Math 233H Supplementary Notes, Fall 2021

Mark Williams

Contents

1 Introduction

These notes were written for the honors section of multivariable calculus (Calc III) taught at UNC during the Fall 2021 semester. The main text was *Calculus: Early transcendentals* by W. Briggs et al.

Dear Math 233H students,

In these notes I will provide some background material, a more rigorous and/or complete discussion of certain topics covered in the text, and some proofs not covered in class. The notes sometimes clarify points in homework sets or topics whose presentation in class was somehow inadequate, whether incomplete or incorrect. Most installments are associated with the topic of a nearby class.[1] A number of the reading assignments will be from these notes. In most cases the readings will preview, review, or supplement what is done in class. In other cases I will expect you to learn the material in the reading mainly from the reading.

An inserted "why?" or "how?" or "check!" indicates a point that is supposed to be obvious, at least after a bit of thought. You should be able to answer most of those questions if you are following well what is written.

Best,
Mark Williams

2 Sets and functions

Here we recall some basic facts and notation concerning sets and functions that we use repeatedly throughout the course. For our purposes a *set* is defined to be a "well-defined" collection of distinct objects. The objects could be numbers, or letters, or apples, or planets, or other sets, or ... In section 14 we give a critique of this definition and discuss some other aspects of set theory. Section 14, like sections 7 and 15, can be read independently of the rest of these notes.

2.1 Some set notation

One way to specify a *finite* set is to *list* its elements. When we do this, each distinct element should be listed just once. But many sets are *infinite*, for example $\mathbb{N}$ (the natural numbers), $\mathbb{Z}$ (the integers), $\mathbb{Q}$ (the rational numbers), and $\mathbb{R}$ (the real numbers).[2] We can't specify an infinite set by

[1] The notes were written as the semester progressed.

[2] These symbols, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ for the natural numbers, integers, etc., are quite standard. It is good to remember them.

listing all its elements, so we adopt other strategies like

$$\mathbb{N} = \{1, 2, 3, ...\}, \quad \mathbb{Q} = \{p/q : p, q \text{ are integers with } q \neq 0\}, \quad \mathbb{R} = \{r : r \text{ is rational or irrational}\}.$$

If A and B are sets we write $A \subset B$ when A is a *subset* of B. This means that whenever $a \in A$ (read: "a is an element of A"), then $a \in B$. We write $A \cup B$ to denote the set which is the *union* of A and B, so $A \cup B = \{c : c \in A \text{ or } c \in B\}$. We write $A \cap B$ to denote the set which is the *intersection* of A and B, so $A \cap B = \{c : c \in A \text{ and } c \in B\}$. For example, if $A = \{1, 2, 3, 4, 5\}$ and $B = \{3, 4, 5, 6, 7\}$, then $A \cup B = \{1, 2, 3, 4, 5, 6, 7\}$, while $A \cap B = \{3, 4, 5\}$. The symbol $\emptyset$ is used to denote the set which contains no elements. Thus if $A = \{1, 2\}$ and $B = \{3, 4\}$, then $A \cap B = \emptyset$. On the other hand $\emptyset \cup A = A$ and $\emptyset \cap A = \emptyset$ for any set A. If $A \subset B$, we denote by $B \setminus A$ the *complement* of A relative to B, that is, $B \setminus A = \{b \in B : b \notin A\}$. Thus, for example, $\mathbb{R} \setminus \mathbb{Q} = \{r : r \text{ is irrational}\}$. When the "ambient" set B containing A is clear from the context, we often write A^c for $B \setminus A$.

If A and B are *any* sets, then the *cartesian product* of A and B is the set $A \times B := \{(a, b) : a \in A, b \in B\}$. That is, it is the set of *all* ordered pairs (a, b) with $a \in A$, $b \in B$. In general, $A \times B \neq B \times A$. Observe that $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$, and so on.

2.2 Functions and function notation.

Let A and B be *any* sets. We say that f is a *function* with domain A and target space B, that is $f : A \to B$, provided it is a rule that assigns to each element of A just one element of B.[3] We say that f is *one-to-one or injective* provided no two elements of A are assigned to the same element of B, that is, $f(a_1) = f(a_2) \Rightarrow a_1 = a_2$. We say that f is *onto or surjective* when $f(A) = B$, that is, when every element of B lies in the *range*, $f(A)$, of f. For example, suppose $b \in B$ and suppose f is defined as the rule that assigns *every* element of A to b, that is, $f(a) = b$ for all $a \in A$. This *is* a function, but if A contains more than one element, this function is not one-to-one. If B contains more than one element, this function is not surjective. As another example, let $A = \{1, 2, 3\}$ and $B = \{\text{Mary, Bill, Al}\}$. The rule that assigns 1 to Mary, 2 to Bill, and 3 to Al is a function that is both injective and surjective. For any sets A and B a function $f : A \to B$ that is both injective and surjective is said to be *bijective* or *a one-to-one correspondence from A to B*.

Given a function $f : A \to B$, the *graph* of f is the set $G_f := \{(a, f(a)) : a \in A\}$.[4] The notation "$\{(a, f(a)) : a \in A\}$" is read: "the set of *all* ordered pairs $(a, f(a))$ such that $a \in A$." Check that this definition is consistent with the informal concept of "graph of a function" that you learned in pre-calculus. Observe that $G_f \subset A \times B$. On the other hand if $b \in B$, the *level set* of f corresponding to b is the set $L_f(b) := \{a \in A : f(a) = b\}$. Observe that $L_f(b) \subset A$, and it can be empty even when A and B are nonempty.

If $E \subset A$, the *image* of E under a function $f : A \to B$ is the set $f(E) := \{f(a) : a \in E\} \subset B$. Thus, the *range* of f is $f(A) \subset B$. If $C \subset B$, the *preimage* of C under the function f is the set $f^{-1}(C) := \{a \in A : f(a) \in C\}$. This commonly used notation has the drawback that it is easy to confuse with the *image* of C under the inverse function f^{-1}. Recall that the *inverse* of a given

[3] A more precise, set-theoretic definition of function is given near the end of section 14. That definition avoids use of the vague words "rule" and "assigns".

[4] When $f : \mathbb{R}^2 \to \mathbb{R}$, for example, most people write points on the graph of f in the form $(x, y, f(x, y))$, rather than $((x, y), f((x, y)))$, as our definition of G_f might seem to dictate. Thank heavens!

function $f : A \to B$ exists if and only if f is bijective; however, preimages as just defined always exist. *When* the inverse of f exists and $C \subset B$, then the two possible interpretations of $f^{-1}(C)$ coincide (why?). In a given context the meaning of $f^{-1}(C)$ is usually clear.

In the last part of this section we list a few examples of sets of functions whose definitions will be fully explained later in the course. Let $k \in \mathbb{N}$. If $\mathcal{O} \subset \mathbb{R}^n$ is an open set (defined in section 6), we define the set of functions

$$C^k(\mathcal{O}, \mathbb{R}) = \{f : \mathcal{O} \to \mathbb{R} \text{ such that } f \text{ and all its partials of order } \leq k \text{ are continuous on } \mathcal{O}\}.$$

This is the infinite set consisting of *all* functions with the stated properties. When the target set of the functions is understood to be $\mathbb{R}$, many authors write just $C^k(\mathcal{O})$ instead. Another commonly used set of functions is[5]

$$C^k([a,b]) := \{f : [a,b] \to \mathbb{R} \text{ such that } f^{(j)} \text{ is continuous on } [a,b] \text{ for } 0 \leq j \leq k\}.$$

In this definition it is the *one-sided* derivatives that are continuous at the endpoints.

To describe curves in $\mathbb{R}^n$ we often use the set of functions

$$C^1([a,b], \mathbb{R}^n) = \{r : [a,b] \to \mathbb{R}^n \text{ such that each component function of } r \text{ lies in } C^1([a,b])\}.$$

Later, when we parametrize surfaces in $\mathbb{R}^3$ we will use sets of functions like

$$C^1(\mathcal{O}, \mathbb{R}^3) = \{r : \mathcal{O} \to \mathbb{R}^3 \text{ such that each component function of } r \text{ lies in } C^1(\mathcal{O}, \mathbb{R})\},$$

where $\mathcal{O} \subset \mathbb{R}^2$ is open.

3 Linear independence and linear dependence

In HW 2 we defined the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ in $\mathbb{R}^m$ to be *linearly independent* when the only choice of constants $c_1, \ldots, c_n$ such that $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$ is the choice $c_1 = 0, \ldots, c_n = 0$. If that is not the case, the vectors are said to be *linearly dependent*. So what does it mean to say that it is not the case that the only choice of constants $c_1, \ldots, c_n$ such that $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$ is the choice $c_1 = 0, \ldots, c_n = 0$? Give your own answer to this question before reading on.

This is an example of a situation that comes up often in mathematics - we are asked to negate an assertion in English that is slightly complicated. In principle, all that is required to do this is a sound knowledge of the English language and the ability to think clearly. Still, for many of us it takes a little practice to do this confidently and correctly.

Here is the answer to the question posed above. It means that *there exists* a choice of constants which are not *all* zero such that $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$. Note: To say that the constants are not all zero is completely different from saying that they are all not zero! So now we can reformulate the definition of linear dependence in a more useful way: the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are linearly dependent if there exists a choice of constants $c_1, \ldots, c_n$ which are not all zero such that $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n = \mathbf{0}$.

Definition 3.1. *(a) If $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ in $\mathbb{R}^m$ and $c_1, \ldots, c_n$ are constants, we refer to $c_1\boldsymbol{u}_1 + \cdots + c_n\boldsymbol{u}_n$ as a* linear combination *of the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$.*

[5] When $k = 0$ the functions are merely continuous and we usually write $C([a,b])$.

(b) The set of all possible linear combinations of the $\boldsymbol{u}_i$, that is,

$$\{c_1\boldsymbol{u}_1 + \cdots + c_n\boldsymbol{u}_n : c_i \in \mathbb{R} \text{ for } i = 1, \ldots, n\},$$

is called the linear span *(or simply, the* span*) of the $\boldsymbol{u}_i$. We denote this set by* span $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$.

Proposition 3.2. *The set of vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n, \boldsymbol{0}$ in $\mathbb{R}^m$ is linearly dependent.*

Proof. Observe $0\mathbf{u}_1 + 0\mathbf{u}_2 + \cdots + 0\mathbf{u}_n + 154\,\mathbf{0} = \mathbf{0}$. We have exhibited a choice of constants, which are not all zero, such that the corresponding linear combination of the given vectors is $\mathbf{0}$, so we are done.

□

Now it should be easier to understand the footnote in HW 2 where I proved that the the vectors $\langle 1,1,1\rangle$, $\langle 1,0,1\rangle$, $\langle 3,2,3\rangle$ are linearly dependent. Observe that

$$\text{span}\,\{\langle 1,1,1\rangle, \langle 1,0,1\rangle, \langle 3,2,3\rangle\} = \text{span}\,\{\langle 1,1,1\rangle, \langle 1,0,1\rangle\} := \mathcal{P}, \tag{3.1}$$

and that the set of tips of vectors in $\mathcal{P}$, when tails are placed at the origin, is the plane through the origin determined by $\langle 1,1,1\rangle$ and $\langle 1,0,1\rangle$. We often say, somewhat sloppily, that $\mathcal{P}$ *is* the plane through the origin determined by $\langle 1,1,1\rangle$ and $\langle 1,0,1\rangle$

Here is an exercise for you to check your understanding (not to turn in). Prove the following proposition:

Proposition 3.3. *Suppose the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ in $\mathbb{R}^m$ are linearly dependent. Show that the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{u}_{k+1}, \ldots, \boldsymbol{u}_p$ are linearly dependent.*

4 Notation for points and vectors

Let $n \geq 1$. Recall that $\mathbb{R}^n = \{(x_1, \ldots, x_n) : x_j \in \mathbb{R} \text{ for } j = 1, \ldots, n\}$. When we define $\mathbb{R}^n$ in this way we think of it as a set of *points*. Sometimes we think of it differently as a set of *vectors* whose tails are all placed at the origin, that is, $\mathbb{R}^n = \{\langle x_1, \ldots, x_n\rangle : x_j \in \mathbb{R} \text{ for } j = 1, \ldots, n\}$. The set of *tips* of those vectors is of course $\{(x_1, \ldots, x_n) : x_j \in \mathbb{R} \text{ for } j = 1, \ldots, n\}$. Thus, even though a set of points and a set of vectors are, strictly speaking, completely different objects, it is not unreasonable to use the same notation, $\mathbb{R}^n$, for *both* sets. Moreover, "everybody" does this. The *context* will tell us in any given case whether $\mathbb{R}^n$ is a set of points or a set of vectors. Sometimes it won't matter which interpretation we use.

From now on I will often drop the use of boldface when denoting elements of $\mathbb{R}^n$, just as in class we have started to drop the arrow when denoting vectors. So instead of writing $\mathbf{x} \in \mathbb{R}^n$, I will write simply $x \in \mathbb{R}^n$, and instead of $\mathbf{x} = (x_1, \ldots, x_n)$, I will write $x = (x_1, \ldots, x_n)$. Again, we will use *context* to tip us off that the symbol x means an element of $\mathbb{R}^n$ for some specific n rather than, say, the first component of a point $(x, y) \in \mathbb{R}^2$.

If $x = (x_1, \ldots, x_n)$ is a *point* in $\mathbb{R}^n$, then by $|x|$ we mean $\sqrt{x_1^2 + \cdots + x_n^2}$, the distance of that point from the origin $0 \in \mathbb{R}^n$. Similarly, if $x = \langle x_1, \ldots, x_n\rangle$ is a *vector* in $\mathbb{R}^n$ with its tail at the origin, then by $|x|$ we mean $\sqrt{x_1^2 + \cdots + x_n^2}$, the length of that vector. Of course, $x \in \mathbb{R}^n$ thought of as a point is just the tip of $x \in \mathbb{R}^n$ thought of as a vector. Similarly, if x, y are points in $\mathbb{R}^n$, the distance between those points is $|x - y| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}$. If x, y are vectors in $\mathbb{R}^n$

with their tails at the origin, then $|x-y|$ is the length of the vector $x-y$, which is the distance between the tips of x and y.

Finally, note that if $x \in \mathbb{R}^n$ where $n=1$, then $|x| = \sqrt{x^2}$, and $\sqrt{x^2}$ *is* the absolute value of x.[6] If x is thought of as a point in $\mathbb{R}^n$ then $|x|$ *always* means the distance of x from the origin, even when $n=1$.

5 Proof of the "two path test" for nonexistence of limits

In class on Tuesday, Sept. 7 we stated (without proof) the following proposition.

Proposition 5.1 (Two-path test). *Let $f : D \subset \mathbb{R}^n \to \mathbb{R}$, let $a \in \mathbb{R}^n$, and suppose there exist two paths γ_1, γ_2 in D such that*[7]

$$\lim_{x \to a \text{ on } \gamma_i} f(x) = L_i, \text{ where } L_1 \neq L_2.$$

Then $\lim_{x\to a} f(x)$ does not exist.

Proof. We give a proof by contradiction. Suppose $L \in \mathbb{R}$ and $\lim_{x\to a} f(x) = L$. Recall that this means: given any $\epsilon > 0$ there exists a $\delta > 0$ such that if $0 < |x-a| < \delta$ and $x \in D$, then $|f(x) - L| < \epsilon$. Since $L_1 \neq L_2$, the number L must differ from at least one of L_1, L_2, say $L \neq L_1$. Define $\epsilon = |L - L_1| > 0$. There exists $\delta_1 > 0$ such that if $0 < |x-a| < \delta_1$ and $x \in D$, then $|f(x) - L| < \frac{\epsilon}{2}$ (why?). On the other hand there exists $\delta_2 > 0$ such that if $0 < |x-a| < \delta_2$ and $x \in \gamma_1$, then $|f(x) - L_1| < \frac{\epsilon}{2}$. Choose $x_0 \in \gamma_1$ such that $0 < |x_0 - a| < \min\{\delta_1, \delta_2\}$. Then by the triangle inequality, we have:[8]

$$|L - L_1| \leq |L - f(x_0)| + |f(x_0) - L_1| < \frac{\epsilon}{2} + \frac{\epsilon}{2} < \epsilon,$$

a contradiction.

□

Remark 5.2. *The way to come up with a proof like this is first to draw pictures. The $\epsilon - \delta$ definition of limit tells us that if $\lim_{x\to a} f(x) = L$, then $f(x)$ will be as close as we like to L provided x is close enough to (but not equal to) a. Now suppose there are two paths with the property stated in Proposition 5.1. A* picture *showing $f(x)$ getting close to L_i as $x \to a$ on γ_i makes it obvious to the person on the street that if such paths exist, there is no way that simply making x close enough to a will guarantee that $f(x)$ is close to some* fixed *L. Having this picture in mind, it was not hard to write the above proof.*

[6]Here we use the standard convention that if $a > 0$, $\sqrt{a}$ denotes the *positive* square root of a.

[7]To say "$\lim_{x\to a \text{ on } \gamma_1} f(x) = L_1$" means: given $\epsilon > 0$ there exists a $\delta > 0$ such that if $0 < |x-a| < \delta$ and $x \in \gamma_1$, then $|f(x) - L_1| < \epsilon$.

[8]The triangle inequality says $|a+b| \leq |a| + |b|$ for any $a, b \in \mathbb{R}$. The inequality remains true when $a, b \in \mathbb{R}^n$; see Prop. 6.6 below.

6 Open and closed sets in $\mathbb{R}^n$.

You are familiar with open intervals $(a,b) \subset \mathbb{R}$, closed intervals $[a,b]$, and intervals that are neither open nor closed such as $[a,b)$. We will now define what it means for subsets of $\mathbb{R}^n$ to be open or closed. The cases $n = 1, 2, 3$ are the most important for this course, but it requires no extra effort to define these concepts for any $\mathbb{R}^n$.

Definition 6.1 (ball)**.** *Let $a \in \mathbb{R}^n$ and $r > 0$. We define the* ball $B(a,r) \subset \mathbb{R}^n$ *centered at a of radius r to be the set $B(a,r) = \{x \in \mathbb{R}^n : |x-a| < r\}$. Here, as always, $|x-a| = \sqrt{(x_1-a_1)^2 + \cdots + (x_n-a_n)^2}$ is the euclidean distance between a and x in $\mathbb{R}^n$.*

For example, if $n = 1$, the ball $B(a,r)$ is the open interval $(a-r, a+r)$ (why?). If $n = 2$ the ball $B(a,r)$ is the set of points in $\mathbb{R}^2$ that lie strictly inside the circle of radius r centered at a (why?). If $n = 3$, the ball $B(a,r)$ is the set of points in $\mathbb{R}^3$ that lie strictly inside the sphere of radius r centered at a. Before going any further, draw these balls and make sure you understand these examples.

Definition 6.2 (open set)**.** *A subset $S \subset \mathbb{R}^n$ is said to be* open *if, for* every *point $a \in S$,* there exists *a corresponding $r > 0$ such that the ball $B(a,r)$ is completely contained in S, that is, $B(a,r) \subset S$.*[9]

Draw a picture and convince yourself that the set of points

$$S_1 = \{(x,y) \in \mathbb{R}^2 : \sqrt{(x-1)^2 + (y-2)^2} < 5\}$$

is open. To do this you should consider an arbitrary point $b \in S_1$, and "see" that one can find a ball of *some* positive radius centered at b that is completely contained in S_1. This "argument" is not a *proof*, but that's ok. At this point, I just want you to be able to recognize quickly when a given set is open or not. A *proof* of openness is given in Proposition 6.8 below.

On the other hand the set

$$S_2 = \{(x,y,z) \in \mathbb{R}^3 : \sqrt{(x-1)^2 + (y-5)^2 + (z+10)^2} \leq 20\}$$

is not open (why?).[10] The set $S_3 = \{(x,y) \in \mathbb{R}^2 : \frac{x^2}{9} + \frac{y^2}{16} < 1\}$ is open, but the ellipse which is the boundary curve of S_3 is not an open set in $\mathbb{R}^2$ (why?). The set $S_4 = \{(x,y) \in \mathbb{R}^2 : 1 < x < 5,\ 8 < y < 10\}$ is open, but $S_5 = \{(x,y) \in \mathbb{R}^2 : 1 < x < 5, 8 < y \leq 10\}$ is not open (why? Draw!).

Definition 6.3 (Interior point of a set)**.** *(a) Let $S \subset \mathbb{R}^n$ and $p \in S$. We say that p is an* interior point *of S if there exists an $r > 0$ such that $B(p,r) \subset S$. In other words, p is an interior point of S if there is some ball centered at p that is completely contained in S.*

(b) If $S \subset \mathbb{R}^n$, we denote by $\mathring{S}$ the set of all interior points of S,"

$$\mathring{S} = \{p \in S : p \textit{ is an interior point of } S\}.$$

We usually refer to $\mathring{S}$ as the "interior of S."

[9] As always, pay close attention to the quantifiers here: "for every", "there exists an". A choice of radius r that "works" for a given a, when such an r exists, usually depends on the choice of a.

[10] Answer: If $c \in S_2$ is a point on the sphere which defines the boundary of this set, it is impossible to find a ball centered at c that is completely contained in S_2.

Remark 6.4. *We can now rephrase the definition of an* open *set: the set $S \subset \mathbb{R}^n$ is open if every point of S is an interior point. Thus, S is open if and only if $S = \mathring{S}$.*

Definition 6.5 (Closed sets)**.** *A set $T \subset \mathbb{R}^n$ is said to be* closed *if its complement $T^c := \mathbb{R}^n \setminus T$ is open.*

Observe that the set $S_2 \subset \mathbb{R}^3$ defined above is closed, since its complement

$$S_2^c = \{(x, y, z) \in \mathbb{R}^3 : |(x, y, z) - (1, 5, -10)| > 20\}$$

is open (check!). The set $S_5 \subset \mathbb{R}^2$ is neither open nor closed. The set $T_1 = \{(x, 0, 0) \in \mathbb{R}^3 : x \in \mathbb{R}\}$, that is, the $x-$axis in $\mathbb{R}^3$, is closed. Let $a_1, \dots, a_n$ denote any collection of n distinct points in $\mathbb{R}^n$. Then $T_2 = \{a_1, \dots, a_n\} \subset \mathbb{R}^n$ is closed (check!).

Next we wish to *prove* that any ball $B(a, r) \subset \mathbb{R}^n$ is an open set. For this it will be helpful to state a few properties of the function $|\cdot| : \mathbb{R}^n \to \mathbb{R}$ given by $x \to |x|$, where $|x| = \sqrt{x_1^2 + \cdots + x_n^2}$.

Proposition 6.6 (Properties of $|\cdot|$)**.** *The function $|\cdot| : \mathbb{R}^n \to \mathbb{R}$ has the following properties:*
(a) for all $x \in \mathbb{R}^n$ we have $|x| \geq 0$, and $|x| = 0 \Leftrightarrow x = 0$.
(b) for $c \in \mathbb{R}$, $|cx| = |c||x|$.
(c) for $x, y \in \mathbb{R}^n$, $|x + y| \leq |x| + |y|$ (the "triangle inequality")

Proof. Properties (a) and (b) are obvious. To prove (c) we write

$$|x + y|^2 = (x + y) \cdot (x + y) = |x|^2 + |y|^2 + 2x \cdot y \leq (|x| + |y|)^2,$$

where we used the Cauchy-Schwarz inequality (proved in HW 1) for the last inequality. □

Remark 6.7. *Any function $N : \mathbb{R}^n \to \mathbb{R}$ having the above three properties is called a* norm *on $\mathbb{R}^n$. There are many other norms on $\mathbb{R}^n$. For example, you can check that the function N_1 defined by $N_1(x) = |x_1| + \cdots + |x_n|$ is a norm on $\mathbb{R}^n$. Another norm is given by $N_2(x) = \max(|x_1|, \dots, |x_n|)$.*[11]

Now we can prove that balls are open.

Proposition 6.8. *Let $a \in \mathbb{R}^n$ and $r > 0$. The ball $B(a, r) = \{x \in \mathbb{R}^n : |x - a| < r\}$ is open.*

Proof. Let $b \in B(a, r)$. We must find a radius $s > 0$ such that $B(b, s) \subset B(a, r)$. We claim that $s = r - |b - a|$ works.[12] Let $q \in B(b, s)$. We will be done if we show $q \in B(a, r)$ (why?). We have[13]

$$|q - a| \leq |q - b| + |b - a| < (r - |b - a|) + |b - a| = r.$$

So $q \in B(a, r)$ (why?). □

In view of Proposition 6.8 we will often refer to $B(a, r)$ as the "open ball" centered at a of radius r. The "closed ball" of radius a centered at r is the set $C(a, r) = \{x \in \mathbb{R}^n : |x - a| \leq r\}$. Clearly, $C(a, r)$ is a closed set.

[11] However, $x \to \min(|x_1|, \dots, |x_n|)$ does not define a norm (why?).
[12] Draw a picture in $\mathbb{R}^2$ to see how we came up with this.
[13] For the first inequality we used the triangle inequality of Prop. 6.6 (how?).

7 Infinite sets: countable versus uncountable

In this section we introduce a precise way of comparing the "size" of infinite sets. The main tool is the idea of a "one-to-one correspondence".

Suppose A and B are sets. We say that there exists a "one-to-one correspondence" between A and B when there exists a function $f : A \to B$ that is both one-to-one (or injective) and onto (or surjective). A map that is both injective and surjective is called bijective. We now recall what all these words mean.

First, we say that f is a *function* with domain A and target space B, that is $f : A \to B$, provided it is a rule that assigns to each element of A just one element of B. We say that f is *one-to-one or injective* provided no two elements of A are assigned to the same element of B, that is, $f(a_1) = f(a_2) \Rightarrow a_1 = a_2$. Finally, we say that f is *onto or surjective* when $f(A) = B$, that is, when every element of B lies in the range of f. For example, suppose $b \in B$ and suppose f is defined as the rule that assigns *every* element of A to b, that is, $f(a) = b$ for all $a \in A$. This *is* a function, but if A contains more than one element, this function is not one-to-one. If B contains more than one element, this function is not surjective. As another example, let $A = \{1, 2, 3\}$ and $B = \{\text{Mary, Bill, Al}\}$. The rule that assigns 1 to Mary, 2 to Bill, and 3 to Al is a function that is both injective and surjective, a one-to-one correspondence.

Clearly, if A and B are finite sets, there exists a one-to-one correspondence between A and B if and only if they contain the same number of elements. But what can we say when the sets in question are infinite sets? Now things get more interesting. Consider the set $\mathbb{N}$ of natural numbers and the subset $\mathbb{N}_{even}$ of even natural numbers $\mathbb{N}_{even} = \{2, 4, 6, \dots\}$. Ask yourself: does there exist a one-to-one correspondence between $\mathbb{N}$ and $\mathbb{N}_{even}$? Since $\mathbb{N}_{even}$ is a proper subset of $\mathbb{N}$, you might be tempted to answer no, but the answer is yes.[14] A one-to-one correspondence is given by the function $f : \mathbb{N} \to \mathbb{N}_{even}$ defined by $f(n) = 2n$ (check). This example shows that it is possible for an infinite set to be in one-to-one correspondence with a *proper* subset of itself! Whenever a one-to-one correspondence exists between two sets, we say that the sets have the same *cardinality*. Thus, $\mathbb{N}$ and $\mathbb{N}_{even}$ have the same cardinality.

Definition 7.1. *A set A that has the same cardinality as $\mathbb{N}$ is said to be a* countable *set. That is, a set A is countable if and only if it can be put into one-to-one correspondence with $\mathbb{N}$.*

Are the integers $\mathbb{Z}$ countable? To see that the answer is yes, observe that there is an obvious way to indicate a listing of the elements of $\mathbb{Z}$: $0, 1, -1, 2, -2, 3, -3, \dots$. Now if we define $f : \mathbb{N} \to \mathbb{Z}$ to be the function that assigns 1 to 0, 2 to 1, 3 to -1, 4 to 2, etc., then f gives a one-to-one correspondence between $\mathbb{N}$ and $\mathbb{Z}$. Are the rational numbers $\mathbb{Q}$ countable? It may surprise you but here again the answer is yes. Every rational is the quotient of two integers (p, q). Let's just focus on seeing that the *positive* rationals are countable. Consider the set $\mathcal{Q}$ of all possible pairs (p, q)

[14] If $A \subset B$ we say that A is a *proper* subset of B when $A \neq B$.

where $p, q \in \mathbb{N}$.[15] There is an obvious way to indicate a listing of the elements of $\mathcal{Q}$:

$$\begin{aligned}&(1,1),(1,2),(1,3),(1,4),\ldots\\&(2,1),(2,2),(2,3),(2,4),\ldots\\&(3,1),(3,2),(3,3),(3,4),\ldots\\&\text{etc.},\end{aligned} \tag{7.1}$$

Now define $f:\mathbb{N}\to\mathcal{Q}$ to be the map that assigns 1 to $(1,1)$, 2 to $(2,1)$, 3 to $(1,2)$, 4 to $(3,1)$, 5 to $(2,2)$, 6 to $(1,3)$, 7 to $(4,1)$, etc. This gives a one-to-one correspondence between $\mathbb{N}$ and $\mathcal{Q}$, and it is not hard to use this map to obtain a one-to-one correspondence between $\mathbb{N}$ and the positive rationals, or between $\mathbb{N}$ and $\mathbb{Q}$.

Are there infinite sets that are not countable? We show next that the set $\mathbb{R}$ cannot be put into one-to-one correspondence with $\mathbb{N}$.

Definition 7.2. *An infinite set that cannot be put into one-to-one correspondence with* $\mathbb{N}$ *is called an* uncountable *set.*

To see that $\mathbb{R}$ is uncountable we use a classical "diagonal argument" due to Georg Cantor and published in 1891. Let us focus on showing just that the reals in the interval $[0,1]$ are uncountable. We argue by contradiction. Suppose the set $[0,1]$ is countable. Then we can use decimal expansions along with any function f that gives a one-to-one correspondence of $\mathbb{N}$ with $[0,1]$ to indicate a listing of all the elements of $[0,1]$:

$$\begin{aligned}f(1)&=.a_{1,1}a_{1,2}a_{1,3}a_{1,4}\ldots\\f(2)&=.a_{2,1}a_{2,2}a_{2,3}a_{2,4}\ldots\\f(3)&=.a_{3,1}a_{3,2}a_{3,3}a_{3,4}\ldots\\&\text{etc.},\end{aligned} \tag{7.2}$$

where the $a_{i,j}\in\{0,1,2,\ldots,9\}$. We will now show how to construct an element of $[0,1]$ that is not in the list, thereby obtaining a contradiction. The element will have the form $.b_1b_2b_3\ldots$. To construct b_1, look at $a_{1,1}$ and set $b_1=\begin{cases}4, & \text{if } a_{1,1}\neq 4\\ 7 & \text{if } a_{1,1}=4\end{cases}$. Thus, no matter how we choose $b_2,b_3,\ldots$ the number $.b_1b_2b_3\ldots$ will not equal $f(1)$. Similarly, set $b_2=\begin{cases}4, & \text{if } a_{2,2}\neq 4\\ 7 & \text{if } a_{2,2}=4\end{cases}$, so no matter how we choose $b_3,b_4,\ldots$ the number $.b_1b_2b_3\ldots$ will differ from both $f(1)$ and $f(2)$. In general, for any n set $b_n=\begin{cases}4, & \text{if } a_{n,n}\neq 4\\ 7 & \text{if } a_{n,n}=4\end{cases}$.[16] Clearly, $.b_1b_2b_3\ldots$ is not equal to any element of the above list. Contradiction.

Since $\mathbb{R}$ contains countable subsets, but cannot be put into one-to-one correspondence with $\mathbb{N}$, we can think of $\mathbb{R}$ as exhibiting a "higher level of infinity" than $\mathbb{N}$.[17] Are there sets that exhibit a higher level of infinity than $\mathbb{R}$?

[15] We think of (p,q) as determining the rational $\frac{p}{q}$. For now ignore the fact that a given rational can be determined by more than one pair (p,q).

[16] The sequence $(a_{n,n})_{n=1}^{\infty}$ is referred to as the "diagonal sequence".

[17] Our informal term, "higher level of infinity", is not standard.

To answer this we need the notion of "power set". If A is any set, the *power set* of A, $P(A)$, is the set of all subsets of A. Thus, if $A = \{1, 2\}$, we have $P(A) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.[18] It is clear for finite sets A that $P(A)$ will always contain more elements than A. Cantor showed that this remains true for infinite sets in the following sense. If A is an infinite set, $P(A)$ contains a subset that can be put into one-to-one correspondence with A, but $P(A)$ itself cannot be put into one-to-one correspondence with A. Thus, in our informal language, $P(A)$ exhibits a higher level of infinity than A. For example, $P(\mathbb{R})$ cannot be put into one-to-one correspondence with $\mathbb{R}$, and so exhibits a higher level of infinity than $\mathbb{R}$. Similarly, $P(P(\mathbb{R}))$ exhibits a higher level of infinity than $P(\mathbb{R})$, and so on. Thus, there is an infinite hierarchy of distinct levels of infinity! Infinite sets A and B "exhibit two different levels of infinity" when they have different cardinalities.

8 Proof of the chain rule

When we proved the Chain Rule in class, I called attention to a point at the end of the proof that I had swept under the rug. Here I repeat the proof we did in class, but now show the proper way to handle the point at the end I referred to. The notation is the same as in class, except here I write t_0 instead of a.

Proposition 8.1 (Chain Rule)**.** *(a) Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is differentiable at $(x(t_0), y(t_0))$, where $(x, y) : \mathbb{R} \to \mathbb{R}^2$ is differentiable at t_0. Then $z(t) := f(x(t), y(t))$ is differentiable at t_0 and $z'(t_0) = \nabla f(x(t_0), y(t_0)) \cdot (x'(t_0), y'(t_0))$.*

(b) The statement in part (a) extends in an obvious way to functions $f : \mathbb{R}^n \to \mathbb{R}$ differentiable at $x(t_0) \in \mathbb{R}^n$, where $x : \mathbb{R} \to \mathbb{R}^n$ is differentiable at t_0.[19] *Of course, it is not necessary here that the domain of f be all of $\mathbb{R}^n$, or that the domain of x be all of $\mathbb{R}$.*

Proof. Using the assumption that f is differentiable at $(x(t_0), y(t_0))$, we write[20]

$$\begin{aligned} &\frac{z(t_0+h) - z(t_0)}{h} = \frac{f(x(t_0+h), y(t_0+h)) - f(x(t_0), y(t_0))}{h} = \\ &f_x(x(t_0), y(t_0))\frac{\Delta x}{h} + f_y(x(t_0), y(t_0))\frac{\Delta y}{h} + \frac{r(x(t_0+h), y(t_0+h))}{h}. \end{aligned} \tag{8.1}$$

Observe that the sum of the first two terms in the last line of (8.1) approaches the desired limit as $h \to 0$.

To see that the final term of (8.1) approaches 0 as $h \to 0$ we write

$$\frac{r(x(t_0+h), y(t_0+h))}{h} = \frac{r(x(t_0+h), y(t_0+h))}{|(\Delta x, \Delta y)|} \cdot \frac{|(\Delta x, \Delta y)|}{h}. \tag{8.2}$$

But notice that the denominator of the first factor on the right may be zero for some $h \neq 0$. To fix this, observe that if $|(\Delta x, \Delta y)| = 0$, then

$$\begin{aligned} &r(x(t_0+h), y(t_0+h)) := \\ &\quad f(x(t_0+h), y(t_0+h)) - f(x(t_0), y(t_0)) - f_x(x(t_0), y(t_0))\Delta x - f_y(x(t_0), y(t_0))\Delta y = 0. \end{aligned} \tag{8.3}$$

[18]For any set A we regard the empty set $\emptyset$ and A itself as subsets of A.

[19]What is that extension?

[20]Here Δx, for example, is $x(t_0 + h) - x(t_0)$.

Thus, for $h \neq 0$ we have

$$\frac{r(x(t_0+h),y(t_0+h))}{h} = \begin{cases} \frac{r(x(t_0+h),y(t_0+h))}{|(\Delta x,\Delta y)|} \cdot \frac{|(\Delta x,\Delta y)|}{h}, & \text{if } |(\Delta x,\Delta y)| \neq 0 \\ 0, & \text{if } |(\Delta x,\Delta y)| = 0 \end{cases}, \tag{8.4}$$

and the right side is well-defined and approaches 0 as $h \to 0$.[21]

□

9 Remarks on "surfaces" and the implicit function theorem

First let's discuss what we mean by a "surface" in $\mathbb{R}^3$. Would we want to call the graph of our standard pathological example $f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{if } (x,y) \neq (0,0) \\ 0, & \text{if } (x,y) = 0 \end{cases}$ over an open set containing the origin a "surface"? We will use the following definition of "surface", which forces the answer to be *no.*

Definition 9.1 (Surface in $\mathbb{R}^3$)**.** *We call a set $S \subset \mathbb{R}^3$ a* surface *if S is* locally *the graph of a C^1 function $g : V \subset \mathbb{R}^2 \to \mathbb{R}$, where $V \subset \mathbb{R}^2$ is open. This means that for* any *$P \in S$, there is a ball $B(P,r) \subset \mathbb{R}^3$ such that $S \cap B(P,r)$ is the graph of a C^1 function of two variables.* [22]

For example, if $g_i : V_i \subset \mathbb{R}$ are C^1 functions defined on open sets $V_i \subset \mathbb{R}^2$ for $i = 1, 2$, the sets $S_1 = \{(x, y, g_1(x, y)) : (x, y) \in V_1\}$ and $S_2 = \{(g_2(y, z), y, z) : (y, z) \in V_2\}$ are both surfaces in $\mathbb{R}^3$. A sphere $\{x \in \mathbb{R}^3 : |x - a| = R\}$ is also a surface. Even though it is *not* the graph of any single C^1 function, it is locally the graph of a C^1 function. One can similarly define a "surface" in $\mathbb{R}^n$ for any $n \geq 2$.

Observe that a set S satisfying the condition in Definition 9.1 is "well-behaved" in the following sense. The surface can't have jumps because the function g is continuous; moreover, the surface is *very* well approximated near any one of its points by the tangent plane at that point because g is C^1.[23] These conditions imply that S resembles an object that we would naturally think of as a "surface".

In class on Thursday, Sept. 16 we stated the Implicit Function Theorem for a function $f : \mathcal{O} \subset \mathbb{R}^2 \to \mathbb{R}$. The statement for a function $f : \mathcal{O} \subset \mathbb{R}^3 \to \mathbb{R}$ is quite similar.

Theorem 9.2 (Implicit Function Theorem)**.** *(a) Let $\mathcal{O} \subset \mathbb{R}^3$ be an open set, and suppose $f \in C^1(\mathcal{O}, \mathbb{R})$. For some constant C define the level set $S = \{(x, y, z) \in \mathcal{O} : f(x, y, z) = C\}$. Suppose $(a, b, c) \in S$, and suppose $f_z(a, b, c) \neq 0$. Then there is an open set $U \subset \mathcal{O}$ containing (a, b, c), an open set $V \subset \mathbb{R}^2$ containing (a, b), and a C^1 function $g : V \to \mathbb{R}$ such that*

$$S \cap U = \{(x, y, g(x, y)) : (x, y) \in V\}. \tag{9.1}$$

(b) The obvious analogues of part (a) hold if we assume instead that $f_x(a, b, c) \neq 0$, or that $f_y(a, b, c) \neq 0$.[24]*. The theorem extends in the expected way to the case of C^1 functions $f : \mathcal{O} \subset \mathbb{R}^n \to \mathbb{R}$, $n \geq 2$, such that $f_{x_i}(a) \neq 0$ for some $a \in \mathcal{O}$ and some $i \in \{1, 2, \ldots, n\}$.*

[21] We use here the special property of r, the fact that $|(\Delta x, \Delta y)| \to 0$ as $h \to 0$ (differentiability implies continuity), and $\frac{|(\Delta x,\Delta y)|}{h} \to |(x'(t_0), y'(t_0))|$.

[22] Because of the C^1 condition, S is often referred to as a "smooth surface". On the other hand the surface of a cube and the surface of a coffee can are often described as "piecewise smooth" surfaces.

[23] Theorem 15.5 of the text, proved in class, shows that a function $g \in C^1(V, \mathbb{R})$ is differentiable at any point of V.

[24] What are those analogues?

The equality (9.1) states that near (a, b, c) the level set S is given by, or coincides with, the graph of the C^1 function $g : V \to \mathbb{R}$. Thus, by Definition 9.1 the level set $S \cap U$ is a *surface*, so it is natural to call it a *level surface* of f.

When (9.1) is satisfied it is a simple matter to find C^1 curves on S passing through $P = (a, b, g(a, b)) = (a, b, c)$. For example, let $r_1(t) = (a+t, b, g(a+t, b))$ and $r_2(t) = (a, b+t, g(a, b+t))$, where $t \in I$, some interval containing 0. These curves have tangent vectors at P given by $r_1'(0) = (1, 0, g_x(a, b))$ and $r_2'(0) = (0, 1, g_y(a, b))$. Since $r_1'(0) \times r_2'(0) = \langle -g_x(a, b), -g_y(a, b), 1\rangle$, these tangent vectors determine the plane through P with equation

$$z = g(a, b) + g_x(a, b)(x - a) + g_y(a, b)(y - b). \tag{9.2}$$

Since g is differentiable at (a, b), this is the tangent plane to the graph of g at P. In view of (9.1) we can equally well refer to this plane as the tangent plane to the level surface S at P. We showed in class that $\nabla f(P)$ is perpendicular to this plane.[25] Thus, the equation of this plane can also be written

$$\nabla f(a, b, c) \cdot \langle x - a, y - b, z - c\rangle = 0. \tag{9.3}$$

Summary: If $S \subset \mathbb{R}^3$ is a level set of a function $f \in C^1(\mathcal{O}, \mathbb{R})$, $P = (a, b, c) \in S$, and $\nabla f(P) \neq 0$, then S is a surface near P and its tangent plane at P is given by (9.3). Moreover, near P the level surface S can be written as the graph of a C^1 function of two variables. In the case where $f_z(P) \neq 0$, the tangent plane to S at P can also be expressed in the form (9.2). If $f_x(P) \neq 0$ or $f_y(P) \neq 0$, the same tangent plane can be expressed in a form similar to, but different from, (9.2) (how?).

10 The second derivative test

Here we give the proof of the second derivative test, fixing the problem that arose near the end of class on Sept. 23. The fix is mainly in steps 2 and 3 of the proof.

Proposition 10.1 (Second derivative test)**.** *Let $f \in C^2(\mathcal{O}, \mathbb{R})$, where $\mathcal{O} \subset \mathbb{R}^2$ is open. Let $(a, b) \in \mathcal{O}$ and suppose $\nabla f(a, b) = (0, 0)$. Define*

$$D(a, b) = \det \begin{pmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{yx}(a, b) & f_{yy}(a, b) \end{pmatrix}.$$

(a) If $D(a, b) < 0$, then (a, b) is a saddle point; (b) if $D(a, b) > 0$ and $f_{xx}(a, b) > 0$, then (a, b) is a local minimum; (c) if $D(a, b) > 0$ and $f_{xx}(a, b) < 0$, then (a, b) is a local maximum; (d) if $D(a, b) = 0$, then the test is inconclusive.

Proof. **1.** For $h = (h_1, h_2) \neq (0, 0)$ set $g(t) := f(a + th_1, b + th_2)$. Then, since $g'(0) = 0$ we have[26]

$$\begin{aligned} f(a + h_1, b + h_2) - f(a, b) = g(1) - g(0) = g'(0) + \frac{g''(c)}{2} = \frac{g''(c)}{2}, \text{ where} \\ g''(c) = (f_{xx}h_1^2 + 2f_{xy}h_1h_2 + f_{yy}h_2^2)|_{(a+ch_1, b+ch_2)} \text{ with } c \in (0, 1). \end{aligned} \tag{10.1}$$

[25]This follows from the fact that $\frac{d}{dt}|_{t=0} f(r_i(t)) = \nabla f(P) \cdot r_i'(0) = 0$ for $i = 1, 2$ (why?).

[26]Here we used the Calc II Taylor's formula to rewrite $g(1) - g(0)$ and the chain rule to compute $g'(0) = \nabla f(a, b) \cdot h = 0$ and $g''(t)$. Remember that g depends on h.

Now

$$g''(0) = (f_{xx}h_1^2 + 2f_{xy}h_1h_2 + f_{yy}h_2^2)|_{(a,b)} \text{ and thus}$$
$$f_{xx}(a,b)g''(0) = (f_{xx}h_1 + f_{xy}h_2)^2 + (f_{xx}f_{yy} - f_{xy}^2)h_2^2 := A_1 + A_2. \tag{10.2}$$

Observe that if $(h_1, h_2) \neq (0,0)$, $D(a,b) > 0$, and $f_{xx}(a,b) > 0$, then $A_1 + A_2$ is *strictly positive.* So $g''(0)$ is strictly positive. We claim that $g''(c)$ is therefore strictly positive if $(h_1, h_2) \neq (0,0)$ is small enough. This is *not* obvious, but we show this in the next step.

2. The quantity $g''(0)$ is a quadratic polynomial in (h_1, h_2) that we can write for $h \neq (0,0)$ as

$$g''(0) = \alpha_1 h_1^2 + \alpha_2 h_1 h_2 + \alpha_3 h_2^2 := P(0,h) = P(0, h/|h|)|h|^2. \tag{10.3}$$

Similarly,

$$g''(c) = \tilde{\alpha}_1 h_1^2 + \tilde{\alpha}_2 h_1 h_2 + \tilde{\alpha}_3 h_2^2 := P(c,h) = P(c, h/|h|)|h|^2. \tag{10.4}$$

Now for $|h| \neq 0$, the point $h/|h|$ lies in the *closed and bounded* set $S^1 = \{x \in \mathbb{R}^2 : |x| = 1\}$. Moreover, the map $h/|h| \to P(0, h/|h|)$ is continuous on S^1, so by the Extreme Value Theorem, this map attains an absolute minimum on S^1.[27] Since $g''(0)$ is strictly positive for $(h_1, h_2) \neq (0,0)$ by step 1, this minimum must be a positive number m. Thus, we have

$$g''(0) = P(0, h/|h|)|h|^2 \geq m|h|^2, \text{ where } m > 0 \text{ is independent of } h. \tag{10.5}$$

On the other hand by taking $|h|$ small enough, which makes the $\tilde{\alpha}_j$ as close as desired to the α_j by continuity of the second order partials of f at (a,b), we can arrange so that[28]

$$|[P(c, h/|h|) - P(0, h/|h|)]|\,|h|^2 < \frac{m}{2}|h|^2. \tag{10.6}$$

With (10.5) this implies[29]

$$g''(c) = P(c, h/|h|)|h|^2 \geq \frac{m}{2}|h|^2 \text{ for } h \neq (0,0) \text{ small enough.}$$

With (10.1) this shows that (a,b) is a local minimum and concludes the proof of the second derivative test in the case where $D(a,b) > 0$ and $f_{xx}(a,b) > 0$. The case where $D(a,b) > 0$ and $f_{xx}(a,b) < 0$ is proved in the same way.[30]

3. The argument when $D(a,b) < 0$ is slightly different. Now one can use (10.2) to find *particular* choices of $h \neq (0,0)$, say h_a and h_b, and positive constants m_a and m_b such that $P(0, h_a) = m_a$ and $P(0, h_b) = -m_b$.[31] So for $\delta > 0$ we have

$$P(0, \delta h_a) = m_a \delta^2 \text{ and } P(0, \delta h_b) = -m_b \delta^2. \tag{10.7}$$

[27] The Extreme Value Theorem says that a continuous real-valued function on a closed and bounded set attains an absolute max and an absolute min on that set.

[28] Whenever we say "for $|h|$ small enough", we mean for $|h| < \delta$ for some small enough $\delta > 0$.

[29] In fact, we can arrange (10.6) with any small $\epsilon > 0$ in place of $\frac{m}{2}$.

[30] The use of the Extreme Value Theorem in this step is a key element of the proof. Try to understand this step well.

[31] For example, if $f_{xx}(a,b) > 0$ we can make a choice of $h_a = (h_1, h_2)$ with $h_2 = 0$, $h_1 \neq 0$ in (10.2) to force $P(0, h_a) > 0$, and we can make a choice of $h_b = (h_1, h_2)$ so that $A_1 = 0$, $h_2 \neq 0$ to force $P(0, h_b) < 0$. When $f_{xx}(a,b) = 0$, use the first line of (10.2) instead of the second to choose m_a and m_b.

Let $m = \min(m_a, m_b)$. For δ small enough, similarly to (10.6) we have for $j = a, b$,

$$|P(c, \delta h_j) - P(0, \delta h_j)| < \frac{m}{2}\delta^2,$$

and with (10.7) that implies

$$P(c, \delta h_a) \geq \frac{m}{2}\delta^2 \text{ and } P(c, \delta h_b) \leq -\frac{m}{2}\delta^2 \text{ for } \delta > 0 \text{ small enough.}$$

So (a, b) is a saddle point.

4. To see that when $D(a, b) = 0$ the test is inconclusive, consider the functions $f_1(x, y) = x^4+y^4$, $f_2(x, y) = -(x^4+y^4)$, and $f_3(x, y) = x^4-y^4$, which have respectively a minimum, maximum, and saddle point at $(0, 0)$. In each case $D(0, 0) = 0$.

□

11 Lagrange multipliers: the case of one constraint

Here we recall the proof given in class of the following proposition.

Proposition 11.1. *Suppose $f, g \in C^1(\mathbb{R}^3, \mathbb{R})$, $P = (a, b, c) \in \mathbb{R}^3$, and $\nabla g(P) \neq 0$. Let $S = \{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = C\}$ for some $C > 0$. If $f|_S$ has a local max (or min) at P, then there exists a real number λ such that $\nabla f(P) = \lambda \nabla g(P)$.*

Proof. Without loss of generality we may suppose $g_z(P) \neq 0$. The Implicit Function Theorem implies that near P, the level set S is the graph of a function $h \in C^1(V, \mathbb{R})$, where $V \subset \mathbb{R}^2$ is open and $(a, b) \in V$. That is, near P the set S is given by $\{(x, y, h(x, y)) : (x, y) \in V\}$. Define $F(x, y) := f(x, y, h(x, y))$. If f has a local max at P, then F has a local max at (a, b), so

$$\langle 0, 0\rangle = \nabla F(a, b) = \langle f_x(P) + f_z(P)h_x(a, b), f_y(P) + f_z(P)h_y(a, b)\rangle.$$

We can regard each component of the vector on the right as a dot product, so this equation implies that $\nabla f(P)$ is orthogonal to both $\langle 1, 0, h_x(a, b)\rangle$ and $\langle 0, 1, h_y(a, b)\rangle$ (why?). We recognize the latter two vectors as ones that determine the tangent plane to S at P. Thus, $\nabla f(P)$ is normal to that plane. But recall that $\nabla g(P)$ is also normal to that plane (why?), so the vectors $\nabla f(P)$ and $\nabla g(P)$ must be parallel.

□

12 Boundary points of sets in $\mathbb{R}^n$

In this section we define "boundary points of sets in $\mathbb{R}^n$" and give another characterization of closed sets in $\mathbb{R}^n$. A clear understanding of boundaries is important in Riemann integration theory. We will use notation and terms introduced earlier in section 6.

Definition 12.1. *Let $S \subset \mathbb{R}^n$. A point $p \in \mathbb{R}^n$ is a* boundary point *of S if* every *open ball $B(p, r)$ contains at least one point in S and at least one point not in S. Let bS denote the* boundary *of S, that is, the set of all boundary points of S.*

Observe that every point of S is either an interior point (Definition 6.3) or a boundary point (why?).

For $p \in \mathbb{R}^3$ consider, for example, the closed and open balls $S_1 = \overline{B(p,r)}$ and $S_2 = B(p,r)$. In each case the set of all boundary points is, as one might expect, the sphere of radius r centered at p:

$$bS_1 = bS_2 = \{q \in \mathbb{R}^3 : |q-p| = r\}.$$

Thus, we see that, in general, a boundary point of a set S may or may not belong to S.

Let $S_3 = B(p,r) \cup \{q\} \subset \mathbb{R}^3$, where $|q-p| > r$. Then q is a boundary point of S_3 (check); however, q is not a boundary point of $S_2 = B(p,r)$. Let S_4 be the plane $\{(x,y,z) \in \mathbb{R}^3 : z = x+y\}$. Then $bS_4 = S_4$.

The next example arises in integration theory. Let

$$S_5 = \{(x,y) \in \mathbb{R}^2 : (x,y) \in [0,1] \times [0,1] \text{ and both } x, y \in \mathbb{Q}\},$$

the set of all points with rational coordinates in the closed rectangle $[0,1] \times [0,1]$. Then $bS_5 = [0,1] \times [0,1]$ (check). Thus, we have $S_5 \subset bS_5$. This shows that the boundary of a set $S \subset \mathbb{R}^2$ can strictly contain S *and* have positive area! Similar examples are easily given in $\mathbb{R}^n$ for any n.

The ambient space of S, that is, $\mathbb{R}^n$ in Definition 12.1, is important for determining bS, since it determines the open balls that are used in that definition. Let $S_6 \subset \mathbb{R}^1$ be the open interval $(0,1)$. Then $bS_6 = \{0,1\}$. Now let $S_7 = \{(x,0) \in \mathbb{R}^2 : x \in (0,1)\}$. Then $bS_7 = \{(x,0) \in \mathbb{R}^2 : x \in [0,1]\}$. These examples illustrate the need for a precise definition of boundary point; make sure you understand each of them (or ask!).

We have defined a set $S \subset \mathbb{R}^n$ to be closed if its complement $S^c := \mathbb{R}^n \setminus S$ is open. The next proposition characterizes closed sets in terms of boundary points.

Proposition 12.2. *Let $S \subset \mathbb{R}^n$. Then S is closed if and only if it contains all its boundary points, that is, if and only if $bS \subset S$.*

Proof. 1. ($\Rightarrow$) Suppose S is closed. Let p be a boundary point of S. We show $p \in S$. If $p \notin S$, then there exists $r > 0$ such that $B(p,r) \subset S^c$ (why?). Contradiction (why?).

2. ($\Leftarrow$). Suppose $bS \subset S$. We show S^c is open. Let $p \in S^c$. Then p is not a boundary point of S (why?). So there exists $r > 0$ such that $B(p,r) \subset S^c$ (why?). Thus, S^c is open. ☐

12.1 Related notions: isolated points and limit points

Definition 12.3. *Let $S \subset \mathbb{R}^n$. A point $p \in S$ is an* isolated *point of S if there exists a ball $B(p,r)$ such that $B(p,r) \cap S = \{p\}$.*

A point $p \in \mathbb{R}^n$ is a limit point *of S if* every *open ball $B(p,r)$ satisfies $B(p,r) \cap (S \setminus \{p\}) \neq \emptyset$, that is, every open ball centered at p contains a point of S different from p.*

Observe that if p is an isolated point of S, then $p \in bS$. A limit point of S need not belong to S. A limit point of S can be either an interior point or a boundary point of S. See if you can prove the following proposition. (Hint: Start by drawing pictures.)

Proposition 12.4. *A point $p \in \mathbb{R}^n$ is a limit point of S if and only if every open ball centered at p contains* infinitely many *points of S.*

Remark 12.5 (Limit points and limits)**.** *Let $D \subset \mathbb{R}^n$ and suppose $f : D \to \mathbb{R}$. If $a \in \mathbb{R}^n$ and $L \in \mathbb{R}$, we defined $\lim_{x\to a} f(x) = L$ to mean: given any $\epsilon > 0$ there exists a $\delta > 0$ such that if $0 < |x-a| < \delta$ and $x \in D$, then $|f(x) - L| < \epsilon$. Until now we have avoided applying this definition in cases where for some δ the set $\{x \in D : 0 < |x-a| < \delta\}$ is empty. In other words, we have applied this definition* only *in cases where a is a limit point of D.*[32] *Henceforth, we will explicitly require that a be a limit point of D in our definition of $\lim_{x\to a} f(x)$.*[33]

Definition 12.6 (Limit)**.** *Let $D \subset \mathbb{R}^n$ and suppose $f : D \to \mathbb{R}$. If $a \in \mathbb{R}^n$ is a limit point of D and $L \in \mathbb{R}$, we define $\lim_{x\to a} f(x) = L$ to mean: given any $\epsilon > 0$ there exists a $\delta > 0$ such that if $0 < |x-a| < \delta$ and $x \in D$, then $|f(x) - L| < \epsilon$.*

13 The Riemann integral in n dimensions.

In this section we summarize the definition of the Riemann integral given in class in a way that tries to make clear that there is really a *single* definition that works in all dimensions, that is, for integrals on domains in $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$,..., $\mathbb{R}^n$,...[34] It will be helpful to use the words "rectangle" and "volume" in the slightly generalized sense given by the next definition.

Definition 13.1. *(a) A* rectangle in $\mathbb{R}^1$ *is an interval of the form $[a, b]$ for some reals $a < b$. A* rectangle in $\mathbb{R}^2$ *is a set of the form $[a_1, b_1] \times [a_2, b_2]$ for some reals $a_i < b_i$ (that is, a rectangle in the classical sense). For any n a* rectangle in $\mathbb{R}^n$ *is a set of the form $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ for some reals $a_i < b_i$ (an $n-$dimensional "box"). Note that rectangles in $\mathbb{R}^n$ are bounded and always have their edges* parallel *to the standard coordinate axes.*

(b) If $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ is a rectangle in $\mathbb{R}^n$, we define the volume *of R by*

$$V(R) := (b_1 - a_1)(b_2 - a_2)\cdots(b_n - a_n). \tag{13.1}$$

Thus, the volume of a rectangle in $\mathbb{R}^1$ is its length, in $\mathbb{R}^2$ its classical area, in $\mathbb{R}^3$ its classical volume, and so on.

Next we give a definition of "partition" that works in all dimensions.

Definition 13.2. *(a) Fix a rectangle R in $\mathbb{R}^n$. A* partition *of R is a decomposition of R into a finite number of subrectangles R_j in $\mathbb{R}^n$ with disjoint interiors $\mathring{R}_j$. We will denote a partition of R by $\mathcal{P} = \{R_1, \ldots, R_N\} = \{R_j, j = 1, \ldots, N\}$, where N, the total number of subrectangles in the partition, can vary from partition to partition. The subrectangles R_j satisfy*

$$\cup_{j=1}^N R_j = R,\ \mathring{R}_j \cap \mathring{R}_k = \emptyset \text{ if } j \neq k. \tag{13.2}$$

[32]To see why, try to apply this definition when there exists a δ such that $\{x \in D : 0 < |x-a| < \delta\}$ is empty.

[33]On p. 933 our text tries to define $\lim_{x\to a} f(x)$ in the case where a is a boundary point of D. This works only if a is also a limit point of D. Boundary points of D are not always limit points of D (why?), and if a is not a limit point of D, the set $\{x \in D : 0 < |x-a| < \delta\}$ is empty for δ small enough.

[34]We often say that $\mathbb{R}^n$ "has dimension n." This is a way of saying that to specify the location of a point in $\mathbb{R}^n$ we need n scalar coordinates. For the same reason, we would say that the dimension of the spherical surface $S^2 := \{x \in \mathbb{R}^3 : |x| = 1\}$ is *two*. We need latitude and longitude to specify a point. Similarly, a line (or C^1 curve) in $\mathbb{R}^3$ has dimension one, and a plane (or C^1 surface) in $\mathbb{R}^3$ has dimension two. If $r : [a, b] \to \mathbb{R}^3$ defines a C^1 curve in $\mathbb{R}^3$, we can specify a point on the curve by specifying the single scalar $t \in [a, b]$. An open set in $\mathbb{R}^3$ has dimension three.

(b) If R is any rectangle in $\mathbb{R}^n$, the diameter *of R is defined by*

$$\text{diam } R := \max\{|x - y| : x, y \in R\},$$

that is, the length of the longest "diagonal".

(c) If $\mathcal{P} = \{R_j, j = 1, \ldots, N\}$ is a partition of a rectangle R in $\mathbb{R}^n$, we define the mesh of $\mathcal{P}$ *to be*

$$\|\mathcal{P}\| = \max\{\text{diam } R_j, j = 1, \ldots, N\},$$

the maximum of the diameters of the subrectangles in $\mathbb{P}$.

Thus, an efficient way to state that the diameter of every rectangle in a given partition $\mathcal{P}$ is $< \delta$ is to say $\|\mathcal{P}\| < \delta$. Below we will sometimes speak of "a sufficiently fine" partition" $\mathcal{P}$. This is equivalent to saying $\|\mathcal{P}\|$ is sufficiently small. We can now define Riemann sums in any dimension.

Definition 13.3 (Riemann sums)**.** *Let R be a rectangle in $\mathbb{R}^n$ and suppose $f : R \to \mathbb{R}$. Let $\mathcal{P} = \{R_1, \ldots, R_N\}$ be a partition of R. A* Riemann sum *for f associated to $\mathcal{P}$ is any sum of the form*

$$RS(f, \mathcal{P}) = \sum_{j=1}^{N} f(x_j)V(R_j), \text{ where } x_j \in R_j \text{ for each } j. \tag{13.3}$$

Observe that given a function $f : R \to \mathbb{R}$ and given a partition $\mathcal{P}$ of R, many different Riemann sums for f associated to $\mathcal{P}$ can be formed (why?).[35]

Roughly, one defines a function $f : R \to \mathbb{R}$ to be (Riemann) integrable if there exists a number L such that Riemann sums for f associated to sufficiently fine partitions of R are as close as desired to L. The next definition makes this precise.

Definition 13.4 (Integrability)**.** *Let R be a rectangle in $\mathbb{R}^n$ and suppose $f : R \to \mathbb{R}$. We say that f is* integrable on R *if*[36]

$$\lim_{\|\mathcal{P}\| \to 0} RS(f, \mathcal{P}) = L \text{ for some } L \in \mathbb{R}. \tag{13.4}$$

This means *the following: given $\epsilon > 0$ there exists a corresponding $\delta > 0$ such that for* any *partition $\mathcal{P}$ of R with $\|\mathcal{P}\| < \delta$ and* any *Riemann sum for f (13.3) associated to $\mathcal{P}$, we have $|RS(f, \mathcal{P}) - L| < \epsilon$. In this case we write*

$$\int_R f(x)dV = L. \tag{13.5}$$

When $n \geq 2$, we refer to $\int_R f(x)dV$ as a multiple integral*.*

The following proposition, stated without proof, gives us a large class of functions for which $\int_R f(x)dV$ exists.[37]

[35] A particular Riemann sum, $RS(f, \mathcal{P})$, depends of course on the choices $x_j \in R_j$, but we do not indicate that dependence in the notation.

[36] The limit in (13.4) is *not* a special case of the types of limits we defined earlier.

[37] We give a proof of Proposition 13.5 in section 19.

Proposition 13.5. *Let R be a rectangle in $\mathbb{R}^n$ and suppose $f : R \to \mathbb{R}^n$ is continuous. Then $\int_R f(x)dV$ exists.*

Remark 13.6 (Other notations)**.** *When $R \subset \mathbb{R}^1$, one usually writes $\int_R f(x)dx$ or $\int_a^b f(x)dx$ for the integral in* (13.5)*. When $R \subset \mathbb{R}^2$, many authors (such as the authors of our text) write $\int\int_R f(x,y)dA$ (a "double integral") for that integral. When $R \subset \mathbb{R}^3$ these same authors write $\int\int\int_R f(x,y,z)dV$ (a "triple integral").*

Next we define the integral of a function f on a set that is not necessarily a rectangle. It is easy to do this by making use of Definition 13.4.

Definition 13.7. *Let $D \subset \mathbb{R}^n$ be a* bounded *set, and suppose $f : D \to \mathbb{R}$. Choose a rectangle R in $\mathbb{R}^n$ such that $D \subset R$. Let*

$$\tilde{f}(x) = \begin{cases} f(x), & x \in D \\ 0, & x \in R \setminus D \end{cases} .$$

Now define $\int_D f(x)dV := \int_R \tilde{f}(x)dV$, when *the integral on the right exists.*

One easily checks that this definition is *independent* of the choice of rectangle R containing D; thus, this definition is *well-defined*. Observe that when $f \geq 0$ is continuous, $\int_D f(x)dV$ gives what we would normally call the "volume under the graph of f and over D".[38]

We will occasionally use the next definition.[39]

Definition 13.8 (Partition *on* a set $D \subset \mathbb{R}^n$)**.** *Suppose $D \subset \mathbb{R}^n$ is a bounded set. We construct a* partition on D *by first choosing a rectangle in R^n, R, such that $D \subset R$, partitioning R, and then retaining the rectangles which have nonempty intersection with D.*

Note that a rectangle R_j in a "partition on D" need not satisfy $R_j \subset D$.

Remark 13.9 (Well-defined definitions)**.** *The issue of* well-definedness *came up earlier when we defined the length of a curve given by $r \in C^1([a,b],\mathbb{R}^3)$ as $L = \int_a^b |r'(t)|dt$. This definition involves a* choice *of parametrization, just as Definition 13.7 involved a choice of rectangle R containing D. In a homework problem you verified that our definition of the length of a curve is independent of the choice of parametrizing function r. Thus, that definition is well-defined. Well-definedness is an issue in that case because we* want *the length of a curve to depend just on the set of points that constitute the curve, and not on the particular way we represent the curve. The issue of well-definedness arises whenever a definition involves a choice,* and *one wants that definition to be independent of the choice.*[40]

Can you think of other definitions we have given where there is an issue of well-definedness? I'd be interested to hear about any other examples you think of.

Next we give a precise, general definition of *volume* for bounded sets in $\mathbb{R}^n$.

[38] I added the phrase "what we would normally call", because we actually have *no* prior, *clearly defined* notion of volume to invoke here. Indeed, in this case ($f \geq 0$ and continuous) we can take $\int_D f(x)dV$ as our first, precise *definition* of the "volume under the graph of f and over D."

[39] The phrase "partition on D" is not standard, but the procedure described in Definition 13.8 is commonly used.

[40] I am not claiming this is the only way the issue of well-definedness can arise.

Definition 13.10 (Volume of a bounded set $S \subset \mathbb{R}^n$)**.** *Let $S \subset \mathbb{R}^n$ be a bounded set and choose a rectangle in R^n, R, such that $S \subset R$. Define $\chi_S : R \to \mathbb{R}$, the* characteristic function of S on R, *by*

$$\chi_S(x) = \begin{cases} 1, \ x \in S \\ 0, \ x \in R \setminus S \end{cases} .$$

We say that the set S has volume *if $\int_R \chi_S(x)dV$ exists. In that case we write $V(S) = \int_R \chi_S(x)dV$. If the integral does not exist, we say that* the volume of S is undefined. *This definition of the volume of S is often referred to as the* Jordan content *of S.*

Remark 13.11. *(a) Using Definition 13.7, we see that $\int_S 1dV = \int_R \chi_S(x)dV$ when this integral exists. This is because if $f : S \to \mathbb{R}$ is the constant function equal to one, then $\tilde{f} = \chi_S$. Thus, one can equally well take $\int_S 1dV$ as the definition of the volume of a bounded set $S \subset \mathbb{R}^n$.*

(b) If $S \subset \mathbb{R}^n$ is a bounded set, one can show that

$$S \text{ has volume if and only if } V(bS) = 0, \tag{13.6}$$

where bS is the boundary of S (Definition 12.1). In section 12 we gave the example

$$S_5 = \{(x, y) \in \mathbb{R}^2 : (x, y) \in [0, 1] \times [0, 1] \text{ and both } x, y \in \mathbb{Q}\},$$

and saw that $bS_5 = [0, 1] \times [0, 1]$. Thus, $V(bS_5) = 1$, so the volume of S_5 is undefined. In problem 5 of HW 8 you showed directly, that is, without using (13.6)*, that the volume of S_5 is undefined (how?).*

Similarly, the volume of the set

$$\tilde{S}_5 = \{(x, y) \in \mathbb{R}^2 : (x, y) \in [0, 1] \times [0, 1] \text{ and both } x, y \text{ are irrational}\}$$

is also undefined.

(c) It is unsatisfying that our definition of volume fails to assign a volume to some bounded subsets of $\mathbb{R}^n$. It is a rather subtle fact that there is no way *to give a definition of volume that is "realistic", and which assigns to* every *bounded subset of $\mathbb{R}^n$ a volume.*[41] *All attempts to produce such a definition of volume* must *lead to inconsistent theories. Some subsets of $\mathbb{R}^n$ are just too weird to have volume.*

(d) Some definitions of volume are "better" than others, though, in the sense that they realistically assign volumes to a larger *class of bounded subsets of $\mathbb{R}^n$; call this wider "applicability". For example, the theory of* Lebesgue measure *is better than Definition 13.9 in being more widely applicable. The Lebesgue measure of a set S is defined and equal to the Jordan content of S whenever the latter is defined; in particular, the Lebesgue measure of a rectangle in $\mathbb{R}^n$ is its usual volume. It turns out that the Lebesgue measure of S_5 in part (b) is 0, while that of $\tilde{S}_5$ is 1. A better theory of volume leads to a better theory of integration; for example, more functions are integrable.*[42] *It can*

[41] To be "realistic" a proposed definition of volume $\mathcal{V}(S)$ should: assign the usual volume to rectangles, have the property that $\mathcal{V}(\cup_{j=1}^{\infty} S_j) = \mathcal{V}(S_1) + \mathcal{V}(S_2) + \ldots$ whenever both sides are defined and the S_j are mutually disjoint, and assign equal volumes to sets that differ only by translation. We *could* consistently define the volume of every subset of $\mathbb{R}^n$ to be 0, but that would be not be "realistic"; it would also be useless.

[42] The characteristic function of S_5 is integrable on $[0, 1] \times [0, 1]$ in the Lebesgue theory, but not in the Riemann theory.

be proved that Lebesgue measure is the optimal *choice, that is, the most widely applicable possible choice, of a definition of volume for subsets of* $\mathbb{R}^n$*.*

The theory of Lebesgue measure and its associated theory of Lebesgue integration are studied in our real analysis course, Math 653, as well as in our measure theory course, Math 753. It helps to understand Riemann integration before studying the Lebesgue theory.

We conclude this section with a couple of extensions of Proposition 13.5 that provide us with a *much* larger class of integrable functions. The first is stated without proof.

Proposition 13.12. *Let* R *be a rectangle in* $\mathbb{R}^n$*. If the function* $f : R \to \mathbb{R}$ *is bounded and has discontinuities only on a set of volume zero, then* $\int_R f(x)dV$ *exists.*[43]

The idea behind the proof of Proposition 13.12 is similar to that behind the solution of problem 2 in HW 8, and is easy to describe. Because the set of discontinuities S has volume zero, one can enclose S in a union of rectangles whose total area is as small as desired. Then, because f is bounded, the total contribution of those rectangles to an associated Riemann sum can be made as small as desired. In the part of R away from S the function f is continuous, so an argument similar to the proof of Prop. 13.5 can be used to treat this part. I hope to give a proof of Prop. 13.5 later in the course if we can fit it in.

Proposition 13.13. *Let* $D \subset \mathbb{R}^n$ *be a closed and bounded set with a well-defined volume.*[44] *Suppose* $f : D \to \mathbb{R}$ *is continuous. Then* $\int_D f(x)dV$ *exists.*

Proof. Enclose D in a rectangle R and note that the function $\tilde{f} : R \to \mathbb{R}$ as in Definition 13.7 is discontinuous only at points of bD, a set with volume zero. The result now follows from Proposition 13.12. (How do we know $\tilde{f}$ is bounded on R?)

□

14 Set theory and the foundations of mathematics

In section 2 we defined a set as a "well-defined collection of distinct objects". We begin with a critique of this definition, and then discuss a crisis it led to in the early twentieth century and how that crisis was resolved.

14.1 Naive set theory, the crisis of Russell's paradox, and the development of axiomatic set theory.

The above definition, which emerged well before the end of the 19th century and is often referred to as the "naive" definition of set, is clearly rather vague. The definition has the embarrassing feature that it defines the word "set" in terms of the word "collection"; it also uses the ill-defined terms "well-defined" and "objects". Here is another try at a definition, possibly even more vague, by Felix Hausdorff (1868-1942) in the early 1900s: "A set is formed by the grouping together of single objects into a whole. A set is a plurality thought of as a unit." Even as late as 1915, Georg Cantor (1845-1918) gave a similar, equally unclear definition in a published paper.

[43] To say f is bounded means: there exists M such that $|f(x)| \le M$ for all $x \in R$.

[44] Recall that this is equivalent to saying $V(bD) = 0$; see Remark 13.11(b).

In 1901 Bertrand Russell shocked the mathematical world by finding a paradox, or logical inconsistency, associated with the naive definition of set. **Russell's Paradox** is as follows: Consider the set S of all sets which do not contain themselves as elements.[45] This *seems* to be "a well-defined collection of distinct objects". Does S contain itself as an element? If yes, then S does not contain itself as an element. Contradiction. If no, then S does contain itself as an element. Again, a contradiction. Set theory had been proposed as a foundation for all of mathematics (see section 14.2). Who would build mathematics on a foundation that gives rise to such paradoxes? This crisis made it clear that set theory had to be clarified and made rigorous.

It is a subtle problem in the foundations of mathematics to characterize sets precisely, and we shall not attempt to do so. In rigorous set theory (also known as "axiomatic set theory"), which developed in the wake of Russell's paradox, the concept of "set" is taken as a "primitive", or undefined, notion whose properties are restricted by a list of formally stated axioms. One such set of axioms, due to Zermelo and Fraenkel, was mostly in place by about 1922 and continues to serve as probably the most widely used axiomatization of set theory. There are about ten ZF axioms, the number depending on how they are formulated. One of the axioms, *informally* stated, is: "Two sets are equal if they have the same elements". Another says that the union of two sets is a set. Other axioms guarantee the existence of the empty set, power sets, and of sets that are infinite. One of the ZF axioms implies that the "set" S in the statement of Russell's paradox is *not* a set in ZF set theory, and thereby eliminates the paradox from that theory. That same axiom implies that the "set of all sets" is not a set in the ZF theory.[46]

An aside on Bertrand Russell (1872-1970). Bertrand Russell worked on logic and the foundations of mathematics in his early years. This work culminated in Principia Mathematica (1913), a three-volume work written with his teacher Alfred N. Whitehead, which was one of the earliest attempts to provide a foundation for mathematics, or at least arithmetic, in terms of set theory and logic. Russell was one of the founders of analytic philosophy, and wrote many works for general readers on logic, mathematics, science, and philosophy. His book "A History of Western Philosophy" was a best-seller. He was a pacifist who did jail time more than once for expressing his beliefs; six months in 1918 for publicly lecturing against the entrance of the US into WWI on the side of the UK, 7 days in 1961 at the age of 89 for participating in an anti-nuclear demonstration. He wrote more than 60 books and over 2000 articles on many different subjects, and won the Nobel Prize for Literature in 1950. [47] On youtube you can see him being interviewed.)

What is a formal axiomatic system? Zermelo-Fraenkel set theory is an example of a formal axiomatic system. The axiom that eliminates Russell's paradox from ZF set theory is called the Axiom of Regularity. Here is its formal statement, just to give you an example of what a formal axiom looks like:

$$A \neq \emptyset \Rightarrow (\exists x) : [x \in A \ \& \ (\forall y)(y \in x \Rightarrow y \notin A)].$$

This axiom implies that *no* set can contain itself as an element. So if some seemingly "well-defined collection of distinct objects" does contain itself as an element, then that collection is not a *set* in

[45]For example, the set $\{1, 2, 3\}$ does not contain itself as an element, but the set of sets that were described by me today does contain itself as an element.

[46]Russell himself, in work done partly with Alfred N. Whitehead, proposed a systematic way to avoid Russell's paradox, the "theory of types", not long after he found the paradox.

[47]Russell's paternal grandfather, John Russell, was twice prime minister of the United Kingdom in the mid-1800s.

the ZF theory. As you can see, a formally stated axiom involves no English words. The symbols that appear in a formal axiom are drawn from a specific list of allowed symbols and are primitive (or undefined). There are precise formal rules for combining the symbols into well-formed "formulas". There are also formally stated rules of logic for "deducing" new formulas from old ones.[48] The procedures of formula formation and "deduction" of formulas can be programmed or mechanized. No thinking, with its attendant human errors, needs to be involved in the construction of "proofs".

When mathematicians construct such a formal system, they generally have a specific purpose in mind, for example, to make set theory rigorous. To make such use of a formal system, they must attach *interpretations* or *meanings* to the formal symbols. For example, in the Axiom of Regularity we interpret A, x, and y as sets, $\exists$ as "there exists", $\Rightarrow$ as "implies", $\in$ as "is an element of", and so on.[49] But a computer (or a human who knows all the formal rules) does not need such interpretations to construct "proofs" of well-formed formulas. A "proof" in the formal system consists of a list of well-formed formulas where each succeeding formula is derived from the axioms and preceding formulas by the allowed mechanical rules for manipulating formulas. The formula being proved appears at the end of the list. One can imagine the same formal proof being interpreted in two completely different ways, depending on the interpretations attached to the individual symbols.

It is perhaps not surprising that "set" has to be taken as a primitive notion. If all of mathematics is to be built on the foundation of set theory, then there is no mathematical concept more fundamental than set in terms of which we can define "set". One has to start somewhere to avoid the confusion of an infinite regress.[50]

14.2 Axiomatic set theory as a foundation for mathematics

Preamble. Have you ever seen a careful definition of the number "1"? or of the word "number"? I would be surprised if the answer is yes. What would you say if you were asked what "1" actually IS? Most students would probably say something like: "1" is the first natural number, or: "1" is the number that tells us how many objects are in a set with a single element". The first try clearly fails, because if "1" has not been defined, neither have the natural numbers. The second try is much better (it invokes sets!), but it is still inadequate because it uses the word "single" which is another way of saying one, it uses the word "number" which itself needs defining, and it uses the concept "how many" which needs defining. It seems about time to give a much more clear and careful definition of 1. One way to define "1" is this: it is the *set* $\{\emptyset\}$; that is, $1 := \{\emptyset\}$. I explain further below.

If we take advantage of the (apparently) solid foundation provided by the ZF axioms, it is not hard to define real numbers and the operations of arithmetic completely in terms of set theory.[51] The starting point is to define zero and the natural numbers using sets by taking $0 := \emptyset$, $1 := \{\emptyset\}$,

[48]The formula $\exists \& \Rightarrow \in$, for example, is *not* a well-formed (or permissible) formula in the ZF theory. It violates the rules for formula formation.

[49]In axiomatic set theory the elements of sets are usually restricted to be sets as well. We have already seen examples of sets whose elements are sets, for example, the power set of a set.

[50]You might be reassured (or not!) to know that many mathematicians never experience the need in their own work for a definition of set that is more precise than the above naive one. A good place to read more about set theory is the article "Set Theory" in the Stanford Encyclopedia of Philosophy, easily found online.

[51]So far, no one has been able to prove the consistency of the ZF axioms, but it is universally believed that they are consistent.

$2 = \{\emptyset, \{\emptyset\}\}$, $3 := \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, etc.. Here, for example, $\{\emptyset\}$ is the *nonempty* set whose single element is the empty set $\emptyset$. Let $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$. Notice that we can rewrite our definition of $0, 1, 2, 3, \dots$ as follows: $0 := \emptyset$, $1 := \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, etc. If $m, n \in \mathbb{N}_0$ we can define $m \leq n$ to mean $m \subset n$.

Addition of elements of $\mathbb{N}_0$ can then be defined set-theoretically by taking unions in the right way. The sum $1 + 2$ is *not*, as one might hope or guess, just the union of the sets defining 1 and 2; that union would be

$$\{\emptyset\} \cup \{\emptyset, \{\emptyset\}\} = \{\emptyset, \{\emptyset\}\} = 2,$$

which is the *wrong* answer! Rather, to define addition we will use what is called the "disjoint union" of the sets 1 and 2 as part of the following simple procedure. Write $1 = \{\emptyset_a\}$ and $2 = \{\emptyset_b, \{\emptyset\}_b\}$, where the subscripts a, b allow us to *distinguish* $\emptyset$, for example, considered as an element of the set 1, from $\emptyset$ considered as an element of 2. Then to define $1 + 2$ we first take the union[52]

$$\{\emptyset_a\} \cup \{\emptyset_b, \{\emptyset\}_b\} = \{\emptyset_a, \emptyset_b, \{\emptyset\}_b\} := T,$$

and then search the above list of sets $0, 1, 2, 3, \dots$ for the unique set that can be put into one-to-one correspondence with T.[53] That set is 3. This second step, by the way, gives a set-theoretic *definition* of the familiar operation of *counting*. After forming T, we reach the conclusion $1 + 2 = 3$ by "counting" the elements in T. The sum of any two elements of $\mathbb{N}_0$ is defined by this set-theoretic procedure.. Having defined addition of natural numbers, we define multiplication by repeated addition: for example, $3 \times 2 := (2 + 2) + 2$.

Remark 14.1. *One can also check that the operations of addition and multiplication we have just defined on* $\mathbb{N}_0$ *satisfy the usual commutative, associative, and distributive laws. Subtraction in* $\mathbb{N}_0$ *presents a problem. For example, we expect* $1 - 2$ *to lie outside* $\mathbb{N}_0$. *But differences like* $2 - 1$, *or more generally* $n - m$ *when* $m \subset n$, *can be defined within* $\mathbb{N}_0$. *For example, we can* define $n - m$ *to be* $r \in \mathbb{N}_0$ *precisely when* $m + r = n$. *Moreover, if* $p \in \mathbb{N}_0$ *and* $m \subset n$, *we have* $p(n - m) = pn - pm \in \mathbb{N}_0$.

Now that we have defined $\mathbb{N}_0$ set-theoretically, we can define the integers $\mathbb{Z}$ set-theoretically using ordered pairs of elements of $\mathbb{N}_0$. First, we say that two ordered pairs of elements of $\mathbb{N}_0$, (m, n) and (m', n'), are *equivalent* if [54]

$$m + n' = n + m'. \tag{14.1}$$

In that case we write $(m, n) \sim (m', n')$. Check for example that $(0, 7) \sim (3, 10) \sim (5, 12)$. The *infinite* set consisting of *all* ordered pairs equivalent to $(0, 7)$ is called the *equivalence class* of $(0, 7)$ with respect to $\sim$ and is usually denoted $[(0, 7)]$. We can now *define* the integer "-7" as $[(0, 7)] = [(25, 32)]$.[55] Thus, we have

$$-7 := [(0, 7)] := \{(m, n) \in \mathbb{N}_0 \times \mathbb{N}_0 : (m, n) \sim (0, 7)\}. \tag{14.2}$$

Similarly, the integer $+7$ is the equivalence class $[(7, 0)]$ (which equals $[(12, 5)]$ or $[(9, 2)]$). In this way every integer is defined as an equivalence class consisting of ordered pairs of elements of $\mathbb{N}_0$. The relation $\sim$ defined by (14.1) is an example of an *equivalence relation*.

[52]The set T is called the *disjoint union* of the sets 1 and 2.

[53]"One-to-one correspondences" were defined in section 7. They are a special kind of function, and functions can be defined set-theoretically as we'll see below.

[54]Informally, think $m - n = m' - n'$, which would not be acceptable as a definition of equivalent pairs at this point (why?)

[55]Informally, think $-7 = [(0, 7)]$ " = " $0 - 7 = 25 - 32$ " = " $[(25, 32)]$.

Remark 14.2 (Equivalence relations and equivalence classes)**.** *In general, an "equivalence relation" $\sim$ on a set X is a relation between elements of X that satisfies for any a, b, c in X the following properties: (i) $a \sim a$ (reflexivity), (ii) $a \sim b$ if and only if $b \sim a$ (symmetry), and (iii) $a \sim b$ and $b \sim c$ implies $a \sim c$ (transitivity). The* equivalence class *of $a \in X$ is denoted $[a]$ and is defined as*

$$[a] = \{x \in X : x \sim a\}.$$

The particular relation $\sim$ defined by (14.1) *is an equivalence relation on the set $X = \mathbb{N}_0 \times \mathbb{N}_0$ (check). You can easily check that "has the same birthday as" is an equivalence relation on the set of all people. The relation "$\geq$" on $\mathbb{N}_0$ is not an equivalence relation.*[56]

Now you might object that I have defined the integers using the notion of an "ordered pair"; and you might wonder if *that* can be defined set-theoretically. Indeed, we can define the ordered pair $(a, b) := \{a, \{a, b\}\} = \{\{a, b\}, a\}$. The right side makes clear both that there are two components and which component "comes first". Note that $\{a, b\}$ *fails* to work as a definition of (a, b)(why?). In HW 10 you will explore how to define addition and multiplication on $\mathbb{Z}$ when its elements are defined using equivalence classes as in (14.2).

Using our set-theoretic definition of $\mathbb{Z}$, we can give an analogous set-theoretic definition of the rational numbers $\mathbb{Q}$. Each rational can be defined as an equivalence class of ordered pairs (p, q) of integers with respect to an equivalence relation that is an appropriate analogue of the equivalence relation $\sim$ defined by (14.1). Can you see how this new equivalence relation should be defined?

The next step is to define $\mathbb{R}$, and in particular the set of irrational numbers, *in terms of* rational numbers. We won't discuss how to do this in detail here, but instead we give an example of how this can be done. One can define the irrational number $\sqrt{2}$ as

$$\sqrt{2} := \{q \in \mathbb{Q} : q \leq 0\} \cup \{q \in \mathbb{Q} : q^2 < 2\}. \tag{14.3}$$

The right side of (14.3) is an example of a "Dedekind cut", named after Richard Dedekind (1831-1916). If we "cut" the real line at $\sqrt{2}$, the set defined in (14.3) is the set of all rationals to the left of the cut. [57] Similarly, *every* real number can be defined as a Dedekind cut; for example, the cut corresponding to $\frac{1}{2} \in \mathbb{Q}$ is $\{q \in \mathbb{Q} : q < \frac{1}{2}\} \in \mathbb{R}$.

Remark 14.3. *You may have noticed by now that we have given a number like* "8" four *apparently different definitions; after all, "8" is supposed to be an element of $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$, and the individual elements of these sets are defined in four* different *ways. We regard the four different ways of defining "8" as essentially different "names" for the same object. To make this more concrete: if you do your taxes with rational numbers thought of as elements of $\mathbb{Q}$, and then do them again with rational numbers thought of as elements of $\mathbb{R}$, you will end up paying (or receiving back) the same amount of money in the end. More precisely, there are obvious injective maps i_j, $j = 1, 2, 3$ such that*

$$\mathbb{N} \xrightarrow{i_1} \mathbb{Z} \xrightarrow{i_2} \mathbb{Q} \xrightarrow{i_3} \mathbb{R},$$

[56] More generally, a "relation" R on a set X is given by any subset $R \subset X \times X$. If $(a, b) \in R$ we write aRb. Thus, an equivalence relation $\sim$ on X is a special type of relation on X. When discussing equivalence relations, we can usually avoid explicit reference to this set-theoretic definition of the word "relation".

[57] You can read more about how Dedekind cuts are used to define real numbers in the classic text *Principles of Mathematical Analysis* by Walter Rudin, easily found online. It is shown there that the order symbols $\leq$ and $<$ used in (14.3) can be defined set-theoretically. Rudin also shows how to add and multiply cuts. You could do your taxes with them.

and these maps "preserve" order and arithmetic operations.[58] *Thus, for example, if* $a, b \in \mathbb{Z}$ *with* $a < b$, *then* $i_2(a) < i_2(b)$, $i_2(a+b) = i_2(a) + i_2(b)$, *and* $i_2(a \times b) = i_2(a) \times i_2(b)$. *Note that the definition of each of the symbols "* $<, +, \times$*" changes from one side of each of these three assertions to the other.*[59]

The concept of *function* is also easily definable in terms of set theory. A function $f : A \to B$ is a subset $f \subset A \times B$ with the properties:

(a) for all $a \in A$ there exists some $b \in B$ such that $(a, b) \in f$;

(b) if (a, b) and (a, c) both belong to f, then $b = c$.

If $(a, b) \in f$, we write $f(a) = b$.[60]

One can continue in this way to give set-theoretic definitions of more and more complicated mathematical objects in terms of simpler objects that have already been defined set-theoretically.

15 Mathematical induction

Mathematical induction is the standard way to show that a statement depending on $n \in \mathbb{N}$ is true for all n. Proofs by mathematical induction work as follows. Consider a proposition $P(n)$ whose statement depends on $n \in \mathbb{N}$. To prove $P(n)$ for all n, first check that $P(1)$ is true. Then *assume* the truth of $P(n)$ and *use that assumption* to prove the truth of $P(n+1)$. Those two steps establish the truth of $P(n)$ for all n (why?).[61] If $P(n)$ fails to be true for some n, then one of the two steps will fail. This kind of argument is an example of what is sometimes called a "bootstrapping argument".

To see (or recall) how this works, let us use mathematical induction to prove that the statement $P(n)$ given by

$$1 + 2 + 3 + \cdots + n = n(n+1)/2$$

holds for all n. First check $P(1)$ is true. This is the statement $1 = 1 \cdot 2/2$, which is true. Now assume $P(n)$ is true and try to show that $P(n+1)$ is true. If we add $n+1$ to both sides of the equation that is $P(n)$ we get

$$1 + \cdots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)[(n+1)+1]}{2}.$$

This shows $P(n+1)$ is true, so we are done.

In the solution set to HW 9 I indicate how to use math induction to prove Fubini's theorem for multiple integrals in dimension n for all n.

16 Green's theorem, circulation, and flux

Green's theorem is one of three "higher dimensional versions" of the fundamental theorem of Calculus that we study in Chapter 17; moreover, the FTC is used in its proof. The others are Stokes's theorem and the Divergence theorem.

[58] For example, i_3 maps "1/2" considered as an element of $\mathbb{Q}$ to "1/2" considered as an element of $\mathbb{R}$.

[59] Alternatively, one *could* declare by fiat that the definition of 8 as a Dedekind cut is the single, official definition of 8, and ignore, discard, or demote the earlier definitions; this has obvious drawbacks.

[60] Note that a function $f : X \to X$ is a special kind of relation on X. A "relation on X" was defined in an earlier footnote.

[61] Answer: The first step gives $P(1)$. Using the second step, we deduce $P(2)$. Apply the second step again to get $P(3)$, etc.

Theorem 16.1 (Green)**.** *Let $\overrightarrow{C}$ be a piecewise C^1, simple closed curve in $\mathbb{R}^2$ which is oriented counterclockwise.*[62] *Let D be the closed set whose boundary is the curve.*[63] *Suppose $D \subset \mathcal{O} \subset \mathbb{R}^2$, where $\mathcal{O}$ is open, and assume $F = \langle P, Q\rangle \in C^1(\mathcal{O}, \mathbb{R}^2)$. Then $\int_{\overrightarrow{C}} F \cdot dr = \int\int_D (Q_x - P_y) dA$.*

In class we proved this theorem for a special class of sets D. We then applied the theorem to prove the direction $\Leftarrow$ in the third equivalence in the statement of the next proposition. Part (b) of the proposition had been proved earlier.

Proposition 16.2. *(a) Let $\mathcal{O} \subset \mathbb{R}^2$ be an open, simply connected set, and suppose $F = \langle P, Q\rangle \in C^1(\mathcal{O}, \mathbb{R}^2)$. Then F is conservative in $\mathcal{O}$ $\Leftrightarrow$ line integrals of F are independent of path in $\mathcal{O}$ $\Leftrightarrow$ the line integral of F around any piecewise C^1, simple closed curve in $\mathcal{O}$ is zero $\Leftrightarrow$ $Q_x - P_y = 0$ in $\mathcal{O}$.*

(b) The first two equivalences and the direction $\Rightarrow$ in the third equivalence hold if the open set $\mathcal{O}$ is merely connected.

16.1 Circulation integrals in $\mathbb{R}^2$.

Let $\overrightarrow{C}$ be a C^1, simple closed curve in $\mathbb{R}^2$ which is oriented, say, counterclockwise. Suppose $F = \langle P, Q\rangle$ is a continuous vector field defined in some open set containing $\overrightarrow{C}$. Let $r(t) = \langle x(t), y(t)\rangle$, $t \in [a, b]$ parametrize $\overrightarrow{C}$. Then we have[64]

$$\begin{aligned} \int_{\overrightarrow{C}} P dx + Q dy = \int_{\overrightarrow{C}} F \cdot dr = \int_a^b F(r(t)) \cdot r'(t) dt = \int_a^b F(r(t)) \cdot \frac{r'(t)}{|r'(t)|}\, |r'(t)| dt = \\ \int_a^b F(r(t)) \cdot T(t)\, ds = \int_{\overrightarrow{C}} F \cdot T\, ds, \end{aligned} \tag{16.1}$$

where $T(t)$ is the unit tangent vector to $\overrightarrow{C}$ at $r(t)$ in the direction of the orientation of $\overrightarrow{C}$. Since $F(r(t)) \cdot T(t)$ is the component of F at $r(t)$ in the direction of $T(t)$, we say that the integral $\int_{\overrightarrow{C}} F \cdot T\, ds$ gives the net "circulation" of F along $\overrightarrow{C}$. Note that if we restrict F to satisfy the condition $|F| = 1$ along $\overrightarrow{C}$, then the *maximum* possible value of the circulation integral is obtained when $F = T$ along $\overrightarrow{C}$, and that maximum value is the length of $\overrightarrow{C}$. If F is everywhere normal to $\overrightarrow{C}$, then the circulation integral is 0.

Suppose now that $(Q_x - P_y)(P) > 0$ at some point P in an open set where F is C^1. Choose $\overrightarrow{C}$ to be a *small* circle centered at P and oriented counterclockwise. Then by Green's theorem

$$\int_{\overrightarrow{C}} F \cdot T\, ds = \int\int_D (Q_x - P_y) dA > 0, \tag{16.2}$$

so F has positive net circulation around $\overrightarrow{C}$. *Thus, the positivity of $Q_x - P_y$ at a point P allows us to conclude that F tends to circulate in the counterclockwise direction near P.*

[62] Here "simple" means non-self-intersecting and "closed" means $r(a) = r(b)$, where $r : [a, b] \to \mathbb{R}^2$ gives $\overrightarrow{C}$. A circle is a simple closed curve, a figure eight is not.

[63] In this sentence "closed" means the complement of D is open.

[64] Suppose $r'(t)$ is never zero.

16.2 Flux integrals in $\mathbb{R}^2$.

Let $\overrightarrow{C}$ and $F = \langle P, Q\rangle$ be as in (16.1), and let $n(t)$ be the unit *outward* normal vector to $\overrightarrow{C}$ at $r(t)$. Then since $T(t) = \langle x'(t), y'(t)\rangle/|r'(t)|$, we have[65]

$$n(t) = \langle y'(t), -x'(t)\rangle/|r'(t)|. \tag{16.3}$$

Since $F(r(t)) \cdot n(t)$ is the component of F at $r(t)$ in the direction of the outward normal $n(t)$, we say that the integral

$$\int_{\overrightarrow{C}} F \cdot n \, ds := \int_a^b F(r(t)) \cdot n(t)\, |r'(t)| dt \tag{16.4}$$

gives the net "outward flux" of F along $\overrightarrow{C}$. Using (16.3) we see that

$$\int_{\overrightarrow{C}} F \cdot n \, ds = \int_a^b F(r(t)) \cdot \langle y'(t), -x'(t)\rangle dt = \int_{\overrightarrow{C}} P dy - Q dx. \tag{16.5}$$

If we define $\tilde{F} := \langle -Q, P\rangle$ we see from (16.5) that

$$\int_{\overrightarrow{C}} F \cdot n \, ds = \int_{\overrightarrow{C}} \tilde{F} \cdot dr. \tag{16.6}$$

Suppose now that $(P_x + Q_y)(P) > 0$ at some point P in an open set where F is C^1. Choose $\overrightarrow{C}$ to be a *small* circle centered at P and oriented counterclockwise. Then by Green's theorem applied to $\tilde{F}$ we have

$$\int_{\overrightarrow{C}} \tilde{F} \cdot dr = \int_D (P_x + Q_y) dA > 0, \tag{16.7}$$

so F has positive net outward flux along C. *Thus, the positivity of $P_x + Q_y$ at a point P allows us to conclude that F tends to point outward from P near P.*

Definition 16.3. *If $F = \langle P, Q\rangle$, we call $Q_x - P_y$ the* curl *of F and we call $P_x + Q_y$ the* divergence *of F.*

17 Surface elements without and with borders

The purpose of this section is to clarify a few of the definitions that we use frequently in connection with surface integrals and Stokes's theorem.

Definition 17.1 (Surface element vs surface element with border)**.** *(a) Let V be an open, simply connected subset of $\mathbb{R}^2$ and suppose the map $r \in C^1(V, \mathbb{R}^3)$ satisfies:*[66]

(i) $r_u \times r_v \neq 0$ for all $(u, v) \in V$,

(ii) $r : V \to r(V)$ is bijective.

Then we call the pair (r, V) a surface element. *Informally, we often just say "$r(V) := S$ is a surface element".*

[65] Draw a picture to see this is correct.

[66] We write $r(u, v) = \langle x(u, v), y(u, v), z(u, v)\rangle$.

(b) Suppose (r, V) is a surface element, and let $C \subset V$ be a piecewise C^1 simple closed curve enclosing an open set $U \subset V$. Set $D = U \cup C$ and note that $bD = C$.[67] *We call the pair (r, D) a* surface element with border, *and we refer to the curve $r(C)$ as the* border *of $r(D)$.*[68] *Moreover, if $s : [a, b] \to \mathbb{R}^2$ gives the oriented curve $\overrightarrow{C}$, then $r \circ s : [a, b] \to \mathbb{R}^3$ gives the oriented curve $\overrightarrow{r(C)}$.*

Remarks (1) Using a more general version of the implicit function theorem than the one we stated earlier, one can show that a surface element (r, V) is a *surface* in the sense of Definition 8.1. This is not obvious, but we shall not prove it here. We need an assumption like (ii) in Definition 17.1 to avoid situations where r maps to V to more than one copy of $r(V)$. Consider, for example, the parametrizing map for the sphere, $r(\phi, \theta) = \langle a \sin\phi \cos\theta, a \sin\phi \sin\theta, a \cos\phi\rangle$, on the domain $(0, \pi) \times (0, 8\pi)$.

(2). The border of $r(D)$, namely $r(C)$, is *not equal* to the boundary of $r(D)$ considered as a subset of $\mathbb{R}^3$! [69] Indeed we have $br(D) = r(D)$, where $br(D)$ denotes the boundary of $r(D)$. [70] Draw a picture to see this. Sometimes we will write $\partial r(D)$ or ∂S for the border of $r(D) := S$.

(3) On the other hand notice that because D is a subset of $\mathbb{R}^2$, we do have $bD = C$. (Draw)

(4) If (r, V) is a surface element, the unit normal vector $n(u, v) := \frac{r_u(u,v) \times r_v(u,v)}{|r_u(u,v) \times r_v(u,v)|}$ defines one of the two possible orientations of $r(V)$. We'll refer to $n(u, v)$ as the *orientation of $r(D)$ determined by r*. If (r, D) is a surface element with border as in Definition 17.1(b), and if $\overrightarrow{C} = \overrightarrow{bD}$ is oriented counterclockwise, then the orientation of $r(D)$ given by n is consistent with the orientation of $r(C)$ given by $\overrightarrow{r(C)}$. Draw a picture to see this.[71]

(5) We call (r, V) as in Definition 17.1 a surface *element*, because more general surfaces can be constructed by piecing together surface elements in the correct way. For example, a sphere in $\mathbb{R}^3$ can be expressed as the union of a few surface elements $r(V)$. Strictly speaking, the usual parametrizing map $r(\phi, \theta) = \langle a \sin\phi \cos\theta, a \sin\phi \sin\theta, a \cos\phi\rangle$, with $(\phi, \theta) \in [0, \pi] \times [0, 2\pi]$ is *not* a surface element (why?). On the other hand if we restrict the domain to be $(0, \pi) \times (0, 2\pi)$ we do get a surface element.

The next definition will be useful in the following section.

Definition 17.2. *Let $\mathcal{C}$ be a given piecewise C^1, simple closed curve in $\mathbb{R}^3$. Suppose there exists a surface element with border (r, D) as in Definition 17.1 such that $r(bD) = \mathcal{C}$. In that case the curve $\mathcal{C}$ is the border of $r(D)$. In this situation we will refer to $r(D)$ as a* surface element with border given by the curve *$\mathcal{C}$*.[72]

18 Simply connected open sets in $\mathbb{R}^n$, $n \geq 2$.

We would like to use Stokes's theorem to prove an analogue of Proposition 16.2 for vector fields $F = \langle P, Q, R\rangle \in C^1(\mathcal{O}, \mathbb{R}^3)$, where $\mathcal{O}$ is open in $\mathbb{R}^3$. For this we need to define simple connectedness for open sets in $\mathbb{R}^3$.

---

[67] When we write $C \subset V$ or $D = U \cup C$, we are of course thinking of C as being the *set of points* which is the image of some piecewise C^1 map $s : [a, b] \to \mathbb{R}^2$, where s defines the oriented curve $\overrightarrow{C}$.

[68] Informally, we often just say "$r(D)$" is a surface element with border."

[69] Recall Definition 12.1.

[70] The authors of our text never mention this point. They refer to the border of $r(D)$ simply as "the boundary of $r(D)$", but the only careful definition of "boundary" that is ever given in the text is the same as our Definition 12.1.

[71] This is especially easy to see when r has the form $r(x, y) = \langle x, y, h(x, y)\rangle$.

[72] There may well be *other* surface elements with border given by $\mathcal{C}$.

We defined an open subset $\mathcal{O} \subset \mathbb{R}^2$ to be *simply connected* if it is connected and "has no holes". A more precise way of saying this is: a connected open set $\mathcal{O} \subset \mathbb{R}^2$ is simply connected if any piecewise C^1, simple closed curve in $\mathcal{O}$ is the boundary of an open subset $U \subset \mathbb{R}^2$ that is completely contained in $\mathcal{O}$. We might refer to the closed set $U \cup bU \subset \mathcal{O}$ as the "surface element in $\mathbb{R}^2$ with border given by the curve bU."

In $\mathbb{R}^3$ an open set may be simply connected even if it has "holes". For example the open sets $\mathcal{O}_1 := \mathbb{R}^3 \setminus \{x : |x| \leq 1\}$ and $\mathcal{O}_2 = \mathbb{R}^3 \setminus \{0\}$ will be seen to be simply connected. We now define simply connected open sets in $\mathbb{R}^3$ in a way that generalizes the more precise definition we just gave of simply connected open sets in $\mathbb{R}^2$.

Definition 18.1. *Let $\mathcal{O}$ be a connected open subset of $\mathbb{R}^3$. We say $\mathcal{O}$ is simply connected if* every *piecewise C^1 simple closed curve in $\mathcal{O}$ is the border of some "surface element with border" $r(D)$ that is completely contained in $\mathcal{O}$.*

Draw a picture to see that the sets $\mathcal{O}_1$ and $\mathcal{O}_2$ defined just above are simply connected subsets of $\mathbb{R}^3$ in spite of their "holes". On the other hand the set $\mathcal{O}_3 := \mathbb{R}^3 \setminus \{\text{the } z \text{ axis}\}$ is not simply connected. The curve $\{(x, y, 0) : x^2 + y^2 = 1\}$ is not the border of any surface element with border that is completely contained in $\mathcal{O}_3$. The set $\mathcal{O}_4 := \mathbb{R}^3 \setminus \{\text{the nonnegative } z \text{ axis}\}$ *is* simply connected.

An alternative way to characterize simply connected open sets that works in $\mathbb{R}^n$ for any $n \geq 2$ is as follows. A connected open set $\mathcal{O} \subset \mathbb{R}^n$ is simply connected if *any* piecewise C^1 simple closed curve in $\mathcal{O}$ can be continuously deformed to a point "without leaving $\mathcal{O}$". That is, every intermediate curve during the contraction process is completely contained in $\mathcal{O}$. Any hole in a connected open set $\mathcal{O} \subset \mathbb{R}^2$ will prevent continuous deformation within $\mathcal{O}$ to a point for some closed curve in $\mathcal{O}$. In $\mathbb{R}^n$ for $n \geq 3$, there is "more room to move around", so only certain types of holes are an obstruction to such a continuous deformation.

19 Integrability of continuous functions on rectangles.

In this section we finally give a proof of Proposition 13.5, restated here.

Proposition 19.1. *Let R be a rectangle in $\mathbb{R}^n$ and suppose $f : R \to \mathbb{R}$ is continuous. Then $\int_R f(x)dV$ exists.*

Our goal will be to show how this follows from a fundamental property of the real numbers, namely *completeness.*[73]

Remark 19.2. *We made repeated use of Proposition 19.1 in the second half of this course. For example,* every time *we wrote down a line integral, or a surface integral, or a triple integral in connection with Green's theorem, Stokes's theorem, or the divergence theorem, we implicitly used Proposition 19.1 (or its strengthened version, Proposition 13.13) to guarantee the existence of that integral.*

Definition 19.3. *(a) A subset $S \subset \mathbb{R}$ is* bounded above *if there exists $a \in \mathbb{R}$ such that $s \leq a$ for all $s \in S$. We call a an* upper bound *of S.*

[73] The same property is also called the *least upper bound property.*

(b) Suppose $S \subset \mathbb{R}$ is bounded above. If there exists $b \in \mathbb{R}$ such that b is an upper bound of S and $b \leq a$ for any other upper bound a of S, then we call b the least upper bound *or* supremum *of S, and write $b = \sup S$.*

(c) If $S \subset \mathbb{R}$ is bounded below, we define lower bounds *and the* greatest lower bound *or* infimum *of S, $\inf S$, analogously.*

Here is an exercise to check your understanding of sup and inf. We will use this later in the proof of Proposition 19.1.

Exercise A. (a) Let A and B be subsets of $\mathbb{R}$ and suppose $A \subset B$, where B is bounded above. Prove that $\sup A \leq \sup B$.

(b) Let A and B be subsets of $\mathbb{R}$ and suppose $A \subset B$, where B is bounded below. Then $\inf A \geq \inf B$.

Theorem 19.4 (Completeness of $\mathbb{R}$)**.** *Any nonempty subset $S \subset \mathbb{R}$ which is bounded above has a supremum. We refer to this property of $\mathbb{R}$ as the* completeness *or* least upper bound property *of $\mathbb{R}$.*

To prove this theorem one must first give a precise definition of the real numbers. In section 14.2 we gave a precise definition of $\mathbb{Q}$ based on set theory, but we only indicated with a few examples how the elements of $\mathbb{R}$ could be defined in terms of rationals as *Dedekind cuts.*[74] Theorem 19.4 is a simple corollary of this definition of $\mathbb{R}$.

Remark 19.5. *(a) It follows immediately from Theorem 19.4 that any nonempty subset $S \subset \mathbb{R}$ that is bounded below has an infimum (why?).*[75]

(b) The set $\mathbb{Q}$ is not *complete. More precisely, if $S \subset \mathbb{Q}$ is bounded above, it is not necessarily true that S has a least upper bound* in $\mathbb{Q}$*. Consider, for example, $S = \{q \in \mathbb{Q} : q^2 < 2\}$.*

(c) If $S \subset \mathbb{R}$ is bounded above and closed, then $\sup S \in S$; consider the interval $[1, 4]$ for example. On the other hand $\sup[1, 4) \notin [1, 4)$.

In first-year calculus you worked with sequences of real numbers $(a_n)_{n=1}^{\infty}$, which can also be written as: $a_1, a_2, a_3, \ldots$. Given such a sequence, then $a_2, a_4, a_6, \ldots$ is an example of a *subsequence of* $(a_n)_{n=1}^{\infty}$.[76] More generally, the sequence $(b_m)_{m=1}^{\infty} := (a_{n_m})_{m=1}^{\infty}$ is a *subsequence* of $(a_n)_{n=1}^{\infty}$ if the subscripts $n_m \in \mathbb{N}$ satisfy $n_1 < n_2 < n_3 < \ldots$. Also, recall that $\lim_{n\to\infty} a_n = L$ if, given $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $|a_n - L| < \epsilon$.

Definition 19.6 (Sequences and subsequences in $\mathbb{R}^n$)**.** *(a) Let $(x_n)_{n=1}^{\infty}$ be a sequence of elements of $\mathbb{R}^n$ and suppose $x \in \mathbb{R}^n$. We say that x_n* converges to x *and write $\lim_{n\to\infty} x_n = x$ or simply $x_n \to x$ if, given $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $n \geq N$ implies $|x_n - x| < \epsilon$.*

(b) If (x_n) is such a sequence, we say that the sequence $(y_m)_{m=1}^{\infty} := (x_{n_m})_{m=1}^{\infty}$ is a subsequence *of (x_n) if the subscripts $n_m \in \mathbb{N}$ satisfy $n_1 < n_2 < n_3 < \ldots$.*

Remark 19.7. *Observe that if $K \subset \mathbb{R}^n$, a function $f : K \to \mathbb{R}$ is continuous (in the $\epsilon - \delta$ sense) if and only if $x_n \to x$ in K implies $f(x_n) \to f(x)$ in $\mathbb{R}$.*

[74] There are other equivalent ways of defining the reals; for example, one can use equivalence classes of "Cauchy sequences" of rational numbers.

[75] To see this, use the fact that $-S$ is bounded above.

[76] We will often write just (a_n) instead of $(a_n)_{n=1}^{\infty}$.

The next result, which we state without proof, is a (non-obvious) corollary of Theorem 19.4. You will see a proof of Theorem 19.8 if you take Math 521.

Theorem 19.8 (Heine-Borel theorem)**.** *Let $K \subset \mathbb{R}^n$. Then K is closed and bounded if and only if any sequence (x_n) of elements of K has a subsequence which converges to an element of K.*

Using this we can now prove the Extreme Value Theorem.

Theorem 19.9 (Extreme Value Theorem)**.** *Suppose $K \subset \mathbb{R}^n$ is a closed and bounded set and that $f : K \to \mathbb{R}$ is continuous. Then $f(K)$ is closed and bounded, and so f attains an absolute max on K.*

Proof. First we show $f(K)$ is closed and bounded. Suppose $(f(x_n))$ is any sequence in $f(K)$. Since K is closed and bounded, by Heine-Borel the sequence (x_n) has a subsequence $x_{n_m} \to x \in K$. By continuity we have $f(x_{n_m}) \to f(x) \in f(K)$, so Heine-Borel implies $f(K)$ is closed and bounded.[77] Remark 19.5(c) implies $\sup f(K) \in f(K)$, so we are done.

□

Recall that if $S \subset \mathbb{R}^n$, a function $f : S \to \mathbb{R}$ is *continuous* on S if, given any $p \in S$ and any $\epsilon > 0$, there exists $\delta = \delta(\epsilon, p) > 0$ such that $q \in S$ and $|q - p| < \delta$ imply $|f(q) - f(p)| < \epsilon$. There is a stronger notion of continuity, called *uniform continuity*, which is essential in the proof of Proposition 19.1.

Definition 19.10 (Uniform continuity)**.** *Let $S \subset \mathbb{R}^n$. A function $f : S \to \mathbb{R}$ is* uniformly continuous *on S if, given any $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that $q, p \in S$ and $|q - p| < \delta$ imply $|f(q) - f(p)| < \epsilon$.*

The word "uniform" is used here because for a given ϵ, the same $\delta = \delta(\epsilon)$ "works" for *any* pair of points $p, q \in S$ such that $|p - q| < \delta$. The choice of δ depends only on ϵ, not on both p and ϵ as in the definition of continuity at p. For example, the function $\frac{1}{x}$ is continuous on $(0, 1)$, but not uniformly continuous on $(0, 1)$ (draw its graph). Next we show that any continuous function with a closed and bounded domain is actually uniformly continuous on that domain.

Proposition 19.11. *Suppose $K \subset \mathbb{R}^n$ is closed and bounded and $f : K \to \mathbb{R}$ is continuous. Then f is uniformly continuous on K.*

Proof. Suppose not. Then there exists an $\epsilon > 0$ for which no $\delta > 0$ "works" as in the definition of uniform continuity. In particular, $\delta(n) = \frac{1}{n}$ "fails" for all n. More precisely, given any $n \in \mathbb{N}$ there exist points x_n, y_n in K such that $|x_n - y_n| < \frac{1}{n}$ but $|f(x_n) - f(y_n)| \geq \epsilon$. By Heine-Borel there exist subsequences x_{n_m}, y_{n_m} and points $x, y \in K$ such that $x_{n_m} \to x$, $y_{n_m} \to y$ as $m \to \infty$. By continuity we have $f(x_{n_m}) \to f(x)$ and $f(y_{n_m}) \to f(y)$, so $|f(x) - f(y)| \geq \epsilon$. But since $|x_{n_m} - y_{n_m}| < \frac{1}{n_m}$ for all m, we must have $x = y$, a contradiction. □

We will need to use "upper and lower sums" in our proof of Proposition 19.1.

[77] The same argument shows that if $F : K \to \mathbb{R}^m$ is continuous, then $F(K)$ is closed and bounded.

Definition 19.12. *Let R be a rectangle in $\mathbb{R}^n$ and suppose $f : R \to \mathbb{R}$ is bounded. If $P = \{R_1, \ldots, R_N\}$ is a partition of R, we define the* upper and lower sums of f for the partition P *to be respectively*[78]

$$\begin{aligned} U(f,P) &= \sum_{k=1}^{N} M_k V(R_k), \ \textit{where } M_k := \sup\{f(x) : x \in R_k\} := \sup_{R_k} f \\ L(f,P) &= \sum_{k=1}^{N} m_k V(R_k), \ \textit{where } m_k := \inf\{f(x) : x \in R_k\} := \inf_{R_k} f. \end{aligned} \tag{19.1}$$

We need to understand how upper and lower sums change when a partition is "refined". If P and P' are partitions of the rectangle R, we say that P' is a *refinement* of P if every subrectangle of the partition P' is contained in a subrectangle of P.[79] Given two partitions P_1 and P_2, it is always possible to choose a third partition P_3 which is a refinement of both P_1 and P_2. We then refer to P_3 as a *common refinement* of P_1 and P_2.

Exercise B. For f and R as in Definition 19.12, suppose the partition P' is a refinement of P. Show that

$$L(f,P) \leq L(f,P') \text{ and } U(f,P) \geq U(f,P'). \tag{19.2}$$

(Hint: This follows from the result of Exercise A and the fact that each subrectangle of the partition P is a union of subrectangles of P'.)

Proposition 19.13. *Let R be a rectangle in $\mathbb{R}^n$ and suppose $f : R \to \mathbb{R}$ is bounded. Set $m = \inf_R f$ and $M = \sup_R f$. Let P_1 and P_2 be* any *two partitions of R. Then*

$$mV(R) \leq L(f,P_1) \leq U(f,P_2) \leq MV(R). \tag{19.3}$$

Proof. The first and third inequalities are immediate. To see that the second holds, let P_3 be a common refinement of P_1 and P_2. Then by (19.2) we have

$$L(f,P_1) \leq L(f,P_3) \leq U(f,P_3) \leq U(f,P_2).$$

□

Remark 19.14. *Proposition 19.13 shows that* any *upper sum is an upper bound for the set of all possible lower sums, and that* any *lower sum is a lower bound for the set of all possible upper sums. By "the set of all possible lower sums" we mean $\{L(f,P) : P$ is some partition of $R\}$. This implies that for any partitions P, Q of R we have*

$$L(f,P) \leq \sup_P L(f,P) \leq \inf_P U(f,P) \leq U(f,Q). \tag{19.4}$$

Here $\sup_P L(f,P) := \sup\{L(f,P) : P$ is some partition of $R\}$.[80]

[78] Here we use the completeness of $\mathbb{R}$ to define M_k and m_k.

[79] This implies that any subrectangle of the partition P is a union of subrectangles of P'.

[80] To see that the second inequality holds, first use the definition of inf to show that for any partition P we have $L(f,P) \leq \inf_P U(f,P)$ (how?). Then apply the definition of sup to finish (how?).

The next proposition nearly completes the proof of Proposition 19.1.

Proposition 19.15. *Let R be a rectangle in $\mathbb{R}^n$ and suppose $f : R \to \mathbb{R}$ is continuous. Then*

$$\sup_P L(f,P) = \inf_P U(f,P). \tag{19.5}$$

Proof. Fix $\epsilon > 0$. By Proposition 19.11 f is uniformly continuous on R, so there exists $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon/V(R)$. Let P be any partition of R with $\|P\| < \delta$. Then[81]

$$U(f,P) - L(f,P) = \sum_k (M_k - m_k)V(R_k) < \frac{\epsilon}{V(R)} \sum_k V(R_k) = \epsilon. \tag{19.6}$$

Since ϵ is arbitrary, the result now follows from (19.4). □

It is now easy to finish the proof of Proposition 19.1.

Proof of Proposition 19.1. Let $A := \sup_P L(f,P) = \inf_P U(f,P)$ and fix $\epsilon > 0$. In view of Definition 13.4 it will suffice to show that we can choose $\delta > 0$ such that $\|P\| < \delta$ implies

$$|R(f,P) - A| < \epsilon, \tag{19.7}$$

where $R(f,P)$ is any Riemann sum associated to P. Choose δ exactly as in the proof of Proposition 19.15. Then if $\|P\| < \delta$ by (19.6) we have

$$U(f,P) - L(f,P) < \epsilon. \tag{19.8}$$

We also have

$$L(f,P) \le R(f,P) \le U(f,P) \text{ and } L(f,P) \le A \le U(f,P), \tag{19.9}$$

where the first pair of inequalities is immediate, and the second pair follows from (19.4). The inequality (19.7) now follows from (19.8) and (19.9). □

Remark 19.16. *If R is a rectangle in $\mathbb{R}^n$ and $f : R \to \mathbb{R}$ is bounded, we showed in Remark 19.14 that*

$$\sup_P L(f,P) \le \inf_P U(f,P). \tag{19.10}$$

In many presentations of the Riemann integral, the function f is defined *to be integrable on R if and only if equality holds in* (19.10)*. This definition, due to Darboux, can be shown to be* equivalent *to the one we gave in section 13, Definition 13.4, in terms of Riemann sums. Moreover, it is easy to show that f is Darboux integrable on R if and only if:*

$$\textit{given } \epsilon > 0, \textit{there exists a partition } P \textit{ such that } U(f,P) - L(f,P) < \epsilon. \tag{19.11}$$

This criterion for the existence of $\int_R f(x)dV$ is, I think, the simplest one and the easiest to use. For a given $\epsilon > 0$ one need only find a single *partition of R that "works" in the sense of satisfying* (19.11)*. If we had used this criterion, the proof of Proposition 19.1 would have ended with* (19.6).[82]

[81] The extreme value theorem implies $M_k = f(x_k)$ for some $x_k \in R_k$.

[82] The definition of the Riemann integral in terms of Riemann sums, Definition 13.4, is the more convenient one to use when proving Fubini's theorem, the change of variable theorem, and when justifying applications of the integral. Clearly, it is advantageous to have both Riemann's original definition, Definition 13.4, and the Darboux definition available to us.

20 Elementary proof of the implicit function theorem

In this section we give a proof of the version of the implicit function theorem that we used most often in class. The proof uses only the intermediate value theorem (from Calc I) and the definition of differentiability (reformulated slightly).[83]

We'll use the next lemma in step **3** of the proof of Theorem 20.2. Consider a function $F : \mathbb{R}^n_x \times \mathbb{R}_y \to \mathbb{R}$.

Lemma 20.1. *The function F is differentiable at $(a,b) \in \mathbb{R}^n_x \times \mathbb{R}_y$ if and only if there exist real-valued functions $e_i(x,y)$, $e(x,y)$ such that*

$$F(x,y) = F(a,b) + \sum_{i=1}^n F_{x_i}(a,b)(x_i - a_i) + F_y(a,b)(y-b) + \sum_{i=1}^n e_i(x,y)(x_i - a_i) + e(x,y)(y-b), \tag{20.1}$$

where $e_i(x,y) \to 0$ and $e(x,y) \to 0$ as $(x,y) \to (a,b)$.

Proof. ($\Rightarrow$) By definition of differentiability we have

$$F(x,y) = F(a,b) + \sum_{i=1}^n F_{x_i}(a,b)(x_i - a_i) + F_y(a,b)(y-b) + r(x,y), \tag{20.2}$$

where $\frac{r(x,y)}{|(x,y)-(a,b)|} \to 0$ as $(x,y) \to (a,b)$. Recall that for $z \in \mathbb{R}^{n+1}$, we have

$$\frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} |z_i| \le |z| \le \sum_{i=1}^{n+1} |z_i|. \tag{20.3}$$

Thus, $\frac{r(x,y)}{\sum_i |x_i - a_i| + |y-b|} := E(x,y) \to 0$ as $(x,y) \to (a,b)$, and this implies the lemma (why?).[84]

($\Leftarrow$)The function $r(x,y) := \sum_{i=1}^n e_i(x,y)(x_i - a_i) + e(x,y)(y-b)$ satisfies $\frac{r(x,y)}{|(x,y)-(a,b)|} \to 0$ as $(x,y) \to (a,b)$.

□

Theorem 20.2 (Implicit function theorem)**.** *Let $(a,b) \in \mathcal{O} \subset \mathbb{R}^n \times \mathbb{R}$ and suppose $F \in C^1(\mathcal{O}, \mathbb{R})$ satisfies $F(a,b) = 0$ and $F_y(a,b) \neq 0$.*[85] *Then there exist positive constants δ, ϵ such that:*

(a) For each x such that $|x-a| < \delta$, there is a unique y such that $|y-b| < \epsilon$ for which $F(x,y) = 0$. Thus, this correspondence defines a function f on $|x-a| < \delta$ such that for $(x,y) \in B := \{(x,y) : |x-a| < \delta, |y-b| < \epsilon\}$ we have $F(x,y) = 0 \Leftrightarrow y = f(x)$.

(b) The function f is C^1 on $|x-a| < \delta$.

Proof. **1. Existence of f.** We may suppose $F_y(a,b) > 0$. Using continuity of F_y we choose $\delta_1 > 0$, $\epsilon > 0$ such that $F_y(x,y) > 0$ for $|x-a| \le \delta_1$, $|y-b| \le \epsilon$. Since $F(a,b) = 0$ and $F(a,y)$ is strictly increasing in y, we have $F(a, b+\epsilon) > 0$, $F(a, b-\epsilon) < 0$. Using continuity of F choose $\delta < \delta_1$ such that $F(x, b+\epsilon) > 0$ and $F(x, b-\epsilon) < 0$ if $|x-a| < \delta$. This choice of δ and ϵ defines B. For

[83] The proof of Theorem 20.2 is a more detailed version of an unattributed proof found online.

[84] Write $r(x,y) = E(x,y)\left[\sum_i |x_i - a_i| + |y-b|\right]$ to see how to define the functions e_i, e.

[85] There is a more general version of the implicit function theorem for functions $F \in C^1(\mathcal{O}, \mathbb{R}^m)$, where $m < n$.

$|x-a| < \delta$, since $F(x,y)$ is strictly increasing in y, the intermediate value theorem implies there is a unique y with $|y-b| < \epsilon$ such that $F(x,y) = 0$. This defines $f(x) := y$ for $|x-a| < \delta$.

2. Continuity of f. Let a' be such that $|a'-a| < \delta$ and fix $\epsilon_1 > 0$. With $b' := f(a')$ we have $F(a',b') = 0$, $F_y(a',b') > 0$. To prove continuity of f at a', we must choose $\delta_1 > 0$ such that $|x-a'| < \delta_1$ implies $|f(x) - b'| < \epsilon_1$.[86] Choose $\epsilon_2 \le \epsilon_1$ such that $b' \pm \epsilon_2 \in (b-\epsilon, b+\epsilon)$. Then since $F_y > 0$ on B, we have $F(a', b'+\epsilon_2) > 0$, $F(a', b'-\epsilon_2) < 0$. Choose δ_1 such that $|x-a'| < \delta_1$ implies $|x-a| < \delta$ as well as $F(x, b'+\epsilon_2) > 0$ and $F(x, b'-\epsilon_2) < 0$. For $|x-a'| < \delta_1$ the intermediate value theorem implies there is a unique y such that $|y-b'| < \epsilon_2$ and $F(x,y) = 0$. Thus, $|f(x) - b'| < \epsilon_2 \le \epsilon_1$ for such x.

3. f is C^1 on $|x-a| < \delta$. First we show f is differentiable at a. Apply (20.1) with $y = f(x)$ to get

$$\begin{aligned} 0 = F(x, f(x)) = F(a,b) + \sum_i P_i(x, f(x))(x_i - a_i) + Q(x, f(x))(f(x) - f(a)) = \\ \sum_i P_i(x, f(x))(x_i - a_i) + Q(x, f(x))(f(x) - f(a)), \end{aligned} \tag{20.4}$$

where $P_i(x,y) := F_{x_i}(a,b) + e_i(x,y)$ and $Q(x,y) = F_y(a,b) + e(x,y)$. The functions P_i and Q are continuous at (a,b) (why?), so $P_i(x, f(x))$ and $Q(x, f(x))$ are continuous at $x = a$ (why?). Moreover, we have

$$P_i(a, f(a)) = F_{x_i}(a,b) \text{ and } Q(a, f(a)) = F_y(a,b) > 0. \tag{20.5}$$

Using (20.5) and the continuity of $Q(x, f(x))$ at a, we see that we can divide by $Q(x, f(x))$ for x near a to get

$$f(x) = f(a) - \sum_i \frac{P_i(x, f(x))}{Q(x, f(x))}(x_i - a_i) \tag{20.6}$$

Since each function $g_j(x) := -\frac{P_i(x, f(x))}{Q(x, f(x))}$ is continuous at a, (20.6) implies that f is differentiable at a with[87]

$$f_{x_i}(a) = -\frac{F_{x_i}(a,b)}{F_y(a,b)}. \tag{20.7}$$

The above argument can be repeated at other points a' such that $|a'-a| < \delta$ to give

$$f_{x_i}(a') = -\frac{F_{x_i}(a', f(a'))}{F_y(a', f(a'))}. \tag{20.8}$$

Since F is C^1, this formula together with the fact that f is continuous on $|x-a| < \delta$ shows that the f_{x_i} are continuous. Thus, f is C^1 on $|x-a| < \delta$. $\square$

[86] This step is essentially a repetition of step 1, but let's do it anyway.

[87] Write $g_j(x) = g_j(a) + (g_j(x) - g_j(a))$ to see that f is differentiable at a.