of volume* 0*, then* f *is integrable on* A *and* $\int_A f(x)dV(x) = 0$.

Proof. Let R be an aligned rectangle containing A. Let M be an upper bound for $|f|$ on A. Then $-M\chi_A \le f_A \le M\chi_A$ everywhere on R, so since $V(A) = 0$, we have

$$0 = -MV(A) = \underline{\int}_R - M\chi_A(x)dV(x) \le \underline{\int}_R f_A(x)dV(x)$$
$$\le \overline{\int}_R f_A(x)dV(x) \le \overline{\int}_R M\chi_A(x)dV(x) = MV(A) = 0.$$

Hence, $\underline{\int}_R f_A(x)dV(x) = \overline{\int}_R f_A(x)dV(x) = 0$, meaning that f is integrable on A with $\int_A f(x)dV(x) = \int_R f_A(x)dV(x) = 0$. □

14. *Suppose A is a Jordan region in $\mathbb{R}^d$ and g_k is an integrable function on A for $k = 1, 2, \ldots$. Prove that if $g(x) = \sum_{k=1}^{\infty} g_k(x)$, where this series converges uniformly on A, then g is integrable and*

$$\int_A g(x)dV(x) = \sum_{k=1}^{\infty} \int_A g_k(x)dV(x). \tag{2}$$

Proof. As in the hint on Canvas, let $s_n(x)(x) = \sum_{k=1}^{n} g_k(x)$, and let

$$S_n = \sum_{k=1}^{n} \int_A g_k(x)\, dV(x) = \int_A \sum_{k=1}^{n} g_k(x)\, dV(x) = \int_A s_n(x)dV(x).$$

Then $g(x)$ is, by definition, $\lim_{n\to\infty} s_n(x)$, and this limit converges uniformly,. Also, each $s_n(x)$ is a finite sum of integrable functions, hence is integrable. So g is integrable by Theorem 10.3.11.

Now, the left-hand side of (2) is just $\int_A \lim_{n\to\infty} s_n(x)dV(x)$. On the other hand, the right-hand side of (2) is $\lim_{n\to\infty} S_n$. Thus, (2) can be rewritten as

$$\int_A \lim_{n\to\infty} s_n(x)dV(x) = \lim_{n\to\infty} \int_A s_n(x)dV(x).$$

This follows from Theorem 10.3.11. □

HOMEWORK 14 SOLUTIONS, MATH 3220-001, FALL 2017

TRAVIS MANDEL

This assignment is not being collected, but it is recommended as practice for the Final Exam. Note: The first problem (Exercise 10.4.7) is probably too much like a Calc III problem for me to want to put it on your final. Also, I probably won't want computations as difficult as Exercise 10.5.7.

Section 10.4

7. *Find* $\int_A x dV(x,y,z)$ *if* A *is defined by the inequalities*

$$0 \leq x \leq 1, \qquad 0 \leq y \leq x^2 \qquad 0 \leq z \leq x+y$$

Solution: Using Theorem 10.4.10, we can rewrite this integral as

$$\begin{aligned}\int_0^1 \int_0^{x^2} \int_0^{x+y} x dz dy dx &= \int_0^1 \int_0^{x^2} x(x+y) dy dx \\ &= \int_0^1 \left(x^2 y + \frac{1}{2} x y^2 \right) \Bigg|_{y=0}^{y=x^2} dx \\ &= \int_0^1 \left(x^4 + \frac{1}{2} x^5 \right) dx \\ &= \frac{1}{5} x^5 + \frac{1}{12} x^6 \Bigg|_0^1 \\ &= \frac{1}{5} + \frac{1}{12} \\ &= \frac{17}{60}.\end{aligned}$$

□

10. *Use Fubini's Theorem and the previous exercise to prove that if* $A \subset \mathbb{R}^p$ *and* $B \subset \mathbb{R}^q$ *are Jordan regions, then* $V(A \times B) = V(A)V(B)$*. [What the previous Exercise gives you is that* A *and* B *being Jordan regions implies that* $A \times B$ *is a Jordan region.]*

Proof. We know that A, B, and (by Exercise 10.4.9) $A \times B$ are Jordan regions, so their volumes are well-defined. Let S be a rectangle containing A, and let T be a rectangle containing B, so $S \times T$ is a rectangle containing $A \times B$.

We write points of $\mathbb{R}^{p+q}$ as (x,y) for $x \in \mathbb{R}^p$ and $y \in \mathbb{R}^q$. By definition, $(x,y) \in A \times B$ if and only if $x \in A$ and $y \in B$. It follows that $\chi_{A\times B}(x,y) = \chi_A(x)\chi_B(y)$.

Date: September 6, 2019.

We have

$$\begin{aligned}
V(A\times B) :&= \int_{S\times T} \chi_{A\times B}(x,y)dV(x,y) \\
&= \int_S \int_T \chi_A(x)\chi_B(y)dV(y)dV(x) \qquad \text{(By Fubini's Theorem)} \\
&= \int_S \chi_A(x) \int_T \chi_B(y)dV(y)dV(x) \qquad \text{(because } \chi_A(x) \text{ does not depend on } y) \\
&= \int_T \chi_B(y)dV(y) \int_S \chi_A(x)dV(x) \qquad \text{(because } \textstyle\int_T \chi_B(y)dV(y) \text{ does not depend on } x) \\
&= V(B)V(A),
\end{aligned}$$

as desired. □

Section 10.5

7. *Let $A = \{(x,y) \in \mathbb{R}^2 : x \geq 0, y \geq 0, x^2+y^2 \leq 4, x^2-y^2 \geq 1\}$. Compute*

$$\int_A \frac{xy}{x^2+y^2} dV(x,y)$$

by making a change of variables $u = x^2+y^2$, $v = x^2-y^2$ for $x \geq 0$, $y \geq 0$.

Solution: Let us express the change-of-variables as $(u,v) = \phi(x,y) := (x^2+y^2, x^2-y^2)$. Let us restrict to the interior of A (the boundary has volume 0 so this does not affect the integral). Let $B = \phi(A^\circ)$. We have $d\phi(x,y) = \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix}$, and this has determinant $-8xy$. This is non-zero on A°, and ϕ is one-to-one on A, so by the inverse function theorem, we have a smooth inverse $\phi^{-1} : B \to A$. Let $f(x,y) = \frac{xy}{x^2+y^2}$. By the change-of-variables formula, we have

$$\int_A f(x,y)dV(x,y) = \int_B f\circ\phi^{-1}(u,v)|\det d\phi^{-1}(u,v)|dV(u,v).$$

By the Inverse Function Theorem, $|\det d\phi^{-1}(u,v)| = \frac{1}{8xy}$, so $f\circ\phi^{-1}(u,v)|\det d\phi^{-1}(u,v)| = \frac{f(x,y)}{8xy} = \frac{1}{8(x^2+y^2)} = \frac{1}{8u}$.

The set B can be expressed as $\{(u,v) \in \mathbb{R}^2 : u \geq v, u \leq 4, v \geq 1\}$ (to find this, it helps to find that the inverse is given by $x = \sqrt{\frac{1}{2}(u+v)}$ and $y = \sqrt{\frac{1}{2}(u-v)}$. It also helps to draw the regions).

Finally, by Fubini's theorem, the integral becomes:

$$\begin{aligned}\int_1^4 \int_v^4 \frac{1}{8u} du dv &= \int_1^4 \ln(8u)|_v^4 dv \\ &= \int_1^4 [\ln(32) - \ln(8v)] dv \\ &= \int_1^4 [\ln(4) - \ln(v)] dv \text{ (using properties of ln)} \\ &= 3\ln(4) - (v\ln(v) - v)]|_1^4 \\ &= -\ln(4) + 4 - 1 \\ &= 3 - \ln(4).\end{aligned}$$

□

11. *Prove Corollary 10.5.11. That is, assuming Theorem 10.5.10, prove the following: Let U be an open subset of $\mathbb{R}^d$ and let $\phi : U \to \mathbb{R}^d$ be a smooth one-to-one transformation with non-singular differential on U. If R is an aligned rectangle in U, then $V(\phi(R)) = \int_R |\det d\phi(x)| dV(x)$. Furthermore, if $M = \sup_R |\det d\phi|$ and $m = \inf_R |\det d\phi|$, then $mV(R) \le V(\phi(R)) \le MV(R)$.*

Proof. First, the fact that $\phi(R)$ is a Jordan region is actually Theorem 10.5.8. Let S be a rectangle containing $\phi(R)$. Now, by definition,

$$V(\phi(R)) = \int_S \chi_{\phi(R)}(u) dV(u),$$

and by the definition of the integral over a Jordan region, this is the same as

$$\int_{\phi(R)} dV(u).$$

Now by the change-of-variables formula (Theorem 10.5.10), this equals

$$\int_R |\det d\phi(x)| dV(x),$$

as desired.

Finally, since $m \le |\det d\phi(x)| \le M$ for all $x \in R$, we have

$$mV(R) = \int_R m dV(x) \le \int_R |\det d\phi(x)| dV(x) \le \int_R M dV(x) = MV(R),$$

where we have just shown that the middle part of the above inequalities is $V(\phi(R))$. □