APMA SUMMER PROGRAM NOTES

ROSS PARKER

These are the notes for the 2019 summer workshop for incoming graduate students in applied mathematics at Brown University. These are based on the notes for the 2017 and 2018 workshops, which were led by Björn Sandstede. The original notes were typed by Rebecca Santorella and Patrick Liscio.

Contents

1. Metric Spaces

1.1. **Definitions.** A metric space is a nonempty set, together with a function called a metric, which defines a distance between any pair of points in the set. One familiar metric space is $\mathbb{R}^2$, the set of points in the Cartesian plane, with the metric given by the Euclidean distance

formula. The distance between points $p = (x_1, y_1)$ and $q = (x_2, y_2)$ is given by

$$d(p, q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2},$$

and represents the "straight-line distance" between p and q. An arbitrary metric space is defined as follows:

Definition 1.1. *Let X be an arbitrary set. A function $d : X \times X \to \mathbb{R}$ is a **metric** on X if the following conditions hold for all $x, y, z \in X$:*

(i) $d(x, y) \geq 0$.
(ii) $d(x, y) = 0$ *if and only if $x = y$.*
(iii) $d(x, y) = d(y, x)$ *(symmetry).*
(iv) $d(x, y) \leq d(x, z) + d(z, y)$ *(triangle inequality).*

*We call the pair (X, d) a **metric space**.*

Property (iv) is called the triangle inequality, since it generalizes the familiar statement from geometry: "The length of a side of a triangle must be less than or equal to the sum of the lengths of the other two sides". (If equality holds, the "triangle" has been squished down to a line segment, so it's not really a triangle anymore.) From the triangle inequality, we can derive the so-called "reverse triangle inequality",

$$d(x, y) \geq |d(x, z) - d(z, y)|$$

which generalizes the statement from geometry: "The length of a side of a triangle must be greater than or equal to the difference of the lengths of the other two sides". From now on, we will assume we are working in a metric space (X, d).

1.2. **Examples.** Here are some examples of metrics and metric spaces.

(1) The *Euclidean distance* on $\mathbb{R}^n$ is defined by

$$d(x, y) = |x - y|.$$

If $x = (x_1, \ldots x_n)$ and $y = (y_1, \ldots, y_n)$, then this is given by

$$d(x, y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}.$$

(2) The *discrete metric* on any set X is defined by

$$d(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{otherwise} . \end{cases}$$

(3) Compare three metrics on $\mathbb{R}^2$. For $x = (x_1, x_2)$ and $y = (y_1, y_2)$, we define
   (a) Manhattan distance / taxicab metric / ℓ^1 metric:

$$d_1(x, y) = |x_1 - y_1| + |x_2 - y_2|.$$

   (b) Euclidean distance / ℓ^2 metric:

$$d_2(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

   (c) Maximum distance / chessboard metric / ℓ^∞ metric:

$$d_\infty(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}.$$

(4) The maximum (supremum) metric on $C([a, b])$, the space of continuous, real-valued functions on the closed interval $[a, b]$, is defined by

$$d(f, g) = \max_{x \in [a,b]} |f(x) - g(x)|.$$

Since a continuous function attains its maximum and minimum on a closed interval (extreme value theorem), this is well-defined.

(5) Information theory:

    (a) Distance between two strings of identical length (Hamming distance): number of positions at which the two strings are different.

    (b) Distance between two arbitrary strings (Levenshtein distance): minimum number of single-character edits (insertions, deletions or substitutions) required to change one string into the other.

    (c) Damerau-Levenshtein distance: same as Levenshtein distance, except swaps of adjacent characters are also allowed operations. "More than 80% of all human misspellings can be expressed by a single error of one of the four types" (Damerau, 1964).

(6) Geodesic distance for connected graphs: number of edges in a shortest path connecting two vertices. (There may not be a unique shortest path, but the geodesic distance is unique). The graph must be connected to avoid having infinite distances.

1.3. **Product metrics.** Recall that the (Cartesian) product of two sets X and Y, denoted $X \times Y$, is the set of all ordered pairs (x, y), where $x \in X$ and $y \in Y$. The product of a finite sequence of sets $X_1, \ldots, X_n$ is the set of ordered n-tuples

$$\prod_{k=1}^{n} X_k = X_1 \times \cdots \times X_n := \{(x_1, \ldots, x_n) : x_k \in X_k\}.$$

Take a finite sequence of metric spaces $(X_1, d_1), \ldots, (X_n, d_n)$. Note that each set can have a different metric if we like. Then any of the following are metrics on $X_1 \times \cdots \times X_n$. We call these product metrics.

(1) $d(x, y) = \displaystyle\sum_{k=1}^{n} d(x_k, y_k)$.

(2) $d(x, y) = \sqrt{\displaystyle\sum_{k=1}^{n} d(x_k, y_k)^2}$.

(3) $d(x, y) \displaystyle\max_{k=1,\ldots,n} d(x_k, y_k)$.

We can also construct metrics on *countable* products of metric spaces. The countable product of sets $\{X_k\}$ is the set of all sequences

$$\prod_{k=1}^{\infty} X_k := \Big\{ \{x_k\} : x_k \in X_k \Big\}.$$

In order to define a metric on this countable product, we use the following trick, which allows us to construct metrics from other metrics. If $f : [0, \infty) \to [0, \infty)$ is an increasing concave function such that $f(x) = 0$ if and only if $x = 0$, then $f(d(x, y))$ is also a metric. Note that f does not have to be smooth. Important examples are:

(1) $f(x) = \min\{x, 1\}$.

(2) $f(x) = \dfrac{x}{1+x}$.

Both of these functions "cut off" the metric at 1, so that the greatest possible distance between any two points is 1. We can use the second of these "cutoff" functions to define a metric for a countable product of metric spaces $\{(X_k, d_k)\}$:

$$d(x, y) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{d_k(x_k, y_k)}{1 + d_k(x_k, y_k)}.$$

Alternatively, we could use the "cutoff" metric $\min\{d_k(x_k, y_k), 1\}$. We can also replace $1/2^k$ with the terms from any convergent series of positive terms.

1.4. **Sequences.** First, we define the convergence of a sequence in a metric space. Intuitively, a sequence $\{x_n\}$ converges to a limit L if the elements of the sequence get arbitrarily close to L.

Definition 1.2. *A sequence $\{x_n\}$ converges to L, written $x_n \to L$, if $d(x_n, L) \to 0$.*

The condition $d(x_n, L) \to 0$ is the convergence of a real-valued sequence, which is defined in the standard way. We can also write this as follows: $x_n \to L$ if, for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, L) < \epsilon$ for all $n \geq N$.

Another important type of sequence in a metric space is a Cauchy sequence. Intuitively, a sequence $\{x_n\}$ is a Cauchy sequence if its elements get arbitrarily close to *each other* (rather than approach a limit).

Definition 1.3. *$\{x_n\}$ is a **Cauchy sequence** if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < \epsilon$ for all $n, m \geq N$.*

Since this is annoying to write, we can notate this as $d(x_m, x_n) \to 0$. Every convergent sequence is a Cauchy sequence (this follows from the triangle inequality). The converse is unfortunately not true in general. That being said, there are metric spaces in which every Cauchy sequence is convergent (even if we cannot determine what that limit is!) Since this property is so important, we have the following definition.

Definition 1.4. *A metric space is **complete** if every Cauchy sequence is convergent.*

An example of a metric space which is not complete is $\mathbb{Q}$ with $d(x, y) = |x - y|$. To see this, take the sequence $\{x_n\}$, with $x_1 = 1$ and

$$x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}.$$

This is a Cauchy sequence, but its limit is $\sqrt{2}$, which is not in $\mathbb{Q}$.

An example of a complete metric space is $\mathbb{R}$ (the completeness of $\mathbb{R}$ follows from its construction). Using the product metric on $\mathbb{R}^n$ (the maximum version of the metric is easiest here) and the completeness of $\mathbb{R}$, it follows that $\mathbb{R}^n$ is complete. Since $\mathbb{C}$ is isomorphic to $\mathbb{R}^2$, $\mathbb{C}^n$ is complete as well.

At this point, we have two ways to show a sequence $\{x_n\}$ converges.

(1) Use the definition of convergence to show that $x_n \to L$. This means that we need a guess for what L.
(2) Work in a complete metric space, and show $\{x_n\}$ is a Cauchy sequence. This is often easier, since we do not need a guess for the limit, but it has the drawback of not giving us the actual limit.

1.5. **Limits and Continuity.** First we define the limit of a function between two metric spaces.

Definition 1.5. *Let* $f : (X, d_1) \to (Y, d_2)$. *Then* $f(x) \to L$ *as* $x \to x_0$ *if either of the following equivalent definitions holds:*

(1) *For every* $\epsilon > 0$ *there exists* $\delta > 0$ *(dependent on* ϵ *and* x_0*) such that if* $d_2(x, x_0) < \delta$, $d_2(f(x), L) < \epsilon$.

(2) *For every sequence* $\{x_n\}$ *with* $x_n \to x_0$, $f(x_n) \to L$.

We can use this to define continuity of a function (we will call this the metric space definition of continuity).

Definition 1.6. $f : (X, d_1) \to (Y, d_2)$ *is* **continuous at** x_0 *if* $f(x) \to f(x_0)$ *as* $x \to x_0$. f *is* **continuous** *if* f *is continuous at* x_0 *for all* $x_0 \in X$.

We discussed above that $C([a, b])$ is a metric space with the maximum (supremum) metric. Our goal is to show that $C([a, b])$ is a complete metric space. To do that, first we show that the metric itself is a continuous function.

Proposition 1.1. *The metric d is a continuous function from* $X \times X \to \mathbb{R}$.

Proof. Let $D((a, b), (x, y)) = d(a, x) + d(b, y)$ be the product metric on $X \times X$. Let $(x_n, y_n) \to (x, y)$ in $(X \times X, D)$. We will show that $d(x_n, y_n) \to d(x, y)$, which is equivalent to showing that $|d(x_n, y_n) - d(x, y)| \to 0$. Using the triangle inequality on $\mathbb{R}$ and the reverse triangle inequality on X,

$$\begin{aligned} |d(x_n, y_n) - d(x, y)| &\leq |d(x_n, y_n) - d(x, y_n)| + |d(x, y_n) - d(x, y)| \\ &= d(x_n, x) + d(y_n, y) \\ &= D((x_n, y_n), (x, y)) \\ &\to 0. \end{aligned}$$

$\square$

Next, we define the following two modes of convergence. (There are others, which you will study in courses in real analysis and probability). For simplicity, we will only consider real-valued functions here, but the codomain can be any metric space (with the appropriate adjustments).

Definition 1.7. *Consider a sequence of functions* $\{f_n\}$, $f_n : (X, d) \to \mathbb{R}$.

(1) $f_n \to f$ **pointwise** *if* $|f_n(x) - f(x)| \to 0$ *for all* $x \in X$. *In other words, if you give me* x *and* $\epsilon > 0$, *I can find a natural number* $N = N(x, \epsilon)$ *such that* $|f_n(x) - f(x)| < \epsilon$ *whenever* $n \geq N$.

(2) $f_n \to f$ **uniformly** *if the rate of convergence does not depend on* x. *In other words, if you give me* $\epsilon > 0$, *I can find a natural number* $N = N(\epsilon)$ *(independent of* x*) such that for all* $x \in X$, $|f_n(x) - f(x)| < \epsilon$ *whenever* $n \geq N$. *This is equivalent to*

$$\sup_{x \in X} |f_n(x) - f(x)| \to 0.$$

The supremum metric (also called the uniform metric) on the space of real-valued functions on (X, d) is defined by

$$D(f, g) = \sup_{x \in X} |f(x) - g(x)|.$$

Uniform convergence is convergence with respect to this metric. Next, we prove the uniform limit theorem, which says that the uniform limit of continuous functions is continuous.

Theorem 1.1 (Uniform Limit Theorem). *Let $\{f_n\}$ be a sequence of continuous functions on a metric space (X, d). If $f_n \to f$ uniformly, then f is continuous as well.*

Proof. We will show that f is continuous at $x_0 \in X$. Choose any $\epsilon > 0$.

(1) Since $f_n \to f$ uniformly, we can find $N \in \mathbb{N}$ such that $\sup_{x \in X} |f_n(x) - f(x)| < \epsilon$ for all $n \geq N$.
(2) Since f_N is continuous at x_0, we can find $\delta > 0$ such that if $d(x, x_0) < \delta$, $|f_N(x) - f_n(x_0)| < \epsilon$.
(3) By the triangle inequality, as long as $d(x, x_0) < \delta$,
$$|f(x) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(X_0)| \leq 3\epsilon$$

$\square$

Using all of this, we can show that $C([a, b])$ is a complete metric space, i.e. every Cauchy sequence is convergent.

Theorem 1.2. *The space $C([a, b])$ with the uniform metric is a complete metric space.*

Proof. Let $\{f_n(x)\}$ be a Cauchy sequence in $C([a, b])$. Our goal is to find a function $f \in C([a, b])$ such that $f_n \to f$, i.e $\sup_{x \in X} |f_n(x) - f(x)| \to 0$.

(1) Since for all $x \in [a, b]$, $|f_n(x) - f_m(x)| \leq \sup |f_n(x) - f_m(x)| \to 0$, $\{f_n(x)\}$ is a Cauchy sequence for each $x \in [a, b]$.
(2) By the completeness of $\mathbb{R}$, for each $x \in [a, b]$, $f_n(x)$ converges to some limit in $\mathbb{R}$. Call this limit $f(x)$.
(3) Next, we show that $f_n \to f$ uniformly, i.e. $\sup_{x \in X} |f_n(x) - f(x)| \to 0$. For any $\epsilon > 0$, choose $N \in \mathbb{N}$ such that for all $x \in [a, b]$ and all $m, n \geq N$, $|f_n(x) - f_m(x)| < \epsilon$. Since the metric $d(a, b) = |a - b|$ (i.e. the absolute value on $\mathbb{R}$) is continuous, send $m \to \infty$ to get $|f_n(x) - f(x)| < \epsilon$. Since this is true for all $x \in [a, b]$, $\sup_{x \in X} |f_n(x) - f(x)| < \epsilon$, i.e. $f_n \to f$ uniformly on $[a, b]$.
(4) $f(x)$ is continuous by the uniform limit theorem.

$\square$

1.6. Open and Closed Sets.
In this section, we define open and closed sets of a metric space (X, d).

Definition 1.8. *The ϵ-**neighborhood** or **open $\epsilon-$ball** of x_0 is the set*
$$B_\epsilon(x) = \{x \in X : d(x, x_0) < \epsilon\}$$
You may see this notated as $B(x, \epsilon)$ or $U_\epsilon(x)$.

We use this to define an open set in a metric space.

Definition 1.9. *A subset $U \subset X$ is **open** if for every $x \in U$ there exists $\epsilon > 0$ such that $B_\epsilon(x) \subset U$.*

In other words, we can find an open ball around any point in U which is entirely contained in U. This "open ball property" is so useful, that we can give it a name.

Definition 1.10. *x is an **interior point** of E there exists $\epsilon > 0$ such that $B_\epsilon(x) \subset U$.*

We can then say that a set U is open if it consists entirely of interior points. We also define a closed set.

Definition 1.11. *A subset $K \subset X$ is **closed** if whenever $\{x_n\} \subset K$ with $x_n \to x$, $x \in K$.*

The half-open interval $(0, 1]$ is not closed, since the sequence $\left\{\frac{1}{2^k}\right\}$ converges to 0, but $0 \notin (0, 1]$. Sometimes you will see closed sets defined in terms of limit points.

Definition 1.12. *A point x is a **limit point** of a set E if for every $\epsilon > 0$, B_ϵ contains a point in E which is not x. (Note that x may or may not be in E).*

The point 0 is a limit point of the half-open interval $(0, 1]$, but $0 \notin (0, 1]$. It follows from this definition that if x is a limit point of E, there exists a sequence $\{x_n\} \subset E$ such that $x_n \to x$. We then have the following alternative definition of a closed set.

Definition 1.13. *A subset $E \subset X$ is **closed** if it contains all of its limit points.*

We can show that that if K is closed, its complement $X \setminus K$ is open. In fact, topologists use this as the definition of a closed set, i.e. they say that $K \subset X$ is closed if $X \setminus K$ is open. We can also show the following properties of open and closed sets.

 (1) Arbitrary unions of open sets are open.
 (2) Finite intersections of open sets are open.
 (3) Arbitrary intersections of closed sets are closed.
 (4) Finite unions of closed sets are closed.

Topologists actually use these properties to define open sets! A topology on a set X is a family $\mathcal{T}$ of subsets of X which contains X and the empty set; is closed under arbitrary unions; and is closed under finite intersections. The sets in $\mathcal{T}$ are called then called the open sets.

Finally, we define the closure, interior, and boundary of a set. For all of these, several equivalent definitions are given.

Definition 1.14. *The* closure *of a set E, denoted $\overline{E}$, is defined in any of the following ways:*
 (1) All points in E together with all limit points of E.
 (2) The intersection of all closed sets containing E.
 (3) The smallest closed set containing E (i.e. if K is a closed set containing E, then K contains $\overline{E}$).

Definition 1.15. *The* interior *of a set E, denoted E°, is defined in any of the following ways:*
 (1) The set of all interior points of E.
 (2) The union of all open sets contained in E.
 (3) The largest open set contained in E (i.e. if U is an open set contained in E, then $U \subset E^\circ$).

Definition 1.16. *The* boundary *of a set E, denoted ∂E, is defined in any of the following ways:*
 (1) The closure of E without the interior of E: $\partial E = \overline{E} \setminus E^\circ$.
 (2) The set of points x such that every ball $B_\epsilon(x)$ contains at least one point of E and at least one point of $X \setminus E$.
 We note that the boundary of a set is always closed.

1.7. **Equivalent Metrics.** Sometimes it does not matter much what metric we use, i.e. different metrics give us the same convergent sequences and the same open sets.

Definition 1.17. *Two metrics d_1 and d_2 on X are **strongly equivalent** if there exist constants C_1 and C_2 such that for all $x, y \in X$,*

$$C_1 d_1(x, y) \leq d_2(x, y) \leq C_2 d_1(x, y)$$

If two metrics on X are strongly equivalent, the open sets and convergent sequences of X are the same; there may, however, be other differences between the spaces. For finite dimensional spaces such as $\mathbb{R}^n$, the Euclidean, taxicab, and maximum metrics are all strongly equivalent. For example, for the maximum and taxicab metrics, we have

$$\max_{k=1,\dots,n} |x_k - y_k| \leq \sum_{k=1}^{n} |x_k - y_k| \leq n \max_{k=1,\dots,n} |x_k - y_k|$$

so the definition of strong equivalence is satisfied with with $C_1 = 1$ and $C_2 = n$. It is a good exercise to draw the unit balls in the taxicab, Euclidean, and maximum metric in $\mathbb{R}^2$. The take-home message is that it the shape of the open balls usually don't matter too much.

1.8. **Topological Definition of Continuity.** It is sometimes easier to use the topological definition of continuity rather than the metric space definition. This way you can avoid all of those pesky ϵs and δs! Topologists define continuity as follows.

Definition 1.18. *A function $f : X \to Y$ is **continuous** if $f^{-1}(U)$ is open in X for each open set $U \subset Y$.*

Fortunately, the metric space definition of continuity is equivalent to the topological definition of continuity.

Theorem 1.3. $f : (X, d_1) \to (Y, d_2)$ *is continuous $\iff$ $f^{-1}(U)$ is open in X for each open set $U \subset Y$.*

Proof. We prove this in both directions.
($\implies$) Suppose f is continuous using the metric space definition. Let $U \subset Y$ be an open set. We need to show that $f^{-1}(U)$ is open. Let $x_0 \in f^{-1}(U)$, so that $y_0 = f(x_0) \in U$. Since U is open, $y_0 \in B_\epsilon(y_0) \subset U$. By continuity of f, there exists $\delta > 0$ such that $x \in B_\delta(x_0)$ implies $f(x) \in B_\epsilon(y_0)$. Since $B_\epsilon(y_0) \in U$, $x_0 \in B_\delta(x_0) \subset f^{-1}(U)$.

($\impliedby$) Suppose f is continuous using the topological definition. Choose any x_0 and any $\epsilon > 0$, and let $y_0 = f(x_0)$. Let $U = B_\epsilon(y)$, which is open. Then $f^{-1}(U)$ is open and contains x_0, thus there exists $\delta > 0$ such that $x_0 \in B_\delta(x_0) \subset f^{-1}(U)$. By definition of U, $d(f(x), f(x_0)) < \epsilon$ whenever $d(x, x_0) < \delta$. □

Since the complement of a open set is a closed set, and since the inverse image commutes with all set operations, we also have the following equivalent statement: $f : X \to Y$ is continuous if $f^{-1}(K)$ is closed in X for each closed set $K \subset Y$.

1.9. **Connectedness.** Intuitively, a disconnected set has two (or more!) "separate pieces". We state this mathematically as follows. There are many equivalent definition of connectedness. I like this one.

Definition 1.19. *A subset E is **disconnected** if we can find two disjoint, nonempty open sets A and B such that $E \subset A \cup B$ and both $E \cap A$ and $E \cap B$ are nonempty; E is **connected** if this is not possible.*

Here is another common definition you will see (i.e. the one in Rudin and Munkres). The advantage of this definition is that it defines a disconnected set as actually being split up into two separate pieces.

Definition 1.20. *Two sets A and B are **separated** if both $A \cap \overline{B}$ and $\overline{A} \cap B$ are empty. A subset E of X is **disconnected** if it is the union of two nonempty, separated sets. A subset E of X is **connected** if this is not possible.*

In the case of the real line, we can say exactly what the connected subsets are. Before we do that, we define an interval in the real line.

Definition 1.21. *A subset $I \subset \mathbb{R}$ is an **interval** if, for all x and y in I, every real number between x and y is in I as well.*

Examples of intervals include the open intervals (a, b), the closed intervals $[a, b]$, and unbounded intervals such as $[a, \infty)$. We then have the following proposition.

Proposition 1.2. *The connected subsets of $\mathbb{R}$ are intervals and points.*

Proof. Let E be a subset of $\mathbb{R}$. As an edge case, if E contains a single point, then E is connected. Therefore, we can assume that E contains at least two points. We will show that E is connected $\iff$ E is an interval, by proving the two contrapositives.

($\implies$) Show that if E is not an interval, than E is not connected. Suppose E is not an interval. Then we can find $x < y < z$ with $x, z \in E$ and $y \notin E$. Then $E \subset (-\infty, y) \cup (y, \infty)$, which implies that E is not connected.

($\impliedby$) Show that if E is not connected, then E is not an interval.
(1) Suppose E is not connected. Then we can write $E = A \cup B$, where A and B are separated, i.e. both $A \cap \overline{B}$ and $\overline{A} \cap B$ are empty.
(2) Let $x \in A$, $y \in B$, and without loss of generality, take $x < y$.
(3) Let $z = \sup (A \cap [x, y])$. Then $z \in \overline{A}$. Since A and B are separated, $z \notin B$, which implies $x \le z < y$. There are two possibilities to consider: $z \notin A$ and $z \in A$.
(4) If $z \notin A$, then $x < z < y$ with $z \notin A \cup B = E$, so E is not an interval.
(5) If $z \in A$, then $z \notin \overline{B}$. Therefore we can find $w \in (z, y)$ such that $w \notin B$. (If this were not possible, then z would be in $\overline{B}$). Since $w > z$, $w \notin A$ as well (by the definition of z as the supremum). Since $x \le z < w < y$ and $w \notin A \cup B = E$, E is not an interval.

$\square$

Using this, we can completely characterize the open sets of $\mathbb{R}$.

Proposition 1.3. *Any nonempty open set in $\mathbb{R}$ is the finite or countable union of disjoint open intervals.*

Proof. Let U be an open set in $\mathbb{R}$.
(1) On the set U, define the equivalence relation $x \sim y$ if $x, y \in I \subset U$, where I is connected. This partitions U into disjoint equivalence classes, which are the called the **connected components** of U.

(2) Let U_x be the connected component containing x. This is the largest connected subset of U which contains x. In other words, if I is an interval containing x, then $I \subset U_x$.

(3) By the property of equivalence classes, we can write U as the union of disjoint equivalence classes. In this case, that means we can find a subset $F \subset E$ such that
$$U = \bigcup_{x \in F} U_x,$$
and the elements in the union are all disjoint.

(4) Next, we show each U_x is an open interval. Since U_x is connected, it is an interval (by the previous proposition). Let $y \in U_x$. Since U is open, we can find $\epsilon > 0$ such that $B_\epsilon(y) = (y - \epsilon, y + \epsilon) \subset U$. We will show that $B_\epsilon(y) \subset U_x$.

(5) We note that $B_\epsilon(y)$ and U_y are both intervals containing y. Since U_y is the largest interval containing y, $B_\epsilon(y) \subset U_y$.

(6) Since the element y is shared by the equivalence classes U_y and U_x, it follows that $U_x = U_y$. This implies that $B_\epsilon(y) \subset U_x$, therefore U_x is open.

(7) We have shown that U is the union of disjoint open intervals. Since the rational numbers $\mathbb{Q}$ are dense in $\mathbb{R}$, we can find a rational number inside each of these intervals; furthermore, these rational numbers are distinct, since the open intervals are disjoint. Since $\mathbb{Q}$ is countable, we conclude that U is the disjoint union of at most countably many open intervals.

□

Finally, we show that connectedness is a topological property, i.e. it is preserved by continuous functions.

Theorem 1.4. *If $E \subset X$ is connected and $f : X \to Y$ is continuous, then $f(E)$ is connected.*

Proof. We employ proof by contradiction. Assume that E is connected.

(1) Suppose the conclusion is not true, i.e. $f(E)$ is disconnected. Then, by the definition of a disconnected set, $f(E) \subset A \cup B$, where A, B are disjoint, nonempty open sets with $F(E) \cap A \neq \emptyset$ and $F(E) \cap B \neq \emptyset$.

(2) It follows that $E \subset f^{-1}(A) \cup f^{-1}(B)$, both of which are open by the continuity of f.

(3) Since $F(E) \cap A \neq \emptyset$, $f^{-1}(A) \neq \emptyset$ and $f^{-1}(A) \cap E \neq \emptyset$. Similarly, $f^{-1}(B) \neq \emptyset$ and $f^{-1}(B) \cap E \neq \emptyset$.

(4) Finally, since the inverse image commute with set operations, $f^{-1}(A) \cap f^{-1}(B) = f^{-1}(A \cap B) = \emptyset$.

(5) Thus E is disconnected, which is not true.

□

The intermediate value theorem from calculus follows directly from this.

Corollary 1.1 (Intermediate Value Theorem (IVT)). *Let f be continuous on $[a, b]$. Then for every c (strictly) between $f(a)$ and $f(b)$ there exists $x \in (a, b)$ such that $f(x) = c$.*

What is curious (and really cool!) is that the derivative of a function also has the intermediate value property, regardless of whether the derivative itself is continuous.

Theorem 1.5 (Darboux). *Suppose $f : \mathbb{R} \to \mathbb{R}$ is differentiable, and let $a < b$. Then for every c (strictly) between $f'(a)$ and $f'(b)$, there exists $x \in [a, b]$ such that $f'(x) = c$.*

Proof. We proceed in the following steps.
  (1) Without loss of generality, take $f'(a) < c < f'(b)$ (otherwise replace f with $-f$).
  (2) Define $g(x) = f(x) - cx$. Since g is continuous and $[a, b]$ is closed, g must attain a minimum somewhere on $[a, b]$.
  (3) First, we show that this minimum cannot be at a. If g attains its minimum at a, then $g(x) - g(a) \geq 0$ on $[a, b]$, thus for $x in (a, b]$,
  $$\frac{g(x) - g(a)}{x - a} \geq 0.$$
  This implies from the definition of the derivative that $g'(a) \geq 0$, which cannot be the case since
  $$g'(a) = f'(a) - c < 0.$$
  (4) Similarly, we show that this minimum cannot be at b.
  (5) Therefore, the minimum is at a point $x \in (a, b)$. Since this is a local minimum, $g'(x) = 0$, which implies $f'(x) = c$.

$\square$

1.10. **Compactness.** Subsets of $\mathbb{R}^n$ that are closed and bounded (e.g. closed boxes) have nice properties. The concept of compactness generalizes this idea of "closed and bounded" to arbitrary spaces. A rough overview of the development of the mathematical concept of compactness is as follows. The first important result was proved by Bolzano in 1817 (as a lemma to prove the IVT), and then rediscovered by Weierstrass 50 years later. The theorem (which was originally proved on $\mathbb{R}$) bears both of their names.

Theorem 1.6 (Bolzano-Weierstrass). *Every bounded sequence in $\mathbb{R}^n$ has a convergence subsequence.*

Proof. Bolzano's proof uses the bisection method, a.k.a. "slicey-dicey". For convenience, we will prove the theorem on $\mathbb{R}$, although the proof works the same way for $\mathbb{R}^n$. Let $\{x_n\}$ be a bounded sequence in $\mathbb{R}$. Then $\{x_n\}$ fits inside a closed interval $I_1 = [a, b]$.
  (1) First, suppose $\{x_n\}$ has finite range $\{a_1, \ldots, a_n\}$, i.e. there are only finitely many distinct terms in the sequence. Then one of these, say a_k, must appear infinitely many times. This implies that the constant subsequence consisting only of the element a_k is a convergent subsequence.
  (2) With that out of the way, we may assume $\{x_n\}$ has infinite range. Cut I_1 in half. Then (at least) one half must contain infinitely many terms of $\{x_n\}$. Call that half I_2. Cut I_2 in half, and let I_3 be the half which contains infinitely many terms of $\{x_n\}$. Keep doing this to get a sequence of closed, nested intervals $I_1 \supset I_2 \supset I_3 \supset \ldots$. The lengths of these intervals go to 0, since each interval has half the length of the previous one.
  (3) We can show that the intersection of a sequence of closed, nested intervals $\{I_k\}$ is nonempty (this is sometimes called the Nested Intervals Theorem). Since the interval lengths go to 0, this intersection must be exactly one point x^*. (The same idea works in $\mathbb{R}^n$ for nested boxes).
  (4) It follows that x^* is a limit point of $\{x_n\}$. Since x^* is a limit point of $\{x_n\}$, we can find a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ which converges to x^*.

$\square$

In the late 19th century, a similar idea was applied to spaces of functions, in particular $C([a, b])$. A teaser is the following theorem. We will define equicontinuity in the next section.

Theorem 1.7 (Arzela-Ascoli). *Every sequence in $C([a, b])$ which is uniformly bounded and equicontinuous has a convergence subsequence.*

The term compactness was first used by Fréchet in 1905 to describe a set for which every sequence has a convergent subsequence. Today, this property is usually called sequential compactness.

Definition 1.22. *A subset K of X is **sequentially compact** if every sequence in K contains subsequence which converges to an element of K.*

The topological definition of compactness has its origins with Heine in 1870, who proved that every continuous function on $[a, b]$ is uniformly continuous (we will define that shortly). In the process, he proved that, given a family of countably many open intervals which cover $[a, b]$, you can select a finite number of them which still cover $[a, b]$. This was generalized into the modern, topological definition of compactness.

Definition 1.23. *A subset K of X is **compact** if any open cover of K has a finite subcover. In other words, given any collection of open sets whose union contains K, you can select a finite number of them whose union still contains K.*

It turns out that, in metric spaces, compactness and sequential compactness are equivalent. In fact, as we shall soon see, there are three equivalent criteria. Before we get to that, we need one more definition.

Definition 1.24. *A subset K of X is **totally bounded** if, for every $\epsilon > 0$, we can find a finite set of elements $\{x_1, \ldots, x_n\} \subset K$ such that*

$$K \subset \bigcup_{k=1}^{n} B_\epsilon(x_k).$$

In other words, for any $\epsilon > 0$, K can be covered by a finite number of ϵ-balls.

We start with the following proposition, which is an interesting result in its own right.

Proposition 1.4. *Suppose $\{x_n\}$ is a Cauchy sequence, and contains a subsequence $\{x_{n_k}\}$ which converges to x^*. Then $x_n \to x^*$.*

Proof. By the triangle inequality,

$$d(x_n, x^*) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x^*) \to 0$$

as $n, k \to \infty$. $\square$

We can now state the main equivalence theorem regarding compactness in metric spaces.

Theorem 1.8 (Compact Equivalence Theorem). *Let (X, d) be a metric space with $K \subset X$. Then the following are equivalent:*

  (i) *K is compact.*
 (ii) *K is sequentially compact.*
(iii) *K is complete and totally bounded.*

Proof. We will prove that (i) $\implies$ (ii), (ii) $\implies$ (iii), and (iii) $\implies$ (i).

(i) $\implies$ (ii): Suppose K is compact. We employ proof by contradiction. Let $\{x_n\}$ be a sequence in K, and assume (for a contradiction) that $\{x_n\}$ does not have a convergent subsequence, i.e. no subsequence converges to any element of K. Then for every $x \in K$, we can find a radius $\epsilon(x)$ such that $B_{\epsilon(x)}(x)$ contains only finitely many elements of $\{x_n\}$. Since the set of all open balls $B_{\epsilon(x)}(x)$ is an open cover for K, and K is compact, we can find a finite subcover for K. In other words, we can find a finite set of points $x_1, \ldots, x_n \in K$ such that

$$K \subset \bigcup_{k=1}^{n} B_{\epsilon(x_k)}(x_k).$$

Since each $B_{\epsilon(x)}(x)$ contains finitely many elements of the sequence $\{x_n\}$, this implies that K only contains finitely many elements of the sequence $\{x_n\}$, which is impossible.

(ii) $\implies$ (iii): Suppose that K is sequentially compact.

(1) First, we show that K is complete, i.e. all Cauchy sequences converge. Let $\{x_n\}$ be a Cauchy sequence in K. By sequential compactness, $\{x_n\}$ has a subsequence which converges to $x^* \in K$. By the prior proposition, $x_n \to x^*$, and so K is complete.
(2) We employ proof by contradiction. Suppose K is not totally bounded. Then there exists $\epsilon > 0$ such that K cannot be covered by finitely many open balls $B_\epsilon(x_k)$. Define a sequence $\{x_n\}$ as follows. Start by choosing any $x_1 \in K$. Then choose

$$x_2 \in K \setminus B_\epsilon(x_1)$$
$$x_3 \in K \setminus (B_\epsilon(x_1) \cup B_\epsilon(x_2))$$
$$x_3 \in K \setminus (B_\epsilon(x_1) \cup B_\epsilon(x_2)) \cup B_\epsilon(x_3))$$
$$\vdots$$

In other words, each element x_k in the sequence lies outside all of the previous ϵ-balls. This process never terminates, otherwise K could in fact be covered by finitely many ϵ-balls. By sequential compactness, $\{x_n\}$ has a convergent subsequence, but this is impossible since $d(x_j, x_k) \geq \epsilon$ for all $j \neq k$.

(iii) $\implies$ (i): Suppose K is complete and totally bounded. Once again, we employ proof by contradiction. Let $\{U_\alpha\}_{\alpha \in \mathcal{A}}$ be an open cover of K, and assume that there is no finite subcover. We construct the following sequence of sets.

(1) Take $\epsilon_1 = 1/2$. Since K is totally bounded, we can find points $y_1^1, \ldots, y_{n(1)}^1$ such that

$$K \subset \bigcup_{k=1}^{n(1)} B_{1/2}(y_k^1).$$

Since no finite subcover of $\{U_\alpha\}$ covers K, no finite subcover can cover at least one of the open balls $B_{1/2}(y_k^1)$ (otherwise the finite subcover would cover K). We will call this "uncoverable" open ball the "bad ball". Rearrange the $\{y_k^1\}$ to put the "bad ball" at the beginning, i.e. the "bad ball" is labeled $B_{1/2}(y_1^1)$. Let

$$B_1 = B_{1/2}(y_1^1) \cap K.$$

Note that B_1 cannot be covered by finitely many U_α.

(2) Repeat this for $\epsilon_2 = 1/2^2$. Again, since K is totally bounded, we can find points $y_1^2, \ldots, y_{n(2)}^2$ such that

$$K \subset \bigcup_{k=1}^{n(2)} B_{1/2^2}(y_k^2).$$

Since B_1 cannot be covered by finitely many U_α, no finite subcover can cover at least one of the sets $B_1 \cap B_{1/2^2}(y_k^2)$. Again, rearrange the $\{y_k^2\}$ so that $B_{1/2^2}(y_1^2)$ is the "bad ball", and let

$$B_2 = B_1 \cap B_{1/2^2}(y_1^2).$$

Again, B_2 cannot be covered by finitely many U_α.

(3) Repeat this process with $\epsilon_n = 1/2^n$ to get a nested sequence of nonempty sets

$$K \supset B_1 \supset B_2 \supset B_3 \supset \ldots$$

such that $B_n \subset B_{1/2^n}(y_1^n)$ for some $y_1^n \in K$, and none of the B_n can be covered by finitely many U_α. We note that each set B_n is contained in a ball of radius $1/2^n$.

(4) For each $n \in \mathbb{N}$, choose $x_n \in B_n$. Since the sets $\{B_n\}$ are nested, and each B_n is contained in a ball of radius $1/2^n$, $\{x_n\}$ is a Cauchy sequence. Since K is complete, $x_n \to x^* \in K$. Furthermore, $x^* \in B_n$ for all $n \in \mathbb{N}$.

(5) Since $\{U_\alpha\}$ covers K, $x^* \in U_{\alpha_0}$ for some α_0. Since U_{α_0} is open and the B_n are nested, shrinking, and contain x^*, $B_n \subset U_{\alpha_0}$ for sufficiently large n, which contradicts the fact that no B_n can be covered by finitely many U_α.

$\square$

As a corollary, closed subsets of compact sets are compact.

Corollary 1.2. *Let K be a compact subset of a metric space X. If $A \subset K$ is closed, then A is compact.*

Proof. Let $\{x_n\}$ be a sequence in A. Then there is a subsequence $x_{n_k} \to x^* \in K$, since K is sequentially compact. Since A is closed, $x^* \in A$. It follows that A is sequentially compact, thus A is compact by the compact equivalence theorem. $\square$

Next, we show that compact subsets of metric spaces are closed.

Proposition 1.5. *Compact subsets of metric spaces are closed.*

Proof. Let K be a compact subset of metric space X. We will show that $X \setminus K$ is open, which implies that K is closed. Choose any $x \in X \setminus K$, i.e. $x \neq K$. For all $y \in K$, let $r(y)$ be the distance $r(y) = \frac{1}{2}d(y,x) > 0$, since $y \neq x$. The collection of open balls $\{B_{r(y)}(y)\}_{y \in K}$ is an open cover for K, and none of them contain x by our definition of $r(y)$. By compactness, we can find a finite subcover $\{B_{r(y_1)}(y_1), \ldots, B_{r(y_n)}(y_n)\}$ of K. Let $r = \min\{r(y_1), \ldots, r(y_n)\}$. Then $B_r(x)$ does not intersect this finite subcover, which means that $B_r(x)$ lies outside of K. It follows that $B_r(x) \subset X \setminus K$, from which we conclude that $X \setminus K$ is open. $\square$

Although it is nice to have three equivalent criteria for compactness in metric spaces, they are still annoying to check. Fortunately, we have a nice criterion for compactness in $\mathbb{R}^n$, which generalizes the idea that closed, bounded intervals are "special".

Theorem 1.9 (Heine-Borel)**.** *A subset $K \subset \mathbb{R}^n$ is compact $\iff$ K is closed and bounded.*

Proof. We prove both directions.

($\implies$) Since $\mathbb{R}^n$ is a metric space and K is compact, K is closed and totally bounded. Since a totally bounded set is bounded, K is bounded.

($\impliedby$) Since K is bounded, it fits inside a closed box B in $\mathbb{R}^n$. Since K is closed, and closed subsets of compact sets are compact, it suffices to show that B is compact. We will show that B is sequentially compact. Since B is bounded, it follows from the Bolzano-Weierstrass theorem that any sequence in B has a convergent subsequence, whose limit must be in B since B is closed. Since B is sequentially compact, it is compact by the compact equivalence theorem. $\qquad\square$

Next, we show that compactness is also a topological property, i.e. it is preserved by continuous functions. Before we do this, we recall two important relations involving f and the inverse image operation f^{-1}.

(1) $f(f^{-1}(E)) \subset E$, with equality if f is surjective.
(2) $f^{-1}(f(E)) \supset E$, with equality if f is injective.

Theorem 1.10. *Let $f : (X, d_1) \to (Y, d_2)$ be continuous, and K compact in X. Then $f(K)$ is compact in Y. In other words, continuous images of compact sets are compact.*

Proof. Here, it makes sense to use the topological definition of compactness.

(1) Let $\{U_\alpha\}_{\alpha \in \mathcal{A}}$ be an open cover of $f(K)$.
(2) Since f is continuous, $f^{-1}(U_\alpha)$ is open in X, thus $\{f^{-1}(U_\alpha)\}_{\alpha \in \mathcal{A}}$ is an open cover for K.
(3) Since K is compact, we can find a finite subcover $\{f^{-1}(U_1), \ldots, f^{-1}(U_n)\}$ for K.
(4) Sending the finite subcover back through f, $\{f(f^{-1}(U_1)), \ldots, f(f^{-1}(U_n))\}$ covers K.
(5) Since $f(f^{-1}(U_k)) \subset U_k$, $\{U_1, \ldots, U_n\}$ is a finite subcover for $f(K)$.

$\qquad\square$

The extreme value theorem is a direct consequence of this theorem.

Theorem 1.11 (Extreme Value Theorem). *Let $f : (X, d) \to \mathbb{R}$ continuous and $K \subset X$ compact. Then f attains an absolute maximum and an absolute minimum on K.*

Proof. Since K is compact, $f(K) \subset \mathbb{R}$ is compact, thus closed and bounded by Heine-Borel. $\qquad\square$

Finally, we will prove a generalization of the theorem proved by Heine that continuity implies uniform continuity on a compact set. Uniform continuity is defined as follows.

Definition 1.25. *A function $f : (X, d_1) \to (Y, d_2)$ is **uniformly continuous** if, for every $\epsilon > 0$, there exists $\delta > 0$ (dependent on ϵ, but not on x) such that whenever $d_1(x, y) < \delta$, $d_2(f(x), f(y)) < \epsilon$.*

The main difference between uniform continuity and continuity is that, for a given $\epsilon > 0$, the same δ must work for all x, i.e. δ is independent of x. For simplicity, we will prove the uniform continuity theorem for real-valued functions, although it same result holds (with the same proof) for any pair of metric spaces.

Theorem 1.12 (Uniform Continuity Theorem). *Let $f : (X, d) \to \mathbb{R}$ be continuous and let $K \subset X$ compact. Then f is uniformly continuous on K.*

Proof. We present two proofs here. The first is the standard constructive proof.

(1) Let $\epsilon > 0$. Since f is continuous on K, for every $x_0 \in K$ we can find $\delta(x_0)$ such that $|f(x) - f(x_0)| < \epsilon/2$ whenever $d(x, x_0) < \delta(x_0)$. In general, the $\delta(x_0)$ will be different.

(2) The collection of open balls $\left\{ B_{\delta(x)/2}(x) \right\}_{x \in K}$ is an open cover for K. By compactness, we can find a finite subcover. In other words, we can find points $x_1, \ldots, x_n \in K$ such that
$$K \subset B_{\delta(x_1)/2}(x_1) \cup \cdots \cup B_{\delta(x_n)/2}(x_n).$$

(3) Let $\delta = \min\{\delta(x_1)/2, \ldots, \delta(x_n)/2\} > 0$. (We need this minimum to be well-defined and positive, which is why we need compactness to give us a finite subcover).

(4) Choose any $x, y \in K$ with $d(x, y) < \delta$. We will show that $|f(x) - f(y)| < \epsilon$.

(5) Because of the finite subcover, x must be inside one of the finite set of open balls $B_{\delta(x_k)/2}(x_k)$. It follows that $d(x, x_k) < \delta(x_k)/2$ for some $k \in \{1, dots, n\}$.

(6) By the triangle inequality,
$$d(y, x_k) \leq d(y, x) + d(x, x_k) < \delta + \delta(x_k)/2 \leq \delta(x_k),$$
since $\delta \leq \delta(x_k)/2$.

(7) Finally, by the triangle inequality and the continuity of f,
$$|f(y) - f(x)| \leq |f(y) - f(x_k)| + |f(x_k) - f(x)| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$
since both y and x are within a distance $\delta(x_k)$ of x_k.

Alternatively, we may use proof by contradiction.

(1) Suppose the conclusion is not true. Then for a specific $\epsilon > 0$, we can find sequences $\{x_n\}, \{y_n\} \subset K$ such that $d(x_n, y_n) \to 0$, but $|f(x_n) - f(y_n)| \geq \epsilon$.

(2) Since K is compact, thus sequentially compact, $\{x_n\}$ has a convergent subsequence $x_{n_k} \to x^*$, where $x^* \in K$, since K is closed.

(3) Since $d(x_n, y_n) \to 0$, $y_{n_k} \to x^*$ as well.

(4) By the triangle inequality,
$$0 < \epsilon \leq |f(x_{n_k}) - f(y_{n_k})| \leq \underbrace{|f(x_{n_k}) - f(x^*)| + |f(x^*) - f(y_{n_k})|}_{\text{both} \to 0 \text{ by continuity of } f} \to 0,$$

which is a contradiction.

$\square$

2. Arzela-Ascoli theorem

The Arzela-Ascoli theorem is the analogue of the Heine-Borel theorem for the metric space $C([a, b])$, the space of real-valued functions on $[a, b]$. It gives criteria for compactness that are (hopefully!) easier to check. We will prove the theorem on the more general space of real-valued, continuous functions on a compact subset K of a metric space (X, d). This space is denoted $C(K)$ (sometimes you will see $C^0(K)$ or $C(K, \mathbb{R})$), and we use the supremum (maximum) metric, which is defined by
$$d(f, g) = \max_{x \in K} |f(x) - g(x)|.$$

Since f and g are continuous, and K is compact, the maximum is well-defined by the Extreme Value Theorem. In addition, $C(K)$ is complete; the proof is identical to that for $C([a, b])$. Before we can state the theorem, we will need the following two definitions.

Definition 2.1. *A subset $A \subset C(K)$ is **uniformly bounded** if there exists $M \geq 0$ such that $|f(x)| \leq M$ for all $f \in A$ and $x \in K$. This can also be written as*

$$\sup_{f \in A, x \in K} |f(x)| \leq M.$$

Next, we define equicontinuity of a set of functions. This is essentially the same as uniform continuity, except the same δ must work for every function in the set.

Definition 2.2. *A subset $A \subset C(K)$ is **equicontinuous** if, for every $\epsilon > 0$, there exists $\delta > 0$ (dependent on ϵ, but not on f or x) such that whenever $d(x,y) < \delta$, $|f(x) - f(y)| < \epsilon$ for all $f \in A$. Alternatively, we can write this as*

$$\sup_{f \in A} |f(x) - f(y)| \to 0 \ \text{as } d(x,y) \to 0.$$

Before we can prove the theorem, we will need a result which is interesting in its own right. You will see the concept of separability (which is unfortunately named!) again in real analysis.

Definition 2.3. *A metric space is **separable** if it contains a countable dense subset.*

Here are some examples of separable metric spaces.
   (1) $\mathbb{R}$ contains $\mathbb{Q}$ as countable dense subset. In general $\mathbb{R}^n$ contains $\mathbb{Q}^n$ as a countable dense subset.
   (2) $C([0,1])$ is separable. It follows from the Weierstrass Approximation Theorem and the fact that $\mathbb{Q}$ is countable that $\mathbb{Q}[x]$, the set of polynomials with rational coefficients, is a countable dense subset of $C([0,1])$.

Next, we show that every compact metric space is separable.

Proposition 2.1. *Every compact metric space K contains a countable dense subset S.*

Proof. Choose any positive integer n. Then the collection of open balls $\{B_{1/n}(x)\}_{x \in K}$ with radius $1/n$ is an open cover for K. By compactness of K, we can find a finite collection of points $S_n = \{x_1^n, \ldots, x_{k(n)}^n\} \subset K$, where $k(n)$ depends on n, such that

$$K \subset \bigcup_{j=1}^{k(n)} B_{1/n}(x_j^n).$$

The set S_n is a finite set of points in K, and every point in K is "$1/n$-close" to a point in S_n. Perform this procedure for all $n \in N$, and take the union

$$S = \bigcup_{n=1}^{\infty} S_n.$$

S is dense in K, since, for any $\epsilon > 0$ and $x \in K$, every open ball $B_\epsilon(x)$ contains a point in S. S is countable, since it is the countable union of finite sets. $\square$

We can now state and prove the Arzela-Ascoli theorem.

Theorem 2.1 (Arzela-Ascoli). *Let K be a compact subset of a metric space (X, d). If a sequence of functions $\{f_n\} \in C(K)$ is uniformly bounded and equicontinuous, it has a uniformly convergent subsequence. Thus if a subset $A \subset C(K)$ is closed, uniformly bounded, and equicontinuous, it is compact.*

Proof. Let $\{f_n\}$ be an equicontinuous and uniformly bounded, with uniform bound M. Since K is a compact metric space, it is separable by the above proposition, therefore we can find a countable dense subset $S = \{x_n\} \subset K$. The first step is to use a diagonal argument to find a subsequence of $\{f_n\}$ which converges pointwise on S.

(1) We start by applying the sequence of functions $\{f_n\}$ to x_1. Consider the real-valued sequence $\{f_n(x_1)\}$. Since $|f_n(x_1)| \leq M$ for all n, $\{f_n(x_1)\}$ contains a convergent subsequence by Bolzano-Weierstrass. We will denote this convergent subsequence by $\{f_{1,n}(x_1)\}$. The corresponding functions form a subsequence $\{f_{1,n}\}$ of $\{f_n\}$.

(2) Next, we apply the sequence of functions $\{f_{1,n}\}$ from the previous step to x_2, and form the real-valued sequence $\{f_{1,n}(x_2)\}$. This sequence is also bounded by M, so it contains a convergent subsequence, which we will denote $\{f_{2,n}(x_2)\}$. The corresponding functions $\{f_{2,n}\}$ converge at both x_1 and x_2. (We note that $\{f_{2,n}\}$is a subsequence of $\{f_{1,n}\}$, which is in turn a subsequence of the original sequence $\{f_n\}$, i.e we have constructed a subsequence of a subsequence!)

(3) Iterate this procedure to get a countable collection of subsequences of the original sequence $\{f_n\}$, which we can depict in the following grid.

$$
\begin{array}{cccc}
f_{1,1} & f_{1,2} & f_{1,3} & \cdots \\
f_{2,1} & f_{2,2} & f_{2,3} & \cdots \\
f_{3,1} & f_{3,2} & f_{3,3} & \cdots \\
\vdots & \vdots & \vdots & \vdots
\end{array}
$$

By construction, each row in the grid is a subsequence of the previous row, and the sequence of functions in row n converges at the points $x_1, \ldots x_n$.

(4) Take the diagonal sequence $\{g_n\} = \{f_{n,n}\}$. This sequence is a subsequence of the original sequence $\{f_n\}$, and it converges at every point $x_n \in S$. (It turns out that we don't care what these limits actually are).

We have constructed a subsequence $\{g_n\}$ of $\{f_n\}$ which converges on a dense subset of K. The last step is to show that $\{g_n\}$ converges uniformly on K.

(1) Let $\epsilon > 0$. By the definition of equicontinuity, we can find $\delta > 0$ such that whenever $d(x, y) < \delta$, $|g_n(x) - g_n(y)| < \epsilon$ for all $n \in \mathbb{N}$.

(2) Since $S = \{x_n\}$ is dense in K, every point $x \in K$ lies inside $B_\delta(x_n)$ for some x_n. Thus the collection of open balls $\{B_\delta(x_n)\}_{x_n \in S}$ is an open cover for K. Since K is compact, we can find a finite subcover, i.e. there exists a positive integer n_0 such that

$$K \subset \bigcup_{k=1}^{n_0} B_\delta(x_k).$$

Let $S_0 = \{x_1, \ldots, x_{n_0}\}$. Then for every $x \in K$, $d(x, x_k) < \delta$ for some $x_k \in S_0$. (Alternatively, we can construct S_0 directly by choosing n such that $1/n < \delta$, covering K with $(1/n)-$balls at each point, extracting a finite subcover, and letting S_0 be the set of the centers of the balls in the finite subcover).

(3) The sequence of functions $\{g_n\}$ converges at the *finite* set of points S_0. This means that each real-valued sequence $\{g_n(x_k)\}$, for $k = 1, \ldots n_0$, is convergent, and therefore is a Cauchy sequence. Since this is a finite set of sequences, we can find a positive integer N such that for all $m, n \geq N$, $|g_m(x_k) - g_n(x_k)| < \epsilon$ for all $x_k \in S_0$.

(4) Choose any $x \in K$, and select $x_k \in S_0$ such that $d(x, x_k) < \delta$ according to step (2). For $m, n \geq N$, by the triangle inequality,

$$|g_n(x) - g_m(x)| \leq |g_n(x) - g_n(x_k)| + |g_n(x_k) - g_m(x_k)| + |g_m(x_k) - g_m(x)| < 3\epsilon,$$

where the first and third terms on the RHS are less than ϵ by equicontinuity, and the second term is less than ϵ by step (3). This implies that $\{g_n(x)\}$ is a (real-valued) Cauchy sequence for all x.

(5) Since $\{g_n(x)\}$ is a Cauchy sequence for all x, it follows from the completeness of $\mathbb{R}$ that $g_n(x) \to g(x)$ for all $x \in K$. Taking $m \to \infty$ above, we have

$$|g_n(x) - g(x)| \leq 3\epsilon$$

for all $x \in K$ and $n \geq N$. Since this is independent of x, the convergence is uniform.

We have successfully found a subsequence $\{g_n\}$ of $\{f_n(x)\}$ which converges uniformly on K. To prove the second statement of the theorem, let $A \subset C(K)$ be closed, uniformly bounded, and equicontinuous. By what we have already proved, any sequence $\{f_n\} \subset A$ must contain a convergent subsequence $f_{n_k} \to f$. Since A is closed, the limit $f \in A$, so A is sequentially compact. It follows that A is compact by the compact equivalence theorem. $\square$

We also have a partial converse to the Arzela-Ascoli theorem.

Theorem 2.2. *Let $K \subset (X, d)$ be compact. Then if $A \subset C(K)$ is totally bounded, it is uniformly bounded and equicontinuous.*

Proof. The idea behind the proof is to take an arbitrary function $f \in A$ and show that f is bounded by a constant which is independent of f, and f is uniformly continuous with δ depending on ϵ only, i.e. independent of f. We proceed as follows.

(1) Let $\epsilon > 0$. Using the definition of totally bounded, there exists a finite collection of functions $f_1^\epsilon, \ldots, f_{n(\epsilon)}^\epsilon \in A$, where the number of functions $n(\epsilon)$ depends on ϵ, such that

$$A \subset \bigcup_{k=1}^{n(\epsilon)} B_\epsilon(f_k^\epsilon).$$

The open balls $B_\epsilon(f_k^\epsilon)$ are defined using the maximum (supremum) metric.

(2) Choose any $f \in A$. Then $f \in B_\epsilon(f_k^\epsilon)$ for some $k \in \{1, \ldots, n(\epsilon)\}$. In particular, this means that $\sup_{x \in K} |f(x) - f_k^\epsilon(x)| < \epsilon$.

(3) First, we show uniform boundedness. By the triangle inequality

$$\sup_{x \in K} |f(x)| \leq \sup_{x \in K} |f(x) - f_k^\epsilon(x)| + \sup_{x \in K} |f_k^\epsilon(x)|$$

$$\leq \underbrace{\epsilon + \max_{j=1,\ldots,n(\epsilon)} \sup_{x \in K} |f_j^\epsilon(x)|}_{\text{define this to be } M, \text{ which is independent of } f}$$

$$= M < \infty,$$

where $\sup_{x \in K} |f_j^\epsilon(x)|$ is finite by the extreme value theorem, and we are taking the maximum over a finite set. Since M is independent of f, we conclude that A is uniformly bounded.

(4) Next, we show equicontinuity. Using the triangle inequality again,

$$|f(x) - f(y)| \leq |f(x) - f_k^\epsilon(x)| + |f_k^\epsilon(x) - f_k^\epsilon(y)| + |f_k^\epsilon(y) - f(y)|$$
$$\leq 2\epsilon + \max_{j=1,\ldots,n(\epsilon)} |f_j^\epsilon(x) - f_j^\epsilon(y)|.$$

(5) By the uniform continuity theorem, each function $f_j^\epsilon(x)$ is uniformly continuous on K. This means that for all $j = 1, \ldots, n(\epsilon)$, we can find $\delta_j > 0$ such that if $|x - y| < \delta_j$, $|f_j^\epsilon(x) - f_j^\epsilon(y)| < \epsilon$. Let $\delta = \min\{\delta_1, \ldots, \delta_{n(\epsilon)}\}$, which is positive since we are taking the minimum of a finite set of positive real numbers. Thus, whenever $|x - y| < \delta$,

$$\max_{j=1,\ldots,n(\epsilon)} |f_j^\epsilon(x) - f_j^\epsilon(y)| < \epsilon.$$

(6) Combining the previous two steps, if $|x - y| < \delta$, then

$$|f(x) - f(y)| < 3\epsilon.$$

Since δ is independent of f, we conclude that A is equicontinuous.

$\square$

We would like criteria which are easier to check than uniform boundedness and equicontinuity. Here are some results of that nature. First, we define Lipschitz continuity, which shows up in many contexts, including existence and uniqueness of solutions to ODEs. Again, we define this for real-valued functions, but the definition extends to any metric space.

Definition 2.4. *A function $f : (X, d) \to \mathbb{R}$ is **Lipschitz continuous** (or just Lipschitz) if there exists a positive constant $L > 0$ (the Lipschitz constant) such that for all $x, y \in X$,*

$$|f(x) - f(y)| \leq L d(x, y).$$

It is not difficult to show that Lipschitz continuity implies uniform continuity. For function from $\mathbb{R}$ to $\mathbb{R}$, the Lipschitz condition reduces to $|f(x) - f(y)| \leq L|x - y|$. This is trivially true for $x = y$, thus for $x \neq y$ we can divide both sides by $|x - y|$ to obtain the condition

$$\frac{|f(x) - f(y)|}{|x - y|} \leq L.$$

Intuitively, the Lipschitz constant L puts a bound on the slopes of all of the possible secant lines of f. A Lipschitz function is "nice" in the sense that it "does not change too fast". If $f : \mathbb{R} \to \mathbb{R}$ is continuously differentiable (which is denoted C^1) and $f'(x)$ is bounded by L, it follows from the mean value theorem that f is Lipschitz continuous with constant L.

Lemma 2.1. *Suppose $f : \mathbb{R} \subset \mathbb{R} \to \mathbb{R}$ is continuously differentiable, and $|f'(x)| \leq L$. Then f is Lipschitz continuous, with Lipschitz constant L.*

Proof. Choose any distinct points x and y in $\mathbb{R}$. Then by the mean value theorem, there exists a point c between x and y such that

$$\frac{f(x) - f(y)}{x - y} = f'(c).$$

Taking absolute values and rearranging, this becomes

$$|f(x) - f(y)| = |f'(c)||x - y| \leq L|x - y|,$$

which is the desired result.

$\square$

In particular, this holds when $f'(x)$ is continuous and we restrict ourselves to a closed interval $[a, b]$, in which case the Lipschitz constant is $L = \max_{x \in [a,b]} |f'(x)|$. In the next lemma, we give a criterion for equicontinuity in terms of Lipschitz continuity.

Lemma 2.2. *Let $K \subset \mathbb{R}^d$ be compact and $A \subset C(K)$. If every function in A is Lipschitz continuous with the same Lipschitz constant L, then A is equicontinuous.*

Proof. Let $\epsilon > 0$, and choose $\delta = \epsilon/2L$. Then for all $x, y \in K$ with $|x - y| < \delta$ and for all $f \in A$,

$$|f(x) - f(y)| \leq L|x - y| \leq L\delta = \frac{\epsilon}{2} < \epsilon.$$

$\square$

The same result holds for Hölder continuous functions. Hölder continuity is a weaker condition than Lipschitz continuity, and it shows up in the study of PDEs.

Definition 2.5. *A function $f : \mathbb{R}^d \to \mathbb{R}$ is **Hölder continuous** with exponent $\alpha \in (0, 1]$ (sometimes this is called α-Hölder continuous) if there exists a positive constant $C > 0$ such that for all $x, y \in X$,*

$$|f(x) - f(y)| \leq L|x - y|^\alpha.$$

The case $\alpha = 1$ is Lipschitz continuity. We exclude $\alpha > 1$, because it can be shown that any function on an interval $[a, b]$ satisfying the Hölder condition with $\alpha > 1$ is constant. Hölder continuity also implies uniform continuity, and the result of the previous lemma holds if every function in A is Hölder continuous with the same α and C. We mentioned above that the mean value theorem implies that if $f : [a, b] \to \mathbb{R}$ is continuously differentiable, f is Lipschitz with Lipschitz constant $\max_{x \in [a,b]} |f'(x)|$. Unfortunately, this does not work in higher dimensions, since there is no $n-$dimensional analogue to the mean value theorem. We can obtain a similar result, however, if f is continuously differentiable on a compact, convex set. We recall that since f' is continuous, it is bounded on all compact sets. First, we define a convex set.

Definition 2.6. *A set E in $\mathbb{R}^d$ is **convex** if for all $x, y \in E$, the line segment joining them is also in E. In other words, for all $x, y \in E$,*

$$tx + (1 - t)y \in E \qquad\qquad t \in [0, 1].$$

We then have the following lemma.

Lemma 2.3. *Let $K \subset \mathbb{R}^d$ convex and compact, and $K \subset U \subset \mathbb{R}^d$, where U is open. Let $f : U \to \mathbb{R}$ be continuously differentiable, and let $\sup_{x \in K} \|Df(x)\| = L$. Then $f : K \to \mathbb{R}$ is Lipschitz with constant L.*

Proof. Let $x, y \in K$. Since K is convex, $tx + (1-t)y \in K$ for all $t \in [0, 1]$. Next, we use the fundamental theorem of calculus to write $f(x) - f(y)$ in the following integrated form:

$$\begin{aligned}
f(x) - f(y) &= \int_0^1 \frac{d}{dt} f\left(tx + (1-t)y\right) dt \\
&= \int_0^1 Df(tx + (1-t)y) \cdot (x - y) dt \\
&= \left(\int_0^1 Df(tx + (1-t)y) dt\right) \cdot (x - y),
\end{aligned}$$

where the second line follows from the multivariable chain rule, and the dot represents the dot product in $\mathbb{R}^d$. Taking absolute values, we have

$$|f(x) - f(y)| \leq \left| \int_0^1 Df(tx + (1-t)y)dt \right| |x - y|$$

$$\leq \left(\int_0^1 \underbrace{\|Df(tx + (1-t)y)\|}_{\leq L \text{ since } tx+(1-t)y \in K} dt \right) |x - y|$$

$$\leq L|x - y|.$$

$\square$

2.1. Application to Existence of Solutions to ODEs.

In this section, we will use the Arzela-Ascoli theorem to prove the existence of solution to first order ODEs. Before we get too theoretical, let's look at several examples of initial value problems on $\mathbb{R}$.

(1) Exponential growth/decay:

$$\frac{du}{dt} = ku$$
$$u(0) = u_0$$

By separation of variables, this has a solution $u(t) = u_0 e^{kt}$, which exists for all time t.

(2) "Superexponential" growth:

$$\frac{du}{dt} = u^2$$
$$u(0) = 1$$

By separation of variables, this has solution $u(t) = \frac{1}{1-t}$. Since we start at $t = 0$, $u(t) \to \infty$ as $t \to 1$ from below. In other words, the solution blows up to infinity in finite time. (This is not good!)

(3) Non-uniqueness:

$$\frac{du}{dt} = u^{1/3}$$
$$u(0) = 0$$

We can see by inspection that $u(t) = 0$ is a solution. By separation of variable, we can also find another solution:

$$u(t) = \left(\frac{2t}{3} \right)^{3/2} \qquad\qquad t \geq 0.$$

We actually have an infinite family of solutions, which can be written piecewise as

$$u(t) = \begin{cases} 0 & t \leq T \\ \left(\frac{2(t-T)}{3} \right)^{3/2} & t > T, \end{cases}$$

where $T \geq 0$. In other words, we start at the zero solution, and at time T, we start following the nonzero solution. (This piecewise function is continuously differentiable but not smooth).

With these examples in hand, we can discuss existence of solutions to ODEs. From the second example, we will in general only be able to show local existence, i.e. existence in an interval around the starting point. The most basic result is due to Peano, where only continuity is assumed. The proof uses a compactness argument featuring the Arzela-Ascoli theorem. We note that while the procedure is relatively straightforward, some of the estimates in the proof are quite cumbersome.

Theorem 2.3 (Cauchy-Peano Existence Theorem). *Consider the initial value problem on* $\mathbb{R}^n$

$$\frac{du}{dt} = f(t, u)$$
$$u(t_0) = u_0.$$

If f is continuous in a neighborhood of (t_0, u_0), then there exists at least one solution $u(t)$ defined in a neighborhood of t_0. There is no guarantee of uniqueness.

Proof. For simplicity, we will only consider functions $f : \mathbb{R} \to \mathbb{R}$, although the proof is the same for functions on $\mathbb{R}^n$. For convenience and without loss of generality (since we can always translate the function via a change of variables), we will take $t_0 = 0$ and $u_0 = 0$. The strategy of the proof is:

(1) Rewrite the problem in integral form.
(2) Construct a sequence of approximate solutions $\{u_n(t)\}$ using the forward Euler method, where the mesh size $h \to 0$ as $n \to \infty$.
(3) Show this sequence $\{u_n(t)\}$ is uniformly bounded and equicontinuous.
(4) By Arzela-Ascoli, $\{u_n(t)\}$ must have a subsequence which converges to a function $u(t)$.
(5) Show that the limit $u(t)$ is what we want, i.e. it solves the initial value problem.

We proceed as follows.

(1) Since f is continuous in a neighborhood of $(0, 0)$, f is continuous on a box $B = [-R, R] \times [-R, R]$, for some $R > 0$. Since f is continuous, f is bounded on B, thus we can find a constant $M \geq 1$ such that $|f(t, u)| \leq M$ on B.
(2) Next, we rewrite the problem in integral form. This is necessary because we will use piecewise linear functions as approximate solutions, and these are not differentiable at the points where the pieces join together. We can show that $u(t)$ satisfies the original initial value problem if and only if $u(t)$ satisfies the integral equation

$$u(t) = u_0 + \int_{t_0}^{t} f(s, u(s)) ds.$$

(To see this, differentiate this expression with respect to t). Since we are taking $t_0 = 0$ and $u_0 = 0$, this becomes

$$u(t) = \int_{0}^{t} f(s, u(s)) ds.$$

(3) Let $T = R/M$, and consider the sequence of forward Euler approximations $u_n(t)$ on $[-T, T]$, which are defined as follows. Rather than write formulas for the approximations $u_n(t)$, which is annoying, we will give the procedure used to construct them.

(a) Choose any $n \in \mathbb{N}$, and let $h_n = T/n$ be the mesh size for the time grid, so that the grid for t is given by

$$[-t_n^n, -t_{n-1}^n, \ldots, -t_1^n, 0, t_1^n, t_2^n, \ldots, t_n^n]$$
$$= [-nh_n, -(n-1)h_n, \ldots, -h_n, 0, h_n, 2h_n, \ldots, nh_n].$$

The time grid contains $2n + 1$ grid points, and $nh_n = T$.

(b) We start with the initial condition $u_0^n = 0$, and then we compute u_n on the rest of the grid using the forward Euler method:

$$u_0^n = 0$$
$$u_1^n = 0 + f(0,0)h_n$$
$$u_2^n = u_1^n + f(t_1^n, u_1^n)h_n$$
$$\vdots$$
$$u_{m+1}^n = u_m^n + f(t_m^n, u_m^n)h_n$$
$$\vdots$$

and similarly for going backwards in t. (Is going backwards in t a form of time travel?)

(c) Define $u_n(t)$ to be the piecewise linear interpolation of these grid values, i.e. "connect the dots" by joining the points (t_m^n, u_m^n) with line segments. Since $u_n(t)$ is piecewise linear, $u_n(t)$ is differentiable on the open intervals (t_m^n, t_{m+1}^n), with the derivative given by

$$u_n'(t) = f(t_m^n, u_m^n).$$

It follows from the bound on f that $|u_n'(t)| \le M$ (where this is defined).

(4) Next, we show that the sequence $\{u_n(t)\}$ is uniformly bounded and equicontinuous. For uniform boundedness, we note that, for each Euler step, we have the bound

$$|u_{m+1}^n - u_m^n| \le Mh_n.$$

Since the function $u_n(t)$ involves (at most) n Euler steps in each direction, and $u_n(t)$ is linear between these steps, we have the uniform bound

$$|u_n(t)| \le nMh_n = MT.$$

For equicontinuity, since $u_n(t)$ is differentiable everywhere except at a finite number of grid points, we have for $-T \le s, t \le T$,

$$|u_n(t) - u_n(s)| \le \int_s^t |u_n'(r)| dr \le \int_s^t M dr = M|t - s|,$$

where we evaluate the integral by splitting it up into pieces at the grid points. Since the functions $\{u_n(t)\}$ are Lipschitz continuous with the same Lipschitz constant M, the sequence $\{u_n(t)\}$ is equicontinuous.

(5) By the Arzela-Ascoli theorem, $\{u_n(t)\}$ has a uniformly convergent subsequence, which we will denote by $\{v_k(t)\} = \{u_{n_k}(t)\}$, so that $v_k(t) \to u(t)$ uniformly.

(6) The only thing left is to show that $u(t)$ is a solution to the integrated form of the ODE. This is the really technical part. Using the triangle inequality

$$\left| u(t) - \int_0^t f(s, u(s))ds \right|$$

$$\leq |u(t) - v_k(t)| + \left| \int_0^t f(s, v_k(s))ds - \int_0^t f(s, u(s))ds \right| + \left| v_k(t) - \int_0^t f(s, v_k(s))ds \right|.$$

(7) The first term on the RHS $\to 0$ by the uniform convergence of $\{u_n(t)\}$ from the Arzela-Ascoli theorem. For the second term on the RHS,

$$\left| \int_0^t f(s, v_k(s))ds - \int_0^t f(s, u(s))ds \right| \leq \int_0^t |f(s, v_k(s))ds - f(s, u(s))|ds,$$

which also $\to 0$ by the uniform convergence of $v_k(t)$, together with the uniform continuity of f on the closed interval $[-T, T]$.

(8) All that remains is to show the third term on the RHS $\to 0$. This is unfortunately really annoying to do, since it quickly becomes a subscript nightmare. The idea is as follows. Since we only care about the subsequence $\{v_k(t)\} = \{u_{n_k}(t)\}$, we will always have $n = n_k$. For convenience, we will denote the grid points by t_j^k, where $t_j^k = t_j^{n_k}$. Let $t \in [0, T]$. Then t is always between two grid points, i.e. $t \in [t_m^k, t_{m+1}^k]$ for some m. (If t is actually one of the grid points, we will take $t = t_{m+1}^k$).

(9) Let v_j^k be the values of $v_k(t)$ on the grid t_j^k, i.e. $v_j^k = v_k(t_j^k)$. Then we can write $v_k(t)$ as the following telescoping sum involving grid points:

$$v_k(t) = \sum_{j=0}^{m-1}(v_{j+1}^k - v_j^k) + (v_k(t) - v_m^k),$$

where we recall that $v_k(0) = 0$. Substituting this into the 3rd term on the RHS of part (6), splitting the integral up at the grid points, and using the fundamental theorem of calculus, this becomes

$$v_k(t) - \int_0^t f(s, v_k(s))ds$$

$$= \sum_{j=0}^{m-1}(v_{j+1}^k - v_j^k) - \sum_{j=0}^{m-1}\int_{t_j^k}^{t_{j+1}^k} f(s, v_k(s))ds + (v_k(t) - v_m^k) - \int_{t_m^k}^t f(s, v_k(s))ds$$

$$= \sum_{j=0}^{m-1}\int_{t_j^k}^{t_{j+1}^k} v_k'(s, v_k(s))ds - \sum_{j=0}^{m-1}\int_{t_j^k}^{t_{j+1}^k} f(s, v_k(s))ds + \int_{t_m^k}^t v_k'(s, v_k(s))ds - \int_{t_m^k}^t f(s, v_k(s))ds$$

$$= \sum_{j=0}^{m-1}\int_{t_j^k}^{t_{j+1}^k} (f(t_j^k, v_j^k) - f(s, v_k(s)))ds + \int_{t_m^k}^t f(t_m^k, v_m^k) - f(s, v_k(s)))ds.$$

Taking absolute values, we obtain

$$\left| v_k(t) - \int_0^t f(s, v_k(s)) ds \right|$$

$$\leq \sum_{j=0}^{m-1} \int_{t_j^k}^{t_{j+1}^k} |f(t_j^k, v_j^k) - f(s, v_k(s))| ds + \int_{t_m^k}^t |f(t_m^k, v_m^k) - f(s, v_k(s))| ds$$

$$\leq \sum_{j=0}^{m} \int_{t_j^k}^{t_{j+1}^k} |f(t_j^k, v_j^k) - f(s, v_k(s))| ds.$$

(10) All that remains is to estimate this nasty integral! To to this, we use the uniform continuity of f on B. Let $\epsilon > 0$. Then we can find $\delta > 0$ such that if $\max(|s-t|, |u-v|) < \delta$, $|f(s,u) - f(t,v)| < \epsilon$. Since as n increases we are refining the grid, choose k sufficiently large so that $h_{n_k} < \delta/M$. Then, for all $s \in [t_m^k, t_{m+1}^k]$,

$$|s - t_m^k| \leq |t_{m+1}^k - t_m^k| < \delta,$$

and

$$|v_k(s) - v_m^k| \leq |v_{m+1}^k - v_m^k| \leq M h_{n_k} < \delta,$$

since $v_k(s)$ is a piecewise interpolation between v_k^k and v_{m+1}^k. For all the integrands involved in the sum, we have the bound

$$|f(t_j^k, v_j^k) - f(s, v_k(s))| \leq \epsilon.$$

Putting all of this together, we have

$$\left| v_k(t) - \int_0^t f(s, v_k(s)) ds \right| \leq \sum_{j=0}^{m} \int_{t_j^k}^{t_{j+1}^k} \epsilon \, ds$$

$$\leq \epsilon \sum_{j=0}^{m} (t_{j+1}^j - t_j^k) ds \leq \epsilon \sum_{j=0}^{m} h_{n_k} \epsilon \, ds$$

$$\leq \epsilon(m+1) h_{n_k} \leq \epsilon n_k \frac{T}{n_k} = \epsilon T,$$

since m is at most $n_k - 1$.

$\square$

Since the solution we obtained is only guaranteed to exist on $[-T, T]$, we call $u(t)$ a local solution to the initial value problem. We can extend this result to uniqueness by adding the condition that the function f is Lipschitz. In order to do this, we need a very useful (but technical) result known as the Gronwall inequality. There are seemingly hundreds of versions of the Gronwall inequality. (I think there is a whole book on them). Here is one version.

Theorem 2.4 (Gronwall Inequality)**.** *Let $u(t)$ and $g(t)$ be non-negative, real-valued functions defined on $t \in [t_0, t_1]$. Let $C \geq 0$ be a constant so that*

$$u(t) \leq C + \int_{t_0}^t g(s) u(s) ds \qquad\qquad t \in [t_0, t_1].$$

Then

$$u(t) \leq C \exp \left(\int_{t_0}^t g(s)ds \right) \qquad\qquad t \in [t_0, t_1].$$

It follows that if $C = 0$, $u = 0$ for $t \in [t_0, t_1]$.

Proof. We first consider the case where $C > 0$. Let

$$v(t) = C + \int_{t_0}^t g(s)u(s)ds.$$

Then $u(t) \leq v(t)$ on $[t_0, t_1]$ (by assumption), and $v(t) \geq C > 0$ on $[t_0, t_1]$ (since $g, u \geq 0$). Differentiating v with respect to t, we obtain

$$v'(t) = g(t)u(t) \leq g(t)v(t).$$

Since $v(t) > 0$ on $[t_0, t_1]$, we can divide by $v(t)$ to obtain

$$\frac{v'(t)}{v(t)} \leq g(t) \qquad\qquad t \in [t_0, t_1].$$

Next, we integrate the function g to obtain

$$\int_{t_0}^t g(s)ds \geq \int_{t_0}^t \frac{v'(s)}{v(s)}ds = \int_{t_0}^t \frac{d}{ds}\left(\log v(s)\right) ds = \log \frac{v(t)}{v(t_0)} = \log \frac{v(t)}{C}.$$

Finally, we exponentiate both sides and multiply by C to get

$$u(t) \leq v(t) \leq C \exp \left(\int_{t_0}^t g(s)ds \right) \qquad\qquad t \in [t_0, t_1],$$

which is the bound in the statement of the theorem. If $C = 0$, then for any $\epsilon > 0$,

$$u(t) \leq \epsilon + \int_{t_0}^t g(s)u(s)ds \qquad\qquad t \in [t_0, t_1]$$

By the first result (taking $C = \epsilon$), we have for $t \in [t_0, t_1]$

$$u(t) \leq \epsilon \exp \left(\int_{t_0}^t g(s)ds \right) \leq \epsilon \exp \left(\int_{t_0}^{t_1} g(s)ds \right) \leq M\epsilon,$$

which is independent of t. (To get the second line, we use the fact that $g(s)$ is nonnegative). Since ϵ is arbitrary, we conclude that $u = 0$ on $[t_0, t_1]$. $\qquad\square$

At first glance, this inequality appears to a useless result: why are we bounding a function by something which is growing exponentially? Let's take a closer look. We are taking a function $u(t)$ which is bounded by an exponential involving u itself and obtaining a bound which involves an exponential *independent* of u. Essentially, this says that $u(t)$ can grow at most exponentially fast. We can use the Gronwall inequality to prove the uniqueness of solutions to the initial value problem $\frac{du}{dt} = f(t, u)$ in the case where the function f is Lipschitz continuous.

Theorem 2.5 (Local Uniqueness of Solutions). *Consider the initial value problem on $\mathbb{R}^n$*

$$\frac{du}{dt} = f(t, u)$$
$$u(t_0) = u_0$$

Suppose that f is Lipschitz continuous in u in a neighborhood of (t_0, u_0), with Lipschitz constant L which is independent of t. Then then there exists a unique *solution $u(t)$ defined in a neighborhood of t_0.*

Proof. As in the proof of the Cauchy-Peano existence theorem, we can without loss of generality take $t_0 = 0$. Local existence follows from the Cauchy-Peano existence theorem. For uniqueness, suppose that two functions $u_1(t)$ and $u_2(t)$ are local solutions to the initial value problem, both of which exist on a time interval $[-T, T]$. If necessary, shrink T so that the Lipschitz condition holds on all of $[-T, T]$. We will show that $u_1(t) = u_2(t)$ on $[-T, T]$. Let $u(t) = u_1(t) - u_2(t)$. Since both u_1 and u_2 solve the integrated form of the problem,

$$
\begin{aligned}
|u(t)| &= \left| \left(u_0 + \int_0^t f(s, u_1(s)) ds \right) - \left(u_0 + \int_0^t f(s, u_2(s)) ds \right) \right| \\
&= \int_0^t |f(s, u_1(s)) - f(s, u_2(s))| ds \leq \int_0^t L|u_1(s) - u_2(s)| ds \\
&= 0 + \int_0^t Lu(s) ds
\end{aligned}
$$

We have satisfied the conditions of the Gronwall inequality with $g(t) = L$ and $C = 0$. It follows that $u(t) = 0$ for $t \in [0, T]$, from which we conclude that $u_1(t) = u_2(t)$ for $t \in [0, T]$. We can similarly obtain the result for $t \in [-T, 0]$. $\square$

We have successfully obtained conditions for local existence and local uniqueness for solutions to initial value problems. We will return to this topic and obtain more uniform results when we discuss contraction mappings.

3. Normed Vector Spaces and Banach Spaces

3.1. **Definitions.** A norm on a vector space is a function which measures the "length" of a vector.

Definition 3.1. *A **norm** on a vector space V is a function $\|\cdot\| : V \to \mathbb{R}$ with the following properties:*

> *(1) $\|x\| \geq 0$*
> *(2) $\|x\| = 0 \iff x = 0$*
> *(3) $\|\lambda x\| = |\lambda| \|x\|$ for all scalars λ*
> *(4) $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)*

*A vector space paired with a norm is a **normed vector space**.*

Every normed vector space is a metric space, since a norm induces a metric, which is given by

$$
d(x, y) = \|x - y\|.
$$

The converse, however, is not true. There are vector spaces on which there is a metric, but no norm can be found. Next, we define a bounded linear map between normed vector spaces. This is different from the concept of a bounded function that we use in, say, the Bolzano-Weierstrass theorem.

Definition 3.2. *Let X and Y be normed vector spaces, with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. Let $L : X \to Y$ be a linear map. Then L is **bounded** if there exists a constant $C \geq 0$ such that for all $u \in X$,*

$$\|Lu\|_Y \leq C\|u\|_X.$$

In a normed vector space, boundedness and continuity of linear operators is equivalent, which we show in the next proposition.

Proposition 3.1. *Let X and Y be normed vector space, with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively. A linear operator $L : X \to Y$ is bounded if and only if L is continuous.*

Proof. If L is bounded, then

$$\|Lu - Lv\|_Y = \|L(u - v)\|_Y \leq L\|u - v\|_X,$$

and so L is Lipschitz, thus continuous. For the other direction, assume L is continuous. Taking $\epsilon = 1$, since L is continuous at 0, we can find $\delta > 0$ such that, for all $u \in X$ with $\|u\|_X \leq \delta$, $\|Lu\|_Y \leq 1$. Let $x \in X$ with $x \neq 0$. Then

$$\left\| \frac{\delta}{\|x\|} x \right\|_X = \delta \left\| \frac{x}{\|x\|} \right\|_X = \delta,$$

from which it follows that

$$\left\| L\left(\frac{\delta}{\|x\|} x \right) \right\|_Y \leq 1.$$

Using the properties of the norm and the linearity of L, we can rearrange this to get

$$\|Lx\|_Y \leq \frac{1}{\delta} \|x\|_X,$$

thus L is a bounded linear operator with bound $C = 1/\delta$. $\qquad\square$

Let $\mathcal{L}(X, Y)$ be the space of bounded linear maps from X to Y. If $X = Y$, we usually denote this $\mathcal{L}(X)$. We define the operator norm of $L \in \mathcal{L}$ as follows.

Definition 3.3. *Let $L : X \to Y$ be a bounded linear operator. Then the **operator norm** of L is defined as one of the following, all of which are equivalent.*

(1)
$$\|L\| = \sup_{u \neq 0,\, \|u\|_X \leq 1} \|Lu\|_Y$$

(2)
$$\|L\| = \sup_{\|u\|_X = 1} \|Lu\|_Y$$

(3)
$$\|L\| = \inf\{C \geq 0 : \|Lu\|_Y \leq C\|u\|_X \text{ for all } u \in X\}$$

Definition 3.4. *A **Banach space** is a complete normed vector space, where completeness is with respect to the metric induced by the norm.*

We can show that the space $\mathcal{L}(X, Y)$ is a normed vector space using the operator norm. In addition, as long as Y is a Banach space, $\mathcal{L}(X, Y)$ is a Banach space.

Proposition 3.2. *$\mathcal{L}(X, Y)$ is a normed vector space with the operator norm. If Y is a Banach space, then $\mathcal{L}(X, Y)$ is also a Banach space.*

Proof. The proof is left as an exercise. To show $\mathcal{L}(X, Y)$ is complete, recall the proof for the completeness of $C([a, b])$. □

The next lemma is incredibly useful and gives a criterion for a specific linear operator on a Banach space to be invertible with bounded inverse. I use this all the time in my research (no joke!)

Lemma 3.1 (Neumann Series)**.** *Let X be a Banach space, and let $S \in \mathcal{L}(X)$ with $\|S\| < 1$. Then $I - S$ is invertible, and $(I - S)^{-1} \in L(X)$, where I is the identity operator on X.*

Proof. Define the *Neumann series* for S as

$$L = \sum_{n=0}^{\infty} S^n = I + S + S^2 + ...$$

which is the operator analogue of the ordinary geometric series. To show that this is well-defined, we note that the sequence of partial sums of L is a Cauchy sequence, thus the sum converges since X is complete. In addition, L is bounded, with

$$\|L\| \le \sum_{n=0}^{\infty} \|S\|^n \le \frac{1}{1 - \|S\|}.$$

Finally, since

$$(I - S)L = (I - S)\underbrace{\sum_{n=0}^{N} S^n}_{\to (I-S)L} = \sum_{n=0}^{N}(S^n - S^{n+1}) = \underbrace{I - S^{N+1}}_{\to I},$$

$(I - S)L = I$. Similarly, $L(I - S) = I$. □

3.2. **Differentiation in Banach Spaces.** We will now extend the concept of differentiation to general Banach spaces. To do that, we will first look at the derivative of a functions on $\mathbb{R}$ from a different perspective. From calculus, the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at $x = a$ is defined as

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

provided the limit exists. We can rearrange this to get

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

Note that $f'(a)$ is a real number, and as such is (trivially) a linear operator on $\mathbb{R}$. The term $f'(a)h$ can be thought of as applying this linear operator to h. The original definition of the derivative makes no sense in higher dimensions (unless we are taking a directional derivative, in which case h is still a real number). The second definition, however, is easily extended to higher dimensions.

Definition 3.5. *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is **differentiable** at a point $a \in R^n$ if there exists a linear transformation $L : \mathbb{R}^n \to \mathbb{R}^m$ such that*

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

Although it can be annoying to use this definition, if all partial derivatives of f exist and are continuous in a neighborhood of a, then f is differentiable, and the derivative is given by the Jacobian matrix

$$DF(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_m} & \cdots & \frac{\partial f_1}{\partial x_n} \end{pmatrix}_{x=a}.$$

There are analogues of the chain rule and the Taylor theorem in higher dimensions. Finally, we extend this definition to arbitrary Banach spaces.

Definition 3.6. *Let X and Y be Banach spaces, $U \subset X$ open, and $f : U \to Y$. Then f is **differentiable** at $u \in U$ if there exists a bounded linear transformation $L \in \mathcal{L}(X,Y)$ such that*

$$\lim_{h \to 0} \frac{\|f(u+h) - f(u) - Lh\|_Y}{\|h\|_X} = 0.$$

The map L is sometimes called the Fréchet derivative. If f is differentiable at $u_0 \in U$, we use the notation $Df(u_0)$ or $f_u(u_0)$ for the derivative.

We can show that that the Fréchet derivative, if it exists, is unique. In general, to show that the Fréchet derivative exists, we choose a candidate for the derivative, and then use the definition about to show that our candidate works. If f is differentiable for all $u \in U$, then the map $Df : U \to \mathcal{L}(X,Y)$ defined by $u \mapsto Df(u)$ is well-defined. A function f is C^1 if this map is continuous. Finally, we note that the chain rule remains valid in Banach spaces (provided the appropriate derivatives exist), and that higher order derivatives can be defined by considering the differentiability of $Df : U \to \mathcal{L}(X,Y)$, etc.

4. Fixed Point Theorems

This section covers fixed point theorems, which guarantee the existence of a unique fixed point of a function, i.e. a unique x such that $f(x) = x$. Fixed point methods are a powerful analytical tool, as we will soon discover. As motivation, we will look at Newton's method, which is used to find isolated zeros of a differentiable, real-valued function $f(x)$. Newton's method works as follows.

(1) Start with an initial guess x_0 (ideally near an actual zero of f) with $f'(x_0) \neq 0$.
(2) Iterate the algorithm

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Geometrically, the new value of x is where the tangent line of f at $(x_n, f(x_n))$ hits the x-axis.
(3) ????
(4) PROFIT

We would like this method to converge to a unique zero of f, which (we hope!) is the zero near our initial guess. To do this, we define a "Newton function" g by

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

If x is a fixed point of g, i.e. $g(x) = x$, then, as long as $f'(x) \neq 0$, $f(x) = 0$, which is exactly what we want. Our goal is to show that a fixed point of g exists, and that Newton's Method

converges to it. It turns out that the criterion we will need for this to work is that the map g is a contraction.

Definition 4.1. *Let D be a subset of a normed vector space $(X, \|\cdot\|)$. Then $F : D \to D$ is a **contraction** if there exists a constant L with $0 < L < 1$ such that*

$$\|F(x) - F(y)\| \le L\|x - y\|$$

for all $x, y \in D$. In other words, F is Lipschitz continuous with Lipschitz constant $L < 1$.

We will now state the Banach Fixed Point Theorem, which is also called the Contraction Mapping Principle. This theorem not only gives criteria for the existence of a unique fixed point, but also tells that that we can find it using successive iteration, which is exactly what we do in Newton's method.

Theorem 4.1 (Banach Fixed Point Theorem). *Let $(X, \|\cdot\|)$ be a Banach space, and $D \subset X$ a closed, nonempty subset of X. Suppose that the map $F : D \to D$ is a contraction, i.e. there exists a positive constant $L < 1$ such that for all $u, v \in D$,*

$$\|F(u) - F(v)\| \le L\|u - v\|.$$

Then F has a unique fixed point u^* in D, i.e. there exists a unique $u^* \in D$ such that $F(u^*) = u^*$.

Proof. As with many proofs involving complete metric spaces, the idea is to construct a Cauchy sequence in D, which must then converge since X is complete. Since D is closed, the limit must be in D. We construct the sequence by iterating the map F from any point in D.

(1) Start with an arbitrary $u_0 \in D$ and iteratively apply F to obtain a sequence $\{u_n\}$, which is defined as follows. Let

$$u_1 = F(u_0)$$
$$u_2 = F(u_1) = F(F(u_0))$$
$$\vdots$$

so that $u_n = F(u_{n-1})$ for all $n \ge 1$. Since $F : D \to D$, $u_n \in D$ for all n.

(2) Since F is a contraction,

$$\|F(u_1) - F(u_0)\| \le L\|u_1 - u_0\|$$
$$\|F(u_2) - F(u_1)\| \le L\|u_2 - u_1\| = L\|F(u_1) - F(u_0)\| \le L^2\|u_1 - u_0\|$$
$$\vdots$$

so that we have the bound

$$\|F(u_n) - F(u_{n-1})\| \le L^n\|u_1 - u_0\|.$$

Using the definition of F, this becomes

$$\|u_{n+1} - u_n\| \le L^n\|u_1 - u_0\|.$$

(3) Show that $\{u_n\}$ is a Cauchy sequence. For arbitrary $n, k \geq 1$, using the triangle inequality, we have

$$\|u_{n+k} - u_n\| \leq \sum_{j=0}^{k-1} \|u_{n+j+1} - u_{n+j}\| \leq \sum_{j=0}^{k-1} L^{n+j} \|u_1 - u_0\|$$

$$= \|u_1 - u_0\| L^n \sum_{j=0}^{k-1} L^j \leq \|u_1 - u_0\| L^n \sum_{j=0}^{\infty} L^j$$

$$= \|u_1 - u_0\| \frac{L^n}{1 - L}$$

$$\to 0 \text{ as } n \to \infty,$$

where, in the third line, since $0 < L < 1$, we used the sum of the infinite geometric series.

(4) Show $\{u_n\}$ converges to an element of D. Since $\{u_n\}$ is a Cauchy sequence and X is complete, u_n converges to some element u^* in X. Since D is closed and $\{u_n\} \subset D$, we must have $u^* \in D$.

(5) Show u^* is a fixed point of F. Since F is Lipschitz, it is continuous, thus by the definition of u_n, and the continuity of F,

$$F(u^*) = F\left(\lim_{n \to \infty} u_n\right) = \lim_{n \to \infty} F(u_n) = \lim_{n \to \infty} u_{n+1} = u^*.$$

(6) Show u^* is the unique fixed point of F in D. We follow the standard procedure, which is to suppose there are two fixed points and look at their difference. If $\tilde{u}^*$ is another fixed point, then

$$|u^* - \tilde{u}^*| = |F(u^*) - F(\tilde{u}^*)| < L|u^* - \tilde{u}^*|,$$

which is impossible since $L < 1$.

$\square$

It is important to note that to use the Banach fixed point theorem, we require a closed subspace D, and we have to verify two things:

(1) $F : D \to D$.
(2) F is a contraction on D.

It turns out that the first criterion is often the harder one to verify. Once that is done, we sometimes get the second one for free. Since we discussed Newton's method above, let's revisit it using the Banach fixed point theorem. We prove the following lemma, which gives us a criterion for when Newton's method converges.

Lemma 4.1 (Convergence of Newton's Method). *Let $f : [a, b] \to \mathbb{R}$ be C^2. Suppose for some $x \in [a, b]$ that $f(x) = 0$ and $f'(x) \neq 0$. Then there exists an interval $I = [x - \delta, x + \delta] \subset [a, b]$ such that Newton's method converges to x starting at any $x_0 \in I$.*

Proof. Define the "Newton function"

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

We will show that there is an interval containing x on which g is a contraction.

(1) Since f'' is continuous on the closed interval $[a, b]$, f'' is bounded on $[a, b]$ by the extreme value theorem, i.e. we can find a constant $M > 0$ such that $|f''(x)| \leq M$ for all $x \in [a, b]$.

(2) For some $\delta > 0$ (to be chosen later), let $I = [x - \delta, x + \delta] \subset [a, b]$. Then for any $y_1, y_2 \in I$, since g is continuously differentiable, it follows from the mean value theorem that

$$|g(y_1) - g(y_2)| \leq \sup_{y \in I} |g'(y)| \, |y_1 - y_2|.$$

Note that we have not yet guaranteed that $g'(y)$ does not blow up on I. Our goal is to choose δ sufficiently small to control $g'(y)$. Differentiating $g(y)$, we obtain

$$g'(y) = 1 - \frac{f'(y)^2 - f(y)f''(y)}{f'(y)^2} = \frac{f(y)f''(y)}{f'(y)^2}.$$

(3) Now comes the technical part. Since f and f' are continuous, $f(x) = 0$, and $f'(x) \neq 0$, choose δ sufficiently small so that $I \subset [a, b]$, and, for all $y \in I$,

(i) $|f'(y)| \geq \frac{1}{2}|f'(x)|$.

(ii) $|f(y)| \leq \frac{|f'(x)|^2}{8M}$.

(4) For all $y \in I$, using the expression for $g'(y)$ from step (2),

$$|g(y_1) - g(y_2)| \leq \frac{|f(y)||f''(y)|}{|f'(y)|^2}|y_1 - y_2| \leq \frac{|f'(x)|^2}{8M}M\frac{4}{|f'(x)|^2} \leq \frac{1}{2}|y_1 - y_2|.$$

(5) Since the Newton map g is a contraction on I, by the Banach Fixed Point Theorem it has a unique fixed point x^* in I. Since x is also a fixed point of g in I, by uniqueness we have $x^* = x$.

(6) Since Newton's method uses the technique of successive approximations, which is what was done in the proof of the Banach Fixed Point Theorem, Newton's method must converge to x.

$\square$

Another consequence of the Banach fixed point theorem is the inverse function theorem. For intuition, we will look at the one-dimensional case. Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuously differentiable at $x = a$. We know from calculus that if $f'(a) \neq 0$, f is invertible in a neighborhood of a. This makes sense, since if f' is continuous, and $f'(a) \neq 0$, f is either strictly increasing or strictly decreasing in a neighborhood of a. We can then determine the inverse from of the graph of f near a. Using the chain rule on $f^{-1}(f(x)) = x$ at $x = a$, if $b = f(a)$, then

$$(f^{-1})'(b) = \frac{1}{f'(a)}$$

The inverse function theorem extends this to higher dimensions. We will prove this later as a corollary of the Implicit Function Theorem (this is much asier to do!), but we will state the theorem here and outline how the proof uses the Banach fixed point theorem.

Theorem 4.2 (Inverse Function Theorem). *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable, and suppose the derivative $DF(x_0)$ is invertible at x_0. Then F is invertible in a neighborhood of x_0. Precisely,*

(1) *There exist neighborhoods U of x_0 and V of $y_0 = F(x_0)$, such that the restriction $F|_U : U \to V$ is a bijection.*
(2) *The inverse function $G : V \to U$ is also continuously differentiable, and for $y \in V$,*

$$DG(y) = DF(G(y))^{-1}.$$

Proof. For convenience, take $x_0 = 0$. An outline of the proof is as follows.

(1) Since $DF(0)^{-1}DF(x)$ is continuous at $x = 0$ and $DF(0)^{-1}DF(0) = I$, we can find $\delta > 0$ such that for all $\|x\| \leq \delta$,
    (a) $\|I - DF(0)^{-1}DF(x)\| \leq \frac{1}{2}$.
    (b) $DF(x)$ is nonsingular.
    Let $B = \overline{B_\delta(0)}$ (closed ball of radius δ about 0).
(2) For a fixed y (which we take as a parameter), define the "Newton map"

$$N(x; y) = x - DF(0)^{-1}(F(x) - y).$$

Note the resemblance of this map to that used in Newton's method. Since $DF(0)^{-1}$ is nonsingular, x is a fixed point of $N(\cdot; y)$ if and only if $y = F(x)$.
(3) Verify the hypotheses of the Banach fixed point theorem:
    (a) Show $N(\cdot; y) : B \to B$ for all y in a neighborhood V of $f(0)$.
    (b) Show $N(\cdot; y)$ is contraction on B.
(4) For all $y \in V$, use the Banach fixed point theorem to find a unique $x \in B$ such that $f(x) = y$. Let $f^{-1}(y)$ be this unique x.
(5) Show $f^{-1}(y)$ is continuous and differentiable on V. This is annoying, but it can be done. Once we have proved the Uniform Contraction Mapping Principle, this will be much easier to do.

$\square$

The next theorem we will prove is the Uniform Contraction Mapping Principle. The idea here is that we have a family of contraction maps $F(x; \mu)$ indexed by a parameter μ. For each value of the parameter μ, $F(\cdot; \mu)$ has a unique fixed point by the Banach fixed point theorem. Let $G(\mu)$ map each value of μ to that unique fixed point. The Uniform Contraction Mapping Principle says that the map G is as smooth as the original map F. The proof incorporates elements from Humpherys, Jarvis, and Evans, *Foundations of Applied Mathematics, Volume 1: Mathematical Analysis* (2017). We will first need one technical result involving operator norm bounds on derivatives of Lipschitz functions.

Proposition 4.1. *Let X and Y be Banach spaces, U an open subset of X, and $F : U \subset X \to Y$ differentiable. If F is Lipschitz with constant L, then, for all $x \in U$, the operator norm of $DF(x)$ has bound $\|DF(x)\| \leq L$.*

Proof. Let $\epsilon > 0$ and $x \in U$. Since F is differentiable at x, by the definition of differentiability in Banach spaces, we can find $\delta > 0$ such that whenever $\|h\|_X < \delta$,

$$\frac{\|F(x + h) - F(x) - DF(x)h\|_Y}{\|h\|_X} < \epsilon.$$

We use the "supremum of unit vectors" definition of the operator norm (definiton 2 above). Let $u \in U$ be a unit vector, and let $h = \frac{\delta}{2}u$, so that $\|h\|_X < \delta$. Then since $u = h/\|h\|_X$, we

have from the triangle inequality

$$
\begin{aligned}
\|DF(x)u\| &= \frac{\|DF(x)h\|_Y}{\|h\|_X} \\
&\leq \frac{\|DF(x)h - (F(x+h) - F(x))\|_Y + \|F(x+h) - F(x)\|_Y}{\|h\|_X} \\
&= \frac{\|F(x+h) - F(x) - DF(x)h\|_Y}{\|h\|_X} + \frac{\|F(x+h) - F(x)\|_Y}{\|h\|_X} \\
&\leq \epsilon + \frac{L\|x+h-x\|_X}{\|h\|_X} = \epsilon + \frac{L\|h\|_X}{\|h\|_X} \\
&= \epsilon + L,
\end{aligned}
$$

where in the fourth line we used the fact that F is Lipschitz. Since ϵ is arbitrary and this is independent of u, we conclude that $\|DF(x)\| \leq L$ for all $x \in U$. $\square$

We will now state and prove the Uniform Contraction Mapping Principle. For the next two theorems, we will use the notation $|\cdot|$ for the norms on Banach spaces, and $\|\cdot\|$ for the operator norm.

Theorem 4.3 (Uniform Contraction Mapping Principle). *Let*

(1) *X be a Banach space.*
(2) *$D \subset X$ a closed, nonempty subset of X.*
(3) *B an open subset of a Banach space Y (the "parameter space").*
(4) *$F : D \times B \to D$ a map which is **uniform contraction**, i.e. there exists a constant $L < 1$ such that, for all $\mu \in B$ and $u, v \in D$,*

$$|F(u, \mu) - F(v, \mu)| \leq L|u - v|.$$

Let $G : B \to D$ be the map which associates every $\mu \in B$ with the unique fixed point of $F(\cdot; \mu)$ from the Banach fixed point theorem. Then

(i) *If F is **uniformly Lipschitz** in μ, i.e. there exists a constant $M > 0$ such that for all $u \in D$ and $\mu_1, \mu_2 \in B$,*

$$|F(u, \mu_1) - F(u, \mu_2)| \leq M|\mu_1 - \mu_2|,$$

then G is Lipschitz continuous, with Lipschitz constant $M/(1 - L)$.
(ii) *If $F \in C^k(D \times B, X)$ for $k \geq 0$, then $G \in C^k(B, X)$, i.e. G has the same smoothness as F ($k = \infty$ is allowed).*

Proof. Define $G(\mu)$ as in the statement of the theorem. By the Banach fixed point theorem, $G : D \to B$ is the unique function such that $F(x; \mu) = x$ if and only if $x = G(\mu)$. We proceed in the following steps.

(1) Since we would like to show continuity of G, we first derive an expression (and estimate) for $|G(\mu_1) - G(\mu_2)|$. Using the fact that $G(\mu)$ is a fixed point of $F(\cdot; \mu)$ and the triangle inequality,

$$
\begin{aligned}
|G(\mu_1) - G(\mu_2)| &= |F(G(\mu_1); \mu_1) - F(G(\mu_2); \mu_2)| \\
&\leq |F(G(\mu_1); \mu_1) - F(G(\mu_1); \mu_2)| + |F(G(\mu_1); \mu_2) - F(G(\mu_2); \mu_2)| \\
&\leq |F(G(\mu_1); \mu_1) - F(G(\mu_1); \mu_2)| + L|G(\mu_1) - G(\mu_2)|.
\end{aligned}
$$

Since $0 < L < 1$, subtract the last term on the RHS from both sides to obtain
$$(1 - L)|G(\mu_1) - G(\mu_2)| \leq |F(G(\mu_1); \mu_1) - F(G(\mu_1); \mu_2)|.$$

(2) Finally, divide by $(1 - L)$ to get
$$|G(\mu_1) - G(\mu_2)| \leq \frac{1}{1 - L}|F(G(\mu_1); \mu_1) - F(G(\mu_1); \mu_2)|.$$

If F is continuous in both variables (which is part (ii) with $k = 0$, then the RHS above $\to 0$ as $\mu_2 \to \mu_1$, thus G is continuous.

(3) For part (i), if F is uniformly Lipschitz in μ, then the RHS in step (2) becomes
$$|G(\mu_1) - G(\mu_2)| \leq \frac{M}{1 - L}|\mu_1 - \mu_2|.$$

(4) All that remains is consider part (ii) with $k > 0$. We first consider the case $k = 1$, i.e. F is continuously differentiable. The way we will do this is to devise a candidate for $DG(\mu)$ and then use the definition of the derivative to show that this candidate is indeed the derivative.

(5) Recall that $G(\mu) = F(G(\mu), \mu)$. If G were differentiable, then, by the chain rule and differentiability of F, we would have
$$DG(\mu) = DF(G(\mu), \mu) = D_X F(G(\mu), \mu)DG(\mu) + D_B F(G(\mu), \mu).$$

This means that $DG(\mu)$ would be a fixed point of the mapping $\Phi : \mathcal{L}(B, X) \times B \to \mathcal{L}(B, X)$, defined by
$$\Phi(A; \mu) = D_X F(G(\mu), \mu)A + D_B F(G(\mu), \mu).$$

The map Φ is a uniform contraction, since for $A_1, A_2 \in \mathcal{L}(B, X)$,

$|\Phi(A_1; \mu) - \Phi(A_2; \mu)|$
$\quad = |D_X F(G(\mu), \mu)A_1 + D_B F(G(\mu), \mu) - (D_X F(G(\mu), \mu)A_2 + D_B F(G(\mu), \mu))|$
$\quad = |D_X F(G(\mu), \mu)(A_1 - A_2)| \leq \|D_X F(G(\mu), \mu)\| \, |A_1 - A_2|$
$\quad \leq L|A_1 - A_2|,$

where the last line follows from the previous proposition and the fact that $F(\cdot; \mu)$ is Lipschitz with constant L.

(6) Since $L < 1$, by the Banach fixed point theorem, there exists a function $Z : B \to \mathcal{L}(B, X)$ which maps each $\mu \in B$ to the unique fixed point $Z(\mu)$ of $\Phi(\cdot; \mu)$. Since F is C^1, Φ is continuous, thus by the $k = 0$ case of part (ii) of the Uniform Contraction Mapping Principle (which we have already proved!), the map $Z(\mu)$ is continuous.

(7) The function $Z(\mu)$ is our candidate for $DG(\mu)$. All that remains is to use the definition of the derivative to show that $Z(\mu)$ is actually the derivative. This is very technical and time-consuming, so we will omit it. If you are interested in how this is done, you can look the proof of Lemma 7.2.9 on pages 284-285 of Humpherys, Jarvis, and Evans (2017). We then repeat this argument for $k > 1$ for higher order derivatives.

$\square$

We will use the uniform contraction mapping principle to prove the implicit function theorem, which is one of the most important tools in analysis and dynamical systems. We will start with a motivating example.

Consider the unit circle, $x^2 + y^2 = 1$. This is not a function (of y in terms of x), since, other than the right endpoint $(1,0)$ and left endpoint $(-1,0)$, there are two values of y for every x. That being said, at any point other than these endpoints, we can solve *locally* for y as a function of x. To do this, draw a small box around that point, ignore everything outside of the box, and what is left depicts y as a function of x. In this case, we can obtain an explicit formula for the function in the box. (We will usually not be able to do this!) Near any point on the upper semicircle, for example, we can write $y = \sqrt{1 - x^2}$. Why does this fail at the left and right endpoints? Let's rewrite the unit circle as the zero level set of the function $f(x,y) = x^2 + y^2 - 1$. At $(1,0)$, we have $\frac{\partial f}{\partial y} = 0$ and $\frac{\partial f}{\partial y} = 0$. This means that, near $(1,0)$, the zero set of f resembles a vertical line. At $(1,0)$, the zero set has a vertical tangent. Essentially, we cannot solve for y as as function of x at $(1,0)$ because $\frac{\partial f}{\partial y} = 0$. We will now state and prove the implicit function theorem. The proof uses the Uniform Contraction Mapping Principle.

Theorem 4.4 (Implicit Function Theorem). *Let*

    (1) X, Y, and Z be Banach spaces.
    (2) $U \subset X$, $V \subset Y$ be open sets.
    (3) $F : U \times V \to Z$ be a C^k map, with $k \geq 1$.
    (4) $(x_0, y_0) \in U \times V$ with $F(x_0, y_0) = 0$.
    (5) The partial derivative $D_X F(x_0, y_0) : X \to Z$ be invertible, with bounded inverse.

Then we can solve for x as a function of y near (x_0, y_0). That is, there is a neighborhood $U_0 \times V_0 \subset U \times V$ of (x_0, y_0) and a unique C^k function $f : V_0 \to U_0$ with $f(y_0) = x_0$, such that $F(x,y) = 0$ for $(x,y) \in U_0 \times V_0$ if and only if $x = f(y)$. Furthermore, the derivative of f satisfies

$$Df(y) = -D_X F(f(y), y)^{-1} D_Y F(f(y), y).$$

Proof. We proceed in the following steps.

    (1) First, we define a "Newton map". Since $D_X F(x_0, y_0)$ is invertible, define the map $G : U \times V \to X$ by

$$G(x, y) = x - [D_X F(x_0, y_0)]^{-1} F(x, y).$$

    This looks a lot like the map use use for Newton's method, except the derivative is always evaluated at (x_0, y_0). For fixed $y \in V$, $G(x, y) = x$ if and only if $F(x, y) = 0$. Since F is C^k, G is C^k as well, and

$$D_X G(x, y) = I - [D_X F(x_0, y_0)]^{-1} D_X F(x, y).$$

    In particular, $D_X G(x_0, y_0) = 0$. Our goal is to show that G is a uniform contraction on a neighborhood of (x_0, y_0). In the next two steps, we derive some bounds on $D_X G$ and F. We will use these bounds in step (4).

    (2) First, we control $D_X G(x, y)$ near (x_0, y_0). Since $D_X G(x, y)$ is continuous in (x, y), and $D_X G(x_0, y_0) = 0$, we can find open balls $U_0 = B_\delta(x_0)$ and $V_0 = B_\epsilon(y_0)$ such that for $(x, y) \in \overline{U_0} \times V_0$,

$$\|D_X G(x, y)\| < \frac{1}{2}.$$

    Since

$$[D_X F(x_0, y_0)]^{-1} D_X F(x, y) = I - D_X G(x, y),$$

and $\|D_X G(x,y)\| < \frac{1}{2}$ on $\overline{U_0} \times V_0$, it follows from the Neumann series that $D_X F(x,y)$ is invertible for $(x,y) \in \overline{U_0} \times V_0$.

(3) Next, we control $F(x_0, y)$ near (x_0, y_0). For convenience, let $L = [D_X(x_0, y_0)]^{-1}$. Since the map $y \mapsto F(x_0, y)$ is continuous in y, and $F(x_0, y_0) = 0$, we can (if needed) decrease ϵ (i.e. shrink the ball $V_0 = B_\epsilon(y_0)$) so that, for $y \in V_0$,

$$|F(x_0, y)| \leq \frac{\delta}{2\|L\|}.$$

(4) Next, we show that $G : \overline{U_0} \times V_0 \to \overline{U_0}$. To do this we will use the following result: if $f : X \to Y$ is a continuously differentiable map between Banach spaces, $U \subset X$ is convex, and $\sup_{u \in U} \|Df(u)\| \leq M$, then f is Lipschitz on U with constant M. We proved this result for subsets of $\mathbb{R}^n$; the proof for the more general case is identical. Using this result and the triangle inequality, for $x \in \overline{U_0}$ and $y \in V_0$, we have

$$\begin{aligned}
|G(x,y) - x_0| &\leq |G(x,y) - G(x_0, y)| + |G(x_0, y) - x_0| \\
&\leq \sup_{x \in \overline{U_0}} \|D_X G(x,y)\| \, |x - x_0| + |x_0 - LF(x_0, y) - x_0| \\
&< \frac{1}{2}|x - x_0| + |LF(x_0, y)| \leq \frac{\delta}{2} + \|L\| \frac{\delta}{2\|L\|} \\
&\leq \delta.
\end{aligned}$$

Thus we conclude that $|G(x,y) - x_0| < \delta$ for all $x \in \overline{U_0}$ and $y \in V_0$.

(5) Finally, we show that G is a uniform contraction. For $x_1, x_2 \in \overline{U_0}$ and any $y \in V_0$,

$$|G(x_1, y) - G(x_2, y)| \leq \sup_{x \in \overline{U_0}} \|D_X G(x,y)\| \, |x_1 - x_2| \leq \frac{1}{2}|x_1 - x_2|.$$

(6) By the Uniform Contraction Mapping Principle, there is a unique C^k function $f : V_0 \to \overline{U_0}$ which maps $y \in V_0$ to the unique fixed point x of $G(\cdot, y)$. In other words, $G(f(y), y) = f(y)$ for all $y \in V_0$. For all $y \in V_0$, since $f(y) \in \overline{U_0}$, using the result from step (5),

$$|f(y) - x_0| = |G(f(y), y) - x_0| < \delta$$

Thus $f(y) \in U_0$ (not just $f(y) \in \overline{U_0}$), and so we can restrict the codomain of f to U_0, i.e. we can define $f : V_0 \to U_0$.

(7) Since $F(x,y) = 0$ if and only if x is a fixed point of $G(\cdot, y)$, this implies that
   (a) $f(y_0) = x_0$.
   (b) For $(x,y) \in U_0 \times V_0$, $F(x,y) = 0$ if and only if $x = f(y)$.
   which is the result we want!

(8) To compute the derivative of f, use the chain rule on $F(f(y), y) = 0$ to get

$$0 = D_X F(f(y), y)Df(y) + D_Y F(f(y), y).$$

Since $D_X F(x,y)$ is invertible on $U_0 \times V_0$, we can rearrange this to get

$$Df(x) = -D_X F(f(y), y)^{-1} D_Y F(f(y), y).$$

$\square$

The inverse function theorem is a corollary of the implicit function theorem.

Theorem 4.5 (Inverse Function Theorem). *Let*

   (1) X *and* Y *be Banach spaces.*
   (2) $U \subset X$, $V \subset Y$ *be open sets.*
   (3) $G : U \to V$ *a* C^k *map with* $G(x_0) = y_0$.
   (4) The derivative $DF(x_0)$ *be invertible with bounded inverse.*

Then we can invert G *near* y_0. *That is, there exist neighborhoods* $U_0 \subset U$ *of* x_0 *and* $V_0 \subset V$ *of* y_0 *and a unique* C^k *function* $G^{-1} : V_0 \to U_0$ *such that* $G(G^{-1}(y)) = y$ *for all* $y \in V_0$ *and* $G^{-1}(G(x)) = x$ *for all* $x \in U_0$. *Furthermore, the derivative of* G^{-1} *is given by*

$$DG^{-1}(y) = DG(G^{-1}(y))^{-1}.$$

Proof. Take $F(x, y) = G(x) - y$, and use the implicit function theorem.   □

   As a final application of the uniform contraction mapping principle, we will revisit existence and uniqueness of solutions to ODEs. We will only consider the case where the ODE is written as $\dot{u} = f(u)$, with f independent of t.

Theorem 4.6 (Picard-Lindelöf Existence and Uniqueness Theorem). *Consider the initial value problem (IVP) on* $\mathbb{R}^n$

$$\frac{du}{dt} = f(u)$$
$$u(0) = u_0.$$

Suppose that f *is* **locally Lipschitz**, *i.e. for every* $\tilde{u} \in \mathbb{R}^n$ *there exists a radius* δ *and a Lipschitz constant* L *(both depending on* $\tilde{u}$*) such that for all* $u_1, u_2 \in B(\tilde{u}, \delta)$,

$$|f(u_1) - f(u_2)| \leq L|u_1 - u_2|.$$

Then, for every $\tilde{u} \in \mathbb{R}^n$ *there exists a radius* δ *and a time interval* $[-r, r]$ *such that:*

   (i) For each initial condition $u_0 \in B_\delta(\tilde{u})$, *the initial value problem has a unique solution* $u(t; u_0)$ *on* $[-r, r]$.
   (ii) The map $u_0 \mapsto u(\cdot; u_0)$ *is Lipschitz in* u_0.
   (iii) If f *is* C^k *for* $k \geq 1$, *then*
        (a) The solution $u(t; u_0)$ *is* C^{k+1} *in* t.
        (b) The map $u_0 \mapsto u(\cdot; u_0)$ *is* C^k *in* u_0.

Proof. The steps of the proof are as follows: write the IVP in integrated form, use the integrated form to construct a linear operator between Banach spaces, show that this operator is a uniform contraction, and then apply the Uniform Contraction Mapping Principle.

   (1) First, we reformulate the problem as an integral equation. By the fundamental theorem of calculus, if u is differentiable, then

$$u(t) = u(0) + \int_0^t u'(\tau)d\tau.$$

   Thus $u(t)$ is a solution to the IVP if an only if it is a solution to the integral equation

(1) $$u(t) = u(0) + \int_0^t f(u(\tau))d\tau.$$

   (2) Next, we define some constants. Choose any $\tilde{u} \in \mathbb{R}^n$. Since f is locally Lipschitz, we can find a radius $\delta > 0$ and a Lipschitz constant $L > 0$ such that, for $u_1, u_2 \in B_{2\delta}(\tilde{u})$,

$$|f(u_1) - f(u_2)| \leq L|u_1 - u_2|,$$

Since f is continuous, it is bounded in the closure of this ball, so we define

$$C = \sup\{f(u) : u \in \overline{B_{2\delta}(\tilde{u})}\}.$$

Finally, for convenience, let

$$B = B_\delta(\tilde{u})$$

be the set of initial conditions we will consider.

(3) We now define our Banach spaces. Consider the closed interval $[-r, r]$, where $r > 0$ is small and will be chosen later. We will show that we have a unique solution to the IVP for $t \in [-r, r]$. Let

$$X = C^0([-r, r], \mathbb{R}^n),$$

which is a Banach space equipped with the supremum (maximum) norm. Let $k \in X$ be the constant function $\tilde{u}$, i.e.

$$k(t) = \tilde{u} \text{ for all } t \in [-r, r].$$

Finally, let D be the closed ball in X of radius 2δ about the constant function k, i.e.

$$D = \overline{B_{2\delta}(k(t))} = \{u \in X : \sup_{t \in [-r,r]} |u(t) - \tilde{u}| \leq 2\delta\}.$$

(4) Next, we define our mapping between Banach spaces. Define $F : D \times B \to X$ by

$$[F(u, u_0)](t) = u_0 + \int_0^t f(u(\tau))d\tau,$$

which is the RHS of the integrated form of the initial value problem. As we noted above, a fixed point of F is a solution to the the IVP. Since f is continuous (at minimum), the RHS of F is in X.

(5) Show that for sufficiently small r, $F : D \times B \to D$, i.e. the RHS of F from step (4) is actually in D. For $u \in D$ and $u_0 \in B$,

$$\sup_{|t| \leq r} |F(u, u_0) - \tilde{u}| = \sup_{|t| \leq r} \left| u_0 + \int_0^t f(u(\tau))d\tau - \tilde{u} \right|$$

$$\leq |u_0 - \tilde{u}| + \int_0^r |f(u(\tau))|d\tau.$$

Since we are taking $u \in D$, $u(\tau) \in B_{2\delta}(\tilde{u})$ for all $\tau \in [-r, r]$, and so $|f(u(\tau)| \leq C$ for all $\tau \in [-r, r]$. Thus we have

$$\sup_{|t| \leq r} |F(u, u_0) - \tilde{u}| < \delta + Cr.$$

If we take $r \leq \delta/C$, we have $\sup_{|t| \leq r} |F(u, u_0) - \tilde{u}| < 2\delta$, which is what we want.

(6) Show that for sufficiently small r, F is a contraction. For $u, v \in D$ and $u_0 \in B$,

$$\sup_{|t|\leq r}|F(u, u_0) - F(v, u_0)|$$

$$= \sup_{|t|\leq r} \left| u_0 + \int_0^t f(u(\tau))d\tau - \left(u_0 + \int_0^t f(v(\tau))d\tau \right) \right|$$

$$\leq \sup_{|t|\leq r} \int_0^t |f(u(\tau)) - f(v(\tau))|d\tau \leq Lr \sup_{|t|\leq r} |u(t) - v(t)|$$

$$\leq Lr|u - v|.$$

If we take $r \leq 1/2L$, then F is a contraction. Thus we choose $r = \min\{\delta/C, 1/L\}$, and we are all set!

(7) Show that the map F is uniformly Lipschitz in the initial condition u_0. For $u_0, u_1 \in B$,

$$|F(u, u_0) - F(u, u_1)| = |u_0 - u_1| \leq |u_0 - u_1|,$$

thus F is uniformly Lipschitz in u_0 with Lipschitz constant 1.

(8) Finally, we show that F is C^k (in both variables) whenever f is C^k. Since F is linear in u_0, F is smooth in u_0, thus it remains to show that $F(\cdot, u_0)$ is C^k. First, we will show that $F(\cdot, u_0)$ is C^1 if f is C^1. Let $u(t) \in D$. We make an educated guess that the Fréchet (partial) derivative with respect to u at $u(t)$ is given by L, where L is defined by

$$Lh(t) = \int_0^t Df(u(\tau))h(\tau)d\tau.$$

Where did we get this from? One way to see it is by expanding $F(u(t) + h(t), u_0)$ in a Taylor series and taking the term which is linear in $h(t)$.

By the properties of integration, L is a linear operator, and it is bounded since f is C^1. Using the definition of the Fréchet derivative,

$$\frac{\|F(u+h) - F(u) - L(h)\|}{\|h\|}$$

$$\leq \frac{1}{\|h\|} \sup_{|t|\leq r} \left| \int_0^t [f(u(\tau) + h(\tau)) - f(u(\tau))]d\tau - \int_0^t Df(u(\tau))h(\tau)d\tau \right|$$

$$= \frac{1}{\|h\|} \sup_{|t|\leq r} \left| \int_0^t [f(u(\tau) + h(\tau)) - f(u(\tau)) - Df(u(\tau))h(\tau)]d\tau \right|$$

$$\leq \int_0^r \frac{|f(u(\tau) + h(\tau)) - f(u(\tau)) - Df(u(\tau))h(\tau)|}{|h(\tau)|}d\tau.$$

As $\|h\| \to 0$, $|h(\tau)| \to 0$, thus the RHS of this goes to 0 since f is continuously differentiable, and we are on a closed interval $[0, r]$. We can repeat this to conclude that F is C^k.

(9) We have now satisfied the hypotheses of the Uniform Contraction Mapping Principle. Thus there exists a C^k map $G : B \to D$ which maps each initial condition u_0 to the unique fixed point of $F(\cdot; u_0)$, which is the unique solution $u(t; u_0)$ on $[-r, r]$ to the IVP. Note that this is a local existence result, but the interval of existence $[-r, r]$ is the same for all initial conditions $u_0 \in B$.

(10) Finally, the solution $u(t; u_0)$ is C^{k+1} whenever f is C^k. Since we now know that $u(t)$ is a solution to the IVP, this follows from repeatedly differentiating $\frac{d}{dt}u(t) = f(u(t))$.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Finally, we look at what happens when a solution to an IVP approaches the boundary of the region where the solution exists. The following theorem shows that such a solution blows up at the boundary.

Theorem 4.7 (Blowup at Boundary). *Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is C^1, and $u(t)$ satisfies the initial value problem*

$$\frac{du}{dt} = f(u)$$
$$u(0) = u_0$$

on an interval $[0, T)$, but there is no solution to the IVP on the interval $[0, T + \epsilon)$ for any $\epsilon > 0$. Then u blows up as it approaches the boundary, i.e. $|u(t)| \to \infty$ as $t \to T$.

Proof. We employ a proof by contradiction. The main idea is that if we assume $u(t)$ remains bounded, we can construct a solution to the IVP which exists on a larger time interval, i.e. at time $t > T$.

(1) Suppose this is not true. Then there is some sequence of times $\{t_n\}$ with $t_n \uparrow T$ such that $\{u(t_n)\}$ remains bounded, i.e. there exists a constant K such that $|u(t_n)| \leq K$ for all n.

(2) Since $\{u(t_n)\}$ is a bounded sequence in $\mathbb{R}^n$, by the Bolzano-Weierstrass theorem, it has convergent subsequence. Passing to this subsequence if needed, we may assume that $u(t_n) \to \tilde{u}$.

(3) Now consider the same initial value problem, but this time we take the initial condition $\tilde{u}_0$ to be close to the limit point $\tilde{u}$. By the existence and uniqueness theorem, there is a radius δ and an interval $[-r, r]$ such that for all initial conditions $\tilde{u}_0 \in B_\delta(\tilde{u})$, there is a unique solution $u(t)$ for $t \in [-r, r]$ with $u(0) = \tilde{u}_0$.

(4) Since f does not depend on t, we can translate these unique solutions in time, i.e shift them to different starting times. In other words, for each $\tilde{u}_0 \in B_\delta(\tilde{u})$, there is a family of unique solutions $u(t; \tau)$ on the interval $[\tau - r, \tau + r]$ with $u(\tau; \tau) = \tilde{u}_0$. The solutions in this family are just translates of each other (in time).

(5) Since $t_n \uparrow T$ and $u(t_n) \to \tilde{u}$, choose an integer m sufficiently large so that $t_m > T - r/2$ and $u(t_m) \in B_\delta(\tilde{u})$. Consider the initial value problem

$$\frac{dv}{dt} = f(v)$$
$$v(t^*) = u^*,$$

, where $t^* = t_m$ and $u^* = u(t_m)$. By what we showed in step (4), this IVP has a unique solution $v(t)$ for $t = [t^* - r, t^* + r]$.

(6) By uniqueness (since f is locally Lipschitz), we can splice these solutions together, since we can stop $u(t)$ at (t^*, u^*), and $v(t)$ starts at (t^*, u^*). Thus we have the following solution to the original IVP:

$$w(t) = \begin{cases} u(t) & T \in [0, t^*] \\ v(t) & T \in [t^*, t^* + r]. \end{cases}$$

Since $t^* + r > T$, this solution exists on a larger interval than $[0, T)$, thus we have a contradiction. This implies that our original assumption must be false, i.e. the solution $u(t)$ must blow up as it hits the boundary.

$\square$

The existence result from the Picard-Lindelöf theorem (or the Cauchy-Peano theorem) is only a *local* existence result. If we have a linear system, however, we can obtain a global existence result.

Theorem 4.8 (Global Existence for Linear Systems). *Consider the system*

$$\frac{du}{dt} = A(t)u$$
$$u(0) = u_0,$$

where $u \in \mathbb{R}^n$ and $A : \mathbb{R} \to \mathbb{R}^{n \times n}$ is continuous. Then there exists a unique solution $u(t)$ which exists for all $t \in \mathbb{R}$.

Proof. We proceed as follows:

(1) Write the ODE as $\dot{u} = f(u, t)$ with $f(u, t) = A(t)u$. Since $f(u, t)$ is Lipschitz in u on every bounded interval $[-T, T]$, the initial value problem has a unique solution $u(t)$ for t in an interval containing 0.
(2) We wish to show that $u(t)$ exists for all $t \in \mathbb{R}$. Suppose this solution exists on $[0, t_0)$, but does not exist for any larger interval $[0, t_0 + \epsilon)$. Then $u(t)$ must blow up as it approaches t_0. We will show this cannot happen.
(3) Write the ODE in integrated form as

$$u(t) = u_0 + \int_0^t A(\tau)u(\tau)d\tau \qquad\qquad t \in [0, t_0)$$

Taking absolute values and using the operator norm (matrix norm) of $A(t)$, this becomes

$$|u(t)| = |u_0| + \int_0^t \|A(\tau)\|\,|u(\tau)|d\tau \qquad\qquad t \in [0, t_0)$$

(4) This satisfies the hypotheses of Gronwall's Inequality. Thus, for $t \in [0, t_0)$

$$|u(t)| \leq |u_0| \exp\left(\int_s^t \|A(\tau)\|d\tau\right) \leq |u_0| \exp\left(\int_s^{t_0} \|A(\tau)\|d\tau\right) \leq C|u_0|,$$

where the bound holds uniformly for $t \in [0, t_0)$ since it does not depend on t.
(5) We conclude that $u(t)$ cannot blow up as $t \to t_0$, thus it follows that $u(t)$ must exist for all $t \geq 0$. A similar argument shows that $u(t)$ must exist for all $t \leq 0$.

$\square$

5. Integration

5.1. **Riemann Integral.** We start with a review of the Riemann integral. In particular, we will discuss its strengths and weaknesses. Let $f : [a, b] \to \mathbb{R}$ be a bounded function. (It is important that the domain is a closed, bounded interval, and that the function is bounded). The Riemann integral is defined as follows.

(1) First, partition the domain $[a, b]$. Let P be the partition

$$P : a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b,$$

where n is arbitrary.

(2) Define the upper and lower sums on P by

$$U_P(f) = \sum_{j=0}^{n-1} (x_{j+1} - x_j) \sup_{x \in [x_j, x_{j+1}]} f(x)$$

$$L_P(f) = \sum_{j=0}^{n-1} (x_{j+1} - x_j) \inf_{x \in [x_j, x_{j+1}]} f(x).$$

The supremum and infimum are both well-defined since f is bounded.

(3) The function f is Riemann integrable on $[a, b]$ if

$$\inf_P U_P(f) = \sup_P L_P(f).$$

If this holds, we denote the Riemann integral by $\int_a^b f(x)dx$.

(4) A convenient integrability criterion is the following: f is Riemann integrable on $[a, b]$ if, for every $\epsilon > 0$, we can find a partition P such that

$$U_P(f) - L_P(f) < \epsilon.$$

Technically, what we just defined is known as the Riemann-Darboux integral, which is equivalent to the Riemann integral, and is often more convenient for analysis. The standard Riemann integral involves the Riemann sums you learned in calculus. We can define that as follows.

(1) Choose a tagged partition (P, t) of $[a, b]$. This is a partition P, given by

$$P : a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b,$$

together with n points $\{t_0, t_1, \ldots, t_{n-1}\}$, with one selected from each subinterval of the partition, i.e. $t_j \in [x_j, x_{j+1}]$. Common choices for the tags t_i are the left endpoints, the right endpoint, and the midpoints of the partition intervals, although the choice of tags does not matter from a theoretical standpoint. The mesh size of the partition is the maximum length of the partition subintervals, i.e.

$$\Delta P = \max_{j=0,\ldots,n-1} (x_{j+1} - x_j).$$

Typically, each partition subinterval is chosen to be the same size, although this need not be the case.

(2) The Riemann sum corresponding to (P, t) is

$$R_{P,t}(f) = \sum_{j=0}^{n-1} (x_{j+1} - x_j) f(t_j).$$

(3) The function f is Riemann integrable with integral S if, for all $\epsilon > 0$, there exists $\delta > 0$ such that

$$|R_{P,t}(f) - S| < \epsilon$$

for all tagged partitions (P, t) with mesh size $\Delta P < \delta$.

The main advantages of the Riemann integral are as follows:

- You can compute them exactly with the fundamental theorem of calculus (as long as you can find an antiderivative!)
- The definition is intuitive, and captures the idea of finding the area under a curve by successive approximation.
- The approximating Riemanns sums are easy to compute numerically.
- Many useful classes of functions are Riemann integrable:
  (1) Continuous functions on $[a, b]$.
  (2) Bounded functions on $[a, b]$ which are continuous except at a finite number of points.
  (3) Bounded, monotonic functions on $[a, b]$.

There are, however, several disadvantages to the Riemann integral.

- It is difficult to extend to domains that are not open subsets of $\mathbb{R}^n$.
- It is difficult to extend to unbounded domains, such as all of $\mathbb{R}$. As in calculus class, you can define an "improper integral" as the limit of integrals on bounded intervals, although the best way to do this is not always clear.
- It can be difficult to figure out which functions are and are not Riemann integrable. Consider the following two examples.
  (1) Since the rational numbers are countable, enumerate the set of rational numbers in $[0, 1]$ as the sequence $\{r_k\}$, i.e.

  $$\mathbb{Q} \cap [0, 1] = \{r_k\}_{k \in \mathbb{N}}.$$

  Then define the function

  $$f(x) = \sum_{k=1}^{\infty} \frac{1}{k^2} H(x - r_k),$$

  where H is the Heaviside step function

  $$H = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0. \end{cases}$$

  The function $f(x)$ is well-defined, since the sum $\sum_{k=1}^{\infty} \frac{1}{k^2}$ is convergent. How can we picture this function? Imagine you are taking a walk from $x = 0$ to $x = 1$ on the number line, and you are keeping a "running total" as you do this. (Maybe this should be called a "walking total"?) The total starts at 0 at $x = 0$, and every time you pass a rational number r_k, you add $1/k^2$ to the total. Although $f(x)$ is discontinuous at every rational number, $f(x)$ is Riemann integrable on $[0, 1]$, since it is bounded and monotonic (increasing).
  (2) Let $\chi_{\mathbb{Q}}(x)$ be the characteristic function of the set of rational numbers on $[0, 1]$, defined by

  $$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q}. \end{cases}$$

  The function $\chi_{\mathbb{Q}}(x)$ is also bounded and discontinuous at every rational number. However, since every partition interval contains both rational and irrational numbers (both of these sets are dense in $[0, 1]$), $\chi_{\mathbb{Q}}(x)$ is not Riemann integrable on $[0, 1]$.

- The limit of a sequence of Riemann integrable functions is not necessarily Riemann integrable. Once again, let $\mathbb{Q} \cap [0,1] = \{r_k\}_{k \in \mathbb{N}}$, and define the sequence of functions $f_n : [0,1] \to \mathbb{R}$ by

$$f_n(x) = \begin{cases} 1 & x \in r_1, \ldots, r_n \\ 0 & \text{otherwise.} \end{cases}$$

Since f_n has only a finite number of discontinuities, f_n is Riemann integrable with

$$\int_0^1 f_n(x)dx = 0.$$

For every $x \in [0,1]$, $f_n(x) \to \chi_{\mathbb{Q}}(x)$, but $\chi_{\mathbb{Q}}(x)$ is not Riemann integrable on $[0,1]$.

A final disadvantage (and perhaps the most important one), is that it is hard to find good criteria that allow us to exchange limits and integration. We would like to find conditions for which

$$\lim_{n \to \infty} \int_a^b f_n(x)dx = \int_a^b \lim_{n \to \infty} f_n(x)dx,$$

i.e. the limit of the integrals is the integral of the limit. The best we can do, in general, is if we have a uniformly convergent sequence of functions on a bounded interval.

Theorem 5.1. *For all $n \in \mathbb{N}$, let $f_n : [a,b] \to \mathbb{R}$ be a Riemann integrable functions, and suppose the sequence of functions $\{f_n\}$ converges uniformly to f. Then f is Riemann integrable, and*

$$\lim_{n \to \infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx.$$

Proof. Let $\epsilon_n = \sup_{x \in [a,b]} |f_n(x) - f(x)|$, so that $f_n(x) - \epsilon_n \leq f(x) \leq f_n(x) + \epsilon_n$ on $[a,b]$. Since $\{f_n\}$ converges to f uniformly, $\epsilon_n \to 0$. Then we have

$$\int_a^b f_n(x)dx - (b-a)\epsilon_n = \int_a^b (f_n(x) - \epsilon_n)dx = \sup_P L_P(f_n(x) - \epsilon_n)$$

$$= \sup_P \sum_{j=0}^{n-1} (x_{j+1} - x_j) \inf_{x \in [x_j, x_{j+1}]} (f(x) - \epsilon_n)$$

$$\leq \sup_P \sum_{j=0}^{n-1} (x_{j+1} - x_j) \inf_{x \in [x_j, x_{j+1}]} f(x)$$

$$= \sup_P L_P(f),$$

where P is an arbitrary partition of $[a,b]$. Similarly,

$$\inf_P U_p(f) \leq \int_a^b f_n(x)dx - (b-a)\epsilon_n.$$

Putting these together

$$\int_a^b f_n(x)dx - (b-a)\epsilon_n \leq \sup_P L_p(f) \leq \inf_P U_p(f) \leq \int_a^b f_n(x)dx - (b-a)\epsilon_n.$$

This rearranges to

$$0 \leq \inf_P U_P(f) - \sup_P L_P(f) \leq 2(b-a)\epsilon_n \to 0,$$

from which it follows that f is Riemann integrable on $[a, b]$.   □

The main issue is that uniform convergence is, in a sense, too strong condition. As an example, consider the sequence of functions $f_n(x) = x^n$ on $[0, 1]$. For the limit, $f_n(x) \to f(x)$, where

$$f(x) = \begin{cases} 0 & x \in [0, 1) \\ 1 & x = 1, \end{cases}$$

but this convergence is not uniform. (Either we can show that directly, or use the fact that if the convergence were uniform, the limit would be continuous by the uniform limit theorem, which it is not in this case). However, we still have

$$\lim_{n \to \infty} \int_0^1 f_n(x)dx = \lim_{n \to \infty} \frac{1}{1 + n} = 0 = \int_0^1 f(x)dx.$$

Even though the convergence is not uniform, the limit of the integrals is equal to the integral of the limit. Something else must be going on, and we would like to have a theory which explains this case.

5.2. **Lebesgue Integral.** The Lebesgue integral is a complete different way of defining the integral. For this discussion (and for comparison purposes), we will only consider nonnegative, bounded, real-valued functions defined on a closed interval $[a, b]$, although we will see that the theory is more general than this. The secret sauce here is that instead of partitioning the domain, we will partition the range.

Assume $f : [a, b] \to \mathbb{R}$ is bounded and nonnegative, and let $M = \sup_{x \in [a,b]} f(x)$. For now, we define the Lebesgue integral as follows. (Note that this is an intuitive construction; ultimately this is not the definition we will use).

(1) Let P be a partition the range of f, i.e.

$$P : 0 = y_0 < y_1 < y_2 < \cdots < y_{n-1} < y_n = M$$

(2) For $j = 0, dots, n - 1$, define the set I_j by

$$I_j = f^{-1}([y_{j+1}, \infty)) = \{x : f(x) \ge y_{j+1}\}.$$

This set is all values of x for which $f(x) \ge y_{j+1}$.

(3) Define the sum

$$I_P(f) = \sum_{j=0}^{n} (y_{j+1} - y_j)m(I_j)$$

where $m(I_j)$ is the "measure" of the set I_j. We have not defined this yet (this is what measure theory is all about!), but think of it as the "length" of I_j. We can think of this as constructing layers of a cake. The first layer has height $y_1 - y_0$. Since we want this layer to include all values of x corresponding to values of $f(x)$ which are y_1 or more, the total "length" of this layer is $m(I_0)$; we note that this layer may consist of many disconnected pieces. The subsequent layers are similarly constructed, and each layer stacks neatly on top of the previous one.

(4) Then we define the Lebesgue integral as

$$\int_a^b f(x)dx = \sup_P I_P(f),$$

if the supremum exists.

The Lebesgue integral resolves most of the weaknesses of Riemann integral. The main drawback is that, in most cases you cannot actually compute an integral using the Lebesgue formulation, and you have to fall back on the fundamental theorem of calculus from Riemann integration theory. Luckily, if a function is Riemann integrable, it is also Lebesgue integrable, and the two integrals are the same! In fact, we can use Lebesgue integration theory to tell us what are the "worst" functions which are still Riemann integrable. In order to use the Lebesgue integral, we need to understand the "measure" term $m(I_j)$ in the definition of $I_p(f)$. The field of measure theory is devoted to exactly this.

6. Measure Theory

What is measure theory, and why should we care? Consider the following two sets: $[0, 1]$ and $\mathbb{Q} \cap [0, 1]$. There are (at least) two different notions of the "size" of these sets.

(1) "Number of elements" (cardinality). Although they both contain an infinite number of elements, set theorists would say that $[0, 1]$ is uncountable and $\mathbb{Q} \cap [0, 1]$ is countable.
(2) "Length" (measure). Intuitively, the interval $[0, 1]$ has a length of 1. If the empty set has a length of 0 (which makes sense), then the "length" of $\mathbb{Q} \cap [0, 1]$ (whatever that means) should be somewhere between 0 and 1, although where in the middle is not clear.

The central goal of measure theory is to generalize our intuitive ideas of length, area, and volume to subsets of arbitrary sets. This will be applied to familiar settings like the line $\mathbb{R}$ and the plane $\mathbb{R}^2$, and we will see that our new theory not only encompasses our commonsense notions of length and area but also allows us to "measure" subsets which we would not be able to do with a ruler. Here are some applications.

(1) Probability. Take any sample spaces you want, and declare that the measure of the entire sample space to be 1. Then the measure of any subset of the sample space is the probability that an event in that subset occurs. Probabilists put measures on all sorts of spaces, including function spaces. As an example, if the sample space is $C([0, 1])$, we could ask ourselves what is the measure of the subset of differentiable functions.
(2) Integration theory. Measure theory lets us define the Lebesgue integral, which allows us to integrate more functions, and gives us more general theorems for when we can exchange limits and integration.

6.1. **Measures.** With this motivation in mind, let us define a measure on an arbitrary set X.

Definition 6.1. *A* measure *on a set X is a function μ on subsets of X with the following properties:*

(i) $\mu(E) \in [0, \infty]$ (∞ is allowed!)
(ii) $\mu(\varnothing) = 0$.
(iii) If the sets $\{E_k\}_{k \in \mathbb{N}}$ are disjoint subsets of X, then

$$\mu\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mu(E_k).$$

   This last property is called countable additivity

Other properties we might like μ to have are the following

(iv) μ is invariant under symmetries such as translations, rotations, and reflections.

(v) μ agrees with our common notion of length, area, and volume. For example, we would like $\mu([0,1]) = 1$.

Notice that I intentionally left the domain of μ vague. We would like the domain of μ to be as large as possible. Ideally, we would like it to be all subsets of X. However, as you will see in the beginning of your analysis course, it is impossible to construct a measure on all subsets of $\mathbb{R}$ that has these five properties. The good news is that the construction involves a "nasty" subset of $\mathbb{R}$ (and requires invoking the axiom of choice), so if we restrict ourselves to a collection of "nice" subsets, we should be able to make this work. Our goal is to find a collection of subsets of on which we can define a measure which is as large as possible and contains everything we care about. The collection we will need is called a σ-algebra.

Definition 6.2. *A collection $\mathcal{M}$ of subsets of X is a σ-**algebra** if*

(i) $\varnothing \in \mathcal{M}$.

(ii) $\mathcal{M}$ is closed under complement, i.e. if $E \in \mathcal{M}$ then $X \setminus E \in \mathcal{M}$.

(iii) $\mathcal{M}$ is closed under countable unions, i.e. if $\{E_k\}_{k \in \mathbb{N}} \subset \mathcal{M}$, then $\bigcup_{k=1}^{\infty} E_k \in \mathcal{M}$.

*An **algebra** is the same thing, except it is only closed under finite unions.*

If we want to define a measure on a $\sigma-$algebra, the fact that it is closed under countable unions is crucial for the countable additivity property of a measure to make sense. Other properties of $\sigma-$algebras which follow from the definition and basic set theory include

(iv) $X \in \mathcal{M}$.

(v) $\mathcal{M}$ is closed under countable intersection. This follows from De Morgan's laws.

(vi) $\mathcal{M}$ is closed under relative complement, i.e. if $E, F \in \mathcal{M}$, then $E \setminus F \in \mathcal{M}$.

Here are some simple examples of σ-algebras on X.

(1) $M = \{\varnothing, X\}$ (the smallest σ-algebra).

(2) $M = \{A, A^c, \varnothing, X\}$ (the σ-algebra generated by the subset A).

(3) $M = \{\text{all subsets of X}\}$ (the largest σ-algebra, also called the power set of X).

The first two are not very interesting. The power set of X is typically too large to define a measure on. An exception is if X is a countable or finite set. In this case, let X be a countable set, and let μ be a measure. As long as we know $\mu(\{x\})$ for all single-point sets $\{x\}$, then, for any subset S of X,

$$\mu(S) = \sum_{x \in S} \mu(\{x\}),$$

where the sum is countable or finite (and can be infinite). The first σ-algebra which is of interest is the σ-algebra generated by a collection of subsets.

Definition 6.3. *Let $\mathcal{E}$ be a collection of subsets of X. Then the σ-**algebra generated by** $\mathcal{E}$, denoted $\mathcal{M}(\mathcal{E})$, is the unique smallest σ-algebra containing $\mathcal{E}$. We can define $\mathcal{M}(\mathcal{E})$ in one of two equivalent ways.*

(i) $\mathcal{M}(\mathcal{E})$ is the intersection of all $\sigma-$algebras containing $\mathcal{E}$.

(ii) If $\mathcal{N}$ is another σ-algebra containing $\mathcal{E}$, then $\mathcal{M}(\mathcal{E}) \subset \mathcal{N}$.

For (i), we can show that the intersection of σ-algebras is a $\sigma-$algebra. It is worth noting that $\sigma-$algebras are huge and unwieldy. There is essentially no way of writing down an arbitrary element of $\mathcal{M}(\mathcal{E})$ in terms of the elements of $\mathcal{E}$. The most important of these σ-algebras is the Borel σ-algebra, which is generated by the open sets.

Definition 6.4. *The **Borel σ-algebra** on X, denoted $\mathcal{B}_X$, is the σ-algebra generated by the collection of open sets of X.*

The Borel $\sigma-$algebra contains all open sets, all closed sets, all countable unions and countable intersections of open and closed sets (some of which may be neither open nor closed), etc. There is no nice way to write down a generic element of $\mathcal{B}_X$. For the special case of the Borel $\sigma-$algebra on $\mathbb{R}$, we have a nice result.

Proposition 6.1. *The Borel $\sigma-$algebra on $\mathbb{R}$ is generated by any of the following:*
  (1) All open intervals, i.e. all intervals of the form (a, b).
  (2) All intervals of one of the following forms: $[a, b], (a, b], [a, b)$.
  (3) All rays of one of the following forms $(a, \infty), [a, \infty), (-\infty, b), (-\infty, b]$.

Proof. Let $\mathcal{E}$ be one of these collections. Since $\mathcal{E} \in \mathcal{B}_\mathbb{R}$, $\mathcal{M}(\mathcal{E}) \in \mathcal{B}_\mathbb{R}$. All we have do now is show that $\mathcal{M}(\mathcal{E})$ contains all open sets. For (i), we showed earlier that every open set in $\mathbb{R}$ is the countable union of disjoint open intervals. For (ii) and (iii), we need to show we can construct an open interval (a, b) using countably many set operations. For example, for the closed intervals in (ii),

$$(a, b) = \bigcap_{n=1}^{\infty} \left[a - \frac{1}{n}, b + \frac{1}{n} \right].$$

$\square$

Now that we have discussed σ-algebras, we will return to measures, which, after all, is the whole point of all of this. From now on, let X be a set and μ a measure defined on a σ-algebra $\mathcal{M}$. This set of three related objects is sometimes written as $(X, \mu, \mathcal{M})$ or $(X, \mathcal{M}, \mu)$, and is called a measure space. For now, we assume that μ is in fact defined on all of $\mathcal{M}$. In applications, we will generally construct μ and $\mathcal{M}$ at the same time, so that everyone plays together nicely. A measure μ has many nice properties, which we summarize in the following proposition. You will prove these in your analysis class.

Proposition 6.2. *Let X be a set, and μ a measure defined on a σ-algebra $\mathcal{M}$. Then μ has the following properties.*
  (1) (Monotonicity) If $E \subset F$, $\mu(E) \leq \mu(F)$. In addition, if $\mu(E) < \infty$, then $\mu(F \setminus E) = \mu(F) - \mu(E)$.
  (2) (Countable Subadditivity) For any sequence of sets $\{E_n\}$ (not necessarily disjoint),

$$\mu\left(\bigcup_{n=1}^{\infty} E_n \right) \leq \sum_{n=1}^{\infty} \mu(E_n).$$

  (3) (Continuity from Below) If $\{E_n\}$ is an increasing sequence of nested sets, i.e. $E_n \subset E_{n+1}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n \right) = \lim_{n \to \infty} \mu(E_n).$$

  (4) (Continuity from Above) If $\{E_n\}$ is an decreasing sequence of nested sets, i.e. $E_n \supset E_{n+1}$, and $\mu(E_1) < \infty$, then

$$\mu\left(\bigcap_{n=1}^{\infty} E_n \right) = \lim_{n \to \infty} \mu(E_n).$$

With that taken care of, it is time to construct our first measure. We will construct the Lebesgue measure, often denoted m, which measures length of subsets of $\mathbb{R}$. It gives the "correct" measure for intervals, but will in addition allow us to measure many other sets. The contruction we will use is one I find intuitive. You will probably do this a little differently in analysis class. That version has the advantage of letting you construct a whole class of measures for which the Lebesgue measure is a special case; however, it requires more definitions and theorems, and is one more step removed from what is actually going on. We will do our construction in the following steps. Proofs will be either outlined, or omitted entirely.

(1) Construct something which is roughly what we want, but which we can use on all subsets of $\mathbb{R}$. This is called the Lebesgue outer measure, denoted m^*.
(2) Show that m^* gives the correct length for intervals.
(3) Find a σ-algebra $\mathcal{L}$ on which m^* is actually a measure. This is the Lebesgue σ-algebra.
(4) Show that this σ-algebra contains:
  (a) All subsets E of X with $m^*(E) = 0$. These are called null sets.
  (b) The Borel $\sigma-$algebra.

First, we define the Lebesgue outer measure. Consider any subset E of $\mathbb{R}$. We can find always find a countable collection of open intervals $\{(a_n, b_n)\}$ such that

$$E \subset \bigcup_{n=1}^{\infty} (a_n, b_n).$$

In addition, it makes intuitive sense that we should have

$$\text{length of } E \leq \sum_{n=1}^{\infty} (b_n - a_n),$$

since E is contained in the union of these intervals. With this in mind, we define the Lebesgue outer measure as follows.

Definition 6.5. *For any subset E of $\mathbb{R}$, the **Lebesgue outer measure** is the function m^*, taking values in $[0, \infty]$ (∞ is allowed!) defined by*

$$m^*(E) = \inf \left\{ \sum_{n=1}^{\infty} (b_n - a_n) : E \subset \bigcup_{n=1}^{\infty} (a_n, b_n) \right\}.$$

Essentially, we are "approximating E from the outside" by open intervals, which is why we call this outer measure. The idea is that we cover E with a countable collection of open intervals and add up their lengths; the outer measure is the infimum over all such covers (which may still be infinite!) Formally, an outer measure is defined as follows.

Definition 6.6. *An **outer measure** on a set X is a function μ^* defined on* all *subsets of X with the following properties:*

(i) $\mu^*(E) \in [0, \infty]$ (∞ is allowed!)
(ii) $\mu(\varnothing) = 0$.
(iii) (Monotonicity) If $E \subset F$ then $\mu^*(E) \leq \mu^*(F)$
(iv) (Countable subadditivity) $\mu\left(\bigcup_{k=1}^{\infty} E_k\right) \leq \sum_{k=1}^{\infty} \mu(E_k)$.

It is important to note that the outer measure μ^* is defined on all subsets of X. Going back to the Lebesgue outer measure, we can show the following.

(1) m^* is an outer measure
(2) The m^* outer measure of any interval (open, closed, half-open) is equal to its length.

Next, we show that the Lebesgue outer measure of any countable set is 0.

Proposition 6.3. *Let $E \subset \mathbb{R}$ be a countable set. Then $m^*(E) = 0$.*

Proof. Since E is countable, we can write it as the sequence $\{x_n\}$. We now use the "countable epsilon trick". Let $\epsilon > 0$. Cover each x_n with the open interval $I_n = (x_n - r_n, x_n + r_n)$, where

$$r_n = \frac{1}{2}\frac{\epsilon}{2^n}$$

By construction, each interval I_n has length $2r^n = \epsilon/2^n$. Since the outer measure is the infimum of the total length of all such covers,

$$m^*(E) \le \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon,$$

where we used the infinite sum of the geometric series. Since ϵ is arbitrary, $m^*(E) = 0$. $\square$

In particular, this implies that $m^*(\mathbb{Q}) = 0$ (which is kind of crazy, if you think about it!) Next, for any outer measure μ^*, we define the concept of μ^*-measurability. Inituitively, a set A is μ^*-measurable if it "behaves nicely" by splitting any arbitrary "test set" E into two "nice pieces". The formal definition is as follows, and is due to Carathéodory.

Definition 6.7. *Let μ^* be an outer measure on X. Then a subset $A \subset X$ is μ^*-measurable if for every test set $E \subset X$,*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

A set A splits any set E into a part which overlaps A and a part which does not. A is "nice" if, whenever we do this, the outer measure of E is the sum of the outer measures of these two pieces. You will sometimes see the criterion written

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A).$$

To complete the construction of the Lebesgue measure, we will use the Carathéodory Extension Theorem, which we will not prove here but will be proved (hopefully!) in your analysis class.

Theorem 6.1 (Carathéodory Extension Theorem). *Let μ^* be an outer measure on X, and let $\mathcal{M}$ be the collection of μ^*-measurable sets. Then*

(1) $\mathcal{M}$ is a σ-algebra.
(2) μ^*, when restricted to $\mathcal{M}$, is a measure, which we designate μ.
(3) $\mathcal{M}$ contains all μ^*-null sets, i.e. all sets N with $\mu^*(N) = 0$.

Proof. Here is an outline of the proof, which you may find useful. The basic tool you use is the definition of μ^* measurability, which is used over and over. Here is the order in which you show things.

(1) Show $\mathcal{M}$ is closed under complement and finite unions, thus is an algebra. (It suffices to show if $A, B \in \mathcal{M}$ then $A \cup B \in \mathcal{M}$).
(2) Show μ^* is finitely additive on $\mathcal{M}$. (It suffices to show if A and B are disjoint sets in $\mathcal{M}$, then $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$).

(3) Show $\mathcal{M}$ is closed under countable disjoint unions. This implies it is closed under countable unions, and thus is a σ-algebra. (To see why this works, take any sequence of sets $\{E_n\}$. Construct F_n from E_n by deleting anything which overlaps with $E_1, \ldots, E_{n-1}$. We can get from E_n to F_n, or vice versa, by using a finite number of set operations. In addition, by construction, $\cup E_n = \cup F_n$).
(4) Show μ^* is countably additive on $\mathcal{M}$, thus is a measure on $\mathcal{M}$.
(5) Show if $\mu^*(N) = 0$ then $N \in \mathcal{M}$.

$\square$

The last property is particularly important. If a set has measure 0, we can, in general, ignore it. (You could argue that the entire point of measure theory is figuring out which sets are null sets so you can ignore them!) This property guarantees that if we construct a measure this way, all μ^*-null sets are actually in the σ-algebra.

We are almost done. First, we use the Carathéadory Extension Theorem on the Lebesgue outer measure m^*. Let $\mathcal{L}$ be the resulting $\sigma-$algebra of m^*-measurable sets, and let m be the measure we obtain by restricting m^* to $\mathcal{L}$. (We use a different letter for the Lebesgue σ-algebra, since the Lebesgue measure is the one of the most important measures, and we want to distinguish it from generic σ-algebras). All that is left is to show that $\mathcal{L}$ contains the entire Borel $\sigma-$algebra. To do this, it suffices to show that $\mathcal{L}$ contains any collecton of sets which generates the $\sigma-$algebra. Since $\mathcal{L}$ is a $\sigma-$algebra, it must then contain the entire Borel $\sigma-$algebra. The easiest way to do this is to show that all open intervals are μ^*-measurable (although we will not actually do this here). We can also show that the Lebesgue measure m is translation-invariant.

6.2. **Measurable Functions and Integration.** One of the advantages of the Lebesgue formulation of the integral is that we can use it on a large class of functions, as long as we have defined a measure. In order to do make this work, we need to define a measurable function between two measure spaces.

Definition 6.8. *Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be sets together with $\sigma-$algebras. Then $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-**measurable** (or just **measurable**) if, for all $F \in \mathcal{N}$, $f^{-1}(F) \in \mathcal{M}$.*

This is analogous to our definition of continuous functions. If you define a "measurable" set to be a set in the appropriate σ-algebra, a measurable function is one where preimages of measurable sets are measurable. Note that although both of these spaces likely have measures associated with their σ-algebras, the measures themselves do not figure into this definition. In addition, we note that as long as the $\sigma-$algebras of the sets involved "line up" appropriately, compositions of measurable functions are measurable.

Although the definition of measurability can be applied to functions between any two arbitrary measure spaces, we will only consider real-valued functions from now on. Let $f : (X, \mathcal{M}) \to \mathbb{R}$ be a real-valued function. Unless otherwise specified, we always use the Borel σ-algebra on the codomain $\mathbb{R}$. Thus we say that $f : (X, \mathcal{M}) \to \mathbb{R}$ is $\mathcal{M}$-measurable (or just measurable) if it is $(\mathcal{M}, \mathcal{B})$-measurable. For functions on $\mathbb{R}$, you will see the following terms.

(1) $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable if f is $(\mathcal{B}_{\mathbb{R}}, \mathcal{B}_{\mathbb{R}})$-measurable.
(2) $f : \mathbb{R} \to \mathbb{R}$ is Lebesgue measurable if f is $(\mathcal{L}, \mathcal{B}_{\mathbb{R}})$-measurable, where $\mathcal{L}$ is the Lebesgue σ-algebra on $\mathbb{R}$.

Since $\mathcal{B}_{\mathbb{R}} \subset \mathcal{L}$, every Borel measurable function is Lebesgue measurable, but the converse is not true. The main issue with the definition of measurability is that it is hard to verify, since we don't have a good way to characterize what sets belong to a given σ-algebra. Luckily, it suffices to check the measurability criterion on a set which generates a σ-algebra

Proposition 6.4. *Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measure spaces, and suppose $\mathcal{N}$ is the σ-algebra generated by the collection of sets $\mathcal{E}$. Then $f : X \to Y$ is measurable if and only if $f^{-1}(E) \subset \mathcal{M}$ for all $E \in \mathcal{E}$.*

Proof. An outline of the proof is as follows. ($\implies$) This follows from the definition of measurability, since $\mathcal{E} \subset \mathcal{N}$.
($\impliedby$) Define the set

$$\mathcal{H} = \{A \in \mathcal{N} : f^{-1}(A) \in \mathcal{M}\}.$$

We want to show that $\mathcal{H}$ contains $\mathcal{N}$. From our initial assumption, $\mathcal{H}$ contains $\mathcal{E}$. Next, we show that $\mathcal{H}$ is a σ-algebra, by using the definition of a σ-algebra and the fact the the inverse image operator f^{-1} commutes with set operations. Since $\mathcal{H}$ is a σ-algebra containing $\mathcal{E}$, it must also contain $\mathcal{N}$, since $\mathcal{N}$ is the σ-algebra generated by $\mathcal{E}$. $\qquad\square$

We have the following important corollary.

Corollary 6.1. *Every continuous function $f : (X, \mathcal{B}_X) \to (Y, \mathcal{B}_Y)$ is measurable.*

We can also define measurability in terms of the generators of the Borel σ-algebra on $\mathbb{R}$. The infinite rays are more useful than the intervals in this case.

Corollary 6.2. *A function $(X, \mathcal{M}) \to \mathbb{R}$ is measurable if and only if one of the following is true.*

    (1) $f^{-1}((a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.
    (2) $f^{-1}([a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.
    (3) $f^{-1}((-\infty, a)) \in \mathcal{M}$ for all $a \in \mathbb{R}$.
    (4) $f^{-1}((-\infty, a]) \in \mathcal{M}$ for all $a \in \mathbb{R}$.

I find the fourth one to be the most useful. If we call the set $f^{-1}(a)$ the level set of a, then the set

$$f^{-1}((-\infty, a]) = \{x : f(x) \leq a\}$$

can be called the sublevel set of a. Thus, for measurability of real-valued functions, we only have to check the sublevel sets.

Finally, the set of measurable functions is nice, in that pretty much however we slice and dice them, we wind up with a measurable function.

Proposition 6.5. *Let f, g and $\{f_n\}_{n \in \mathbb{N}}$ be measurable functions from $(X, \mathcal{M})$ to $\mathbb{R}$. Then the following are measurable.*

    (1) $f + g$ and fg.
    (2) $\max(f, g)$ and $\min(f, g)$.
    (3) $\sup_n f_n(x)$ and $\inf_n f_n(x)$.
    (4) $\limsup_{n \to \infty} f_n(x)$ and $\liminf_{n \to \infty} f_n(x)$.
    (5) $\lim_{n \to \infty} f_n(x)$, provided the limit exists.

The limit superior and limit inferior of a real-valued sequence $\{x_n\}$ are defined by

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty}\left(\sup_{k\geq n} x_k\right) \qquad \limsup_{n\to\infty} x_n = \lim_{n\to\infty}\left(\inf_{k\geq n} x_k\right).$$

We can think of the limit superior as the limit of the sequence $\{y_n\}$, given by

$$y_n = \sup_{k\geq n} x_k.$$

This is a decreasing sequence, since we are taking the supremum over fewer and fewer terms. The limit superior of $\{x_n\}$ is the limit of this sequence, which is the infimum since the sequence is decreasing. Similarly, the limit inferior is the limit of an increasing sequence. As long as we allow the values $\pm\infty$, $\limsup_{n\to\infty} x_n$ and $\liminf_{n\to\infty} x_n$ exist for all real-valued sequences, and

$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n.$$

The limit of a sequence $\lim_{n\to\infty} x_n$ exists if and only if $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n$.

6.3. **Lebesgue Integral.** From now on, we will assume all functions are measurable. This is not an unreasonable assumption, because measurable functions are nice, and we like nice things! (In general, a function which is merely measurable is the "least nice" function we are willing to deal with). Probabilists often deal with nonmeasurable functions, but the rest of us can safely use the heuristic that "every reasonable function is measurable".

We will now define the Lebesgue integral of real-valued functions on an arbitrary measure space $(X, \mathcal{M}, \mu)$. This is done in three steps. These same three steps recur in many contexts.

(1) Define the integral for nonnegative, simple functions.
(2) Extend this to nonnegative functions (by taking the supremum).
(3) Extend this to all real-valued functions (by splitting into positive and negative parts).

First, we define a simple function.

Definition 6.9. *A **simple function** is a function whose range is a finite set. Let $\phi : (X, \mathcal{M}) \to \mathbb{R}$ be simple. Then we can write ϕ in standard form as*

$$\phi(x) = \sum_{k=1}^{n} y_k \chi_{E_k}(x),$$

*where the sets $E_k \in \mathcal{M}$ are disjoint, and $\chi_{E_k}(x)$ is the **characteristic function***

$$\chi_{E_k}(x) = \begin{cases} 1 & x \in E_k \\ 0 & x \notin E_k \end{cases}$$

The range of ϕ is the finite set $\{y_1, \ldots, y_n\}$, and we can also see that $E_k = f^{-1}(\{y_k\})$.

We can show that the characteristic function $\chi_{E_k}(x)$ is a measurable function if and only if $E_k \in \mathcal{M}$ (using the sublevel sets criterion). Thus the simple function $\phi(x)$ is measurable, since it is the sum of measurable functions. We the define the Lebesgue integral of a nonnegative, simple function as follows.

Definition 6.10. *Let $\phi : (X, \mathcal{M}, \mu) \to \mathbb{R}$ be a simple function written in standard form as*

$$\phi(x) = \sum_{k=1}^{n} y_k \chi_{E_k}(x),$$

with $y_k \geq 0$. Then the integral of ϕ is defined by

$$\int_X \phi d\mu = \sum_{k=1}^{n} y_k \mu(E_k),$$

where we always take $0 \cdot \infty = 0$. We note that this integral can take a value of ∞. We sometimes write this as $\int \phi$ for shorthand.

We can show that this definition is well-defined, i.e. if we have two different representations of the same simple function, the integral is the same. We can also integrate over a set $A \in \mathcal{M}$ instead of the whole space X (this is analogous to integrating over a bounded interval $[a, b]$ instead of all of $\mathbb{R}$).

Definition 6.11. *Let $A \in \mathcal{M}$ and $\phi : (X, \mathcal{M}, \mu) \to \mathbb{R}$ a simple function written in standard form. Then we define the integral over the set A as*

$$\int_A \phi d\mu = \int_X \phi \, \chi_A d\mu = \sum_{k=1}^{n} y_k \mu(E_k \cap A).$$

This integral has the following nice properties.

Proposition 6.6. *Let ϕ and ψ be simple functions $(X, \mathcal{M}, \mu) \to \mathbb{R}$. Then*

(1) $\int (\phi + \psi) = \int \phi + \int \psi$.
(2) If $c \geq 0$, then $\int c\phi = c \int \phi$.
(3) If $0 \leq \phi \leq \psi$, then $\int \phi \leq \int \psi$.
(4) Define the real-valued function ρ on $\mathcal{M}$ by

$$\rho(E) = \int_E \phi d\mu.$$

Then ρ is a measure on $\mathcal{M}$.

The first three are familiar from calculus. The last one may seem a little strange, but is really useful for proving things (especially the monotone convergence theorem). It also gives us another way of constructing a measure. Now that we have defined the integral for simple functions, the next step is to extend it to nonnegative functions. First, we show that we can approximate nonnegative real-valued functions with simple functions.

Lemma 6.1 (Simple Approximation Lemma). *Let $f : (X, \mathcal{M}) \to \mathbb{R}$ be measurable with $f \geq 0$. Then there is an increasing sequence of nonnegative simple functions $\{\phi_n(x)\}$, i.e. $0 \leq \phi_1 \leq \phi_2 \leq \cdots \leq f$ such that $\phi_n \to f$ pointwise. This convergence is uniform if f is bounded.*

Proof. The rough idea is as follows. If the range of f is bounded, partition the range into finitely many points and make "cake layers", refining the partition mesh at each step. If the range is unbounded, for each n, partition $[0, n]$ using a finer mesh at each step. $\square$

We then define the integral of a nonnegative function in the following way.

Definition 6.12. *Let $f : (X, \mathcal{M}, \mu) \to \mathbb{R}$ be a measurable, nonnegative function. Then we define the integral of f by*

$$\int_X f d\mu = \sup \left\{ \int_X \phi d\mu : \phi \text{ simple}, 0 \leq \phi \leq f \right\}.$$

*Note that this can be infinite. We say that a nonnegative function f is **integrable** if this integral is finite.*

The same properties for the integral of simple functions hold for the integral of nonnegative functions. Finally, we define the integral of real-valued functions. First, define the positive and negative parts of a function f as

$$f^+ = \max\{f, 0\} \qquad\qquad f^- = \max\{-f, 0\}.$$

Note that $f = f^+ - f^-$ and $|f| = f^+ + f^-$. We then have the following definition

Definition 6.13. *Let $f : (X, \mathcal{M}, \mu) \to \mathbb{R}$ be a measurable function. Then we say f is **integrable** if $|f|$ is integrable, i.e. $\int_X |f| d\mu < \infty$. In that case, we define the integral of f by*

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu.$$

The integral has all of the familiar properties, which we summarize in the following proposition.

Proposition 6.7. *Let $f, g : (X, \mathcal{M}, \mu) \to \mathbb{R}$ be integrable. Then the integral has the following properties*

(1) *Linearity:* $\int (cf + g) = c \int f + \int g$.
(2) *Comparison: If $f \leq g$, then $\int f \leq \int g$.*
(3) *Monotonicity: If $f \geq 0$ and $A \subset B$, then $\int_A f \leq \int_B f$, where $\int_A f = \int f \chi_A$.*
(4) $|\int f| \leq \int |f|$.

Before we conclude this section, we note the special role played by sets of measure 0, which are known as null sets. We say that a particular property holds "almost everywhere" (abbreviated "a.e.") if it holds everywhere except a set of measure zero. Examples include: a function which is continuous almost everywhere; a function which is 0 almost everywhere; a function which is bounded almost everywhere; two functions which are equal almost everywhere; and a sequence of functions which converges almost everywhere. The following proposition gives us a way to characterize functions which are 0 almost everywhere.

Proposition 6.8. *Let $f : X \to \mathbb{R}$ be a nonnegative function. Then $f = 0$ almost everywhere if and only if $\int f = 0$.*

Proof.
($\implies$) We first prove this for a simple function ϕ. Write ϕ in standard form as

$$\phi(x) = \sum_{k=1}^n y_k \chi_{E_k}(x).$$

If $\phi = 0$ almost everywhere, either $y_k = 0$ or $\mu(E_k) = 0$ (or both!) for all k. Using the definition of the integral of a simple function, $\int \phi = 0$. Now take any nonnegative function f with $f = 0$ almost everywhere. For any simple function ϕ with $0 \leq \phi \leq f$, $\phi = 0$ almost everywhere as well. Thus, by the definition of the integral of nonnegative functions,

$$\int f = \sup \left\{ \int \phi \text{ simple with } 0 \leq \phi \leq f \right\} = 0.$$

($\Longleftarrow$) Here we prove the contrapositive, i.e. we will show that if f is not 0 almost everywhere, the integral of f must be positive. Let $E = \{x \in X : f(x) > 0\}$ be the set on which f is nonzero. We then employ the following useful decomposition

$$E = \bigcup_{n=1}^{\infty} E_n \qquad E_n = \left\{x \in X : f(x) > \frac{1}{n}\right\}.$$

If f is not 0 almost everywhere, then one of the sets E_n has positive measure, i.e. $\mu(E_n) = r > 0$ for some n. But then we have

$$f \geq \frac{1}{n}\chi_{E_n},$$

from which it follows that

$$\int f \geq \int \frac{1}{n}\chi_{E_n} = \frac{1}{n}\mu(E)n = \frac{r}{n} > 0.$$

$\square$

Similarly, for a real-valued function $f : X \to \mathbb{R}$, we have $f = 0$ almost everywhere if and only if $\int |f| = 0$.

Recall that a function $f : X \to \mathbb{R}$ is integrable if $\int |f| d\mu < \infty$. From the integration properties, the set of integrable, real-valued functions is closed under addition and scalar multiplication, thus is a vector space. We call this vector space L^1 or $L^1(X)$ (if we wish to indicate the domain of the functions in the space). We would like to define a natural norm on this vector space by

$$\|f\| = \int_X |f| d\mu.$$

We call this the L^1 norm, which is sometimes notated $\|f\|_{L^1}$ Since we are integrating $|f|$, this is nonnegative, and is finite since we are restricting ourselves to integrable functions. The triangle inequality follows from the triangle inequality on $\mathbb{R}$, together with the linearity of the integral. Similarly, we can pull constants out by the linearity of the integral. If $f = 0$, then $\int |f| d\mu = 0$. However, $\|f\| = 0$ does not imply $f = 0$, only the weaker result that $f = 0$ almost everywhere. Thus this is not a norm on the vector space of integrable functions.

It seems like we are in a bit of a bind here. However, we can "cheat" our way out of it. If two functions f and g are equal almost everywhere, then they are "essentially the same". They may differ on a null set, but for most purposes we don't care about that. In this particular case, if f and g are integrable and are equal almost everywhere $\int f d\mu = \int g d\mu$. Thus, we shift our discussion from functions to equivalence classes of functions, where we define the equivalence relation $f \sim g$ if $f = g$ almost everywhere. If we consider L_1 to be the vector space of these equivalence classes of integrable functions, the L^1 norm becomes a true norm on this space. Since we can essentially ignore null sets, most of the time we can ignore this equivalence class mumbo jumbo as well. When we write $f \in L^1$, what we are saying is that f is an integrable function which is uniquely defined up to null sets, which is good enough for most purposes.

6.4. **Convergence Theorems.** Finally, we present the three big convergence theorems, which are perhaps the main advantage of the Lebesgue integration theory. You will prove them in the analysis course. The first result is the Monotone Convergence Theorem, which states that limits and integration can be exchanged for increasing sequences of nonnegative functions.

Theorem 6.2 (Monotone Convergence Theorem (MCT)). *Let $f_n : (X, \mathcal{M}, \mu) \to \mathbb{R}$ measurable with $f_n \geq 0$. If $f_n \uparrow f$ pointwise, then*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X \left(\lim_{n\to\infty} f_n\right) d\mu = \int_X f d\mu.$$

where this can be infinite.

The next result is Fatou's Lemma. While it is mainly used to prove the Dominated Convergence Theorem, it is a useful result in its own right.

Theorem 6.3 (Fatou's Lemma). *Let $f_n : (X, \mathcal{M}, \mu) \to \mathbb{R}$ measurable with $f_n \geq 0$. Then*

$$\int \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int f_n.$$

Finally, the Dominated Convergence Theorem gives a criterion for when limits and integration can be exchanged for a sequence of measurable functions (which do not have to be nonnegative).

Theorem 6.4 (Lebesgue Dominated Convergence Theorem (DCT)). *Let $f_n : (X, \mathcal{M}, \mu) \to \mathbb{R}$ measurable with $f_n(x) \to f(x)$ pointwise. If there exists an integrable function g such that $|f_n| \leq g$ for all n, then f is integrable and*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X \left(\lim_{n\to\infty} f_n\right) d\mu = \int_X f d\mu.$$

We can use this theorem to explain what occurs with the sequence of functions $f_n(x) = x^n$ on $[0, 1]$. Recall that $f_n(x) \to f(x)$ pointwise, where

$$f(x) = \begin{cases} 0 & x \in [0, 1) \\ 1 & x = 1, \end{cases}$$

but this convergence is not uniform. Let $g(x) = 1$, i.e. the constant function at 1. Then g is integrable on $[0, 1]$, since $\int_0^1 g(x) dx = 1$, and $|f_n(x)| \leq g(x)$ on $[0, 1]$ for all n. Since the conditions of the Dominated Convergence Theorem are satisfied, we conclude that $\lim_{n\to\infty} \int_0^1 f_n(x) dx = \int_0^1 f(x) dx$, which explains the convergence result we obtained above!