Analysis Notes

Rebekah Dix

March 29, 2019

Contents

1 Sequence and Series of Functions

1.1 Uniform Convergence

Definition 1.1 (Pointwise Convergence)**.** A sequence of functions $(f_n)_n$ **converges pointwise** to a function f on E if for every $\varepsilon > 0$ and for all $x \in E$ there is an integer N (which depends on x) such that for all $n \geq N$

$$|f_n(x) - f(x)| < \varepsilon$$

Definition 1.2 (Uniform Convergence)**.** A sequence of functions $(f_n)_n$ **converges uniformly** to a function f on E if for every $\varepsilon > 0$ there is an integer N such that for all $n \geq N$ and for all $x \in E$ we have

$$|f_n(x) - f(x)| < \varepsilon$$

Theorem 1.1 (Limit of uniformly convergent, bounded sequence of functions is bounded.)**.** If $(f_n)_n$ converges uniformly to f and each f_n is bounded, then f is bounded.

Proof. Fix $\varepsilon > 0$. Uniform convergence implies that $\exists N \in \mathbb{N}$ such that $\forall n \geq N$ and $\forall x \in X$ we have $|f_n(x) - f(x)| < \varepsilon$. Further, boundedness of the sequence of functions implies that $\forall n$ $\exists M_n \in [0, \infty)$ such that $|f_n(x)| \leq M_n$ $\forall x \in X$. Fix an $n \geq N$. Then $\forall x \in X$

$$\begin{aligned} |f(x)| &= |f(x) - f_n(x) + f_n(x)| \\ &\leq |f(x) - f_n(x)| + f_n(x) && \text{(trianlge inequality)} \\ &< \varepsilon + M_n && \text{(uniform convergence and boundedness)} \end{aligned}$$

Thus $|f(x)| < \varepsilon + M_n$ $\forall x \in X$, so $f(x)$ is bounded. □

Theorem 1.2 (Limit of uniformly convergent, continuous sequence of functions is continuous.)**.** If $(f_n)_n$ converges uniformly to f and each f_n is continuous, then f is continuous.

Proof. Fix $\varepsilon > 0$. Uniform convergence implies that $\exists N \in \mathbb{N}$ such that $\forall n \geq N$ and $\forall x \in X$ we have $|f_n(x) - f(x)| < \frac{\varepsilon}{3}$. Continuity of each function $f_n(x)$ at $x \in X$ in the sequence implies that $\exists \delta > 0$ such that $\forall y$ for which $|x - y| < \delta$ we have that $|f_n(x) - f_n(y)| < \frac{\varepsilon}{3}$. Then $\forall y$ for which $|x - y| < \delta$ we have that

$$\begin{aligned} |f(x) - f(y)| &= |f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y)| \\ &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| && (\Delta) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} && \text{(uniform convergence and continuity of } f_n \text{ at } x) \\ &= \varepsilon \end{aligned}$$

Therefore f is continuous at x, $\forall x \in X$. □

Definition 1.3 (Uniformly cauchy). A sequence of functions (f_n) is **uniformly cauchy** if $\forall \varepsilon > 0\ \exists N \in \mathbb{N}$ such that $\forall n, m \geq N$ and $\forall x \in X$, we have $|f_n(x) - f_m(y)| < \varepsilon$.

Theorem 1.3 (Uniform convergence iff uniform cauchy). A sequence $(f_n)_n$ of functions on a metric space X converges uniformly if and only if it is uniformly Cauchy.

Proof. Fix $\varepsilon > 0$.

$\Rightarrow$ Uniform convergence implies $\exists N$ such that $\forall n \geq N$ and $\forall x \in X$ we have that $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$. Then $\forall n, m \geq N$ and $\forall x \in X$ we have that

$$\begin{aligned} |f_n(x) - f_m(x)| &= |f_n(x) - f(x) + f(x) - f_m(x)| \\ &\leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Therefore $(f_n)_n$ is uniformly cauchy.

$\Leftarrow$ Notice that $(f_n)_n$ is a cauchy sequence in a complete metric space ($\mathbb{R}$), therefore it converges pointwise. We need to show uniform convergence. [[Incomplete]] □

Theorem 1.4 (Weierstrass M-test). Let $(f_n)_n$ be a sequence of functions on a metric space X such that there exists a sequence of non-negative real numbers $(M_n)_n$ such that $\forall n$ we have that

$$|f_n(x)| \leq M_n \tag{1.1}$$

If $\sum_{n=1}^{\infty} M_n$ converges, then the series $\sum_{n=1}^{\infty} f_n$ converges uniformly. That is, the sequence of partial sums $(\sum_{n=1}^{m} f_n)_m$ converges uniformly.

Proof. We will show that the sequence of partial sums is uniformly cauchy, which then implies that sequence of partial sums converges uniformly. Fix $\varepsilon > 0$. The Cauchy criterion (adapted for series: sequences of partial sums) implies that since $\sum_{n=1}^{\infty} M_n$ converges, there exists an N such that $\forall m, n\ m \geq n \geq N$ we have that

$$\left| \sum_{k=n}^{m} M_k \right| = \sum_{k=n}^{m} M_k < \varepsilon$$

Therefore $\forall m, n\ m \geq n \geq N$ and $\forall x \in X$ we have that

$$\begin{aligned} \left| \sum_{k=n}^{m} f_k(x) \right| &\leq \sum_{k=n}^{m} |f_k(x)| \qquad (\Delta) \\ &\leq \sum_{k=n}^{m} M_k \\ &\leq \varepsilon \end{aligned}$$

Therefore $(\sum_{n=1}^{m} f_n)_m$ is uniformly cauchy, so the series converges uniformly. □

1.2 Power Series

Definition 1.4 (Power Series)**.** A **power series** is a function of the form

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \tag{1.2}$$

where $c_n \in \mathbb{C}$ are complex coefficients.

Definition 1.5 (Radius of convergence)**.** To a power series we can associate a number $R \in [0, \infty]$ (thus R is an *extended* real number) called its **radius of convergence** such that

(i) $\sum_{n=0}^{\infty} c_n x^n$ converges for every $|x| < R$.

(ii) $\sum_{n=0}^{\infty} c_n x^n$ diverges for every $|x| > R$.

Theorem 1.5 (Power series continuous on interval of convergence)**.** A power series with radius of convergence R converges uniformly on $[-R+\varepsilon, R-\varepsilon]$ for every $0 < \varepsilon < R$, Therefore, power series are continuous on $(-R, R)$.

Theorem 1.6 (Abel summation (summation by parts))**.**

$$\sum_{n=0}^{N} (a_n - a_{n-1}) b_n = a_N b_N + \sum_{n=0}^{N-1} a_n (b_n - b_{n+1}) \tag{1.3}$$

Proof. Assume that $a_{-1} = 0$. We can derive this formula by reordering terms:

$$\begin{aligned}
\sum_{n=0}^{N} (a_n - a_{n-1}) b_n &= a_0 b_0 + a_1 b_1 - a_0 b_1 + a_2 b_2 - a_1 b_2 + \cdots + a_N b_N - a_{N-1} b_N \\
&= a_0(b_0 - b_1) + a_1(b_1 - b_2) + a_{N-1}(b_{N-1} - b_N) + a_N b_N \\
&= a_N b_N + \sum_{n=0}^{N-1} a_n (b_n - b_{n+1})
\end{aligned}$$

□

Theorem 1.7 (Abel)**.** Let $f(x) = \sum_{n=0}^{\infty} c_n x^n$ be a power series with radius of convergence $R = 1$. Assume $\sum_{n=0}^{\infty} c_n$ converges. Then

$$\lim_{x \to 1^-} f(x) = \sum_{n=0}^{\infty} c_n \tag{1.4}$$

Proof. We use summation by parts. Set $s_n = \sum_{k=0}^{n} c_k$ and by convention assume $s_{-1} = 0$. □

2 Compactness in Metric Spaces

2.1 Review of Basic Topology

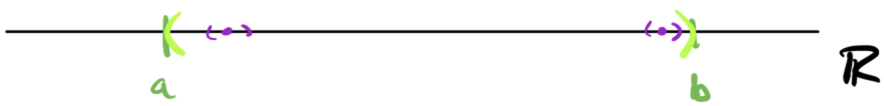
Definition 2.1 (Open, open relative to)**.** Let $E \subset U \subset X$, where X is a metric space.

(i) E is an **open** subset of X if for each point $p \in E$ there exists an $r > 0$ such that for all $q \in X$ for which $d(p, q) < r$ we have that $q \in E$.

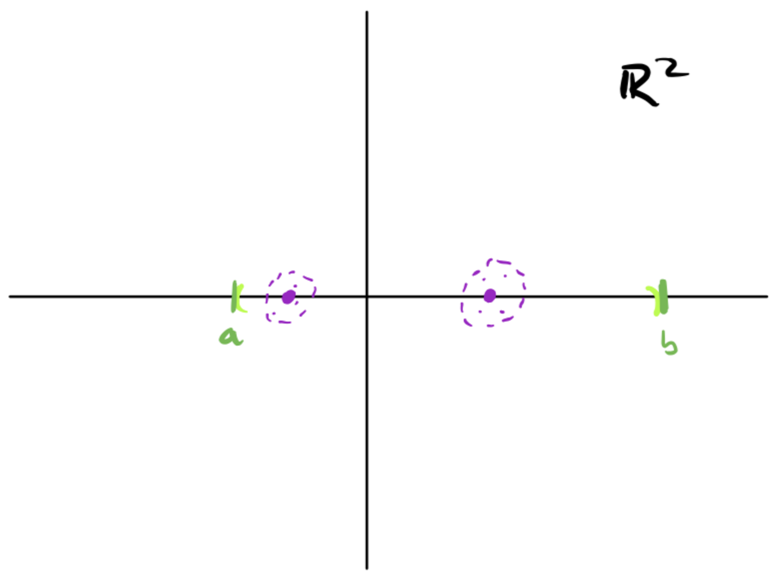
(ii) E is **open relative to** Y if for each point $p \in E$ there exists an $r > 0$ such that for all $q \in Y$ for which $d(p, q) < r$ we have that $q \in E$.

Example 2.1 (Relative open sets)**.** Let (a, b) be an interval on the real line. Notice that $(a, b) \subset \mathbb{R} \subset \mathbb{R}^2$.

(i) (a, b) is an open subset of (or open relative to) $\mathbb{R}$.



(ii) (a, b) is *not* an open subset of $\mathbb{R}^2$. Indeed, any ball around a point $x \in (a, b) \subset \mathbb{R}$ will leave the x-axis and intersect with points in the second dimension. Thus no point of (a, b) is interior relative to $\mathbb{R}^2$.



Theorem 2.1 (Relative open sets)**.** Let $Y \subset X$. A subset E of Y is open relative to Y if and only if $E = Y \cap G$ for some open subset G of X.

Proof. [[Todo]] □

Example 2.2 (Relative open sets)**.** Let $X = \mathbb{R}$ and $A = [0, 1]$. Observe that $B = [0, \frac{1}{2}) \subset A \subset X$ is open in A, but not open in X. However, there exists a C open in X such that $B = C \cap A$. One example is $C = (-\frac{1}{2}, \frac{1}{2})$.

Definition 2.2 (Continuous)**.** Suppose X and Y are metric spaces, $E \subset X$, $p \in E$, and f maps E into Y. Then f is **continuous** at p if for every $\varepsilon > 0$ there there exists a $\delta > 0$ such that for all $x \in E$ for which $d_X(x, p) < \delta$, we have that $d_Y(f(x), f(p)) < \varepsilon$.

Definition 2.3 (Uniformly Continuous)**.** Let f be a mapping of a metric space X into a metric space Y. We say that f is **uniformly continuous** on X if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $p, q \in X$ for which $d_X(p, q) < \delta$, we have $d_Y(f(p), f(q)) < \varepsilon$.

Definition 2.4 (Dense)**.** TFAE: E is **dense** in X if

(i) Every point of X is a limit point of E or a point of E (or both).

(ii) $\bar{E} = X$.

(iii) $\forall \varepsilon > 0$ and $\forall x \in X$ we have that $B(x, \varepsilon) \cap E \neq \emptyset$.

Theorem 2.2 (Functions, inverses, and subsets)**.** Let $f : X \to Y$

(i) $E \subset Y \Rightarrow f(f^{-1}(E)) \subset E$

(ii) $E \subset X \Rightarrow f(f^{-1}(E)) \supset E$

2.2 Basic Definitions

Definition 2.5 (Open Cover)**.** Let I be an arbitrary index set. A collection $(G_i)_{i \in I}$ of open sets $G_i \subset X$ is called an **open cover** of X if $X \subset \cup_{i \in I} G_i$.

Definition 2.6 (Compact)**.** X is **compact** if every open cover of X contains a finite subcover. More explicitly, for every open cover $(G_i)_{i \in I}$ there exists $m \in \mathbb{N}$ and $i_1, i_2, \ldots, i_m \in I$ such that $X \subset \cup_{j=1}^{m} G_{i_j}$.

Remark. This is also called the Heine-Borel property.

Definition 2.7 (Compact subset). A subset $A \subset X$ is called **compact subset** if $(A, d|_{A\times A})$ is a compact metric space. $d|_{A\times A}$ is the restriction of d to $A \times A$.

Theorem 2.3 (Heine-Borel). A subset $A \subset \mathbb{R}^n$ is compact if and only if A is closed and bounded.

Definition 2.8 (Relatively compact or precompact). A subset $A \subset X$ is called **relatively compact** or **precompact** if the closure $\bar{A} \subset X$ is compact.

Example 2.3 (Any finite metric space is compact). Let $X = \{x_i\}_{i=1}^n$ be a finite metric space. Let J be an arbitrary index set, and let $(G_j)_{j\in J}$ be an open cover of X. Therefore, each x_i must be in at least one G_j. Suppose that $x_i \in G_{j_i}$, $j_i \in J$. Then $\cup_{i=1}^n G_{j_i}$ is a finite subcover.

Example 2.4 ($K = \{0\} \cup \{1/n\}_{n=1}^\infty \subset \mathbb{R}$ is compact.). Let J be an arbitrary index set, and let $(G_j)_{j\in J}$ be an open cover of K. First note that 0 must be contained in some element of the open cover; call it G_i. Since G_i is open, each element of G_i is an interior element, so there exists a ball around 0 of radius $\varepsilon > 0$ contained in G_i. The ball also contains all the elements of K for which $n > \frac{1}{\varepsilon}$. Then, each for each of the finitely many $n \leq \frac{1}{\varepsilon}$ there exists a G_{j_n} that contains $\frac{1}{n}$. Therefore every open cover has a finite subcover.

Example 2.5 (Compactness and relative compactness in $\mathbb{R}$). Any closed and bounded interval $[a, b]$ in $\mathbb{R}$ is compact. Half open intervals $[a, b)$, $(a, b]$ and open intervals (a, b) in $\mathbb{R}$ are relatively compact, since their closures are closed and bounded intervals (assuming b finite).

Example 2.6 (Unit circle in $\mathbb{R}^n$ compact). The set $C = \left\{x \in \mathbb{R}^n \,\middle|\, \sum_{i=1}^n |x_i|^2 = 1\right\} \subset \mathbb{R}^n$ is compact. To show closedness, consider the map $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(\mathbf{x}) = \sum_{i=1}^n x_i^2$, $\mathbf{x} \in \mathbb{R}^n$. This map is continuous (since it is the sum of continuous functions). Then $C = f^{-1}(\{1\})$, and the singleton $\{1\}$ is a closed set in $\mathbb{R}$. To show boundedness, note that no single element of a vector can have an absolute value greater than one. Therefore $C \subset [-1, 1]^n$. C is closed and bounded, and by Heine-Borel, compact.

Example 2.7 (Orthogonal matrices). The set of orthogonal $n \times n$ matrices with real entries (call this $O(n, \mathbb{R})$) is compact as a subset in $\mathbb{R}^2$. To see this, let M_n be the set of all $n \times n$ matrices with real entries. Define a function $f : M_n \to M_n$ where $f(A) = A^T A$. This mapping is continuous. To see this, first note that $f : A \to A$ is the identity map and hence continuous. Also $f : A \to A^T$ is continuous: this follows since $\|A\| = \|A^t\|$, so

$$\begin{aligned} \|f(A) - f(B)\| &= \|B^t - A^t\| \\ &= \|(B - A)^t\| \\ &= \|B - A\| \end{aligned}$$

Thus whenever $\|B - A\| < \varepsilon$, we have $\|f(A) - f(B)\| < \varepsilon$ (hence we set $\delta = \varepsilon$). The product of two continuous functions is continuous.

Since an orthogonal matrix O has inverse O^T, we have that $f^{-1}(I) = O(n, \mathbb{R})$. The continuity of f implies that $O(n, \mathbb{R})$ is closed since it is the preimage of a singleton (which is a closed set).

To show boundedness, recall that the columns of orthogonal matrices are orthonormal. Therefore, the absolute value of an element cannot be larger than 1 in absolute value. Thus $O(n, \mathbb{R}) \subset [-1, 1]^{n^2}$. Then by Heine-Borel, $O(n, \mathbb{R})$ is compact.

Theorem 2.4 (Compact subsets of metric spaces are closed.).

Proof. Let X be a metric space and K a compact subset of X. To show K is closed, we will show that its complement K^c is open. To do this, we must show that for all $p \in K^c$, there exists a neighborhood of p completely contained in K^c (and hence does *not* intersect K).

Let $q \in K$. We will construct two types of neighborhoods:

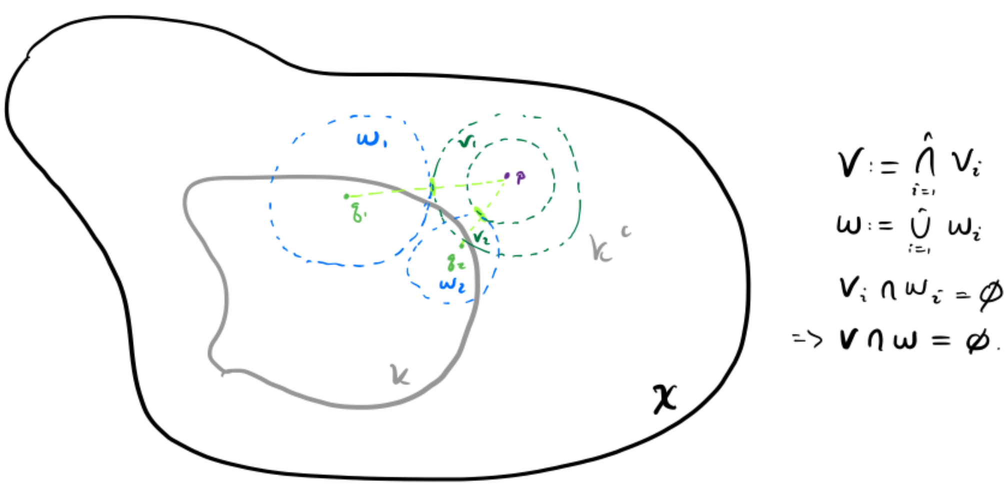
$$W_q = \left\{ x \in X \,\middle|\, d(x, q) < \frac{1}{2} d(p, q) \right\} \qquad \text{(neighborhood of } q\text{)}$$

$$V_q = \left\{ x \in X \,\middle|\, d(x, p) < \frac{1}{2} d(p, q) \right\} \qquad \text{(neighborhood of } p\text{)}$$

Notice that the union of all W_q forms an open cover of K. Since K is compact, this open cover must have a finite subcover, $W = \bigcup_{i=1}^n W_{q_i}$. Define $V = \bigcap_{i=1}^n V_i$, which is still a neighborhood of p. By construction (since we have used open balls), $V_{q_i} \cap W_{q_i} = \emptyset$. Since $V \subseteq V_{q_i}$, $V \cap W_{q_i} \subseteq V_{q_i} \cap W_{q_i} = \emptyset$. Therefore

$$V \cap W = (V \cap W_1) \cup \cdots \cup (V \cap W_n) = \emptyset$$

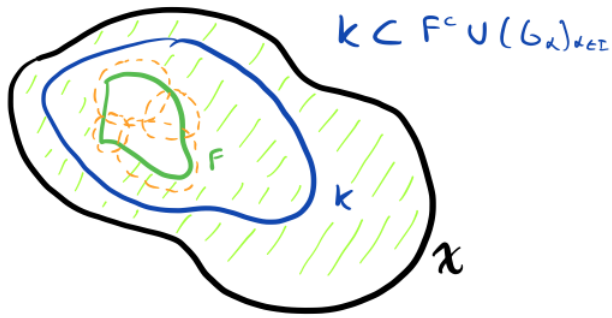
Since $K \subseteq W$, we have that $V \subseteq K = \emptyset$. Thus V is a neighborhood of p completely contained in K^c. Since p was arbitrary, K^c is open, and K is closed.



$\square$

Theorem 2.5 (Closed subsets of compact sets are compact.)**.**

Proof with compactness. Let $F \subset K \subset X$, where F is closed (relative to X) and K is compact. Let $(G_\alpha)_{\alpha\in I}$ be an open cover of F. Since F is closed, F^c is open and $F^c \cup (G_\alpha)_{\alpha\in I}$ is an open cover of K. K is compact, so every open cover has a finite subcover: $K \subset F^c \cup (G_{\alpha_i})_{i=1}^n$. Since $F \subset K$, this finite open cover also covers F, but clearly we don't need F^c in the cover. Therefore F has a finite subcover: $F \subset (G_{\alpha_i})_{i=1}^n$.



□

Proof with sequential compactness. Let $F \subset K \subset X$, where F is closed (relative to X) and K is compact. Let (x_n) be a sequence of points in F (so is also in X). Since X is sequentially compact, there is a subsequence (x_{n_k}) converging to some point $x \in X$. Since F is closed, $x \in X$. Therefore F is sequentially compact since every sequence in F has a convergent subsequence. □

Theorem 2.6 (The intersection of a nested sequence of compact nonempty sets is compact and nonempty.)**.**

Proof. Let (A_n) be a sequence of compact nonempty sets. **Compact:** We know that each A_n is closed. The intersection of closed sets is closed, so $\cap A_n$ is closed. Notice that $\cap A_n \subset A_1$ is a closed subset of a compact set, so is also compact.

Nonempty: Since each A_n is nonempty, fix an $a_n \in A_n$. The sequence $(a_n) \subset A_1$. Sequential compactness of A_1 implies that there is a subsequence (a_{n_k}) converging to some point $a \in A_1$. Notice that this limit must also be in A_2, since the sequence (a_{n_k}) is also in A_2 (except for potentially the first term, which does not affect convergence). This holds for all A_n, so $p \in \cap A_n$, which shows the intersection is nonempty. □

Theorem 2.7. Let X be a compact metric space. Then there exists a countable, dense set $E \subset X$.

Proof. We will show that $\forall \varepsilon > 0$ and $\forall x \in X$ we can find a a set E such that for each open ball in X we have that $B(x, \varepsilon) \cap E \neq \emptyset$. Fix $\varepsilon > 0$ and define an open cover of X as follows:

$$B_n = \bigcup_{x\in X} B(x, \frac{1}{n}) \tag{2.1}$$

X is compact so the open cover B_n must have a finite subcover. Let E be the union of the centers of the balls of each finite subcover. E is the countable union of finite sets, so it is countable. Now fix $x \in X$. Choose n such that $\varepsilon < \frac{1}{n}$. Since B_n covers X, there must be some ball centered at a point of E, call it y, that contains x. Thus $d(x,y) < \frac{1}{n} < \varepsilon$. Thus $y \in B(x,\varepsilon) \cap E$. □

2.3 Compactness and Continuity

Theorem 2.8 (Continuous mappings on compact sets are uniformly continuous.)**.** Let X, Y be metric spaces and assume X is compact. If $f : X \to Y$ is continuous, then it is uniformly continuous.

Proof. Fix $\varepsilon > 0$. **Goal:** We need to find a $\delta > 0$ such that for all $x, y \in X$ for which $d_X(x,y) < \delta$, we have $d_Y(f(x), f(y)) < \varepsilon$.

Use continuity to create balls of points (which cover) X and are mapped by the function: Since f is continuous, we know that for each $x \in X$, there exists a number $\delta_x > 0$ such that for all $y \in X$ for which $d_X(x,y) < \delta_x$, we have that $d_Y(f(x), f(y)) < \varepsilon/2$. Now, let B_x be a ball of radius $\delta_x/2$ centered at x. Formally, we can write

$$B_x = B(x, \delta_x/2) = \{y \in X | d_X(x,y) < \delta_x/2\}$$

Then $(B_x)_{x \in X}$ is an open cover of X.

Use compactness to select a finite subcover of these balls and use the smallest radius to create uniformity: By compactness, there exists a finite subcover of X. Let $x_1, \ldots, x_m$ be the x which generate this finite subcover. Define $\delta > 0$ as follows

$$\delta = \frac{1}{2} \min(\delta_{x_1}, \ldots, \delta_{x_m}) \tag{2.2}$$

(notice that δ is indeed positive, since we are taking the minimum of finitely many positive numbers).

Show that our choice of δ works: Fix $x, y \in X$ such that $d_X(x,y) < \delta$. There exists and index $i \in \{1, \ldots, m\}$ such that $x \in B_{x_i}$ (i.e., an element of X is in some ball of the finite subcover). Then

$$d_X(x_i, y) \leq d_X(x_i, x) + d_X(x, y) < \frac{1}{2}\delta_{x_i} + \delta < \delta_{x_i} \tag{2.3}$$

Now, using the definition of δ_{x_i}, we have that

$$d_Y(f(x), f(y)) \leq d_Y(f(x), f(x_i)) + d_Y(f(x_i), f(y)) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon \tag{2.4}$$

since $d_X(x_i, x) < \delta_x/2 < \delta_x$ so continuity implies $d_Y(f(x), f(x_i)) < \varepsilon/2$, and $d_X(x_i, y) < \delta_{x_i}$ so continuity implies $d_Y(f(x_i), f(y)) < \varepsilon/2$. □

Theorem 2.9 (The image of a continuous function which maps from a compact set is compact)**.** Let X, Y be metric spaces and assume X is compact. If $f : X \to Y$ is continuous, then $f(X) \subset Y$ is compact.

Proof with compactness. Let $(V_i)_{i\in I}$ be an open cover of $f(X)$. Since f is continuous, we have that $U_i = f^{-1}(V_i) \subset X$ is open for each i. Note that $U \subset f^{-1}(f(U))$ for every $U \subset X$. Therefore

$$X \subset f^{-1}(f(X)) \subset \cup_{i\in I} f^{-1}(V_i) = \cup_{i\in I} U_i \tag{2.5}$$

which shows that $\cup_{i\in I} U_i$ is an open cover of X (the second subset relation uses that applying f^{-1} preserves unions). Since X is compact, there are finitely many indices, say up to m, such

$$X \subset \bigcup_{j=1}^{m} U_{i_j} \tag{2.6}$$

Applying f preserves inclusions and $V \supset f(f^{-1}(V))$, so we have that

$$f(X) \subset \bigcup_{j=1}^{m} f(U_{i_j}) \subset \bigcup_{j=1}^{m} V_{i_j} \tag{2.7}$$

Thus every open cover of $f(X)$ has a finite subcover, so $f(X)$ is compact. □

Proof with sequential compactness. Let (y_n) be a sequence in $f(X) \subset Y$. For each $n \in \mathbb{N}$ choose a point x_n such that $f(x_n) = y_n$. By sequential compactness of X, there is some subsequence (x_{n_k}) converging to a point $x \in X$. The continuity of f implies that $f(x_{n_k})$ converges to $f(p) \in f(X)$. Therefore every sequence in $f(X)$ has a convergent subsequence, so $f(X)$ is sequentially compact. □

Theorem 2.10 (Functions from compact sets to the reals achieve their minimum and maximum values)**.** Let X be a compact metric space and $f : X \to \mathbb{R}$ a continuous function. Then there exists $x_0 \in X$ such that $f(x_0) = \sup_{x\in X} f(x)$.

Proof. Since X is compact, we know that the image $f(X)$ is also compact. Since $f(X) \subset \mathbb{R}$, the Heine-Borel theorem tells us that $f(X)$ is closed and bounded. Completeness of the real numbers (every nonempty subset of the real numbers that is bounded above has a supremum in $\mathbb{R}$). Since $f(X)$ is closed, this supremum must be contained in $f(X)$. Therefore there exists $x_0 \in X$ such that $f(x_0) = \sup_{x\in X} f(x)$. □

2.4 Sequential Compactness and Total Boundedness

Definition 2.9 (Sequentially compact)**.** A metric space X is **sequentially compact** if every sequence in X has a convergent subsequence.

Remark. This is also called the Bolzano-Weierstrass property. Recall that the Bolzano-Weierstrass theorem states that every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.

Example 2.8 (Finite sets are sequentially compact)**.** Suppose (x_n) is a sequence contained in a finite set. Then this sequence must repeat a term infinitely often. Define the subsequence which is constant at this term: this subsequence clearly converges.

Theorem 2.11 ((Sequentially) compact sets are closed and bounded).

Proof. Let A be a sequentially compact subset of a metric space X.

Closed: Let p be a limit point of A. Then there is a sequence (a_n) in A that converges to p. Sequential compactness of A implies that there is a subsequence of (a_{n_k}) which converges to some $q \in A$. However every subsequence of a convergent sequence must converge to the same limit as the full sequence, which implies $p = q \in A$. Therefore A is closed.

Bounded: Fix a point $x \in X$. Suppose A is not bounded. Then for each $n \in \mathbb{N}$ there must be some point a_n such that $d(x, a_n) \geq n$. However, sequential compactness implies that some subsequence (a_{n_k}) must converge. Convergent sequences are bounded, so we cannot have that $d(x, a_{n_k}) \to \infty$ as $k \to \infty$. Thus the sequence (a_n) cannot behave as assumed, and it must be that A is bounded for some $r \in \mathbb{R}$.

□

Definition 2.10 (Bounded metric space). A metric space X is **bounded** if it fits in a single fixed ball. More precisely, there exists some $x_0 \in X$ and $r > 0$ such that $X \subseteq B(x_0, r)$.

Definition 2.11 (Totally bounded). A metric space X is **totally bounded** if for every $\varepsilon > 0$ there exist finitely many balls of radius ε that cover X.

Claim 2.1 (Totally bounded implies bounded). Let X be a totally bounded metric space. Then it is bounded.

Proof. Fix $\varepsilon > 0$ and $x, y \in X$. Total boundedness implies there exists points $(x_i)_{i=1}^n$ such that $X \subseteq \bigcup_{i=1}^n B(x_i, \varepsilon)$. Suppose $x \in B(x_i, \varepsilon)$ and $y \in B(x_j, \varepsilon)$. Then

$$d(x, y) \leq d(x, x_i) + d(x_i, x_j) + d(x_j, y) < d(x_i, x_j) + 2\varepsilon \tag{2.8}$$

There are only finitely many values of $d(x_i, x_j)$ we can set $M = \max_{i,j} d(x_i, x_j)$. Therefore $d(x, y) < M + 2\varepsilon$. Therefore X is bounded. □

Example 2.9 (Closed and bounded interval in $\mathbb{R}$ is totally bounded). A closed and bounded interval $I = [a, b] \subset \mathbb{R}$ is totally bounded.

Example 2.10 (Bounded but not totally bounded). The following are examples of metric spaces which are bounded but not totally bounded.

(i) ℓ^1-space is defined to be the collection of all sequences $(a_n)_n$ with $\sum_{n=1}^{\infty} |a_n| < \infty$. We define a distance d on ℓ^1 as follows: $d((a_n)_n, (b_n)_n) = \sum_{n=1}^{\infty} |a_n - b_n|$. The unit ball of ℓ^1 centered at the zero-element (call this call B_1) (i.e., the zero sequence) is bounded but not totally bounded. More precisely,

$$B_1 = \left\{ (a_n)_n \,\middle|\, \sum_{n=1}^{\infty} |a_n| \leq 1 \right\} \tag{2.9}$$

Boundedness is clear. Consider the the set of sequences $A = \{a_n\}$ where each a_n is zero except for a 1 in the n-th entry. Each $a_n \in B_1$. However,

$$d(a_n, a_m) = \begin{cases} 2 & n \neq m \\ 0 & n = m \end{cases} \tag{2.10}$$

If we take balls of radius $\varepsilon = \frac{1}{2}$, then a ball can cover at most one sequence in A. Therefore we can't cover A with finitely ε-balls and hence can't cover B_1 with finitely many ε-balls, so that B_1 is not totally bounded.

(ii) ℓ^∞-space is defined to be the collection of all sequences $(a_n)_n$ with $\sup_n |a_n| < \infty$. We define a distance d on ℓ^∞ as follows: $d\left((a_n)_n, (b_n)_n\right) = \sup_n |a_n - b_n|$. Using the same set of sequences as the previous example, note that for all n we have that

$$d\left(a_n, \{0\}_{n=1}^{\infty}\right) = 1 \tag{2.11}$$

so that $A \subset B_1$ (with the sup norm). Further,

$$d(a_n, a_m) = \begin{cases} 1 & n \neq m \\ 0 & n = m \end{cases} \tag{2.12}$$

But again, if we take balls of radius $\varepsilon = \frac{1}{2}$, then a ball can cover at most one sequence in A. Therefore we can't cover A with finitely ε-balls and hence can't cover B_1 with finitely many ε-balls, so that B_1 is not totally bounded.

Theorem 2.12 (Characterizations of compactness)**.** Let X be a metric space. The following are equivalent:

(i) X is compact.

(ii) X is sequentially compact.

(iii) X is totally bounded and complete.

Proof of X is compact $\Rightarrow$ X is sequentially compact. We argue by contradiction. Suppose X is compact but not sequentially compact. Thus there must exist some sequence $(x_n)_n \subset X$ without a convergence subsequence. Let A be the range of the sequence (more explicitly, $A = \{x_n \mid n \in \mathbb{N}\}$). Note that A has to be an infinite (if A were finite, then there would be a constant subsequence, which is convergent).

Further, A cannot have any limit points (by definition, a point is a limit point of a sequence if there exists a subsequence converging to it). Therefore for each x_n, we can find an open ball centered at x_n that only intersects A at x_n: $B_n \cap A = \{x_n\}$. A is also closed (since it has no limit points), so that $X - A$ is open. We can construct an open cover of X as:

$$X \subseteq \bigcup_n B_n \cup \{X - A\} \tag{2.13}$$

Compactness of A implies this open cover has a finite subcover. Therefore the finite subcover can only contain finitely many of the sets in $\bigcup_n B_n$, so can only contain finitely many of the points in A, which is a contradiction to the covering. $\square$

Proof of X is sequentially compact $\Rightarrow$ X is totally bounded and complete. Suppose X is sequentially compact.

Complete: Fix $\varepsilon > 0$. Let $(x_n)_n \subset X$ be a Cauchy sequence. Sequential compactness implies $(x_n)_n$ has a convergent subsequence. Call this subsequence (x_{n_k}) and its limit x. Then there exists an N such that $d(x_{n_k}, x) < \varepsilon$ for all $n_k \geq N$. Since $(x_n)_n$ is Cauchy, there exists an N such that $n, m \geq M$ implies that $d(x_n, x_m) < \varepsilon$. But then

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < 2\varepsilon \tag{2.14}$$

for all $n, n_k \geq \max\{N, M\}$. Therefore $(x_n)_n$ converges, so X is complete. **In words, Cauchy sequences with convergence subsequences also converge.**

Totally Bounded: We argue by contradiction. Suppose X is not totally bounded. Then there exists an $\varepsilon > 0$ such that X cannot be covered by finitely many balls of radius ε. We know that 1 ball of radius ε cannot cover X. Therefore there must be a point of X outside of this ball: call it x_1. Similarly, 2 balls of radius ε cannot cover X, so pick a point outside of these two balls and call it x_2. Proceeding in this manner generates an infinite sequence. Now consider the n-th term of this sequence. When we choose $n+1$, we must be able to choose a point such that $d(x_i, x_{n+1}) \geq \varepsilon$ for all $i \in \{1, \ldots, n\}$. Otherwise, we would have that $X \subseteq \bigcup_{i=1}^n B(x_i, \varepsilon)$, which would be a contradiction to X not being totally bounded. Thus we can construct such a sequence.

Next, since X is sequentially compact, we must be able to find a convergence subsequence. However, the sequence we've constructed has $d(x_i, x_j) \geq \varepsilon$ for all $i, j \in \mathbb{N}$ (hence no subsequence can be Cauchy, and we know that convergence sequences are Cauchy). Therefore we have reached a contradiction and it holds that X is totally bounded. $\square$

Proof of X is totally bounded and complete $\Rightarrow$ X is sequentially compact. Suppose X is totally bounded and complete. Let $(x_n)_n \subset X$ be a sequence. **We will construct an convergent subsequence.** By the definition of total boundedness, for all $\varepsilon > 0$, X can be covered by finitely many balls of radius ε. **Observation:** One of these balls must contain infinitely many points of $(x_n)_n$. This inspires the following process:

(i) **Step 1:** Cover X with balls of radius 1. One of these balls must contain infinitely many points of $(x_n)_n$. These infinitely many points form a subsequence, call it $(x_n^{(0)})_n$.

(ii) **Step 2:** Cover X with balls of radius $\frac{1}{2}$. One of these balls must contain infinitely many points of $(x_n^{(0)})_n$. These infinitely many points form a subsequence, call it $(x_n^{(1)})_n$.

(iii) **Step n:** Cover X with balls of radius $\frac{1}{2^n}$. One of these balls must contain infinitely many points of $(x_n^{(n-1)})_n$. These infinitely many points form a subsequence, call it $(x_n^{(n)})_n$.

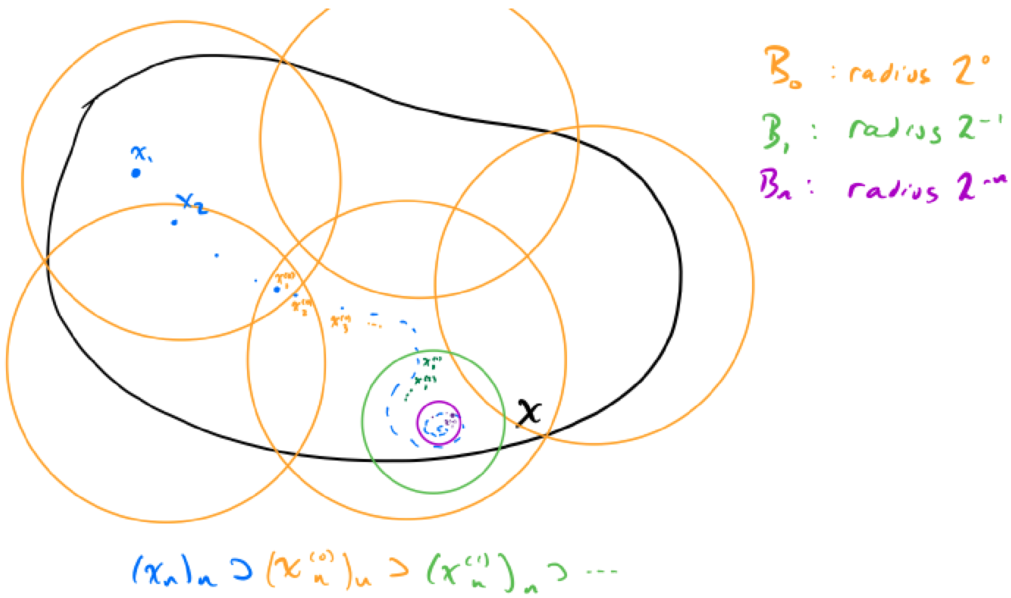
Therefore we have founded nested sequences: $(x_n^{(0)})_n \subset (x_n^{(1)})_n \subset \cdots \subset (x_n^{(n-1)})_n \cdots$. Set $a_n = x_n^{(n)}$ (which is a subsequence of $(x_n)_n$).

Now we show that a_n is a Cauchy sequence. Fix $\varepsilon > 0$ and choose N such that $2^{-N+1} < \varepsilon$. Now for $m > n \geq N$ we have that

$$d(a_m, a_n) < 2 \cdot 2^{-n} < 2^{-N+1} < \varepsilon \tag{2.15}$$

since a_m and a_n are contained in the same ball of radius 2^{-n}. Therefore $(a_n)_n$ is a Cauchy sequence, and completeness implies it converges.

We have found a convergence subsequence, so X is sequentially compact.



□

2.5 Equicontinuity and the Arzela-Ascoli Theorem

[[To prove relatively compact, we construct a sequence of functions that converges uniformly.]]

3 Approximation Theory and Fourier Series

3.1 Orthonormal Systems

Theorem 3.1 (Best L^2-approximation)**.** Let

- $(\phi_n)_n$ be an orthonormal system
- $f \in pc([a,b])$
- $c_n = \langle f, \phi_n \rangle$

Define

- $s_N(x) = \sum_{n=1}^{N} c_n \phi_n(x)$
- $t_N(x) = \sum_{n=1}^{N} b_n \phi_n(x)$ for arbitrary coefficients $b_1, \ldots, b_N \in \mathbb{C}$

Then

$$\|f - s_N\|_2 \leq \|f - t_N\|_2 \tag{3.1}$$

and equality holds iff $b_n = c_n$ for all $n = 1, \ldots, N$.

Proof. We compute the following elements:

$$\langle f, t_N \rangle = \sum_{n=1}^{N} \overline{b}_n \langle f, \phi_n \rangle = \sum_{n=1}^{N} \overline{b}_n c_n$$

$$\langle f, s_N \rangle = \sum_{n=1}^{N} \overline{c}_n \langle f, \phi_n \rangle = \sum_{n=1}^{N} \overline{c}_n c_n = \sum_{n=1}^{N} |c_n|^2$$

$$\begin{aligned}
\langle t_N, t_N \rangle &= \left\langle \sum_{n=1}^{N} b_n \phi_n, \sum_{m=1}^{N} b_m \phi_m \right\rangle \\
&= \sum_{n=1}^{N} \sum_{m=1}^{N} b_n \overline{b}_m \langle \phi_n, \phi_m \rangle \\
&= \sum_{n=1}^{N} |b_n|^2 && ((\phi_n)_n \text{ orthonormal})
\end{aligned}$$

Then

$$\begin{aligned}
\|f - t_N\|_2^2 &= \langle f - t_N, f - t_N \rangle \\
&= \langle f, f \rangle - \langle f, t_N \rangle - \langle t_N, f \rangle + \langle t_N, t_N \rangle \\
&= \langle f, f \rangle - \sum_{n=1}^{N} \overline{b}_n c_n - \sum_{n=1}^{N} b_n \overline{c}_n + \sum_{n=1}^{N} |b_n|^2
\end{aligned}$$

$$
\begin{aligned}
|b_n - c_n|^2 &= (b_n - c_n)(\overline{b_n - c_n}) \\
&= (b_n - c_n)(\overline{b_n} - \overline{c_n}) \\
&= b_n\overline{b_n} - b_n\overline{c_n} - c_n\overline{b_n} + c_n\overline{c_n} \\
&= |b_n|^2 - b_n\overline{c_n} - c_n\overline{b_n} + |c_n|^2
\end{aligned}
$$

Therefore

$$
\begin{aligned}
\|f - t_N\|_2^2 &= \langle f, f\rangle - \sum_{n=1}^{N} \bar{b}_n c_n - \sum_{n=1}^{N} b_n \bar{c}_n + \sum_{n=1}^{N} |b_n|^2 \\
&= \langle f, f\rangle + \sum_{n=1}^{N} |b_n - c_n|^2 - \sum_{n=1}^{N} |c_n|^2
\end{aligned}
$$

And

$$
\begin{aligned}
\|f - s_N\|_2^2 &= \langle f - s_N, f - s_N\rangle \\
&= \langle f, f\rangle - \langle f, s_N\rangle - \langle s_N, f\rangle + \langle s_N, s_N\rangle \\
&= \langle f, f\rangle - 2\sum_{n=1}^{N} |c_n|^2 + \sum_{n=1}^{N} |c_n|^2 \\
&= \langle f, f\rangle - \sum_{n=1}^{N} |c_n|^2
\end{aligned}
$$

Finally

$$
\|f - t_N\|_2^2 = \|f - s_N\|_2^2 + \sum_{n=1}^{N} |b_n - c_n|^2
$$

Thus the claim follows. Indeed, for equality to be achieved, we must have that $b_n = c_n$ for $n = 1, \ldots, N$.

□

Theorem 3.2 (Bessel's Inequality). If $(\phi_n)_n$ is an orthonormal system on $[a, b]$ and $f \in pc([a, b])$, then

$$
\sum_n |\langle f, \phi_n\rangle|^2 \leq \|f\|_2^2 \tag{3.2}
$$

Proof. In the proof of the previous theorem, we calculated that

$$
0 \leq \|f - s_N\|_2^2 = \langle f, f\rangle - \sum_{n=1}^{N} |c_n|^2 \tag{3.3}
$$

Therefore the following holds for all N

$$\sum_{n=1}^{N} |c_n|^2 \leq \langle f, f \rangle \tag{3.4}$$

Take the limit $N \to \infty$ to get the claim

$$\sum_n |c_n|^2 = \sum_n |\langle f, \phi_n \rangle|^2 \leq \|f\|_2^2 \tag{3.5}$$

(Notice that the sum $\sum_{n=1}^{N} |c_n|^2$ converges.) □

Corollary 3.3 (Riemann-Lebesgue Lemma)**.** Let $(\phi_n)_n$ be an orthonormal system and $f \in pc([a,b])$, then

$$\lim_{n\to\infty} \langle f, \phi_n \rangle = 0 \tag{3.6}$$

Proof. In the proof of Bessel's inequality, we showed that the series $\sum_n |\langle f, \phi_n \rangle|^2$ converges. Therefore, a necessary condition of convergence, is that the terms go to zero. □

Definition 3.1 (Complete Orthonormal System)**.** An orthonormal system $(\phi_n)_n$ is called **complete** if

$$\sum_{n=1}^{\infty} |\langle f, \phi_n \rangle|^2 = \|f\|_2^2 \quad \forall f \in pc([a,b]) \tag{3.7}$$

Theorem 3.4. Let

- $(\phi_n)_n$ be an orthonormal system
- $s_N(x) = \sum_{n=1}^{N} c_n \phi_n(x)$

Then $(\phi_n)_n$ is complete if and only if $(s_N)_N$ converges to f in the L^2-norm.
Precisely, $\lim_{N\to\infty} \|f - s_N\|_2 = 0$ for every $f \in \text{pc}([a,b])$).

Proof. Again, from the best approximation theorem, we calculated that

$$\|f - s_N\|_2^2 = \langle f, f \rangle - \sum_n |\langle f, \phi_n \rangle|^2 \tag{3.8}$$

Then this converges to 0 if and only if $(\phi_n)_n$ is complete. □

3.2 Trigonometric Polynomials

Definition 3.2 (Trigonometric Polynomial, Degree)**.** A **trigonometric polynomial** is a function of the form

$$f(x) = \sum_{n=-N}^{N} c_n e^{2\pi i n x} \quad x \in \mathbb{R} \tag{3.9}$$

where $N \in \mathbb{N}$ and $c_n \in \mathbb{C}$. The largest N for which either c_N or c_{-N} is non-zero is called the **degree** of f.

Claim 3.1 (Alternative form of trigonometric polynomial)**.** Every trigonometric polynomial can also be written in the form

$$f(x) = a_0 + \sum_{n=1}^{N} a_n \cos(2\pi n x) + b_n \sin(2\pi n x) \tag{3.10}$$

for coefficients a_n, b_n.

Proof. Using Euler's identity, recall that

$$e^{2\pi i n x} = \cos(2\pi i n x) + i \sin(2\pi i n x) \tag{3.11}$$

Then, for $x \in \mathbb{R}$,

$$\begin{aligned} f(x) &= \sum_{n=-N}^{N} c_n e^{2\pi i n x} \\ &= \sum_{n=-N}^{N} [\cos(2\pi i n x) + i \sin(2\pi i n x)] \\ &= c_0 + \sum_{n=1}^{N} [c_n(\cos(2\pi i n x) + i \sin(2\pi i n x)) + c_{-n}(\cos(2\pi i(-n)x) + i \sin(2\pi i(-nx))] \\ &= c_0 + \sum_{n=1}^{N} [(c_n + c_{-n}) \cos(2\pi i n x) + i(c_n - c_{-n}) \sin(2\pi i n x)] \end{aligned}$$

Therefore the claim alternative holds for $(1 \leq n \leq N)$

$$\begin{aligned} a_0 &= c_0 \\ a_n &= c_n + c_{-n} \\ b_n &= i(c_n - c_{-n}) \end{aligned}$$

□

Definition 3.3 (Partial sums)**.** For a 1-periodic function $f \in \text{pc}$ we define the **partial sums**

$$S_N f(x) = \sum_{n=-N}^{N} \hat{f}(n) e^{2\pi i n x} \tag{3.12}$$

Definition 3.4 (Convolution)**.** For two 1-periodic functions $f, g \in$ pc we define their **convolution** by

$$f * g(x) = \int_0^1 f(t)g(x-t)dt \tag{3.13}$$

Definition 3.5 (Approximation of Unity)**.** A sequence of 1-periodic continuous functions $(k_n)_n$ is called **approximation of unity** if for all 1-periodic continuous functions f we have that $f * k_n$ converges uniformly to f on $\mathbb{R}$. That is

$$\sup_{x\in\mathbb{R}} |f * k_n(x) - f(x)| \to 0 \tag{3.14}$$

as $n \to \infty$

Theorem 3.5 (Sufficient conditions for approximation of unity)**.** Let $(k_n)_n$ be a sequence of 1-periodic continuous functions such that

(i) Non-negative: $k_n(x) \geq 0$.

(ii) Integrates to 1: $\int_{-1/2}^{1/2} k_n(t)dt = 1$.

(iii) "Mass" of k_n concentrated near the origin: For all $1/2 \geq \delta > 0$ we have

$$\int_{-\delta}^{\delta} k_n(t)dt \to 1 \text{ as } n \to \infty \tag{3.15}$$

Proof. Let f be a function that is 1-periodic and continuous. **What can we say about f? Using continuity:** Consider the interval $[-1/2, 1/2]$. f is continuous on this compact set so that on $[-1/2, 1/2]$ it is

(i) Bounded

(ii) Uniformly continuous

Using periodicity: by periodicity, on all of $\mathbb{R}$, f is

(i) Bounded

(ii) Uniformly continuous

Thus there exists a $\delta > 0$ such that

$$|f(x-t) - f(x)| \leq \varepsilon/2 \text{ for all } |t| < \delta, x \in \mathbb{R} \tag{3.16}$$

Then we can write

$$f * k_n(x) - f(x) = \int_{-1/2}^{1/2} (f(x-t) - f(x))\, k_n(t)dt \tag{3.17}$$

Let's break this integral up into two pieces:

$$A = \int_{|t|\leq\delta} (f(x-t) - f(x))\, k_n(t)dt$$

$$B = \int_{1/2 \geq \delta > 0} (f(x-t) - f(x))\, k_n(t) dt$$

To bound A, we use uniform continuity and (ii):

$$|A| = \left| \int_{|t| \leq \delta} (f(x-t) - f(x))\, k_n(t) dt \right|$$

Incomplete.

□

Theorem 3.6 (Fejér)**.** For every 1-periodic, continuous function f we have

$$\sigma_N f \to f \text{ as } N \to \infty \tag{3.18}$$

uniformly on $\mathbb{R}$.

In words, the sequence σ_n of Cesáro means of the sequence (s_n) of partial sums of the Fourier series of f converges uniformly to f on $[0, 1]$.

Corollary 3.7. Every 1-periodic continuous function can be uniformly approximated by trigonometric polynomials.

Proof. □

Corollary 3.8 (Fejér kernel an approximation of unity)**.**

Proof. We must verify the hypotheses of the above theorem.

Incomplete.

□

Theorem 3.9 (Partial sum convoluted with 1-periodic, continuous function converges in 2-norm)**.** Let f be a 1-periodic and continuous function. Then

$$\lim_{N \to \infty} \|S_N f - f\|_2 = 0 \tag{3.19}$$

Proof. Let $\varepsilon > 0$. By Fejer's theorem, there exists a trigonometric polynomial p that uniformly approximates f: That is, $|f(x) - p(x)| < \varepsilon/2$. Then

$$\|f - p\|_2 = \left(\int_0^1 |f(x) - p(x)|^2 \right)^{1/2} \leq \varepsilon/2 \tag{3.20}$$

Suppose p has degree N. Then

$$S_N f - f = S_N f - S_N p + S_N p - f = S_N(f-p) + p - f \tag{3.21}$$

Then by Minkowski's inequality

$$\|S_N f - f\|_2 \le \|S_N(f-p)\|_2 + \|p-f\|_2 \tag{3.22}$$

Bessel's inequality implies that $\|S_N f\|_2 \le \|f\|_2$. Therefore $\|S_N(f-p)\|_2 \le \|f-p\|_2 = \|p-f\|_2$. Then

$$\|S_N f - f\|_2 \le 2\|f-p\|_2 \le \varepsilon \tag{3.23}$$

which proves the claim. $\square$

Corollary 3.10 (Parseval's Theorem). If f, g are 1-periodic, continuous functions, then

$$\langle f, g\rangle = \sum_{n=-\infty}^{\infty} \hat{f}(n)\overline{\hat{g}(n)} \tag{3.24}$$

As a special case

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} \left|\hat{f}(n)\right|^2 \tag{3.25}$$

Proof. We compute

$$\begin{aligned}
\langle S_N f, g\rangle &= \left\langle \sum_{n=-N}^{N} \hat{f}(n) e^{2\pi i n x}, g \right\rangle \\
&= \sum_{n=-N}^{N} \left\langle \hat{f}(n) e^{2\pi i n x}, g \right\rangle \\
&= \sum_{n=-N}^{N} \hat{f}(n) \left\langle e^{2\pi i n x}, g \right\rangle \\
&= \sum_{n=-N}^{N} \hat{f}(n) \int_0^1 e^{2\pi i n x} \overline{g(x)} dx \\
&= \sum_{n=-N}^{N} \hat{f}(n)\overline{\hat{g}(n)}
\end{aligned}$$

Now notice that

$$\langle S_N f, g\rangle \to \langle f, g\rangle \tag{3.26}$$

because

$$\begin{aligned}
|\langle S_N f, g\rangle - \langle f, g\rangle| &= |\langle S_N f - f, g\rangle| \\
&\le \|S_N f - f\|_2 \|g\|_2 && \text{(Cauchy-Schwarz)} \\
&\to 0 \text{ as } N \to \infty && \text{(previous theorem)}
\end{aligned}$$

The special case follows from setting $f = g$. □

Theorem 3.11. Let f be a 1-periodic continuous function and let $x \in \mathbb{R}$. Suppose f is differentiable at x. Then

$$S_N f(x) \to f(x) \text{ as } N \to \infty \tag{3.27}$$

Proof. First by definition

Proof incomplete.

□

3.3 Weierstrass Theorem

Theorem 3.12 (Weierstrass Theorem). For every $f \in C([a,b])$ there exists a sequence of polynomials that converges uniformly to f.

Proof. Let $\varepsilon > 0$.

Incomplete.

□

Remark. This shows that polynomials are dense in $C([a,b])$.

3.4 Stone-Weierstrass Theorem

Theorem 3.13 (Stone-Weierstrass Theorem (Sufficient conditions for a subset of $C(K)$, K compact, to be dense)). Suppose

- K is a compact metric space.
- $\mathcal{A} \subset C(K)$ such that
  - (i) $\mathcal{A}$ is a self-adjoint algebra: for $f, g \in \mathcal{A}$, $c \in \mathbb{C}$, we have

$$f + g \in \mathcal{A}, f \cdot g \in \mathcal{A}, c \cdot f \in \mathcal{A}, \bar{f} \in \mathcal{A} \tag{3.28}$$

  - (ii) $\mathcal{A}$ separates points: for all $x, y \in K$ with $x \neq y$, there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.
  - (iii) $\mathcal{A}$ vanishes nowhere: for all $x \in K$, there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$.

Then $\mathcal{A}$ is dense in $C(K)$ (that is, $\bar{\mathcal{A}} = C(K)$).

Remark. Polynomials and trigonometric polynomials satisfy the conditions of the Stone-Weierstrass Theorem (Theorem 3.13).

Show this.

We prove a sequence of lemmas before proving the theorem.

Lemma 3.14 (Sequence of polynomial (zero intercept) uniformly converging to absolute value function)**.** For every $a > 0$ there exists a sequence of polynomials $(p_n)_n$ with real coefficients such that $p_n(0) = 0\ \forall n$ and

$$\sup_{x\in[-a,a]} |p_n(x) - |x|| \to 0 \text{ as } n \to \infty \tag{3.29}$$

Proof. $f(x) = |x|$ is a continuous function on $[-a, a]$. By Weierstrass' theorem, there exists a sequence of polynomials $(q_n)_n$ converging uniformly to $f(x) = |x|$ on $[-a, a]$. Now set $p_n(x) = q_n(x) - q_n(0)$. It is clear that $p_n(0) = 0$. Further, $p_n(x)$ converges uniformly to $|x|$, since $q_n(x)$ uniformly converges to $|x|$ and $q_n(0)$ converges to 0.

More formally: fix $\varepsilon > 0$. Then $\exists N$ such that $|q_n(x) - |x|| < \varepsilon$ for all $n > N$ and $x \in [-a, a]$. Thus, for this N, $|q_n(0)| < \frac{\varepsilon}{2}$ and $|q_n(x)| < \frac{\varepsilon}{2}$ for all $x \neq 0$. Therefore

$$|p_n(x)| = |q_n(x) - q_n(0)| \leq |q_n(x)| + |q_n(0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \tag{3.30}$$

□

Lemma 3.15. If $f \in \bar{A}$, then $|f| \in \bar{A}$.

Proof. By the previous lemma, there exists a sequence of polynomials (with zero intercept) converging uniformly to the absolute value function. Thus fix $\varepsilon > 0$. Let $a = \max_{x\in K} |f(x)|$. Then there exist coefficients $c_1, c_2, \dots, c_N \in \mathbb{R}$ such that

$$\left|\sum_{i=1}^{n} c_i y^i - |y|\right| \leq \varepsilon \quad \forall\, |y| \leq a \tag{3.31}$$

Note: This sum does not have an intercept/constant term. Since $f \in \bar{A}$ and A is a self-adjoint algebra, we have that

$$g = \sum_{i=1}^{n} c_i f^i \in \bar{A} \tag{3.32}$$

But since Equation 3.31 holds for all values $y \in [-a, a]$, and we have that $|f(x)| \leq a$, the same inequality holds for $y^i = f^i(x)$, $x \in K$. **Note:** This holds for every $x \in K$.

$$|g(x) - |f(x)|| = \left|\sum_{i=1}^{n} c_i f^i(x) - |f(x)|\right| \leq \varepsilon \quad x \in K \tag{3.33}$$

This shows that $|f|$ can be uniformly approximated by functions in $\bar{A}$. Since $\bar{A}$ is closed, we have that $|f| \in \bar{A}$. □

Lemma 3.16 ($\bar{A}$ closed under max and min operations)**.** If $f_1, \dots, f_m \in \bar{A}$ then $\min\{f_1, \dots, f_n\} \in \bar{A}$ and $\max\{f_1, \dots, f_n\} \in \bar{A}$

Proof. We prove for $m = 2$ (and the general case follows by induction). Let $f, g \in \bar{A}$. We can write $\min\{f, g\}$ and $\max\{f, g\}$ as linear combinations of functions in $\bar{A}$. Indeed, observe that

$$\min\{f,g\} = \frac{f-g}{2} - \frac{|f-g|}{2}$$
$$\max\{f,g\} = \frac{f-g}{2} + \frac{|f-g|}{2}$$

Therefore, since $\bar{A}$ is a self-adjoint algebra and is closed under taking the absolute value, we have that $\bar{A}$ is also closed under taking the max and min of finitely many functions. □

Lemma 3.17 (Any two points that could lie on the graph of a function in $\bar{A}$ do lie on the graph of a function in $\bar{A}$)**.** For every $x_0, x_1 \in K$, $x_0 \neq x_1$ and $c_0, c_1 \in \mathbb{R}$, there exists $f \in \bar{A}$ such that $f(x_i) = c_i$ for $i = 0, 1$.

Proof. Using conditions (ii) and (iii) we have that there exist functions $g, h_0, h_1 \in \bar{A}$ such that

(i) g separates points: $g(x_0) \neq g(x_1)$

(ii) h_0 and h_1 don't vanish: $h_0(x_0) \neq 0$ and $h_1(x_1) \neq 0$.

Define

$$u_0(x) = (g(x) - g(x_1))h_0(x) \Rightarrow u_0(x_1) = 0, u_0(x_0) \neq 0$$
$$u_1(x) = (g(x) - g(x_0))h_1(x) \Rightarrow u_1(x_0) = 0, u_1(x_1) \neq 0$$

Now let

$$f(x) = \frac{c_0 u_0(x)}{u_0(x_0)} + \frac{c_1 u_1(x)}{u_1(x_1)} \tag{3.34}$$

It is clear that

$$f(x_0) = c_0$$
$$f(x_1) = c_1$$

and these terms are well-defined (i.e., no issues with zero denominators). □

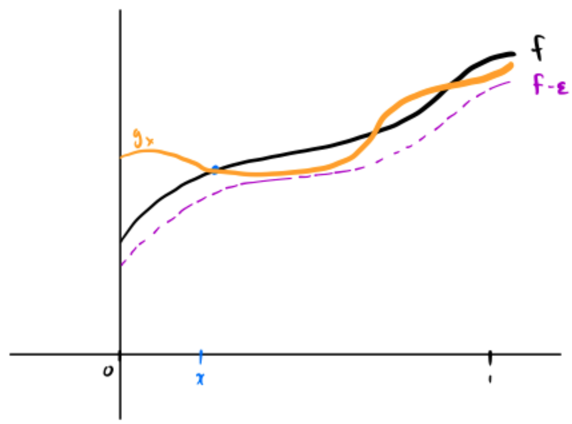
Remark. We can extend this lemma to finitely many points. Thus if K were finite, we would have proved the Stone-Weierstrass theorem.

Claim 3.2. Let $f \in C(K)$ and $\varepsilon > 0$. For every $x \in K$ there exists $g_x \in \bar{A}$ such that

(i) $g_x(x) = f(x)$

(ii) $g_x(t) > f(t) - \varepsilon$

We are looking for g_x like shown in the following figure:



Proof. Take another point $y \in K$, $y \neq x$. By the previous lemma, we can find an $h_y \in \bar{A}$ such that

(i) $h_y(x) = f(x)$

(ii) $h_y(y) = f(y)$

Since h_y is continuous, there exists a $\delta_y > 0$ such that

$$t \in B_{\delta_y}(y) \Rightarrow |h_y(t) - f(t)| < \varepsilon \tag{3.35}$$

Notice that this implies that $h_y(t) > f(t) - \varepsilon$. Now notice that $\left(B_{\delta_y}(y)\right)_{y \in K}$ is an open cover of K. Compactness implies the existence of a finite subcover, identified by points $y_1, \ldots, y_m$. Now let

$$g_x = \max\{h_{y_1}, \ldots, h_{y_m}\} \tag{3.36}$$

We have, by a previous lemma, that $g_x \in \bar{A}$. Further, it is clear that

$$g_x(x) = f(x) \tag{3.37}$$

since each $h_{y_i}(x) = x$ and

$$g_x(t) > f(t) - \varepsilon \tag{3.38}$$

since by taking the max (pointwise) of elements for which this is true. This proves the claim. □

Finally ready to complete the whole proof.

Proof of Stone-Weierstrass Theorem. □

4 Linear Operators and Derivatives

In this section

- $\mathbb{K}$ either one of the fields $\mathbb{R}$ and $\mathbb{C}$
- X a vector space over $\mathbb{K}$

Definition 4.1 (Norm)**.** A map $\|\cdot\| : X \to [0, \infty)$ is called a **norm** if for all $x, y \in X$ and $\lambda \in \mathbb{K}$ we have

(i) $\|\lambda x\| = |\lambda|\|x\|$

(ii) $\|x + y\| \leq \|x\| + \|y\|$

(iii) $\|x\| = 0 \iff x = 0$

Definition 4.2 (Normed vector space)**.** A $\mathbb{K}$-vector space equipped with a norm is called a **normed vector space**. On every normed vector space we have a natural metric space structure defined by

$$d(x, y) = \|x - y\| \tag{4.1}$$

Definition 4.3 (Banach space)**.** A complete normed vector space is called a **Banach space**.

Definition 4.4 (Linear map)**.** Let X, Y be normed vector spaces. A map $T : X \to Y$ is called **linear** if

$$T(x + \lambda y) = Tx + \lambda Ty \tag{4.2}$$

for every $x, y \in X$, $\lambda \in \mathbb{K}$.

Definition 4.5 (Bounded)**.** A linear map $T : X \to Y$ is called **bounded** if there exists $C > 0$ such that

$$\|Tx\|_Y \leq C\|x\|_X \quad \forall x \in X \tag{4.3}$$

Remark. Linear maps between normed vector spaces are also referred to as **linear operators**.

Theorem 4.1. Let $T : X \to Y$ be a linear map. TFAE

(i) T is bounded

(ii) T is continuous

(iii) T is continuous at 0

(iv) $\sup_{\|x\|_X=1}\|Tx\|_Y < \infty$

Proof (i) ⇒ (ii). Suppose T is bounded. Thus there exists $C > 0$ such that $\|Tx\|_Y \leq C\|x\|_X \; \forall x \in X$. Then for $x, y \in X$

$$\begin{aligned} \|Tx - Ty\|_Y &= \|T(x-y)\|_Y && \text{(linearity)} \\ &\leq C\|x-y\|_X && \text{(bounded)} \end{aligned}$$

Fix $\varepsilon > 0$. Set $\delta = \frac{\varepsilon}{C}$. Take $x, y \in X$ such that $\|x-y\|_X < \delta$. Then the calculation above gives that $\|Tx - Ty\|_Y < \varepsilon$ for all $x, y \in X$ for which $\|x-y\|_X < \delta$. Thus T is continuous. □

Proof (ii) ⇒ (iii). Immediate. □

Proof (iii) ⇒ (iv). Since T is continuous at 0, there exists a $\delta > 0$ such that for all $x \in X$ for which $\|x\|_X \leq \delta$ we have that $\|Tx\|_Y \leq \varepsilon = 1$. Now fix $x \in X$ for which $\|x\|_X = 1$. By scale preservation of norm, $\|\delta x\|_X = \delta$. Therefore

$$\|T(\delta x)\|_Y \leq 1 \tag{4.4}$$

Further

$$\|T(\delta x)\|_Y = \delta\|Tx\|_Y \tag{4.5}$$

so that $\|Tx\|_Y \leq \delta^{-1}$. This holds for all $x \in X$ for which $\|x\|_X = 1$. Therefore

$$\sup_{\|x\|_X = 1} \|Tx\|_Y \leq \delta^{-1} < \infty \tag{4.6}$$

□

Proof (iv) ⇒ (i). □

5 Practice

5.1 2018 HW4

$$\begin{aligned} \widehat{f * g}(n) &= \int_0^1 (f * g(t)) e^{-2\pi i n t} \, dt \\ &= \int_0^1 \left(\int_0^1 f(x) g(t-x) \, dx \right) e^{-2\pi i n t} \, dt \\ &= \int_0^1 \left(\int_0^1 f(x) g(t-x) e^{-2\pi i n t} \, dx \right) \, dt && \text{(inner integral didn't depend on } t\text{)} \\ &= \int_0^1 \left(\int_0^1 f(x) g(t-x) e^{-2\pi i n t} \, dt \right) \, dx && \text{(change order of integration)} \\ &= \int_0^1 f(x) \left(\int_0^1 g(t-x) e^{-2\pi i n t} \, dt \right) \, dx \\ &= \int_0^1 f(x) \left(\int_{-x}^{1-x} g(y) e^{-2\pi i n (y+x)} \, dy \right) \, dx && \text{(change of vars: } t - x \to y\text{)} \\ &= \int_0^1 f(x) e^{-2\pi i n x} \left(\int_0^1 g(y) e^{-2\pi i n y} \, dy \right) \, dx && (g \text{ periodic}) \end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 f(x)e^{-2\pi inx}\widehat{g}(n)\,dx \\
&= \widehat{g}(n)\int_0^1 f(x)e^{-2\pi inx}\,dx \\
&= \widehat{f}(n)\widehat{g}(n)
\end{aligned}
$$

6 Differential Calculus in $\mathbb{R}^n$

7 The Baire category theorem

A Review From Elementary Analysis

A.1 The Real and Complex Number System

Definition A.1 (Supremum, Infimum)**.** Let S be an ordered set, $E \subset S$, and E be bounded above. Suppose there exists an $\alpha \in S$ with the following properties:

(i) α is an upper bound of E.

(ii) $If \gamma < \alpha$ then γ is not an upper bound of E.

Then α is called the least upper bound of E or **supremum** of E and we write $\alpha = \sup E$. Similarly, α is the greatest lower bound of E or **infimum** of E if

(i) α is a lower bound of E.

(ii) $If \beta > \alpha$ then β is not a lower bound of E.

and we write $\alpha = \inf E$.

Definition A.2 (Limit Superior, Inferior)**.** Let (x_n) be a sequence of real numbers.

(i) The **limit superior** of the sequence is defined by

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty}\left(\sup_{m\geq n} x_m\right) \tag{A.1}$$

or

$$\limsup_{n\to\infty} x_n = \inf_{n\geq 0}\left(\sup_{m\geq n} x_m\right) = \inf\{\sup\{x_m \mid m \geq n\} \mid n \geq 0\} \tag{A.2}$$

Alternatively, the limit superior of the sequence is the smallest $b \in \mathbb{R}$ such that $\forall \varepsilon > 0$ $\exists N$ such that $x_n < b + \varepsilon \; \forall n > N$. Thus, any number larger than the limit superior is an upper bound for the sequence after a finite number of terms (hence, only a finite number of elements are greater that $b + \varepsilon$).

Alternatively, the limit superior of the sequence is the supremum of the set of subsequential limits.

(ii) The **limit inferior** of the sequence is defined by

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty}\left(\inf_{m\geq n} x_m\right) \tag{A.3}$$

or

$$\liminf_{n\to\infty} x_n = \sup_{n\geq 0}\left(\inf_{m\geq n} x_m\right) = \sup\{\inf\{x_m \mid m\geq n\} \mid n\geq 0\} \tag{A.4}$$

Theorem A.1 (Properties of limit superiors)**.** Let (x_n) and (y_n) be sequence of real numbers. Then

(i) $\limsup_{n\to\infty}(x_n+y_n) \leq \limsup_{n\to\infty} x_n + \limsup_{n\to\infty} y_n$ (as long as the RHS is not of the form $\infty - \infty$)

Proof. We prove each item in turn.

(i) We have that

$$\limsup_{n\to\infty}(x_n+y_n) = \lim_{n\to\infty}\left(\sup_{m\geq n}\{x_m+y_m, x_{m+1}+y_{m+1},\dots\}\right) \tag{A.5}$$

Then

$$M_n := \sup_{m\geq n}\{x_m+y_m, x_{m+1}+y_{m+1},\dots\} = \sup_{m\geq n}\{x_m+y_m\} \leq \sup_{m\geq n}\{x_m\} + \sup_{m\geq n}\{y_m\} \tag{A.6}$$

Take the limit of both sides to get

$$\lim_{n\to\infty} M_n \leq \lim_{n\to\infty}\sup_{m\geq n}\{x_m\} + \lim_{n\to\infty}\sup_{m\geq n}\{y_m\} \tag{A.7}$$

Thus

$$\limsup_{n\to\infty}(x_n+y_n) \leq \limsup_{n\to\infty} x_n + \limsup_{n\to\infty} y_n \tag{A.8}$$

□

A.2 Basic Topology

In what follows, assume X is a metric space.

Definition A.3 (Limit point)**.** A point p is a **limit point** of the set E if every neighborhood of p contains a point $q \neq p$ such that $q \in E$.

Definition A.4 (Closed)**.** E is **closed** if every limit point of E is a point of E.

Definition A.5 (Interior)**.** A point p is an **interior point** of E if there is a neighborhood N of p such that $N \subset E$.

Definition A.6 (Open)**.** E is **open** if every point of E is an interior point of E.

Definition A.7 (Bounded)**.** E is **bounded** if there is a real number M and a point $q \in X$ such that $d(p, q) < M$ for all $p \in E$.

Definition A.8 (Separated, Connected)**.** Two subsets A and B of a metric space X are said to be **separated** if both $A \cap \bar{B}$ and $\bar{A} \cap B$ are empty (i.e., if no point of A lies in the closure of B and no point of B lies in the closure of A). A set $E \subset X$ is said to be **connected** if E is not a union of two nonempty separated sets.

A.3 Numerical Sequences and Series

A.3.1 Sequences

Definition A.9 (Convergent Sequence)**.** A sequence (p_n) in a metric space X is said to **converge** if there is a point $p \in X$ such that for every $\varepsilon > 0$ there is an integer N such that $n \geq N$ implies $d(p_n, p) < \varepsilon$.

Definition A.10 (Subsequence, Subsequential Limit)**.** Given a sequence (p_n), consider a sequence (n_k) of positive integers such that $n_1 < n_2 < n_3 < \cdots$. Then the sequence (p_{n_i}) is called a **subsequence** of (p_n). If (p_{n_i}) converges, its limit is called a **subsequential limit** of (p_n).

Observations:

- (p_n) converges to p if and only if every subsequence of (p_n) converges to p.

Theorem A.2 (Sequences in compact metric spaces have a convergent subsequence)**.** If (p_n) is a sequence in a compact metric space X, then some subsequence of (p_n) converges to a point of X.

Theorem A.3 (Bolzano-Weierstrass)**.** Every bounded sequence in $\mathbb{R}^k$ has a convergent subsequence.

Definition A.11 (Cauchy Sequence)**.** A sequence (p_n) in a metric space X is said to be a **Cauchy Sequence** if for every $\varepsilon > 0$ there is an integer N such that $d(p_n, p_m) < \varepsilon$ if $n \geq N$ and $m \geq N$.

Definition A.12 (Diameter)**.** Let E be a nonempty subset of a metric space X, and let S be the set of all real numbers of the form $d(p, q)$ with $p \in E$ and $q \in E$. The sup of S is called the **diameter** of E.

Theorem A.4 (Facts about Cauchy sequences)**.** We have that

(i) In any metric space X, every convergent sequence is a Cauchy sequence.

(ii) If X is a compact metric space and if (p_n) is a Cauchy sequence in X, then (p_n) converges to some point of X.

(iii) In $\mathbb{R}^k$, every Cauchy sequence converges.

Definition A.13 (Complete)**.** A metric space in which every Cauchy sequence converges is **complete**.

Definition A.14 (Monotonically increasing, decreasing)**.** A sequence (s_n) of real numbers is said to be

(i) **Monotonically increasing** if $s_n \leq s_{n_1}$ for all n.

(ii) **Monotonically decreasing** if $s_n \geq s_{n_1}$ for all n.

Theorem A.5 (Convergence of monotonic sequences)**.** Let (s_n) be a monotonic sequence. Then (s_n) converges if and only if it is bounded.

A.3.2 Series

Definition A.15 (Convergent Series)**.** Let $\sum_{n=1}^{\infty} a_n$ be an infinite series. Define $s_n = \sum_{k=1}^{n} a_n$ to be the nth partial sum of the series. If the sequence of partial sums $\{s_n\}$ converges to s, we say the series **converges**.

Theorem A.6 (Cauchy Criterion for Series)**.** $\sum a_n$ converges iff

$$\forall \varepsilon > 0\, \exists N \text{ s.t. } m, n \geq N \Rightarrow \left| \sum_{k=n}^{m} a_k \right| \leq \varepsilon \tag{A.9}$$

Theorem A.7 (Necessary condition for convergence: individual terms of series go to 0.)**.** If $\sum a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

Theorem A.8. A series of nonnegative terms converges if and only if its partial sums form a bounded sequence.

Theorem A.9 (Root Test)**.** Given $\sum a_n$, let $\alpha = \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}$. Then

(i) If $\alpha < 1$, $\sum a_n$ converges.

(ii) If $\alpha > 1$, $\sum a_n$ diverges.

(iii) If $\alpha = 1$, the test gives no information.

A.4 Continuity

Definition A.16 (Limit)**.** Let X and Y be metric spaces, $E \subset X$, f map E into Y, and p be a limit point of E. We write $f(x) \to q$ as $x \to p$ or $\lim_{\mathbf{x}\to\mathbf{p}} \mathbf{f}(\mathbf{x}) = \mathbf{q}$ if there is a point $q \in Y$ such that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in E$ for which $0 < d_X(x,p) < \delta$, we have $d_Y(f(x), q) < \varepsilon$.

Definition A.17 (Continuous)**.** Suppose X and Y are metric spaces, $E \subset X$, $p \in E$, and f maps E into Y. Then f is **continuous** at p if for every $\varepsilon > 0$ there there exists a $\delta > 0$ such that for all $x \in E$ for which $d_X(x,p) < \delta$, we have that $d_Y(f(x), f(p)) < \varepsilon$.

Definition A.18 (Uniformly Continuous)**.** Let f be a mapping of a metric space X into a metric space Y. We say that f is **uniformly continuous** on X if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $p, q \in X$ for which $d_X(p,q) < \delta$, we have $d_Y(f(p), f(q)) < \varepsilon$.

Observations:

- Uniform continuity is a property of a function on a set, whereas continuity can be defined at a single point.
- If f is continuous on X, then for each $\varepsilon > 0$ and each $p \in X$, we can find a $\delta > 0$ that satisfies the condition in the definition. For uniform continuity, we can find one $\delta > 0$ that works for all points $p \in X$.
- Every uniformly continuous function is continuous. The two concepts are equivalent on compact sets.

A.5 Differentiation

Definition A.19 (Differentiable, Derivative). Let f be defined (and real-valued) on $[a, b]$. For any $x \in [a, b]$ define

$$f'(x) = \lim_{t \to x} \frac{f(t) - f(x)}{t - x} \tag{A.10}$$

provided this limit exists. If f' is defined at a point x, we say that f is **differentiable** at x. f' is called the **derivative** of f.

Theorem A.10 (Mean Value Theorem). If f is a real continuous function on $[a, b]$ which is differentiable on (a, b), then there is a point $x \in (a, b)$ at which

$$f(b) - f(a) = (b - a)f'(x) \tag{A.11}$$

References

Humpherys, J., Jarvis, T. J., and Evans, E. J. (2017). *Foundations of Applied Mathematics, Volume 1: Mathematical Analysis*. SIAM.

Roos, J. (2018). Analysis Lecture Notes.

Rudin, W. (1976). *Principles of Mathematical Analysis*, volume 3. McGraw-Hill New York.