Real Analysis

Notes for Pure Math 351

Kenneth R. Davidson

University of Waterloo

CONTENTS

Chapter 1. Metric Spaces 1
- 1.1. Normed Vector Spaces 1
- 1.2. Metric spaces 4
- 1.3. Topology of Metric spaces 9
- 1.4. Continuous functions 13
- 1.5. Finite dimensional normed vector spaces 16
- 1.6. Completeness 19
- 1.7. Completeness of $\mathbb{R}$ and $\mathbb{R}^n$ 22
- 1.8. Limits of continuous functions 25

Chapter 2. More Metric Topology 29
- 2.1. Compactness 29
- 2.2. More compactness 34
- 2.3. Compactness and Continuity 35
- 2.4. The Cantor Set, Part I 37
- 2.5. Compact sets in $C(X)$ 41
- 2.6. Connectedness 43
- 2.7. The Cantor Set, Part II 46

Chapter 3. Completeness Revisited 50
- 3.1. The Baire Category Theorem 50
- 3.2. Nowhere Differentiable Functions 53
- 3.3. The Contraction Mapping Principle 56
- 3.4. Newton's Method 60
- 3.5. Metric Completion 64
- 3.6. The p-adic Numbers 67
- 3.7. The Real Numbers 71

Chapter 4. Approximation Theory 78
- 4.1. Polynomial Approximation 78
- 4.2. Best Approximation 81
- 4.3. The Stone-Weierstrass Theorems 85

Chapter 5. Differential Equations 90
- 5.1. Reduction to first order 92
- 5.2. Global Solutions of ODEs 95
- 5.3. Local Solutions 97
- 5.4. Existence without Uniqueness 102
- 5.5. Stability of DEs 104

Index 109

CHAPTER 1

Metric Spaces

In calculus, we learned about the structure of the real line and $\mathbb{R}^n$ and continuous functions on subsets of $\mathbb{R}^n$ In this course, we find that the same ideas generalize to a much broader context.

1.1. Normed Vector Spaces

A natural generalization of $\mathbb{R}^n$ with its usual Euclidean distance is the notion of a vector space with a norm, so that we can discuss convergence.

1.1.1. DEFINITION. If V is a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, then a *norm* on V is a function $\|\cdot\| : V \to [0, \infty)$ such that

(1) $\|v\| = 0 \iff v = 0$ (*positive definite*).

(2) $\|\lambda v\| = |\lambda|\,\|v\|$ for all $\lambda \in \mathbb{F}$ and $v \in V$ (*positive homogeneous*).

(3) $\|u + v\| \le \|u\| + \|v\|$ for all $u, v \in V$ (*triangle inequality*).

We say that $(V, \|\cdot\|)$ is a *normed vector space*. A *seminorm* satisfies (2) and (3) and $\|0\| = 0$, but possibly some non-zero vectors have zero norm.

1.1.2. EXAMPLES.
(1) $\mathbb{R}^n$ and $\mathbb{C}^n$ for $n \ge 1$ with the *Euclidiean norm*. If $x = (x_1, \dots, x_n)$, then

$$\|x\|_2 = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{1/2}.$$

Recall that this is an inner product space with $\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$. Indeed, any inner product space is a normed vector space with $\|x\| = \langle x, x\rangle^{1/2}$. The triangle inequality follows from the Cauchy-Schwarz inequality: $|\langle x, y\rangle| \le \|x\|\,\|y\|$.

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + 2\,\mathrm{Re}\langle x, y\rangle + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\|\,\|y\| + \|y|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

(2) If $X \subset \mathbb{R}^n$, let $C^b(X)$ denote the space of bounded continuous functions on X with supremum norm

$$\|f\|_\infty = \sup_{x\in X} |f(x)|.$$

If X is closed and bounded, then the Extreme Value Theorem shows that every continuous function on X is bounded and attains it maximum modulus. In this case, we write $C(X)$ for the space of all continuous functions on X with the supremum norm. We will write $C_{\mathbb{R}}(X)$ if we want the real vector space of real valued continuous functions. The norm properties are easy to verify.

(3) For $1 \le p < \infty$, let $l_p^{(n)}$ be $\mathbb{C}^n$ with the norm

$$\|x\|_p = \big(\sum_{i=1}^n |x_i|^p \big)^{1/p}.$$

We also define $l_\infty^{(n)}$ with norm

$$\|x\|_\infty = \max\{|x_1|, \dots, |x_n|\}.$$

Properties (1) and (2) are easy. It is not obvious that the triangle inequality holds except for $p = 1$ and $p = \infty$. This will be established below for $1 < p < \infty$.

(3′) For $1 \le p < \infty$, let l_p denote the set of all infinite sequences with coefficients in $\mathbb{C}$ or $\mathbb{R}$, $x = (x_1, x_2, x_3, \dots)$, for which the series

$$\|x\|_p^p = \sum_{i\ge 1} |x_i|^p < \infty.$$

Likewise we let l_∞ denote the vector space of all bounded sequences with

$$\|x\| = \sup_{i\ge 1} |x_i|.$$

(4) For $1 \le p < \infty$, the L^p norm on $C[a, b]$ is given by

$$\|f\|_p = \Big(\int_a^b |f(x)|^p \, dx \Big)^{1/p}.$$

Again it is not obvious that this satisfies the triangle inequality except for $p = 1$. For $p = 2$, this follows because

$$\langle f, g\rangle = \int_a^b f(x)\overline{g(x)}\, dx$$

is an inner product that yields the 2-norm.

1.1.3. THEOREM (Minkowski's inequality). *For $1 < p < \infty$, the triangle inequality is valid for the L^p norm on $C[a, b]$ and the norm on l_p and $l_p^{(n)}$. Equality holds only when f and g lie in a 1-dimensional subspace.*

PROOF. Let $f, g \in C[a,b]$ be non-zero functions (the case of $f=0$ or $g=0$ is left to the reader). Define $A = \|f\|_p$ and $B = \|g\|_p$. Note that $A > 0$ because $|f(x)| > 0$ on some interval (c,d) by continuity; and similarly $B > 0$. So we may define $f_0 = f/A$ and $g_0 = g/B$. Clearly $\|f_0\|_p = 1 = \|g_0\|_p$.

Consider the function $\varphi(x) = x^p$ on $[0,\infty)$. Note that $\varphi''(x) = p(p-1)x^{p-2} > 0$ on $(0,\infty)$, and thus $\varphi(x)$ is a strictly convex function, meaning that it curves upwards, or that for all $x_1, x_2 \in [0,\infty)$ and $0 \le t \le 1$,

$$\varphi(tx_1 + (1-t)x_2) \le t\varphi(x_1) + (1-t)\varphi(x_2);$$

with equality only when $x_1 = x_2$ or $t = 0$ or $t = 1$. That is every chord between distinct points on the curve $y = \varphi(x)$ lies strictly above the curve.

For us, this means that for any $x \in [a,b]$ that

$$\Big(\frac{A}{A+B}|f_0(x)| + \frac{B}{A+B}|g_0(x)|\Big)^p \le \frac{A}{A+B}|f_0(x)|^p + \frac{B}{A+B}|g_0(x)|^p. \tag{1.1.4}$$

Now we integrate this:

$$\begin{aligned}
\frac{1}{(A+B)^p}\int_a^b |f(x)+g(x)|^p\,dx &\le \int_a^b \Big(\frac{A|f_0(x)| + B|g_0(x)|}{A+B}\Big)^p\,dx \\
&\le \int_a^b \frac{A}{A+B}|f_0(x)|^p + \frac{B}{A+B}|g_0(x)|^p\,dx \\
&= \frac{A}{A+B}\int_a^b |f_0(x)|^p\,dx + \frac{B}{A+B}\int_a^b |g_0(x)|^p\,dx \qquad (1.1.5)\\
&= \frac{A}{A+B}\|f_0\|_p^p + \frac{B}{A+B}\|g_0\|_p^p \\
&= 1.
\end{aligned}$$

Multiplying through by $(A+B)^p$, we get that

$$\|f+g\|_p^p \le (A+B)^p = (\|f\|_p + \|g\|_p)^p.$$

This establishes the triangle inequality.

Finally, note that equation(1.1.4) is a strict inequality unless $|f_0(x)| = |g_0(x)|$. If they differ at some x_0, then by continuity, they differ on an interval (c,d) containing x_0. Thus when we integrate, the inequality will be strict. This shows that $|f_0| = |g_0|$. Also in the first line of equation (1.1.5), the inequality is strict unless $\operatorname{sign}(f_0(x)) = \operatorname{sign}(g_0(x))$. Again strict inequality at a point leads to strict inequality on a whole interval, and thus a strict inequality when we integrate. Combining the two ideas shows that for equality, we require that $f_0 = g_0$, or that $g = Bf/A$. That is, g is a scalar multiple of f.

The proof for l_p and $l_p^{(n)}$ is basically the same, but without any concern about continuity. Suppose that $x, y \in l_p$ are non-zero. Set $A = \|x\|_p$ and $B = \|y\|_p$. By the convexity of $\varphi(x) = x^p$, we obtain that

$$\Big(\frac{|x_i| + |y_i|}{A+B}\Big)^p = \Big(\frac{A}{A+B}\frac{|x_i|}{A} + \frac{B}{A+B}\frac{|y_i|}{B}\Big)^p \le \frac{A}{A+B}\Big(\frac{|x_i|}{A}\Big)^p + \frac{B}{A+B}\Big(\frac{|y_i|}{B}\Big)^p.$$

Sum from 1 to ∞ (or stop at n) and obtain that

$$\begin{aligned}
\frac{1}{(A+B)^p}\sum_{i=1}^{\infty} |x_i + y_i|^p \, dx &\le \sum_{i=1}^{\infty} \Big(\frac{|x_i| + |y_i|}{A+B}\Big)^p \\
&\le \sum_{i=1}^{\infty} \frac{A}{A+B}\Big(\frac{|x_i|}{A}\Big)^p + \frac{B}{A+B}\Big(\frac{|y_i|}{B}\Big)^p \\
&= \frac{A}{A+B}\sum_{i=1}^{\infty}\Big(\frac{|x_i|}{A}\Big)^p + \frac{B}{A+B}\sum_{i=1}^{\infty}\Big(\frac{|y_i|}{A}\Big)^p \\
&= \frac{A}{A+B}\frac{\|x\|_p^p}{A^p} + \frac{B}{A+B}\frac{\|y\|_p^p}{B^p} = 1.
\end{aligned}$$

Thus $\|x+y\|_p^p \le (A+B)^p = (\|x\|_p + \|y\|_p)^p$, which is the triangle inequality. The case of equality is argued in the same manner. ■

1.2. Metric spaces

The idea of a metric space generalizes the notion of distance beyond subsets of Euclidean space. However many ideas such as continuity and completeness extend naturally to this more general context.

1.2.1. DEFINITION. A *metric space* (X, d) is a set X together with a distance function $d : X \times X \to [0, \infty)$ such that

(1) $d(x, y) = 0 \iff x = y$ for $x, y \in X$.

(2) $d(x, y) = d(y, x)$ for $x, y \in X$ (*symmetry*).

(3) $d(x, z) \le d(x, y) + d(y, z)$ for $x, y, z \in X$ (*triangle inequality*).

Any reasonable function that tries to be a distance will satisfy (2), and generally verifying (1) is very easy. But the triangle inequality can be tricky, as we saw in the previous section.

A useful consequence of the triangle inequality, sometimes called the *reverse triangle inequality* is

$$d(x, z) \ge d(x, y) - d(y, z) \quad \text{for all} \quad x, y, z \in X.$$

Try to convince yourself that this is true.

1.2.2. EXAMPLES.

(1) Let $(V, \|\cdot\|)$ be a normed vector space, and let $X \subset V$. Define $d(x, y) = \|x-y\|$ for $x, y \in X$. Then (X, d) is a metric space *induced* from the norm.

(2) Let X be a set. The *discrete metric* is given by

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

(2′) A variant on (2) is the *Hamming distance* on the collection $\mathcal{P}(X)$ of all subsets of a *finite* set X given by

$$\rho(A,B) = |A \triangle B| = |(A \cup B) \setminus (A \cap B)|,$$

the cardinality of the symmetric difference between the two sets $A, B \in \mathcal{P}(X)$. Convince yourself that this satisfies the triangle inequality.

(3) If X is the sphere S_d, namely the surface of the unit ball in $\mathbb{R}^{d+1}$, or indeed any other manifold, the *geodesic distance* between two points $x, y \in X$ is the length of the shortest path on the surface from x to y. On S^2, there is a unique great circle through $x \neq y$, namely the intersection of the plane spanned by x and y in $\mathbb{R}^3$ with S_2. The shortest path follows the great circle from x to y in the shorter direction. Since any path from x to y and on to z has to be at least as long as the shortest path from x to z, the triangle inequality holds.

(4) Let X be a closed subset of $\mathbb{R}^n$; and let $\mathcal{H}(X)$ denote the collection of all non-empty closed bounded subsets of X. If $A \in \mathcal{H}(X)$ and $b \in X$, let

$$d(b,A) = \inf_{a \in A} \|a - b\| = \min_{a \in A} \|a - b\|.$$

The minimum is obtained because $f(a) = \|a - b\|$ is continuous on the closed bounded set A, and so the Extreme Value Theorem guarantees the minimum is attained. In particular, if $b \notin A$, then $d(b,A) > 0$. Indeed, if $d(b,A) = 0$, then there is a sequence $a_n \in A$ such that $\|b - a_n\| \to 0$. Therefore $\lim_{n \to \infty} a_n = b$. Since A is closed, $b \in A$, contrary to our assumption.

The *Hausdorff metric* on $\mathcal{H}(X)$ is given by

$$d_H(A,B) = \max\Big\{ \sup_{a \in A} d(a,B), \;\; \sup_{b \in B} d(b,A) \Big\}.$$

Again these supremums are obtained. Note that $d_H(A,B) < \infty$ because A and B are bounded. If $A \neq B$, then there is a point in A or B not in the other. For definiteness, suppose that $a \in A \setminus B$. Then $d_H(A,B) \geq d(a,B) > 0$.

The symmetry property is obvious.

Let us verify the triangle inequality. Let $A, B, C \in \mathcal{H}(X)$. Fix $a \in A$, and let $b \in B$.

$$\begin{aligned} d(a,C) &= \inf_{c \in C} \|a - c\| \leq \inf_{c \in C} \|a - b\| + \|b - c\| \\ &= \|a - b\| + \inf_{c \in C} \|b - c\| = \|a - b\| + d(b,C) \\ &\leq \|a - b\| + d_H(B,C). \end{aligned}$$

Since this is valid for any $b \in B$, it is true if we take the infimum over B;

$$d(a,C) \le \inf_{b\in B} \|a-b\| + d_H(B,C) = d(a,B) + d_H(B,C).$$

Now take the supremum over all $a \in A$;

$$\sup_{a\in A} d(a,C) \le \sup_{a\in A} d(a,B) + d_H(B,C) \le d_H(A,B) + d_H(B,C).$$

Now reverse the role of A and C to obtain

$$\sup_{c\in C} d(c,A) \le d_H(C,B) + d_H(B,A) = d_H(A,B) + d_H(B,C).$$

Finally taking the maximum of these last two quantities proves the triangle inequality:

$$d_H(A,C) \le d_H(A,B) + d_H(B,C).$$

(5) Here is an even crazier example, which we will explore during this course: the p-adic metric on $\mathbb{Q}$. Fix a prime p. If $x \ne 0 \in \mathbb{Q}$, we can factor $x = p^a \frac{r}{s}$ where $a, r, s \in \mathbb{Z}$ and $\gcd(r,p) = 1 = \gcd(s,p) = \gcd(r,s)$. That is, we pull out all factors of p leaving something relatively prime. We define what is called a norm by number theorists: $|x|_p := p^{-a}$ for $x \ne 0$ and $|0|_p = 0$. Then define $d_p(x,y) = |x-y|_p$. The function $|\cdot|_p$ satisfies

(i) $|x|_p = 0$ if and only if $x = 0$.

(ii) $|xy|_p = |x|_p\,|y|_p$.

(iii) $|x \pm y|_p \le \max\{|x|_p, |y|_p\}$.

Note that (ii) is easy to check. Now (iii) will follow from the p-adic triangle inequality by taking 0 as the intermediate point. Notice that x and y are close if $x-y$ is divisible by a large positive power of p. For example, the sequence $a_n = p^n$ converges to 0 in this metric because

$$d_p(a_n, 0) = |a_n|_p = p^{-n} \to 0.$$

Again it is clear that (1) and (2) hold, so we need to verify the triangle inequality. In fact, it satisfies the strong triangle inequality:

$$d_p(x,z) \le \max\big\{d_p(x,y),\ d_p(y,z)\big\} \quad \text{for} \quad x,y,z \in \mathbb{Q}. \tag{1.2.3}$$

Let us factor $x - y = p^a \frac{r}{s}$ and $y - z = p^b \frac{t}{u}$ where r, s, t, u are integers relatively prime to p, and $a, b \in \mathbb{Z}$. First suppose that $a < b$. Then

$$x - z = (x-y) + (y-z) = p^a\Big(\frac{r}{s} + \frac{p^{b-a}t}{u}\Big) = p^a\Big(\frac{ru + p^{b-a}st}{su}\Big).$$

It is easy to check that ru and su are relatively prime to p, and thus so is $ru+p^{b-a}st$. When the fraction is reduced to lowest terms, this remains the case. Therefore,

$$d_p(x,z) = p^{-a} = d_p(x,y) = \max\{p^{-a}, p^{-b}\} = \max\{d_p(x,y), d_p(y,z)\}.$$

By symmetry, this also holds when $a > b$. So consider the case when $a = b$. Then

$$x - z = (x - y) + (y - z) = p^a\Big(\frac{r}{s} + \frac{t}{u}\Big) = p^a\Big(\frac{ru + st}{su}\Big).$$

As before, the denominator su is relatively prime to p. However the numerator $ru + st$ factors as $ru + st = p^c v$ for a non-negative integer c and an integer v relatively prime to p. Therefore $x - z = p^{a+c}\frac{v}{su}$. Hence

$$d_p(x, z) = p^{-a-c} \le p^{-a} = \max\{d_p(x, y), d_p(y, z)\}.$$

1.2.4. DEFINITION. If (X, d) is a metric space and $Y \subset X$, then (Y, d) has the *induced metric* $d(y_1, y_2)$ obtained by restricting d to $Y \times Y$.

Two metrics d and d' on X are called *equivalent metrics* if there are constants $0 < c \le C < \infty$ so that

$$cd(x_1, x_2) \le d'(x_1, x_2) \le Cd(x_1, x_2) \quad \text{for all} \quad x_1, x_2 \in X.$$

1.2.5. EXAMPLE. Let S^1 denote the unit circle in $\mathbb{C}$, namely

$$S^1 = \{z \in \mathbb{C} : |z| = 1\} = \{e^{i\theta} : 0 \le \theta \le 2\pi\}.$$

Let ρ be the geodesic distance around the circle. This is easily seen to be

$$\rho(e^{i\theta_1}, e^{i\theta_2}) = \min\{|\theta_1 - \theta_2|, 2\pi - |\theta_1 - \theta_2|\}.$$

Now S^1 also has an induced metric d from the Euclidean norm on $\mathbb{C}$. A simple calculation using trigonometry shows that

$$d(e^{i\theta_1}, e^{i\theta_2}) = \big|e^{i\theta_1} - e^{i\theta_2}\big| = 2\sin\tfrac{1}{2}\rho(e^{i\theta_1}, e^{i\theta_2}).$$

Now on the interval $[0, \pi/2]$, the function $f(x) = \sin x$ is concave down because $f''(x) = -\sin x < 0$ on $(0, \pi/2)$. So $\frac{2}{\pi}x \le \sin x \le x$ on $[0, \pi/2]$. Since $\frac{1}{2}\rho(e^{i\theta_1}, e^{i\theta_2})$ lies in $[0, \pi/2]$, we deduce that

$$\frac{2}{\pi}\rho(e^{i\theta_1}, e^{i\theta_2}) \le d(e^{i\theta_1}, e^{i\theta_2}) \le \rho(e^{i\theta_1}, e^{i\theta_2})$$

So ρ and d are equivalent metrics on S^1.

Exercises

1. (a) Show that the l_1 norm satisfies the triangle inequality. When is it an equality?
(b) Show that the l_∞ norm satisfies the triangle inequality. When is it equality in $l_\infty^{(n)}$?

2. Let $C^n[a, b]$ be the vector space of functions on $[a, b]$ with n continuous derivatives. Prove that $\|f\|_{C^n} = \max_{0\le i\le n} \sup_{a\le x\le b} |f^{(i)}(x)|$ is a norm.

3. Prove that the unit ball of a normed vector space is convex, i.e. if $\|x\| \le 1$ and $\|y\| \le 1$, then $\|tx + (1 - t)x\| \le 1$ for all $0 < t < 1$.

4. If (x, d) is a metric space, prove that $\big|d(x, y) - d(x, z)\big| \le d(y, z)$.

5. (a) Prove that the $l_1^{(n)}$ and $l_\infty^{(n)}$ norms on $\mathbb{R}^n$ yield equivalent metrics.
(b) Show that this is not true for l_1 and l_∞ norms on the subspace V of sequences which are non-zero on only finitely many coordinates.

6. Define $d : [0, 2\pi)^2 \to \mathbb{R}_+$ by $d(x, y) = \min\{|x-y|, 2\pi - |x-y|\}$.
(a) Prove that this is a metric.
(b) Show that the map $f(x) = e^{ix}$ (or $(\cos x, \sin x)$ in $\mathbb{R}^2$) is an isometric map of $([0, 2\pi), d)$ onto the unit circle with the geodesic metric.

7. Put a metric ρ on all the words in a dictionary by defining the distance between two distinct words to be 2^{-n} if the words agree for the first n letters and are different at the $(n+1)$st letter. A space is distinct from a letter. E.g., $\rho(\text{car}, \text{cart}) = 2^{-3}$ and $\rho(\text{car}, \text{call}) = 2^{-2}$.
(a) Verify that this is a metric.
(b) Suppose that words w_1, w_2 and w_3 are listed in alphabetical order. Find a formula for $\rho(w_1, w_3)$ in terms of $\rho(w_1, w_2)$ and $\rho(w_2, w_3)$.

8. Let $X = 2^{\mathbb{N}} = \big\{\mathbf{x} = (x_1, x_2, \dots) : x_i \in \{0, 1\}\big\}$ and define

$$d(\mathbf{x}, \mathbf{y}) = 2 \sum_{i \geq 1} 3^{-i} |x_i - y_i|.$$

(a) Prove that this is a metric.
(b) Define $f : X \to [0, 1]$ by $f(\mathbf{x}) = d(\mathbf{0}, \mathbf{x})$, where $\mathbf{0} = (0, 0, 0, \dots)$. Prove that this maps X onto the Cantor set and satisfies $\frac{1}{3} d(\mathbf{x}, \mathbf{y}) \leq |f(\mathbf{x}) - f(\mathbf{y})| \leq d(\mathbf{x}, \mathbf{y})$ for $\mathbf{x}, \mathbf{y} \in 2^{\mathbb{N}}$.

9. Let V be a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ with a metric d. Say that d is *translation invariant* if $d(x+z, y+z) = d(x, y)$ for all $x, y, z \in V$. Say that d is positive homogeneous if $d(\lambda x, \lambda y) = |\lambda| d(x, y)$ for all $x, y \in V$ and $\lambda \in \mathbb{F}$. Prove that there is a norm on V so that $d(x, y) = \|x - y\|$ if and only if d is translation invariant and positive homogeneous.

10. Let X be a closed subset of $\mathbb{R}^n$, and put the Hausdorff metric on $\mathcal{H}(X)$. If $r > 0$ and $A \subset X$, let $A_r = \{x \in X : d(x, A) \leq r\}$. Show that

$$d_H(A, B) = \inf\{r \geq 0 : A \subset B_r \text{ and } B \subset A_r\}.$$

(*Note*: this is actually a minimum.)

11. A *pseudometric* on X is a map $d : X^2 \to \mathbb{R}_+$ which is symmetric, satisfies the triangle inequality and $d(x, x) = 0$ for $x \in X$.
(a) Let d be a pseudometric on X. Define a relation on X by $x \sim y$ if $d(x, y) = 0$. Prove that this is an equivalence relation.
(b) Define Y to be the set of equivalence classes $[x]$. Define $\rho([x], [y]) = d(x, y)$. Show that this is a well-defined metric on Y.
(c) Now suppose that V is a vector space over $\mathbb{R}$ with a seminorm $\|\cdot\|$. Define equivalence classes in the same way as (a). Show that $N = [0]$ is a subspace, and that the space of equivalence classes is just the quotient vector space $Y = V/N$. Also show that the metric on Y comes from a norm.

1.3. Topology of Metric spaces

In this section, we learn about open and closed sets. In the more general context of topological spaces, everything is determined by the collection of open sets. In metric spaces, there is a trade-off between topological notions that only depend on the open sets and the quantitative aspect that comes from the distance function.

1.3.1. DEFINITION. Let (X, d) be a metric space. The *open ball* about $x \in X$ of radius $r > 0$ is

$$b_r(x) = \{y \in X : d(x, y) < r\}.$$

The *closed ball* about $x \in X$ of radius $r \geq 0$ is

$$\bar{b}_r(x) = \{y \in X : d(x, y) \leq r\}.$$

A subset $N \subset X$ is a *neighbourhood* of $x \in X$ if there is some $r > 0$ so that $b_r(x) \subset N$.
A subset $U \subset X$ is *open* if for all $x \in U$, there is an $r > 0$ so that $b_r(x) \subset U$.
A set $C \subset X$ is *closed* if its complement $C^c := X \setminus C$ is open.

1.3.2. PROPOSITION. *$b_r(x)$ is open for $r > 0$ and $\bar{b}_r(x)$ is closed for $r \geq 0$.*

PROOF. Let $y \in b_r(x)$, say $d := d(y, x) < r$. We claim that $b_{r-d}(y) \subset b_r(x)$. Indeed, if $z \in b_{r-d}(y)$, then $d(z, y) < r - d$; whence

$$d(x, z) \leq d(x, y) + d(y, z) < d + (r - d) = r.$$

Thus $z \in b_r(x)$.

If $y \notin \bar{b}_r(x)$, then $d := d(y, x) > r$. We claim that $b_{d-r}(y) \subset X \setminus \bar{b}_r(x)$. Indeed, if $z \in b_{d-r}(y)$, then $d(z, y) < d-r$; so using the reverse triangle inequality,

$$d(x, z) \geq d(x, y) - d(y, z) > d + (d - r) = r.$$

Thus $z \notin \bar{b}_r(x)$. So $X \setminus \bar{b}_r(x)$ is open, and hence $\bar{b}_r(x)$ is closed. ∎

1.3.3. REMARK. Note that a neighbourhood does not need to be open. Indeed an open set is a set which is a neighbourhood of each of its elements. Some books use a different convention, but there is good reason to use this terminology.

Closed is not the opposite of open. Many sets are neither open nor closed. For example, in $\mathbb{R}$, the sets $(a, b]$, $\mathbb{Q}$ and $\{\frac{1}{n} : n \geq 0\}$ are neither open nor closed in the Euclidean metric.

Points are closed sets because $\{x\} = \bar{b}_0(x)$.

1.3.4. EXAMPLES.

(1) $U := \{(x, y) \in \mathbb{R}^2 : xy > 1\}$ is open in $\mathbb{R}^2$. To see this, let us assume that $x > 0$ and $(x, y) \in U$, say $xy = 1 + \varepsilon > 1$ (the case $x < 0$ is basically the same). Define $r = \min\left\{\frac{x}{2}, \frac{y}{2}, \frac{\varepsilon}{2(x+y)}\right\}$. Suppose that $(u, v) \in b_r((x, y))$, i.e.

$(u-x)^2+(v-y)^2<r^2$. Then in particular, $u>x-r>0$ and $v>y-r>0$. Thus

$$uv>(x-r)(y-r)=xy-r(x+y)+r^2>1+\varepsilon-\frac{\varepsilon}{2}=1+\frac{\varepsilon}{2}>1.$$

Therefore $b_r((x,y))\subset U$.

(2) $C:=\{(x,y)\in\mathbb{R}^2 : xy\geq 1\}$ is closed in $\mathbb{R}^2$. To see this, note that $U=C^c=\{(x,y)\in\mathbb{R}^2 : xy<1\}$. This can be shown to be open as in the previous example.

(3) The entire set X is always open. Also the empty set $\emptyset$ is open because there are no points in the set, and so each point in the set is contained in an open ball. Therefore X and $\emptyset$ are also closed sets.

(4) Let $(\mathbb{N},d_2)$ be $\mathbb{N}$ with the 2-adic metric coming from $(\mathbb{Q},d_2)$. The distance between any two points is always a power of 2. For $n\in\mathbb{N}$ and $d\geq 0$, the closed ball $\bar{b}_{2^{-d}}(n)=\{m\in\mathbb{N} : 2^d|m-n\}$. (Here $a|b$ means that a divides b in $\mathbb{N}$.) Thus

$$\begin{aligned}\bar{b}_{2^{-d}}(n)&=\{m\in\mathbb{N} : d_2(m,n)\leq 2^{-d}\}\\&=\{m\in\mathbb{N} : d_2(m,n)<2^{1-d}\}=b_{2^{1-d}}(n).\end{aligned}$$

Thus these closed balls are also open balls. Sets which are both closed and open are called *clopen*. The exception is the singleton $\{n\}$ which is closed but not open.

1.3.5. Proposition. (a) *If $\{U_\lambda : \lambda\in\Lambda\}$ is a collection of open sets, then $\bigcup_{\lambda\in\Lambda}U_\lambda$ is open. Likewise if $\{C_\lambda : \lambda\in\Lambda\}$ is a collection of closed sets, then $\bigcap_{\lambda\in\Lambda}C_\lambda$ is closed.*

(b) *If $U_1,\dots,U_n$ is a finite collection of open sets, then $\bigcap_{i=1}^n U_i$ is open. Likewise if $C_1,\dots,C_n$ is a finite collection of closed sets, then $\bigcup_{i=1}^n C_i$ is closed.*

Proof. (a) Let $x\in U:=\bigcup_{\lambda\in\Lambda}U_\lambda$. Then there is some $\lambda_0\in\Lambda$ so that $x\in U_{\lambda_0}$. Since U_{λ_0} is open, there is an $r>0$ so that $b_r(x)\subset U_{\lambda_0}$. Therefore $b_r(x)\subset U$; whence U is open.

If C_λ is closed, then $U_\lambda=C_\lambda^c$ is open. Since

$$X\setminus\bigcap_{\lambda\in\Lambda}C_\lambda=\bigcup_{\lambda\in\Lambda}U_\lambda$$

is open by the first paragraph, it follows that $\bigcap_{\lambda\in\Lambda}C_\lambda$ is closed.

(b) Let $V=\bigcap_{i=1}^n U_i$ and let $x\in V$. Then $x\in U_i$ for $1\leq i\leq n$. Since U_i is open, there is some $r_i>0$ so that $b_{r_i}(x)\subset U_i$. Define $r=\min\{r_i : 1\leq i\leq n\}$. Then $b_r(x)\subset U_i$ for $1\leq i\leq n$; and thus $b_r(x)\subset V$. Hence V is open.

If C_i are closed, then $U_i:=C_i^c$ is open. Observe that

$$X\setminus\bigcup_{i=1}^n C_i=\bigcap_{i=1}^n U_i$$

is open; whence $\bigcup_{i=1}^n C_i$ is closed. ∎

1.3.6. DEFINITION. A sequence $\{x_n\}_{n\geq 1}$ in a metric space (X, d) *converges to* x_0 if $\lim_{n\to\infty} d(x_n, x_0) = 0$. We write $\lim_{n\to\infty} x_n = x_0$. That is, for any $\varepsilon > 0$, there is an $N \in \mathbb{N}$ so that $d(x_n, x_0) < \varepsilon$ for all $n \geq N$. Symbolically, $\forall_{\varepsilon>0} \exists_{N\in\mathbb{N}} \forall_{n\geq N}\, d(x_n, x_0) < \varepsilon$.

A *subsequence* of $\{x_n\}_{n\geq 1}$ is a sequence $\{x_{n_i}\}_{i\geq 1}$ where $n_i < n_{i+1} \in \mathbb{N}$ for $i \geq 1$.

1.3.7. DEFINITION. If (X, d) is a metric space and $A \subset X$, the *closure of* A, denoted $\overline{A}$, is the smallest closed set containing A.

If there are points $a_n \in A$ such that $\lim_{n\to\infty} a_n = a_0$, say that a_0 is a *limit point of* A. Moreover if one can choose the points a_n to be all distinct, then a_0 is an *accumulation point*. A point $a \in A$ is an *isolated point* if there is an open set U such that $U \cap A = \{a\}$.

1.3.8. PROPOSITION. *Let (X, d) be a metric space. A subset $A \subset X$ is closed if and only if it contains all of its limit points.*

PROOF. Suppose that A is closed and that $\{a_n\}_{n\geq 1} \subset A$ is a sequence which satisfies $\lim_{n\to\infty} a_n = a_0$. If $a_0 \notin A$, then it belongs to the open set $U = A^c$. Hence there is an $r > 0$ so that $b_r(a_0) \subset U$. From the definition of limit, there is an $N \in \mathbb{N}$ so that $d(a_n, a_0) < r$ for all $n \geq N$. This implies that $\{a_n : n \geq N\} \subset b_r(a_0) \subset U$, and thus $a_n \notin A$ for $n \geq N$. This is a contradiction, whence we must have $a_0 \in A$.

Conversely suppose that A is not closed. Then $U = A^c$ is not open. Therefore U contains a point a_0 so that there is no $r > 0$ such that $b_r(a_0) \subset U$. This means that for each $n \geq 1$, $b_{1/n}(a_0) \cap A$ is not empty. Pick $a_n \in A$ so that $d(a_n, a_0) < 1/n$. Then $\{a_n\}_{n\geq 1} \subset A$ and $\lim_{n\to\infty} a_n = a_0 \notin A$. This is the contrapositive of the desired statement. ∎

We now get a more refined look at limit points.

1.3.9. PROPOSITION. *Let (X, d) be a metric space, and let $A \subset X$. Then*

$$\begin{aligned}\overline{A} &= \bigcap\{C : C \supset A,\ C \text{ closed}\}\\ &= \{\text{all limit points of } A\}\\ &= A \cup \{\text{all accumulation points of } A\}\\ &= \{\text{all isolated points of } A\} \cup \{\text{all accumulation points of } A\}.\end{aligned}$$

PROOF. By Proposition 1.3.5, $\bigcap\{C : C \supset A,\ C \text{ closed}\}$ is a closed set. By definition, it contains A and is contained in any closed set containing A; and thus it is the smallest closed set containing A, namely $\overline{A}$.

The next three sets clearly each contain in the next. Indeed, each point $a \in A$ is a limit of the constant sequence $a, a, a, \dots$. But a limit point a_0 of A is either an isolated point of A or for each $r > 0$, $b_r(a_0) \cap A \setminus \{a_0\}$ is non-empty. Choose a sequence $a_n \in A$ so that $0 < d(a_{n+1}, a_0) < d(a_n, a_0)/2$. Then $a_0 = \lim_{n\to\infty} a_n$ is an accumulation point. Thus these three sets coincide.

By Proposition 1.3.8, $\overline{A}$ contains all limit points of $\overline{A}$, and in particular all limit points of A. To finish the cycle, it suffices to show that the set B consisting of all limit points of A is closed. By Proposition 1.3.8 again, it suffices to show that B contains all of its limit points. Let $b_n \in B$ such that $\lim_{n\to\infty} b_n = b_0$. As b_n is a limit point of A, we may write $b_n = \lim_{i\to\infty} a_{n,i}$ for a sequence of points $a_{n,i} \in A$. Pick i_n so that $d(a_{n,i_n}, b_n) < 1/n$; and set $a_n = a_{n,i_n}$. Since $\lim_{n\to\infty} d(a_n, b_n) = 0$, it follows that

$$0 \le \lim_{n\to\infty} d(a_n, b_0) \le \lim_{n\to\infty} d(a_n, b_n) + d(b_n, b_0) = 0.$$

Therefore b_0 is a limit point of A. Hence B is closed and contained in $\overline{A}$. Since $\overline{A}$ is the smallest closed set containing A, $\overline{A} \subset B$; and thus they are equal. ■

The following consequence is immediate.

1.3.10. Corollary. *Let (X, d) be a metric space, and let $A \subset X$. Then $\overline{\overline{A}} = \overline{A}$.*

1.3.11. Definition. The *interior of* A, written $\operatorname{int} A$, is the largest open set contained in A.

It is straightforward to check that $\operatorname{int} A = \bigcup\{b_r(a) : r > 0,\ b_r(a) \subset A\}$. Verify this. Then to check your facility with these ideas, show that

$$\operatorname{int} A = (\overline{A^c})^c =: A^{c-c}.$$

1.3.12. Example. If (X, d) and (Y, ρ) are metric spaces, we can make the product space $X \times Y := \{(x, y) : x \in X, y \in Y\}$ into a metric space with

$$D\big((x_1, y_1), (x_2, y_2)\big) = \max\{d(x_1, x_2), \rho(y_1, y_2)\}.$$

There are other natural choices for the metric, such as

$$D_p\big((x_1, y_1), (x_2, y_2)\big) = \big(d(x_1, x_2)^p + \rho(y_1, y_2)^p\big)^{1/p}$$

for $p \ge 1$, with $p = 1$ and $p = 2$ being popular choices. The reader can check that these metrics are all equivalent to D.

In $(X \times Y, D)$, $b_r((x_0, y_0)) = b_r(x_0) \times b_r(y_0)$. A sequence $\big((x_n, y_n)\big)_{n\ge1}$ converges in $X \times Y$ to (x_0, y_0) if and only if $\lim\limits_{n\to\infty} x_n = x_0$ and $\lim\limits_{n\to\infty} y_n = y_0$.

1.4. Continuous functions

The notion of continuity from calculus readily generalizes to functions between metric spaces. We distinguish between continuity using the ε–δ definition and sequential continuity using convergent sequences. It will be a theorem to show that they coincide.

1.4.1. DEFINITION. Let (X, d) and (Y, ρ) be metric spaces, and let $f : X \to Y$ be a function. Then f is *continuous at* x_0 if for all $\varepsilon > 0$, there is a $\delta > 0$ so that $d(x, x_0) < \delta$ implies that $\rho(f(x), f(x_0)) < \varepsilon$.

Say that f is *continuous* if it is continuous at x for every $x \in X$.

Also f is *uniformly continuous* if for all $\varepsilon > 0$, there is a $\delta > 0$ so that $d(x_1, x_2) < \delta$ implies that $\rho(f(x_1), f(x_2)) < \varepsilon$. Note that here, the δ does not depend on x.

Finally we say that f is *sequentially continuous* if whenever $\lim_{n\to\infty} x_n = x_0$ in X, then $\lim_{n\to\infty} f(x_n) = f(x_0)$ in Y.

The following result compares the ε–δ definition (1), the topological version of continuity (2) and the sequential version (3). Recall that if $V \subset Y$, that

$$f^{-1}(V) := \{x \in X : f(x) \in V\}.$$

1.4.2. PROPOSITION. *Let (X, d) and (Y, ρ) be metric spaces, and let $f : X \to Y$ be a function. The following are equivalent:*

(1) *f is continuous.*

(2) *$f^{-1}(V)$ is open in X for every open set $V \subset Y$.*

(3) *f is sequentially continuous.*

PROOF. $(1) \Rightarrow (2)$. Let $V \subset Y$ be open, and fix $x_0 \in f^{-1}(V)$. Since $f(x_0) =: y_0 \in V$ and V is open, there is an $\varepsilon > 0$ so that $b_\varepsilon(y_0) \subset V$. By the continuity of f at x_0, there is a $\delta > 0$ so that $d(x, x_0) < \delta$ implies that $\rho(f(x), y_0) < \varepsilon$. That means that

$$f(b_\delta(x_0)) \subset b_\varepsilon(y_0) \subset V.$$

Therefore $f^{-1}(V)$ contains $b_\delta(x_0)$. Since x_0 was an arbitrary point in $f^{-1}(V)$, it follows that $f^{-1}(V)$ is open.

$(2) \Rightarrow (3)$. Suppose that $\lim_{n\to\infty} x_n = x_0$ in X. Given any $\varepsilon > 0$, let $V = b_\varepsilon(f(x_0))$ be an open ball in Y. By (2), $f^{-1}(V)$ is open and contains x_0. Therefore there is some $\delta > 0$ so that $b_\delta(x_0) \subset f^{-1}(V)$. Therefore if $d(x_n, x_0) < \delta$, then $f(x_n) \in V$, i.e. $\rho(f(x_n), f(x_0)) < \varepsilon$. Since $\lim_{n\to\infty} x_n = x_0$, there is some integer N so that $d(x_n, x_0) < \delta$ provided that $n \geq N$. Since $\varepsilon > 0$ was arbitrary, we conclude that $\lim_{n\to\infty} f(x_n) = f(x_0)$ in Y.

(3) $\Rightarrow$ (1). Suppose that f is not continuous at some $x_0 \in X$; i.e. assume that (1) is false. Then there must be some $\varepsilon_0 > 0$ so that no $\delta > 0$ works in the definition of continuity at x_0. Hence for $\delta = \frac{1}{n}$, there is some $x_n \in b_{1/n}(x_0)$ so that $\rho(f(x_n), f(x_0)) \geq \varepsilon_0$. This means that $\lim_{n\to\infty} x_n = x_0$ but since $f(x_n)$ is bounded away from $f(x_0)$, we see that $\lim_{n\to\infty} f(x_n)$ either does not exist or it exists but if different from $f(x_0)$. So f is not sequentially continuous. Thus (3) is false. Thus $\neg(1)$ implies $\neg(3)$. The contrapositive is that (3) implies (1). ■

We collect a few easy ways to build more continuous functions.

1.4.3. PROPOSITION. *Let (X, d), (Y, ρ) and (Z, σ) be metric spaces.*
(a) *The composition of continuous functions is continuous; i.e., if $f : X \to Y$ and $g : Y \to Z$ are continuous, then $g \circ f : X \to Z$ is continuous.*
(b) *If $f : X \to Y$ and $g : X \to Z$ are continuous functions, then $h = (f, g) : X \to Y \times Z$ is continuous.*

PROOF. (a) Let $W \subset Z$ be open. Since g is continuous, $g^{-1}(W) =: V$ is open in Y. Then since f is continuous, $f^{-1}(V)$ is open in X. Therefore

$$(g \circ f)^{-1}(W) = f^{-1}(g^{-1}(W)) = f^{-1}(V)$$

is open. Hence $g \circ f$ is continuous.
(b) Recall from Example 1.3.12 that $b_r((y_0, z_0)) = b_r(y_0) \times b_r(z_0)$. Thus

$$h^{-1}(b_r((y_0, z_0))) = f^{-1}(b_r(y_0)) \cap g^{-1}(b_r(z_0))$$

is open. Every open set $V \subset Y \times Z$ is a union of balls, and the union of open sets is open; thus $h^{-1}(V)$ is open. Therefore h is continuous. ■

1.4.4. PROPOSITION. *The set of continuous functions on a metric space (X, d) with values in $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ is an algebra. That is, sums, products and scalar multiples of continuous functions are continuous.*

PROOF. Let f, g be continuous functions from X to $\mathbb{F}$. By Proposition 1.4.3(b), $h(x) = (f(x), g(x))$ is continuous into $\mathbb{F}^2$. Now the maps from $\mathbb{F}^2$ to $\mathbb{F}$ by $a(s,t) = \lambda s + \mu t$ and $m(s,t) = st$ are continuous for any scalars $\lambda, \mu \in \mathbb{F}$. By Proposition 1.4.3(b), $a \circ h$ and $m \circ h$ are continuous. Now $a \circ h(x) = \lambda f(x) + \mu g(x)$ and $m \circ h(x) = f(x)g(x)$. The result follows. ■

Some functions preserve the structure of a metric space, or at least some part of it,

1.4.5. DEFINITION. Let (X, d) and (Y, ρ) be metric spaces, and let $f : X \to Y$ be a function. Say that f is *isometric* or is an *isometry* if

$$\rho(f(x_1), f(x_2)) = d(x_1, x_2) \quad \text{for all} \quad x_1, x_2 \in X.$$

Say that f is *Lipschitz* if there is a constant $C < \infty$ so that

$$\rho(f(x_1), f(x_2)) \le Cd(x_1, x_2) \quad \text{for all} \quad x_1, x_2 \in X.$$

Say that f is *biLipschitz* if there are constants $0 < c \le C < \infty$ so that

$$cd(x_1, x_2) \le \rho(f(x_1), f(x_2)) \le Cd(x_1, x_2) \quad \text{for all} \quad x_1, x_2 \in X.$$

Say that f is a *homeomorphism* if it is a continuous bijection such that f^{-1} is also continuous.

An isometry preserves the distance, and clearly is biLipschitz with $c = C = 1$. In particular it is injective. If f is a surjective isometry, then the inverse map is also an isometry, and in particular f is a homeomorphism. Such a map preserves all of the structure of the metric space. A bijection which is biLipschitz has an inverse which is also biLipschitz with constants C^{-1} and c^{-1}. Again this will be a homeomorphism. It may stretch or contract the distance a limited amount, but it preserves a lot of the structure. The original metric d will be equivalent to the metric $\sigma(x_1, x_2) = \rho(f(x_1), f(x_2))$.

A homeomorphism preserves open sets. That is, if V is open in Y, then $f^{-1}(V)$ is open in X because f is continuous; and if U is open in X, then $f(U) = (f^{-1})^{-1}(U)$ is open because f^{-1} is continuous. However it may overstretch or understretch the metric so that certain quantitative things change.

The following very easy result is left as an exercise.

1.4.6. Proposition. *Lipschitz maps are uniformly continuous.*

1.4.7. Examples.

(1) In Example 1.2.5, we discuss two metrics on the circle S^1. The geodesic distance ρ, and the metric d induced by the Euclidean distance in $\mathbb{C}$. These metrics were shown to be equivalent. It follows that the identity map $f : (S^1, \rho) \to (S^1, d)$ given by $f(e^{i\theta}) = e^{i\theta}$ is a biLipschitz homeomorphism.

(2) Let $f : ([0, 1), d) \to (S^1, \rho)$ map the half-open interval onto S^1 by $f(t) = e^{2\pi it}$. It is clear that this map is Lipschitz with constant 2π, and it is a bijection. However it is not a homeomorphism because the sequence $t_n = \frac{n}{n+1}$ has no limit in $[0, 1)$ while $f(t_n)$ has the limit 1 in S^1. So f^{-1} is not continuous at 1. Another way to see this is that $d(0, t_n) = t_n \to 1$ while

$$\rho(f(0), f(t_n)) = |1 - e^{-2\pi i/(n+1)}| = 2 \sin \tfrac{\pi}{n+1} \to 0.$$

(3) Let X be a *convex* subset of $\mathbb{R}^n$, and let $F : X \to \mathbb{R}^m$ be a differentiable function. If the derivative DF is bounded, then F is Lipschitz. Remember that if $F = (f_1, \dots, f_m)$ where f_j is the jth coordinate of F, then $DF(x)$ is the $n \times m$ matrix $\left[\frac{\partial f_j}{\partial x_i}\right]$. We will say that DF is bounded if each $\frac{\partial f_j}{\partial x_i}$ is bounded for $1 \le i \le n$ and $1 \le j \le m$. Suppose that each coordinate is bounded by a constant C. This occurs, for example, when F is C^1 and X is closed and bounded by the Extreme Value Theorem.

We need to estimate $\|F(x_1) - F(x_2)\|$. To do this, let u be a unit vector in $\mathbb{R}^m$ colinear with $F(x_1) - F(x_2)$. Then

$$\|F(x_1) - F(x_2)\| = |(F(x_1) - F(x_2)) \cdot u|$$

Consider the scalar valued function $f(x) = F(x) \cdot u = \sum_{j=1}^m u_j f_j(x)$. Note that

$$\frac{\partial f}{\partial x_i}(x) = \sum_{j=1}^m u_j \frac{\partial f_j}{\partial x_i}(x).$$

Therefore the gradient of f is given by

$$\nabla f(x) = u \cdot DF(x) = [u_1 \ \dots \ u_m] \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix}$$

If $d_{ij} = \frac{\partial f_j}{\partial x_i}(x)$, then $DF(x) = \begin{bmatrix} d_{ji} \end{bmatrix}$. The Cauchy-Schwarz inequality yields

$$\|\nabla f(x)\|^2 = \sum_{i=1}^n \big(\sum_{j=1}^m u_j d_{ji} \big)^2 \le \sum_{i=1}^n \|u\|^2 \sum_{j=1}^m d_{ji}^2 \le nmC^2.$$

Since X is convex, we can apply the Mean Value Theorem to f restricted to the line segment $[x_1, x_2]$. We obtain a point $\xi \in (x_1, x_2)$ so that

$$|f(x_1) - f(x_2)| = \big|(x_1 - x_2) \cdot \nabla f(\xi)\big| \le \|x_1 - x_2\| \sqrt{nm}C.$$

Hence $\|F(x_1) - F(x_2)\| = |f(x_1) - f(x_2)| \le \sqrt{nm}C\|x_1 - x_2\|$; i.e. it is Lipschitz.

(4) Let $f : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ by $f(x) = \tan x$. Then f is a continuous bijection. Moreover, $f^{-1}(y) = \tan^{-1}(y)$ is also continuous. Hence f is a homeomorphism. The derivative $f'(x) = \sec^2 x$ blows up as $x \to \pm\frac{\pi}{2}$. It follows that f is not Lipschitz. Indeed, if $\frac{\pi}{2} - \varepsilon < x_1 < x_2 < \frac{\pi}{2}$, then the Mean Value Theorem provides some $\xi \in (x_1, x_2)$ so that

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi) > \sec^2(\tfrac{\pi}{2} - \varepsilon) = \csc^2 \varepsilon > \frac{1}{\varepsilon^2}.$$

On the other hand, $(f^{-1})' = \frac{1}{1+y^2} \le 1$. So f^{-1} is Lipschitz with constant 1.

1.5. Finite dimensional normed vector spaces

Recall that a normed vector space $(V, \|\cdot\|)$ has an induced metric $d(u, v) = \|u - v\|$. If $(V, |\!|\!|\cdot|\!|\!|)$ is another norm on V, we say that the two norms are equivalent if there are constants $0 < c \le C < \infty$ so that

$$c\|v\| \le |\!|\!|v|\!|\!| \le C\|v\| \quad \text{for all} \quad v \in V.$$

It is an easy exercise to see that this is the same as saying that the metrics that they induce are equivalent.

1.5.1. EXAMPLE. Consider $\mathbb{R}^n$ with the Euclidean norm $\|v\|_2 = \left(\sum_{i=1}^n v_i^2\right)^{1/2}$ and the $l_1^{(n)}$ norm $\|v\|_1 = \sum_{i=1}^n |v_i|$. Observe that by the Cauchy-Schwarz inequality,

$$\|v\|_1 = \sum_{i=1}^n |v_i| \cdot 1 \le \|v\|_2 \, \|(1, \dots, 1)\|_2 = \sqrt{n}\|v\|_2$$

On the other hand, by the triangle inequality,

$$\|v\|_2 \le \sum_{i=1}^n \|v_i e_i\|_2 = \sum_{i=1}^n |v_i| = \|v\|_1.$$

Therefore

$$\|v\|_2 \le \|v\|_1 \le \sqrt{n}\|v\|_2 \quad \text{for all} \quad v \in \mathbb{R}^n.$$

Hence these two norms are equivalent. The same argument works for $\mathbb{C}^n$.

1.5.2. THEOREM. *If V is a finite dimensional vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, then any two norms on V are equivalent.*

PROOF. Since V is finite dimensional, we have that $V \simeq \mathbb{F}^n$, where $n = \dim V$. Fix a basis $e_1, \dots, e_n$ for V. Then each $v \in V$ has the form $v = \sum_{i=1}^n v_i e_i$, with $v_i \in \mathbb{F}$. Let $\|v\|_2 = \left(\sum_{i=1}^n |v_i|^2\right)^{1/2}$ be the 2-norm. This is the usual Euclidean norm on $\mathbb{F}^n$. It suffice to show that all norms on V are equivalent to $\|\cdot\|_2$.

Let $|\!|\!|\cdot|\!|\!|$ be another norm on V. By the triangle inequality,

$$|\!|\!|v|\!|\!| \le \sum_{i=1}^n |\!|\!|v_i e_i|\!|\!| \le \sum_{i=1}^n |v_i| \, |\!|\!|e_i|\!|\!|.$$

Hence by the Cauchy-Schwarz inequality,

$$|\!|\!|v|\!|\!| \le \left(\sum_{i=1}^n |v_i|^2\right)^{1/2} \left(\sum_{i=1}^n |\!|\!|e_i|\!|\!|^2\right)^{1/2} =: C\|v\|_2,$$

where $C = \left(\sum_{i=1}^n |\!|\!|e_i|\!|\!|^2\right)^{1/2}$ is a constant. This shows that $|\!|\!|v|\!|\!|$ is a Lipschitz function on $(V, \|\cdot\|_2)$, and in particular it is continuous.

Let $S := \{v \in V : \|v\|_2 = 1\}$ be the unit sphere in V. This is a closed and bounded set in $\mathbb{F}^n$, so we can apply the Extreme Value Theorem to the continuous function $|\!|\!|v|\!|\!|$ to conclude that it attains its *minimum* value on S, say

$$c = |\!|\!|v_0|\!|\!| = \inf_{v \in S} |\!|\!|v|\!|\!|.$$

Since $0 \notin S$, we have that $c > 0$. Take any non-zero $v \in V$. Then $v/\|v\|_2$ belongs to S. Hence

$$c \leq \left\|\left| \frac{v}{\|v\|_2} \right\|\right| = \frac{\|\|v\|\|}{\|v\|_2}.$$

Therefore

$$c\|v\|_2 \leq \|\|v\|\| \leq C\|v\|_2 \quad \text{for all} \quad v \in V.$$

Thus these two norms are equivalent. ∎

1.5.3. COROLLARY. *Every vector space norm on $\mathbb{F}^n$ is biLipschitz homeomorphic to $\mathbb{F}^n$ with the Euclidean norm.*

Exercises

1. Let (X, d) be a set with the discrete metric.
(a) Describe all open sets and all closed sets.
(b) Describe all convergent sequences.
(c) Describe all accumulation points of X.

2. Let V be an inner product space with norm $\|x\| = \langle x, x\rangle^{1/2}$. Prove that a *linear* map T from $(\mathbb{F}^n, \|\cdot\|_2)$ into V is an isometry if and only if the set $\{Te_i : 1 \leq i \leq n\}$ is an orthonormal set.

3. (a) If d_1 and d_2 are equivalent, prove that (X, d_1) and (X, d_2) have the same open sets.
(b) Is the converse true? Either prove it or provide a detailed counterexample.

4. Let d_2 be the 2-adic metric on $\mathbb{Q}$.
(a) Find $\displaystyle\lim_{n\to\infty} \frac{1-(-2)^n}{3}$ in $(\mathbb{Q}, d_2)$, and show that this limit is in the closure of $\mathbb{N}$.
(b) Find the closure of $\mathbb{N}$ in $(\mathbb{Q}, d_2)$. HINT: figure out why (a) works.

5. Let (X, d) be a metric space and let $(x_n)_{n\geq 1}$ be a sequence in X, and let $x_0 \in X$.
(a) If $\lim_{n\to\infty} x_n = x_0$, then every subsequence $(x_{n_i})_{i\geq 1}$ converges to x_0.
(b) If every subsequence $(x_{n_i})_{i\geq 1}$ has a subsequence $(x_{n_{i_j}})_{j\geq 1}$ which converges to x_0, then $(x_n)_{n\geq 1}$ converges to x_0.

6. Let (X, d) be a metric space, and let $Y \subset X$ have the induced metric.
(a) Show that a subset $V \subset Y$ is open if and only if there is an open set U in X such that $V = U \cap Y$.
(b) Show that a subset $A \subset Y$ is closed if and only if there is a closed set B in X such that $A = B \cap Y$.

7. Given an example of a metric space (X, d), a point x_0 and $r > 0$ so that $\overline{b_r(x_0)}$ is properly contained in $\bar{b}_r(x_0)$.

8. Let $(V, \|\cdot\|)$ be a normed vector space. Let $A \subset V$, $x \in V$ and $\lambda \in \mathbb{F}$.
(a) Show that $x + \overline{A} = \overline{x+A}$ and $x + \operatorname{int}(A) = \operatorname{int}(x+A)$.
(b) Show that $\lambda\overline{A} = \overline{\lambda A}$. When is $\lambda \operatorname{int}(A) = \operatorname{int}(\lambda A)$?.

9. For this question, let's write A^- instead of $\bar{A}$. Consider the collection of sets obtained by repeated application of closure and complement inside (X, d). E.g., A^{-c-c-},A^{c-}.
(a) Show that if U is open and $B = U^-$, then $B = B^{c-c-}$.
(b) Use (a) to show that starting with a set A, there are at most 14 possible sets, including A itself, obtained by repeated use of closure and complement.
(c) Find a bounded subset $A \subset \mathbb{R}$ with exactly 14 different sets obtained this way.

10. Let (X, d) and (Y, ρ) be a metric spaces. Let $f, g : X \to Y$ be continuous functions. Show that $\{x \in X : f(x) = g(x)\}$ is closed.

11. Let $(V, \|\cdot\|)$ be a normed vector space. Let W be a finite dimensional subspace of V.
(a) Show that for any $v \in V$, there is a point $w \in W$ such that $\|v - w\| = d(v, W)$. HINT: Extreme Value Theorem on some closed bounded subset of W..
(b) Show that if V is an inner product space, then there is a unique closest point.
(c) Let $V = c_0$, the space of all sequences $x = (x_n)_{n\ge1}$ with $\lim_{n\to\infty} x_n = 0$ and norm $\|x\|_\infty = \sup_{n\ge1} |x_n|$. Let $W_n = \{x : x_k = 0 \text{ for } k > n\}$. Find a point $v \in c_0$ for which there are many closest points in W_n.

12. Suppose that (X, d) is a nonempty metric space.
(a) Fix $x_0 \in X$. For each $x \in X$, define $f_x(y) = d(x, y) - d(x_0, y)$ for $y \in X$. Show that f_x is a bounded continuous function on X.
(b) Show that $\|f_x - f_y\|_\infty = d(x, y)$ for all $x, y \in X$.
(c) Hence deduce that the map that takes $x \in X$ to the function f_x is an isometry that identifies X with a subset of $C^b_{\mathbb{R}}(X)$.

13. Let $X = -\mathbb{N} \cup \bigcup_{n\ge0}(2n, 2n + 1)$ and $Y = -\mathbb{N} \cup \bigcup_{n\ge0}[2n, 2n + 1)$. Show that there are continuous bijections $f : X \to Y$ and $g : Y \to X$, but X and Y are not homeomorphic.

1.6. Completeness

By analogy to what we know in $\mathbb{R}^n$, we can define the notions of Cauchy sequence and completeness in arbitrary metric spaces. Recall that intuitively, Cauchy sequences behave like convergent sequences, but the definition avoids naming the limit point.

1.6.1. DEFINITION. Let (X, d) be a metric space. A sequence $(x_n)_{n\ge1}$ of points in X is a *Cauchy sequence* if for every $\varepsilon > 0$, there is an $N \in \mathbb{N}$ so that $d(x_m, x_n) < \varepsilon$ for all $m, n \ge N$. In symbols, $\forall_{\varepsilon>0}\ \exists_{n\in\mathbb{N}}\ \forall_{m,n\ge N}\ d(x_m, x_n) < \varepsilon$.

A metric space (X, d) is *complete* if every Cauchy sequence in X converges to a limit in X. A complete normed vector space is called a *Banach space*.

This easy proposition confirms the intuition.

1.6.2. PROPOSITION. *Convergent sequences are Cauchy sequences.*

PROOF. Suppose that $\lim_{n\to\infty} x_n = x_0$ and let $\varepsilon > 0$. Using $\varepsilon/2$, we may find an integer N so that $d(x_n, x_0) < \varepsilon/2$ for all $n \geq N$. Then if $m, n \geq N$, we have

$$d(x_m, x_n) \leq d(x_m, x_0) + d(x_0, x_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence $(x_n)_{n\geq 1}$ is a Cauchy sequence. ∎

1.6.3. EXAMPLES.
(1) If X has the discrete topology, a Cauchy sequences is eventually constant. Take $\varepsilon = 1$ and find N; then $d(x_n, x_N) < 1$ means that $x_n = x_N$ for all $n \geq N$. So X is complete.

(2) $\mathbb{R}$ and $\mathbb{R}^n$ are complete. We will review this soon.

(3) Let $X = (-1, 1)$ with the Euclidean metric d induced from $\mathbb{R}$. Then (X, d) is not complete because $x_n = \frac{n}{n+1}$ is Cauchy, but has no limit in X.

However, let $f(x) = \tan(\pi x/2)$. Then f is a strictly increasing map of $(-1, 1)$ onto $\mathbb{R}$, and so it is injective. Define a metric $\rho(x, y) = |f(x) - f(y)|$. (Check the triangle inequality!) Suppose that $(x_n)_{n\geq 1}$ is Cauchy in (X, ρ). Then by definition, $f(x_n)$ is a Cauchy sequence in $\mathbb{R}$. By the completeness of $\mathbb{R}$, this sequence converges, say $\lim_{n\to\infty} f(x_n) = y$. Let $x = f^{-1}(y) = \frac{2}{\pi}\tan^{-1}(y)$. Now $x \in (-1, 1)$ and $\rho(x_n, x) = |f(x_n) - y| \to 0$. So $\lim_{n\to\infty} x_n = x$ in the ρ metric. Therefore (X, ρ) is complete.

The identity map from (X, d) to (X, ρ) is a continuous bijection, and the inverse map is also continuous (check!). Thus these two spaces are homeomorphic. However the map is not biLipschitz. Indeed $f'(x) = \frac{\pi}{2}\sec^2(\pi x/2)$ is unbounded, so f is not Lipschitz. The map $f^{-1}(y) = \frac{2}{\pi}\tan^{-1}(y)$ has derivative $f^{-1\prime}(y) = \frac{2}{\pi(1+y^2)}$ is bounded by $\frac{2}{\pi}$, so f^{-1} is Lipschitz.

1.6.4. PROPOSITION. *Suppose that (X, d) is a complete metric space, and that $Y \subset X$ is a subset with the induced metric. Then (Y, d) is complete if and only if Y is closed in X.*

PROOF. $(\Rightarrow)$ Let $x \in \overline{Y}$. Then there is a sequence $(y_n)_{n\geq 1}$ in Y with $x = \lim_{n\to\infty} y_n$. By Proposition 1.6.2, this is a Cauchy sequence. Since Y is complete, the sequence has a limit in Y, namely $x \in Y$. Thus Y is closed.

$(\Leftarrow)$ Let $(y_n)_{n\geq 1}$ be a Cauchy sequence in Y. Since X is complete, and this sequence is also Cauchy in X, $x = \lim_{n\to\infty} y_n$ exists in X. Because Y is closed, $x \in Y$. So the sequence converges in Y. Thus Y is complete. ∎

1.6.5. THEOREM. *The normed vector space l_p for $1 \leq p < \infty$ is complete.*

PROOF. Let $\mathbf{x}_n = (x_{n1}, x_{n2}, x_{n3}, \dots)$ for $n \geq 1$ be a Cauchy sequence in l_p. Given $\varepsilon > 0$, there is an $N = N(\varepsilon)$ so that $\|\mathbf{x}_m - \mathbf{x}_n\|_p < \varepsilon$ for all $m, n \geq N$. For each $j \geq 1$,

$$|x_{mj} - x_{nj}| \leq \|\mathbf{x}_m - \mathbf{x}_n\|_p.$$

Therefore the sequence $(x_{nj})_{n\geq 1}$ is a Cauchy sequence in $\mathbb{R}$. As $\mathbb{R}$ is complete, $\lim_{n\to\infty} x_{nj} =: x_j$ exists. Let $\mathbf{x} = (x_1, x_2, x_3, \dots)$. This is a sequence, but we do not yet know that it lies in l_p.

Now fix an integer J. Then for all $N \leq m \leq n$,

$$\sum_{j=1}^{J} |x_{nj} - x_{mj}|^p \leq \|\mathbf{x}_m - \mathbf{x}_n\|_p^p < \varepsilon^p.$$

Keep m fixed and let $n \to \infty$. Since this is a finite sum and each term converges, we obtain

$$\sum_{j=1}^{J} |x_j - x_{mj}|^p \leq \varepsilon^p \quad \text{for all} \quad m \geq N.$$

Now let $J \to \infty$ to conclude that

$$\|\mathbf{x} - \mathbf{x}_m\|_p = \sum_{j=1}^{\infty} |x_j - x_{mj}|^p \leq \varepsilon^p \quad \text{for all} \quad m \geq N. \tag{1.6.6}$$

In particular, Minkowski's inequality shows that

$$\|\mathbf{x}\|_p \leq \|\mathbf{x}_m\| + \|\mathbf{x} - \mathbf{x}_m\|_p < \infty.$$

So $\mathbf{x} \in l_p$. Finally, (1.6.6) shows that $\mathbf{x} = \lim_{m\to\infty} \mathbf{x}_m$. So l_p is complete. ■

The case of l_∞ will be established in section 1.8 based on the fact that $l_\infty = C^b(\mathbb{N})$, where $\mathbb{N}$ has the discrete topology.

1.6.7. DEFINITION. If $(V, \|\cdot\|)$ is a normed vector space, let $\mathcal{L}(V, \mathbb{F})$ denote the vector space of linear maps of V into the scalars, *linear functionals* , and let V^* denote the *dual space of* V of all continuous linear functionals.

1.6.8. PROPOSITION. *Let* $(V, \|\cdot\|)$ *be a normed vector space, and let* $\varphi \in \mathcal{L}(V, \mathbb{F})$. *The following are equivalent:*

(1) φ *is continuous.*

(2) $\|\varphi\|_* := \sup\{|\varphi(v)| : \|v\| \leq 1\} < \infty$.

(3) φ *is continuous as* $v = 0$.

PROOF. (2) $\Rightarrow$ (1). If $u \neq v \in V$, set $w = \frac{u-v}{\|u-v\|}$ and note that

$$|\varphi(u) - \varphi(v)| = |\varphi(u - v)| = |\varphi(w)|\,\|u - v\| \leq \|\varphi\|_* \|u - v\|.$$

Hence φ is Lipschitz, and in particular is continuous.

(1) $\Rightarrow$ (3) is trivial.

(3) $\Rightarrow$ (2). Assume that (2) fails. Then there are vectors $v_n \in V$ with $\|v_n\| = 1$ and $|\varphi(v_n)| > n^2$. Thus $\frac{1}{n}v_n \to 0$ while $|\varphi(\frac{1}{n}v_n)| > n$ diverges. So φ is discontinuous at 0. The result follows. ■

1.6.9. THEOREM. *Let $(V, \|\cdot\|)$ be a normed vector space. Then $(V^*, \|\cdot\|_*)$ is a Banach space.*

PROOF. First we show that $\|\cdot\|_*$ is a norm. Clearly $\|\varphi\|_* = 0$ if and only if $\varphi(v) = 0$ for all $v \in V$ with $\|v\| \le 1$. This forces $\varphi = 0$ by linearity. Also if $\lambda \in \mathbb{F}$, then

$$\|\lambda\varphi\|_* = \sup_{\|v\|\le 1} |\lambda\varphi(v)| = |\lambda| \sup_{\|v\|\le 1} |\varphi(v)| = |\lambda|\, \|\varphi\|_*.$$

For the triangle inequality, take $\varphi, \psi \in V^*$.

$$\begin{aligned}\|\varphi+\psi\|_* &= \sup_{\|v\|\le 1} |\varphi(v)+\psi(v)| \le \sup_{\|v\|\le 1} |\varphi(v)| + |\psi(v)| \\ &\le \sup_{\|v\|\le 1} |\varphi(v)| + \sup_{\|v\|\le 1} |\psi(v)| = \|\varphi\|_* + \|\psi\|_*.\end{aligned}$$

To establish completeness, let $(\varphi_n)_{n\ge1}$ be a Cauchy sequence in V^*. For each $v \in V$, $|\varphi_m(v)-\varphi_n(v)| \le \|\varphi_m-\varphi_n\|_*\, \|v\|$. It follows that $(\varphi_n(v))_{n\ge1}$ is a Cauchy sequence in $\mathbb{F}$. Since $\mathbb{F}$ is complete, we may define $\varphi(v) = \lim_{n\to\infty} \varphi_n(v)$. Then

$$\begin{aligned}\varphi(\lambda u + \mu v) &= \lim_{n\to\infty} \varphi_n(\lambda u + \mu v) \\ &= \lim_{n\to\infty} \lambda\varphi_n(u) + \mu\varphi_n(v) = \lambda\varphi(u) + \mu\varphi(v).\end{aligned}$$

Therefore φ is linear. Now let $\varepsilon > 0$ and select N so that if $m, n \ge N$, then $\|\varphi_m - \varphi_n\|_* < \varepsilon$. In particular, if $\|v\| \le 1$, we have that $|\varphi_m(v) - \varphi_n(v)| < \varepsilon$. Holding m fixed and letting $n \to \infty$, we obtain that $|\varphi_m(v) - \varphi(v)| \le \varepsilon$. Taking the supremum over all v with $\|v\| \le 1$ yields $\|\varphi_m - \varphi\|_* \le \varepsilon$ when $m \ge N$. In particular, $\|\varphi\|_* \le \|\varphi_m\|_* + \|\varphi_m - \varphi\|_* < \infty$; so $\varphi \in V^*$. Moreover we have shown that $\lim_{m\to\infty} \varphi_m = \varphi$ in $(V^*, \|\cdot\|_*)$. So V^* is complete. ■

1.7. Completeness of $\mathbb{R}$ and $\mathbb{R}^n$

This is a topic covered in earlier courses, so we will review this quickly. Exactly how one gets started depends on how we define the real numbers. We will explore this important issue later.

For now, we will assume that every infinite decimal describes a unique real number. Of course, some numbers like 1 have two such infinite decimals, namely $1.000\ldots$ and $0.999\ldots$.

Another basic property of $\mathbb{R}$ that we need is that it is *Archimedean*, meaning that if $x \in \mathbb{R}$, $x \ge 0$, and $x < 10^{-n}$ for all $n \ge 1$, then $x = 0$. To see this, note that since $x \geq 0$, it has a decimal expansion $x = x_0.x_1x_2x_3\dots$ where $x_0 \geq 0$. If $x \neq 0$, there is a first non-zero digit in its expansion, say $x_n \geq 1$. Then $x \geq 10^{-n}$.

1.7.1. Least Upper Bound Principle. *If S is a non-empty subset of $\mathbb{R}$ which is bounded above (below), then S has a least upper bound (greatest lower bound).*

Proof. We will deal with the case of the lower bound. The case of the least upper bound follows from the fact that $\sup S = -(\inf -S)$.

Since S is bounded below, there is a largest integer a_0 which is a lower bound. That is

$$a_0 \leq s \text{ for all } s \in S \quad \text{and} \quad \exists s_0 \in S \text{ such that } s_0 < a_0 + 1.$$

Think of the point s_0 as a witness to the fact that $a_0 + 1$ is not a lower bound. Note that this is where the fact that S is non-empty is used.

Now consider the numbers $a_0.0, a_0.1, \dots, a_0.9$. From these, pick the largest one, say $a_0.a_1$, which is a lower bound, and pick $s_i \in S$ to witness that $a_0.a_1 + 10^{-1}$ is not:

$$a_0.a_1 \leq s \text{ for all } s \in S \quad \text{and} \quad s_1 < a_0.a_1 + 10^{-1}.$$

Recursively select $a_n \in \{0, 1, \dots, 9\}$ so that $a_0.a_1 \dots a_n$ is a lower bound for S, and $a_0.a_1 \dots a_n + 10^{-n}$ is not, and $s_n \in S$ is a witness, so that

$$a_0.a_1 \dots a_n \leq s \text{ for all } s \in S \quad \text{and} \quad s_n < a_0.a_1 \dots a_n + 10^{-n}.$$

Let $L = a_0.a_1a_2a_3\dots$.

We claim that $L = \inf S$. First, if $s \in S$, we have $a_0.a_1 \dots a_n \leq s$ for all $n \geq 1$, and so $L \leq s$. This follows from the Archimedean property, because if there were an $s \in S$ with $s < L$, then for some $n \geq 1$,

$$s \leq L - 10^{-n} < a_0.a_1 \dots a_n,$$

which contradicts our construction. So L is a lower bound. If $M > L$, then by the Archimedean property, there is some $n \geq 1$ so that

$$M > L + 10^{-n} \geq a_0.a_1 \dots a_n + 10^{-n} > s_n.$$

So M is not a lower bound. Hence $L = \inf S$. ■

1.7.2. Corollary. *A bounded monotone sequence in $\mathbb{R}$ converges.*

Proof. Suppose that $(x_n)_{n\geq 1}$ is a monotone increasing sequence in $\mathbb{R}$ which is bounded above. By the Least Upper Bound Principle, $L = \sup\{x_n : n \geq 1\}$ exists. We claim that $\lim_{n\to\infty} x_n = L$. Take any $\varepsilon > 0$. Since $L - \varepsilon$ is not an upper bound for the sequence, there is some N so that $L - \varepsilon < x_N$. Hence

$$L - \varepsilon < x_N \leq x_n \leq L \quad \text{for all} \quad n \geq N.$$

Thus $|L - x_n| < \varepsilon$ for all $n \geq N$. So $\lim_{n\to\infty} x_n = L$. ■

1.7.3. BOLZANO-WEIERSTRASS THEOREM. *Every bounded sequence in $\mathbb{R}$ has a convergent subsequence.*

PROOF. Let $(x_n)_{n\geq 1}$ be a sequence of real numbers bounded below by a_0 and bounded above by b_0. We use the disection method.

Let $y = (a_0 + b_0)/2$. Either there are infinitely many n's with $x_n \in [a_0, y]$ or there are infinitely many n's with $x_n \in [y, b_0]$ or both. If there are infinitely many in $[a_0, y]$, let $a_1 = a_0$ and $b_1 = y$. Otherwise set $a_1 = y$ and $b_1 = b_0$. In either case, pick n_1 so that $x_{n_1} \in [a_1, b_1]$. Note that $a_0 \leq a_1 < b_1 \leq b_0$ and $b_1 - a_1 = 2^{-1}(b_0 - a_0)$.

We will recursively select a_k and b_k so that $a_{k-1} \leq a_k < b_k \leq b_{k-1}$ and $b_k - a_k = 2^{-k}(b_0 - a_0)$ so that there are infinitely many n's with $x_n \in [a_k, b_k]$. Then pick $n_k > n_{k-1}$ so that $x_{n_k} \in [a_k, b_k]$.

We claim that $(x_{n_k})_{k\geq 1}$ converges. Note that $a_0, a_1, a_2, \ldots$ is a monotone increasing sequence bounded above by b_0. Hence it converges, say to A. Likewise $b_0, b_1, b_2, \ldots$ is a monotone decreasing sequence bounded below by a_0, so it also converges, say to B. However

$$B - A = \lim_{k\to\infty} b_k - a_k = \lim_{k\to\infty} 2^{-k}(b_0 - a_0) = 0.$$

So $B = A$. Now $a_k \leq x_{n_k} \leq b_k$. Therefore $\lim_{k\to\infty} x_{n_k} = A$ by the Squeeze Theorem. ■

1.7.4. LEMMA. *Cauchy sequences are bounded.*

PROOF. Let $(x_n)_{n\geq 1}$ be a Cauchy sequence in (X, d). Let $\varepsilon = 1$, and find N so that $d(x_n, x_N) < 1$ for $n \geq N$. Let

$$R = \max\{1, d(x_i, x_N) : 1 \leq i < N\}.$$

Then $(x_n)_{n\geq 1} \subset \overline{b}_R(x_N)$. ■

1.7.5. THEOREM. $\mathbb{R}$ *is complete.*

PROOF. Let $(x_n)_{n\geq 1}$ be a Cauchy sequence in $\mathbb{R}$. By the Lemma, this sequence is bounded. Hence by the Bolzano-Weierstrass Theorem, there is a subsequence $(x_{n_k})_{k\geq 1}$ converging to a limit L. We claim that $\lim_{n\to\infty} x_n = L$.

Let $\varepsilon > 0$ and pick N so that $|x_n - x_m| < \frac{\varepsilon}{2}$ for all $m, n \geq N$. Pick k so large that $n_k > N$ and $|L - x_{n_k}| < \frac{\varepsilon}{2}$. Then for $n \geq N$,

$$|L - x_n| \leq |L - x_{n_k}| + |x_{n_k} - x_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore $\lim_{n\to\infty} x_n = L$. ■

These results are circular. That is, completeness in turn implies the Least Upper Bound Principle. To see this, let S be a non-empty set bounded below. Repeat the proof of Theorem 1.7.1. But instead of claiming that the infinite decimal number L is a real number, check that $x_n = a_0.a_1 \dots a_n$ is a Cauchy sequence, and thus $L = \lim_{n\to\infty} x_n$ exists. The rest of the proof is the same.

1.7.6. COROLLARY. *$\mathbb{R}^n$ is complete.*

PROOF. Suppose that $x_k = (x_{k1}, \dots, x_{kn})$, $k \geq 1$, is a Cauchy sequence in $\mathbb{R}^n$. Then for each $1 \leq i \leq n$, the sequence $(x_{ki})_{k\geq 1}$ is Cauchy. Indeed, if $\varepsilon > 0$, choose N so that $\|x_k - x_l\| < \varepsilon$ for all $k, l \geq N$. Then, $|x_{ki} - x_{li}| \leq \|x_k - x_l\| < \varepsilon$ for all $k, l \geq N$. So by the completeness of $\mathbb{R}$, $\lim_{k\to\infty} x_{ki} = y_i$ exists for $1 \leq i \leq n$. Thus $\lim_{k\to\infty} x_k = y := (y_1, \dots, y_n)$. ■

1.7.7. COROLLARY. *If V is a normed vector space and M is a finite dimensional subspace, then M is complete and hence closed in V.*

PROOF. Let $n = \dim M$. By Corollary 1.5.3, $(M, \|\cdot\|)$ is equivalent to $\mathbb{F}^n$ with the Euclidean norm. Therefore M is complete by Exercise 1.8 (3) and Corollary 1.7.6. In particular, M must be closed in V. ■

1.8. Limits of continuous functions

1.8.1. DEFINITION. Let (X, d) and (Y, ρ) be metric spaces. A sequence of functions $f_n : X \to Y$ *converge uniformly* to f if for all $\varepsilon > 0$, there is an $N \in \mathbb{N}$ so that

$$\|f - f_n\|_\infty = \sup_{x\in X} \rho(f(x), f_n(x)) < \varepsilon \quad \text{for all} \quad n \geq N.$$

In particular, uniform convergence of bounded continuous functions in $C^b(X)$ is just convergence in the norm $\|f\|_\infty = \sup_{x\in X} |f(x)|$.

1.8.2. THEOREM. *Let (X, d) and (Y, ρ) be metric spaces. Suppose that $f_n : X \to Y$ for $n \geq 1$ is a sequence of continuous functions which converge uniformly to f. Then f is continuous. If $Y = \mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $f_n \in C^b_{\mathbb{F}}(X)$, then $f \in C^b_{\mathbb{F}}(X)$.*

PROOF. Fix $x_0 \in X$ and $\varepsilon > 0$. By uniform convergence, there is an N so that $\|f - f_N\|_\infty < \varepsilon/3$. Since f_N is continuous, there is a $\delta > 0$ so that $d(x, x_0) < \delta$ implies that $\rho(f_N(x), f_N(x_0)) < \varepsilon/3$. Then if $d(x, x_0) < \delta$,

$$\begin{aligned}\rho(f(x), f(x_0)) &\leq \rho(f(x), f_N(x)) + \rho(f_N(x), f_N(x_0)) + \rho(f_N(x_0), f(x_0)) \\ &< \|f - f_N\|_\infty + \frac{\varepsilon}{3} + \|f - f_N\|_\infty < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$

Hence f is continuous.

For $C^b(X)$, it is easy to see that convergence in the supremum norm is precisely uniform convergence. Using $\varepsilon = 1$ and the corresponding N, we see that

$$\|f\|_\infty \le \|f_N\|_\infty + \|f - f_N\|_\infty \le \|f_N\|_\infty + 1.$$

Thus f lies in $C^b(X)$. ∎

1.8.3. EXAMPLES. It is important to distinguish between uniform convergence and *pointwise convergence*. We say that f_n converges pointwise to f if

$$f(x) = \lim_{n\to\infty} f_n(x) \quad \text{for all} \quad x \in X.$$

(1) Let $f_n(x) = x^n$ in $C[0,1]$. Then

$$f(x) := \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{if} \quad 0 \le x < 1 \\ 1 & \text{if} \quad x = 1. \end{cases}$$

This function is discontinuous. Note that convergence is not uniform because

$$\|f - f_n\|_\infty = \sup_{0\le x<1} x^n = 1$$

for all $n \ge 1$.

(2) Let

$$f_n(x) = \begin{cases} n^2 x & \text{if} \quad 0 \le x \le \frac{1}{n} \\ n^2(\frac{2}{n} - x) & \text{if} \quad \frac{1}{n} \le x \le \frac{2}{n} \\ 0 & \text{if} \quad \frac{2}{n} \le x \le 1. \end{cases}$$

Then $f(x) := \lim_{n\to\infty} f_n(x) = 0$ for all $x \in [0,1]$. Indeed if $x > 0$, then $x \ge \frac{2}{n}$ for $n \ge \frac{2}{x}$, and thus $f_n(x) = 0$; and $f_n(0) = 0$ for all n. So the pointwise limit is continuous. However, this limit is definitely not uniform because

$$\|f - f_n\|_\infty = f_n(\tfrac{1}{n}) = n.$$

The set of functions $\{f_n : n \ge 1\}$ is not even bounded!

1.8.4. THEOREM. *If (X,d) is a metric space, the normed vector space space $C^b_{\mathbb{F}}(X)$ is complete for $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.*

PROOF. Let $(f_n)_{n\ge1}$ be a Cauchy sequence in $C^b(X)$. If $\varepsilon > 0$ is given, there is an $N \in \mathbb{N}$ so that $\|f_n - f_m\|_\infty < \varepsilon$ for all $m, n \ge N$. For each $x \in X$, $|f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty < \varepsilon$ when $m, n \ge N$. Therefore $(f_n(x))_{n\ge1}$ is a Cauchy sequence. Since $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ is complete, there is a pointwise limit $f(x) := \lim_{n\to\infty} f_n(x)$. Moreover if $n \ge N$, by fixing n and letting $m \to \infty$,

$$|f_n(x) - f(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \le \varepsilon.$$

This is valid for all $x \in X$, and thus $\|f_n - f\|_\infty \leq \varepsilon$ for all $n \geq N$. That is, f_n converges uniformly to f. By Theorem 1.8.2, the limit f is continuous and bounded, so lies in $C^b(X)$ Therefore, $C^b(X)$ is complete. ∎

One elementary but very useful test for uniform convergence is the Weierstrass M-test. This is used, for example, to study the radius of convergence of a power series.

1.8.5. WEIERSTRASS M-TEST. *Suppose that $f_n \in C^b(X)$ for $n \geq 1$ and $\sum_{n\geq 1} \|f_n\|_\infty \leq M < \infty$. Then the series $\sum_{n\geq 1} f_n$ converges uniformly to a function $s \in C^b_{\mathbb{F}}(X)$.*

PROOF. The sequence involved is the set of partial sums $s_n(x) = \sum_{k=1}^n f_k(x)$. Let $\varepsilon > 0$. From the convergence of $\sum_{n\geq 1} \|f_n\|_\infty$, there is an $N \in \mathbb{N}$ so that $\sum_{n\geq N} \|f_n\|_\infty < \varepsilon$. Thus if $N \leq m < n$,

$$\|s_n - s_m\|_\infty = \Big\| \sum_{k=m+1}^{n} f_k(x) \Big\|_\infty \leq \sum_{k=m+1}^{n} \|f_k(x)\|_\infty < \varepsilon.$$

Hence (s_n) is a Cauchy sequence. By the completeness of $C^b_{\mathbb{F}}(X)$, (s_n) converges uniformly to a function $s \in C^b_{\mathbb{F}}(X)$. ∎

Exercises

1. Suppose that $(V, \|\cdot\|)$ is a complete normed vector space. Let $\{x_n\}_{n\geq 1}$ be a sequence in V and define a series as the sequence of partial sums $s_k = \sum_{n=1}^k x_n$ converges.
(a) Show that if $\sum_{n\geq 1} \|x_n\| < \infty$, then $(s_k)_{k\geq 1}$ converges.
(b) Show that if $(s_k)_{k\geq 1}$ converges, then $\lim_{n\to\infty} \|x_n\| = 0$.
(c) Show by example that the converse of both (a) and (b) fail.

2. If $A \subset (X, d)$, say the *diameter* of A is $\operatorname{diam} A = \sup_{x,y\in A} d(x, y)$. Show that a metric space (X, d) is a complete if and only if (†) whenever A_n are nested non-empty closed sets ($A_n \supseteq A_{n+1}$) with $\operatorname{diam}(A_n) \to 0$, then $\bigcap_{n\geq 1} A_n$ is not empty.

3. Let (X, d) and (Y, ρ) be metric spaces, and suppose that $f : X \to Y$ is a biLipschitz homeomorphism.
(a) Show that (x_n) is a Cauchy sequence in (X, d) if and only if $(f(x_n))$ is a Cauchy sequence in (Y, ρ).
(b) Show that (x_n) is a convergent sequence in (X, d) if and only if $(f(x_n))$ is a convergent sequence in (Y, ρ).
(c) Hence show that (X, d) is complete if and only if (Y, ρ) is complete.

4. Show that a sequence $(x_n)_{n\geq 1}$ in $(\mathbb{Q}, d_2)$ (the 2-adic metric) is a Cauchy sequence if and only if $d_2(x_n, x_{n+1}) \to 0$.

5. Let X be a closed subset of $\mathbb{R}^n$ and consider the Hausdorff metric d_H on the space $\mathcal{H}(X)$ of all non-empty closed bounded subsets of X. Prove that $\mathcal{H}(X)$ is complete.

6. Suppose that $f_n : X \to \mathbb{R}$ are Lipschitz with constant L. Show that if f_n converge pointwise to f, then f is Lipschitz.

7. Let (X, d) be a set X with the discrete metric.
(a) Which functions $f : X \to \mathbb{R}$ are continuous? Which are uniformly continuous?
(b) Which functions from $\mathbb{R}$ to X are continuous? Which are uniformly continuous?

8. Let (X, d) be a complete metric space, and let Y be an open subset of X. Show that there is a metric ρ on Y which has the same open sets as (Y, d) and is also complete. HINT: find a continuous function $f : (Y, d) \to \mathbb{R}_+$ such that $f(y) \to +\infty$ as y approaches Y^c. Use this to help define ρ.

9. Let (X, d) and (Y, ρ) be metric spaces. Show that the product space $X \times Y$ (see Example 1.3.12) is complete if and only if both X and Y are complete.

10. Let (X, d) and (Y, ρ) be metric spaces.
(a) Let $f : X \to Y$ be uniformly continuous. Show that if $(x_n)_{n\geq 1}$ is a Cauchy sequence in X, then $(f(x_n))_{n\geq 1}$ is Cauchy in Y.
(b) Give an example to show that (a) can fail if f is just continuous.

11. Let V be a normed vector space. Show V is complete if and only if $\overline{b_1(0)}$ is complete.

12. Show that $C^b(X, Y)$, the space of bounded continuous functions from (X, d) to (Y, ρ), is complete if and only if Y is complete.

CHAPTER 2

More Metric Topology

2.1. Compactness

The power of the Bolzano-Weierstrass Theorem is that one can extract from every bounded sequence in $\mathbb{R}^n$, a subsequence which converges. This notion will be called sequential compactness. A topological version will be introduced using open sets. It will be what we call compactness. In the metric setting, we will show that these two notions coincide. The reader who has heard of topological spaces should be warned that this equivalence does not extend to this greater generality, where these are different concepts. In that case, it is the topological property which is more important.

2.1.1. DEFINITION. Let (X, d) be a metric space. An *open cover* of $A \subset X$ is a collection of open sets $\{U_\lambda : \lambda \in \Lambda\}$ such that $A \subset \bigcup_{\lambda\in\Lambda} U_\lambda$. A *subcover* is a subset $\{U_\lambda : \lambda \in \Lambda'\}$, where $\Lambda' \subset \Lambda$, which is still a cover of A. A *finite subcover* is a subcover such that Λ' is a finite set.

A set A is *compact* if every open cover has a finite subcover.

A set A is *sequentially compact* if every sequence $(a_n)_{n\geq 1}$ with all $a_n \in A$ has a subsequence which converges to a point in A.

2.1.2. EXAMPLES.
(1) Finite sets are compact and sequentially compact.

(2) The Heine-Borel Theorem, which we will review, says that every closed and bounded subset of $\mathbb{R}^n$ is compact.

The Bolzano-Weierstrass theorem shows that a closed, bounded subset of $\mathbb{R}$ is sequentially compact. The same readily applies in $\mathbb{R}^n$. Suppose that we are given a sequence $(a_n)_{n\geq 1} \subset A$ where $A \subset \mathbb{R}^n$ is closed and bounded. Consider the first coordinates, which form a bounded seqeunce in $\mathbb{R}$ and select a subsequence so that this coordinate converges. Then consider the second coordinate, and find a subsequence of this subsequence where the second coordinate converges. Note that the first coordinate still converges, since a subsequence of a convergent sequence still converges. Repeat this procedure n times. We arrive at a subsequence in which every coordinate converges. Thus the vectors converge. Since A is closed, the limit remains in A. So A is sequentially compact.

(3) Let X be an infinite set with the discrete metric. Then X is closed and bounded, because all subsets of X are both open and closed, and X has diameter 1. However X is not compact or sequentially compact. The open cover consisting of all singletons $\{x\}$ for $x \in X$ is an infinite open cover of X, and there is no proper subcover. So X is not compact. Select a sequence of distinct points $(x_n)_{n\geq 1}$, $x_n \neq x_m$ if $m < n$. This has no convergent subsequence because the only convergent sequences are eventually constant. So the only compact subsets of X are the finite subsets.

The previous example shows that the converse of the following proposition is false. Keep it in mind to keep yourself clear on this point.

2.1.3. PROPOSITION. *Every compact or sequentially compact subset of a metric space is both closed and bounded.*

PROOF. Suppose that A is not bounded in a metric space (X,d). Fix a point $a_0 \in A$ and consider $\{b_n(a_0) : n \geq 1\}$. Then every point of A belongs to $\bigcup_{n\geq 1} b_n(a_0)$ since any $x \in X$ has $d(a_0,x) < \infty$. However since A is unbounded, there is no finite n so that $b_n(a_0)$ contains A. Hence there is no finite subcover.

Similarly, we could choose a sequence $a_n \in A$ so that $d(a_n, a_0) > n$. This sequence has no convergent subsequence because if $n_1 < n_2 < n_3 < \dots$ and $\lim_{i\to\infty} a_{n_i} = b$, then we arrive at the absurd conclusion:

$$d(a_0, b) = \lim_{i\to\infty} d(a_0, a_{n_i}) = \infty.$$

Now suppose that A is not closed, so that $b \in \overline{A} \setminus A$. Then $b_{1/n}(b) \cap A \neq \emptyset$, so we may choose $a_n \in A$ with $d(a_n, b) < \frac{1}{n}$. Clearly $\lim_{i\to\infty} a_n = b$, and this also holds for any subsequence. So no subsequence has a limit in A. Thus A is not sequentially compact.

Let $U_n = \{x \in X : d(x,b) > \frac{1}{n}\}$. Then $\bigcup_{n\geq 1} U_n = X \setminus \{b\} \supset A$. However as noted in the previous paragraph, no U_n can contain A, and thus there is no finite subcover. So A is not compact. ∎

We need two more definitions.

2.1.4. DEFINITION. Let (X,d) be a metric space. A collection $\mathcal{F} = \{F_\lambda : \lambda \in \Lambda\}$ of subsets of X has the *finite intersection property* (FIP) if every finite subset $\Lambda' \subset \Lambda$ has non-empty intersection $\bigcap_{\lambda\in\Lambda'} F_\lambda \neq \emptyset$.

A set A is *totally bounded* if for all $\varepsilon > 0$, there is a *finite* subset $F \subset X$ so that $A \subset \bigcup_{x\in F} b_\varepsilon(x)$. A finite set $F = \{x_1, \dots, x_n\}$ such that $A \subset \bigcup_{x\in F} b_\varepsilon(x)$ is called a ε*-net* for A.

We come to the main result about compact metric spaces.

2.1.5. BOREL-LEBESGUE THEOREM. *Let (X, d) be a metric space. Then the following are equivalent:*

(1) *X is compact.*

(2) *If $\mathcal{F} = \{F_\lambda : \lambda \in \Lambda\}$ is a collection of closed sets with the finite intersection property, then $\bigcap_{\lambda\in\Lambda} F_\lambda$ is non-empty.*

(3) *X is sequentially compact.*

(4) *X is complete and totally bounded.*

PROOF. (1) $\Rightarrow$ (2). Let $\mathcal{F} = \{F_\lambda : \lambda \in \Lambda\}$ be a collection of *closed* sets with FIP. Define open sets $U_\lambda = F_\lambda^c$. If $\bigcap \mathcal{F} := \bigcap_{\lambda\in\Lambda} F_\lambda$ is empty, then

$$\bigcup_{\lambda\in\Lambda} U_\lambda = \Big(\bigcap_{\lambda\in\Lambda} F_\lambda\Big)^c = X.$$

Thus $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$ is an open cover of X. By compactness, there is a finite subcover $U_{\lambda_1}, \dots, U_{\lambda_n}$. So $\bigcap_{i=1}^n F_{\lambda_i} = \big(\bigcup_{i=1}^n U_{\lambda_i}\big)^c = \emptyset$. This contradicts FIP. Hence we must have $\bigcap \mathcal{F} \neq \emptyset$.

(2) $\Rightarrow$ (3). Let $(x_n)_{n\geq1}$ be a sequence in X. Define non-empty closed sets $F_n = \overline{\{x_k : k \geq n\}}$ for $n \geq 1$. Note that $F_n \supset F_{n+1}$. Hence $\mathcal{F} = \{F_n : n \geq 1\}$ has FIP because $F_{n_1} \cap \dots \cap F_{n_k} = F_{\max\{n_1,\dots,n_k\}}$ is non-empty. By (2), $\bigcap_{n\geq1} F_n$ is non-empty, say x_0 is in the intersection. Then $b_r(x_0) \cap F_n \neq \emptyset$ for all $r > 0$ and $n \geq 1$. Suppose that we have choosen $n_1, \dots, n_k$ so that $d(x_{n_j}, x_0) < \frac{1}{j}$ for $1 \leq j \leq k$. Then $b_{\frac{1}{k+1}}(x_0) \cap F_{n_k+1} \neq \emptyset$. Pick $n_{k+1} > n_k$ so that $d(x_{n_{k+1}}, x_0) < \frac{1}{k+1}$. This recursively selects a subsequence such that $\lim_{k\to\infty} x_{n_k} = x_0$. Therefore X is sequentially compact.

(3) $\Rightarrow$ (4). Let $(x_n)_{n\geq1}$ be a Cauchy sequence in X. By sequential compactness, there is a subsequence $(x_{n_i})_{i\geq1}$ such that $\lim_{i\to\infty} x_{n_i} = x_0$ exists. By the Cauchy property, the whole sequence converges to x_0. Indeed, let $\varepsilon > 0$. Select $I \in \mathbb{N}$ so that if $i \geq I$, then $d(x_{n_i}, x_0) < \varepsilon/2$. Use the Cauchy property to find N so that if $N \leq m, n$, then $d(x_n, x_m) < \varepsilon/2$. Pick some $i \geq I$ so that $n_i > N$. Then if $n \geq N$,

$$d(x_n, x_0) \leq d(x_n, x_{n_i}) + d(x_{n_i}, x_0) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

It follows that $\lim_{n\to\infty} x_n = x_0$. So X is complete.

Suppose that X were not totally bounded. Then for some $\varepsilon > 0$, X cannot be covered by finitely many ε-balls. We claim that we can then select points x_n in X recursively so that $d(x_m, x_n) \geq \varepsilon$ for all $m \neq n$. Indeed suppose that $x_1, \dots, x_k$ have been selected. Since $X \setminus \bigcup_{i=1}^k b_\varepsilon(x_i) \neq \emptyset$, pick x_{k+1} in this set. Then $(x_n)_{n\geq1}$ has no convergent subsequences. This contradicts the sequential compactness of X. So X is complete and totally bounded.

(4) $\Rightarrow$ (1). Suppose that (4) holds, but (1) fails; so that there is an open cover $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$ of X with no finite subcover. Use total boundedness with $\varepsilon = \frac{1}{k}$

to select $x_1^k, \dots, x_{n_k}^k$ to be a finite $\frac{1}{k}$-net for X. We will choose a sequence $y_k = x_{i_k}^k$ so that $X_k := \bigcap_{j=1}^k \overline{b_{\frac{1}{j}}(y_j)}$ has no finite subcover. Suppose that $y_1, \dots, y_k$ has this property. Consider the sets

$$X_{k,i} := X_k \cap \overline{b_{\frac{1}{k+1}}(x_i^{k+1})} \quad \text{for} \quad 1 \le i \le n_{k+1}.$$

If they each had a finite subcover, their union would have a finite subcover. But

$$\bigcup_{i=1}^{n_{k+1}} X_{k,i} = \bigcup_{i=1}^{n_{k+1}} X_k \cap \overline{b_{\frac{1}{k+1}}(x_i^{k+1})} = X_k \cap \bigcup_{i=1}^{n_{k+1}} \overline{b_{\frac{1}{k+1}}(x_i^{k+1})} = X_k \cap X = X_k.$$

has no such cover. Therefore for some i_{k+1}, $X_{k,i_{k+1}} =: X_{k+1}$ has no finite subcover. Set $y_{k+1} = x_{i_{k+1}}^{k+1}$.

Observe that the sequence $(y_k)_{k\ge1}$ is Cauchy. Indeed, if $\varepsilon > 0$, choose an integer $N > 2\varepsilon^{-1}$. If $N \le m \le n$, then $X_n \subset \overline{b_{\frac{1}{m}}(y_m)} \cap \overline{b_{\frac{1}{n}}(y_n)}$ is non-empty, say $x \in X_n$. Hence

$$d(y_m, y_n) \le d(y_m, x) + d(x, y_n) \le \frac{1}{m} + \frac{1}{n} \le \frac{2}{N} < \varepsilon.$$

Since X is complete, there is a limit $y_0 = \lim_{n\to\infty} y_n$ in X. Note that

$$d(y_m, y_0) = \lim_{n\to\infty} d(y_m, y_n) \le \lim_{n\to\infty} \frac{1}{m} + \frac{1}{n} = \frac{1}{m}.$$

Since $\mathcal{U}$ is a cover of X, there is some λ_0 so that $y_0 \in U_{\lambda_0}$. Hence there is an $r > 0$ so that $b_r(y_0) \subset U_{\lambda_0}$. Choose m so large that $\frac{1}{m} < \frac{r}{2}$. Then $x \in X_m \subset \overline{b_{1/m}(y_m)}$ satisfies

$$d(x, y_0) \le d(x, y_m) + d(y_m, y_0) \le \frac{2}{m} < r.$$

That is, $X_m \subset U_{\lambda_0}$ does have a finite subcover. This contradicts the assumption that $\mathcal{U}$ has no finite subcover, since that was how the X_k's were constructed. Therefore X is compact. ∎

2.1.6. REMARK. The definition of compactness depends only on the topology, i.e., the collection of open sets, not on the metric. Likewise the property about collections of closed sets with FIP depends only on closed sets, which are the complements of open sets. So this property is also topological. The same is true of sequential compactness, although that is a bit more subtle. We need to show that we can define convergence for a sequence only using open sets.

Prove the following: *a sequence $(x_n)_{n\ge1}$ converges to x_0 if and only if for each open set U containing x_0, there is an integer N so that $x_n \in U$ for all $n \ge N$.*

However the notions of completeness and total boundedness are metric notions. The real line $\mathbb{R}$ is complete but is not totally bounded. But the real line is homeomorphic to $(0, 1)$ which is not complete, but is totally bounded. So neither property is preserved by a homeomorphism. Somehow the two notions are competing and play off of one another in order to jointly characterize compactness.

2.1.7. REMARK. We have stated our theorem about compactness of the whole space (X, d) for simplicity. Suppose that $A \subset X$. There are two slightly different notions: one is compactness of A as a subset of X; and the other is the compactness of (A, d), thinking of A as a metric space in its own right with the induced metric. Fortunately these two notions coincide.

A set in X is open if it is a union of balls, and the same is true of open sets in A. To make things clear, let $B_r(a)$ denote the ball in X and let $b_r(a)$ be the corresponding ball in A. Then

$$b_r(a) = \{y \in A : d(a, y) < r\} = A \cap \{x \in X : d(a, x) < r\} = A \cap B_r(a).$$

Now suppose that U is open in A, and write $U = \bigcup\{b_r(a) : b_r(a) \subset U\}$. Define $V = \bigcup\{B_r(a) : b_r(a) \subset U\}$. Then V is open in X and

$$V \cap A = \bigcup\{B_r(a) \cap A : b_r(a) \subset U\} = \bigcup\{b_r(a) : b_r(a) \subset U\} = U.$$

This shows that every open subset of A is the intersection of A with an open subset of X.

Suppose that (A, d) is compact, and let $\mathcal{V} = \{V_\lambda : \lambda \in \Lambda\}$ be a collection of open sets in X which cover A. Define $\mathcal{U} = \{U_\lambda := A \cap V_\lambda : \lambda \in \Lambda\}$. This is an open cover of (A, d). By compactness, there is a finite subcover $U_{\lambda_1}, \dots, U_{\lambda_n}$. It follows that $V_{\lambda_1}, \dots, V_{\lambda_n}$ is a finite subcover of A in X.

Conversely suppose that A is a compact subset of X, and let $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$ be an open cover of (A, d). Construct the open sets V_λ in X so that $A \cap V_\lambda = U_\lambda$. Then $\mathcal{V} = \{V_\lambda : \lambda \in \Lambda\}$ is an open cover of A in X. By compactness, there is a finite subcover $V_{\lambda_1}, \dots, V_{\lambda_n}$. Therefore $U_{\lambda_1}, \dots, U_{\lambda_n}$ is a finite subcover from $\mathcal{U}$. Thus (A, d) is compact.

Now let's deal with subsets of $\mathbb{R}^n$.

2.1.8. HEINE-BOREL THEOREM. *A subset $A \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

PROOF. **First proof.** If A is compact, then Proposition 2.1.3 shows that A is closed and bounded. Conversely, if A is closed and bounded, Example 2.1.2(2) used the Bolzano-Weierstrass Theorem to deduce that A is sequentially compact. Now the Borel-Lebesgue Theorem 2.1.5 shows that A is compact.

Second proof. Since $\mathbb{R}^n$ is complete, A is complete if and only if it is closed. If A is bounded say by R, then the finite grid

$$\big\{x = (x_1, \dots, x_n) : x_i \in \{\tfrac{k}{pn} : k \in \mathbb{Z},\ |k| \le Rpn\}\big\}$$

is a $\frac{1}{p}$-net for A. So A is totally bounded. Conversely, if A is unbounded, then no finite set is a 1-net. Thus A is complete and totally bounded if and only if it is closed and bounded. Now the Borel-Lebesgue Theorem 2.1.5 shows this is equivalent to the compactness of A. ■

2.2. More compactness

We continue to collect consequences of compactness, and provide more examples.

2.2.1. PROPOSITION. *Let X be a compact metric space. Then a subset $Y \subset X$ is compact if and only if Y is closed.*

PROOF. If Y is compact, then in particular, it is closed by Proposition 2.1.3. Conversely, suppose that Y is closed., and let $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$ be an open cover of Y. Then $\mathcal{U} \cup \{Y^c\}$. is an open cover of X. Since X is compact, there is a finite subcover, say $U_{\lambda_1}, \dots, U_{\lambda_n}, Y^c$. Since Y^c does not help to cover Y, it follows that $U_{\lambda_1}, \dots, U_{\lambda_n}$ is a finite subcover of Y. Hence Y is compact. ■

2.2.2. DEFINITION. A subset A is *dense* in X if $X \subset \overline{A}$. A metric space is *separable* if it contains a countable dense subset.

2.2.3. PROPOSITION. *Compact metric spaces are separable.*

PROOF. If X is a compact metric space, then it is totally bounded by the Borel-Lebesgue Theorem 2.1.5. For $\varepsilon = \frac{1}{n}$, choose a finite $\frac{1}{n}$-net $x_1^n, \dots, x_{k_n}^n$. Then $\{x_i^n : n \geq 1, 1 \leq i \leq k_n\}$ is a countable dense subset. ■

2.2.4. EXAMPLES.

(1) $\mathbb{R}^n$ is separable because the set of vectors with coefficients in $\mathbb{Q}$ is countable and dense. Also $\mathbb{C}^n$ is separable because it is equivalent as a metric space to $\mathbb{R}^{2n}$.

(2) The space l_p, $1 \leq p < \infty$, is separable. The subspaces $V_n = \text{span}\{e_1, \dots, e_n\}$ are each separable by (1). Their union is dense in l_p, and the countable union of countable sets is countable, so l_p is separable.

(3) l_∞ is not separable. For each subset $E \subset \mathbb{N}$, let $\chi_E(n) = \begin{cases} 1 & \text{if } n \in E \\ 0 & \text{if } n \notin E \end{cases}$. Then $\|\chi_E - \chi_F\|_\infty = 1$ is $E \neq F$. The power set $\mathcal{P}(\mathbb{N})$ of all subsets of $\mathbb{N}$ has cardinality $2^{\aleph_0}$, and so is not countable. No point can be within 0.5 of two of these elements, and hence a dense subset must be uncountable.

(4) Let X be a set, and let d be the discrete metric. Then a subset $Y \subset X$ is dense if and only if $Y = X$. Thus X is separable if and only if X is finite or countable. In particular, $\mathbb{R}$ with the discrete metric is not separable.

2.2.5. PROPOSITION. *If (X, d) and (Y, ρ) are compact metric spaces, then $X \times Y$ is also compact.*

PROOF. Recall from Example 1.3.12 that we put a metric D on $X \times Y$ by $D\big((x_1, y_1), (x_2, y_2)\big) = \max\{d(x_1, x_2), \rho(y_1, y_2)\}$. By the Borel-Lebesgue Theorem 2.1.5, it suffices to show that $X \times Y$ is sequentially compact. Let (x_n, y_n) for $n \geq 1$ be a sequence in $X \times Y$. Since X is compact, there is a subsequence x_{n_i} so that $\lim_{i\to\infty} x_{n_i} = x_0$ exists. Now consider the sequence $(y_{n_i})_{i\geq 1}$. Since Y is compact, there is a subsequence $y_{n_{i_j}}$ so that $\lim_{j\to\infty} y_{n_{i_j}} = y_0$ exists. It is still true that $\lim_{j\to\infty} x_{n_{i_j}} = x_0$. Hence $\lim_{j\to\infty}(x_{n_{i_j}}, y_{n_{i_j}}) = (x_0, y_0)$ is a convergent subsequence. Therefore $X \times Y$ is sequentially compact, and thus compact. ■

2.3. Compactness and Continuity

There is an important connection between compactness and the properties of continuous functions. The next three results are fundamental.

2.3.1. THEOREM. *Let (X, d) and (Y, ρ) be metric spaces. Suppose that X is compact and $f : X \to Y$ is continuous. Then $f(X)$ is compact.*

PROOF. Let $\mathcal{V} = \{V_\lambda : \lambda \in \Lambda\}$ be an open cover of $f(X)$. By continuity, $U_\lambda := f^{-1}(V_\lambda)$ are open in X. Moreover, since $\mathcal{V}$ covers $f(X)$, $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$ covers X. By compactness, X has a finite subcover, say $U_{\lambda_1}, \dots, U_{\lambda_n}$. Therefore

$$f(X) \subset \bigcup_{i=1}^{n} f(U_{\lambda_i}) \subseteq \bigcup_{i=1}^{n} V_{\lambda_i}.$$

Therefore, this is a finite subcover; whence $f(X)$ is compact. ■

Note that $f(U) = f(f^{-1}(V)) \subseteq V$, and this containment may be proper, since it omits points in V that are not in the range of f.

The following result is an important special case should be familiar from calculus.

2.3.2. EXTREME VALUE THEOREM. *If X is a compact metric space and $f : X \to \mathbb{R}$ is continuous, then f is bounded and attains its maximum and minimum values.*

PROOF. By Theorem 2.3.1, $f(X)$ is a compact subset of $\mathbb{R}$. Hence it is closed and bounded. So $f(X)$ contains its (finite) supremum and infimum. These are the maximum and minimum values. ■

This shows that if X is compact, then $\|f\|_\infty = \sup |f(x)|$ is a norm on the space $C(X)$.

2.3.3. THEOREM. *Let (X,d) and (Y,ρ) be metric spaces. Suppose that X is compact and $f : X \to Y$ is continuous. Then f is uniformly continuous.*

PROOF. Let $\varepsilon > 0$ be given. Since f is continuous at $x \in X$, there is a $\delta_x > 0$ so that $f(b_{\delta_x}(x)) \subset b_{\varepsilon/2}(f(x))$. Observe that $\{b_{\delta_x/2}(x) : x \in X\}$ is an open cover of X. By compactness of X, there is a finite subcover, say $b_{\delta_{x_1}/2}(x_1), \dots, b_{\delta_{x_n}/2}(x_n)$. Let $\delta = \min\{\frac{1}{2}\delta_{x_i} : 1 \le i \le n\}$. Suppose that $x, x' \in X$ with $d(x,x') < \delta$. Then there is some i_0 so that $x \in b_{\delta_{x_{i_0}}/2}(x_{i_0})$. Therefore

$$d(x', x_{i_0}) \le d(x',x) + d(x, x_{i_0}) < \delta + \tfrac{1}{2}\delta_{x_{i_0}} \le \delta_{x_{i_0}}.$$

So $x, x' \in b_{\delta_{x_{i_0}}}(x_{i_0})$. Thus, we see that $f(x), f(x') \in b_{\varepsilon/2}(f(x_{i_0}))$. Hence

$$\rho(f(x), f(x')) \le \rho(f(x), f(x_{i_0})) + \rho(f(x_{i_0}), f(x')) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

That is, f is uniformly continuous. ■

It was noted in the discussion following Definition 1.4.5 that a biLipschitz bijection between metric spaces is a homeomorphism. However in Example 1.4.7(2), we showed that a continuous (even Lipschitz) bijection of one metric space onto another need not be a homeomorphism. The following easy result is a critical example where compactness is used to deduce that a bijection is a homeomorphism.

2.3.4. PROPOSITION. *Let (X,d) and (Y,ρ) be metric spaces. Suppose that X is compact and that $f : X \to Y$ is a continuous bijection. Then f is a homeomorphism.*

PROOF. We need to show that f^{-1} is continuous. Let U be an open subset of X. Set $C = U^c$. This is a closed subset of the compact space X, and hence is compact by Proposition 2.2.1. Therefore $f(C)$ is compact by Theorem 2.3.1. Since f is a bijection, $f(U) = f(C)^c$ is the complement of a compact, hence closed, set $f(C)$. So $f(U)$ is open. Again since f is a bijection $(f^{-1})^{-1}(U) = f(U)$. This is open, and hence f^{-1} is continuous. So f is a homeomorphism. ■

Exercises

1. Show that a metric space (X,d) is compact if and only if (‡) whenever A_n are nested non-empty closed sets ($A_n \supseteq A_{n+1}$), then $\bigcap_{n\ge 1} A_n$ is not empty.

2. The Hilbert cube is $H = \{\mathbf{x} = (x_n) \in \ell_2 : 0 \le x_n \le \frac{1}{n},\ n \ge 1\}$. Prove that it is compact.

3. Prove that a metric space (X,d) is totally bounded if and only if every sequence in X has a Cauchy subsequence.

4. Show that the closure $\overline{\mathbb{Z}}$ in $(\mathbb{Q}, d_2)$, the 2-adic metric, is totally bounded but not complete.

5. Let X be a closed subset of $\mathbb{R}^n$ and consider the Hausdorff metric d_H on the space $\mathcal{H}(X)$ of all non-empty closed bounded subsets of X. Prove that $\mathcal{H}(X)$ is compact if and only if X is compact.

6. Suppose that (X, d) is a compact metric space, and that $f : X \to X$ is an isometry. Prove that f is surjective.

7. Let (X, d) be a compact metric space. Suppose that $f, f_n \in C_{\mathbb{R}}(X)$ for $n \geq 1$ and that $f_n \leq f_{n+1}$ for $n \geq 1$ and $f(x) = \lim\limits_{n\to\infty} f_n(x)$ pointwise. Prove that the convergence is uniform.

8. Let (X, d) be a compact metric space. Suppose that $f_n : X \to \mathbb{R}$ are Lipschitz with constant L. Show that if f_n converge pointwise to f, then the convergence is uniform.

9. (a) If (X, d) is compact, show that $U = \{f \in C_{\mathbb{R}}(X) : f(x) > 0 \text{ for } x \in X\}$ is open.
(b) Find the interior of $V = \{f \in C^b_{\mathbb{R}}(\mathbb{R}) : f(x) > 0 \text{ for } x \in X\}$.

10. A metric space (X, d) is *second countable* if there is a countable family $\{U_n : n \geq 1\}$ of open sets that generates the topology, i.e. each open set V satisfies $V = \bigcup_{U_n \subset V} U_n$.
(a) Prove that (X, d) is second countable if and only if it is separable.
(b) Let (X, d) be a separable metric space. Show that every open cover of X has a countable subcover.

11. Show that a metric space (X, d) is complete if and only if every infinite totally bounded subset has an accumulation point.

12. (a) Let (X, d) be a compact metric space. Show that X has finite diameter.
(b) Let (X_i, d_i) for $i \geq 1$ be non-empty metric spaces with diameters bounded by D. Define the product $X = \prod_{i\geq 1} X_i$ to be the set of all sequences $x = (x_1, x_2, \dots)$ where $x_i \in X_i$. Define a metric on X by $\delta(x, y) = \sum_{i\geq 1} 2^{-i} d_i(x_i, y_i)$. Show that a sequence in X converges if and only if each coordinate converges in X_i.
(c) Prove that X is compact if and only if each X_i is compact.

13. The *Lebesgue number* of an open cover $\mathcal{U} = \{U_\lambda : \lambda \in \Lambda\}$ of (X, d) is

$$\delta(\mathcal{U}) = \inf_{x\in X} \sup\{r > 0 : b_r(x) \subset U_\lambda \text{ for some } \lambda \in \Lambda\}.$$

Show that if X is compact and $\mathcal{U}$ is an open cover, then $\delta(\mathcal{U}) > 0$.

2.4. The Cantor Set, Part I

We quickly review the construction of the Cantor set, which is obtained from the interval $[0, 1]$ by successively removing the middle third of each segment. That is, let $C_0 = [0, 1]$, and C_{i+1} is constructed from C_i by removing the middle third from each interval in C_i. The first three terms are

$C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$

$C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$

$C_3 = [0, \frac{1}{27}] \cup [\frac{2}{27}, \frac{1}{9}] \cup [\frac{2}{9}, \frac{7}{27}] \cup [\frac{8}{27}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{19}{27}] \cup [\frac{20}{27}, \frac{7}{9}] \cup [\frac{8}{9}, \frac{25}{27}] \cup \frac{26}{27}, 1]$

The *Cantor set* is $C = \bigcap_{n \geq 0} C_n$. This is an intersection of compact subsets of $[0, 1]$ with FIP, and hence C is not empty. It is a closed subset of a compact set, and hence is compact.

Each point in C is determined by a binary decision tree; whether to choose the left or right interval after removing the middle third at each stage. It will be convenient to label the 2^n intervals in C_n by a sequence of 0's and 2's of length n, where 0 indicates the left interval, and 2 indicates the right. For example, $C_2 = C_{00} \cup C_{02} \cup C_{20} \cup C_{22}$. Proceeding from C_n to C_{n+1}, an interval $C_{a_1 \ldots a_n}$ splits into two intervals $C_{a_1 \ldots a_n 0}$ and $C_{a_1 \ldots a_n 2}$. Note that each interval in C_n has length 3^{-n}.

We will show by induction that

$$C_{a_1 \ldots a_n} = [(0.a_1 \ldots a_n)_{\text{base } 3}, (0.a_1 \ldots a_n)_{\text{base } 3} + 3^{-n}].$$

Recall that $t = (0.a_1 a_2 a_3 \ldots)_{\text{base } 3} = \sum_{i=1}^{\infty} a_i 3^{-i}$ makes sense for any sequence with $a_i \in \{0, 1, 2\}$. The number $(0.a_1 \ldots a_n)_{\text{base } 3} + 3^{-n}$ can be written as the ternary number $(0.a_1 \ldots a_n 222 \ldots)_{\text{base } 3}$ ending with an infinite sequence of 2's.

Our claim is easily verified for the first two levels by inspection. Suppose that it is true for n. At the next stage, $C_{a_1 \ldots a_n}$ is split into two intervals of length 3^{-n-1}, one starting with $(0.a_1 \ldots a_n)_{\text{base } 3} = (0.a_1 \ldots a_n 0)_{\text{base } 3}$ and the other beginning with $(0.a_1 \ldots a_n)_{\text{base } 3} + 2(3^{-n-1}) = (0.a_1 \ldots a_n 2)_{\text{base } 3}$. This establishes the inductive step.

Consider an infinite path in the decision tree given by an infinite sequence $a_1 a_2 a_3 \ldots$ of 0's and 2's. This leads us to

$$\begin{aligned} \bigcap_{n \geq 1} C_{a_1 \ldots a_n} &= \bigcap_{n \geq 1} [(0.a_1 \ldots a_n)_{\text{base } 3}, (0.a_1 \ldots a_n)_{\text{base } 3} + 3^{-n}] \\ &= \{(0.a_1 a_2 a_3 \ldots)_{\text{base } 3}\}. \end{aligned}$$

That is, each infinite path determines a unique point in C. Conversely, each point t in C is determined by the infinite path obtained by picking the interval containing t at every stage. It follows that every point in the Cantor set is represented by a ternary expansion using only 0's and 2's. The number of infinite sequences of 0's and 2's is the same as the number of subsets of $\mathbb{N}$, where we identify the infinite sequence with the set $A = \{i : a_i = 2\}$. This is a bijection. Since the power set of $\mathbb{N}$ is uncountable, the Cantor set is also uncountable.

Note that C has no interior, because any non-empty open set contains an interval of positive length, say r. But if $3^{-n} < r$, it is clear that C_n does not contain an interval of length r. So $\operatorname{int} C = \emptyset$. A set A such that $\overline{A}$ has no interior is called *nowhere dense*. Also C has no isolated points. To see this, note that each point in C is the limit of the left (right) endpoints of each interval $C_{a_1 \ldots a_n}$ containing it. At least one of these sequences is not eventually constant. So every point in C is an accumulation point. A closed set with no isolated points is called *perfect*. Also note that within the metric space (C, d), the sets $C_{a_1 \ldots a_n}$ are closed and open, because

$C \setminus C_{a_1 \dots a_n}$ is the finite union of $2^n - 1$ closed sets, and thus $C_{a_1 \dots a_n}$ is open in the relative topology of C.

In this section, we prove the following remarkable property of the Cantor set.

2.4.1. THEOREM. *Let (X, d) be any compact metric space. Then there is a continuous map of the Cantor set* onto X.

PROOF. Since X is compact, it has a finite $\frac{1}{2}$-net. By adding points if necessary, we may suppose for convenience that the number of points is a power of 2, say $x^1_1, \dots, x^1_{2^{n(1)}}$. Define a function $f_1 : C \to X$ by sending $C \cap C_{a_1 \dots a_{n(1)}}$ to the point x^1_j where $j = \sum_{i=1}^{n(1)} \frac{a_i}{2} 2^{i-1}$. This function is constant on each interval $C_{a_1 \dots a_{n(1)}}$.

Now suppose that a function $f_k : C \to X$ has been defined so that the range is a finite 2^{-k} net for X, say $x^k_1, \dots, x^k_{2^{n(k)}}$ obtained by sending each interval $C_{a_1 \dots a_{n(k)}}$ to x^k_j where $j = 1 + \sum_{i=1}^{n(k)} \frac{a_i}{2} 2^{i-1}$, and moreover, for $k \geq 2$, that

$$\|f_{k-1} - f_k\|_\infty := \sup_{t \in C} d(f_{k-1}(t), f_k(t)) \leq 2^{1-k}.$$

This is true for f_1 because the second condition doesn't apply.

For each ball $\bar{b}_{2^{-k}}(x^k_j)$, we select a 2^{-k-1} net. By adding extra points as needed, we can ensure that each net has the same number of points, which is a power of 2, say 2^p and set $n(k+1) = n(k) + p$. For each j, f_k mapped $C_{a_1 \dots a_{2^{n(k)}}}$ to x^k_j. We split $C_{a_1 \dots a_{2^{n(k)}}}$ into the 2^p intervals $C_{a_1 \dots a_{2^{n(k)}}, a_{2^{n(k)}+1}, \dots, a_{2^{n(k+1)}}}$. These intervals will be mapped by f_{k+1} to the 2^p points in the 2^{-k-1} net for $\bar{b}_{2^{-k}}(x^k_j)$. This ensures that $\|f_k - f_{k+1}\|_\infty$ is no more than 2^{-k}, the radius of the balls.

Observe that the functions f_k are continuous because they are constant on clopen sets. Moreover $\sum_{k \geq 1} \|f_k - f_{k+1}\|_\infty < \infty$. So the sequence $(f_k)_{k \geq 1}$ is a Cauchy sequence in $C(C, X)$. Now X is compact, and hence is complete; so $C(C, X)$ is complete. Therefore $f(t) = \lim_{k \to \infty} f_k(t)$ converges uniformly to a continuous function f. More precisely, for any $t \in C$,

$$\begin{aligned} d\big(f(t), f_k(t)\big) &= \lim_{n \to \infty} d\big(f_n(t), f_k(t)\big) \\ &\leq \lim_{n \to \infty} \sum_{j=k}^{n-1} d\big(f_{j+1}(t), f_j(t)\big) \leq \sum_{j=k}^{\infty} 2^{1-j} = 2^{2-k}. \end{aligned}$$

So $\|f - f_k\|_\infty \leq 2^{2-k} \to 0$.

To see that f is surjective, let $x \in X$. Choose a sequence $x_k = x^k_{j(k)}$ which converges to x. Then x_k is in the range of f_k, so there is a point $t_k \in C$ such that $f_k(t_k) = x_k$. Since C is compact, there is a subsequence t_{k_i} with limit $t \in C$. Now

$$d\big(f(t), x\big) = \lim_{i \to \infty} d\big(f(t_{k_i}), x\big)$$

$$\leq \lim_{i\to\infty} d\big(f(t_{k_i}), f_{k_i}(t_{k_i})\big) + d\big(x_{k_i}, x\big)$$
$$\leq \lim_{i\to\infty} \|f - f_{k_i}\|_\infty + d\big(x_{k_i}, x\big) = 0.$$

Hence $f(t) = x$, and f maps C onto X. ■

We can parlay this result into something even more surprising.

2.4.2. DEFINITION. A *path* is a continuous image of $[0, 1]$. A *Peano curve* or *space filling curve* is a path in $\mathbb{R}^n$, for $n \geq 2$, such that the range has interior.

We will prove the existence of Peano curves in a general setting. Various constructions inside a square or cube have been discovered. Check out Hilbert's curve at Hilbert curve wiki where you can slide the cursor and see the iterative stages in its construction.

2.4.3. THEOREM. *Let X be a compact convex subset of a normed vector space. Then there is a continuous map of $[0, 1]$ onto X.*

PROOF. By Theorem 2.4.1, there is a continuous map f of the Cantor set onto X. Write $[0, 1] \setminus C = \bigcup_{n\geq 1}(a_n, b_n)$, where (a_n, b_n) are the disjoint intervals removed from $[0, 1]$ to form C. Define $g : [0, 1] \to X$ by extending f to be linear on each (a_n, b_n) and matching up with f at the endpoints:

$$g(x) = \begin{cases} f(x) & \text{if} \quad x \in C \\ tf(a_n) + (1-t)f(b_n) & \text{if} \quad x = ta_n + (1-t)b_n,\ 0 < t < 1 \end{cases}$$

The reason for insisting that X be convex is to ensure that $tf(a_n) + (1-t)f(b_n)$, which lies on the line segment from $f(a_n)$ to $f(b_n)$, belongs to X. Clearly g maps $[0, 1]$ onto X. We claim that g is continuous, and hence is the desired map.

Let $\varepsilon > 0$ be given. Since C is compact, f is uniformly continuous by Theorem 2.3.3. Hence there is a $\delta_1 > 0$ so that if $x, y \in C$ and $|x - y| < \delta_1$, then $\|f(x) - f(y)\| < \frac{\varepsilon}{3}$. The lengths of the intervals in the complement of C tend to 0, so there are only finitely many with length $b_n - a_n \geq \delta_1$, say for $n \in F$. Let

$$D = \max\Big\{\frac{\varepsilon}{3},\ \|f(b_n) - f(a_n)\| : n \in F\Big\} \quad \text{and} \quad \delta = \frac{\delta_1\varepsilon}{3D} \leq \delta_1.$$

Suppose that $x, y \in [0, 1]$ with $|x - y| < \delta$.

<u>case 1.</u> $x, y \in C$. Since $\delta \leq \delta_1$, we have $\|g(x) - g(y)\| = \|f(x) - f(y)\| < \frac{\varepsilon}{3}$.

<u>case 2.</u> $x, y \in [a_n, b_n]$ and $n \in F$. Since g is linear on this segment,

$$\|g(x) - g(y)\| = \frac{|x-y|}{b_n - a_n}\|f(b_n) - f(a_n)\| < \frac{\delta_1\varepsilon}{3D}\,\frac{1}{\delta_1}D = \frac{\varepsilon}{3}.$$

<u>case 3.</u> $x, y \in [a_n, b_n]$ and $n \notin F$. Then $b_n - a_n < \delta_1$, so
$$\|g(x) - g(y)\| \leq \|f(b_n) - f(a_n)\| < \frac{\varepsilon}{3}.$$
<u>case 4.</u> $x \in (a_n, b_n)$ and $y \in C$. We may assume that $b_n \leq y$. (The case $y < a_n$ is similar.) Then by the previous cases,
$$\|g(x) - g(y)\| \leq \|g(x) - g(b_n)\| + \|f(b_n) - f(y)\| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon.$$
<u>case 5.</u> $x \in (a_n, b_n)$ and $y \in (a_m, b_m)$ for $n \neq m$. By interchanging x and y if necessary, we may assume that $b_n < a_m$. Then by the previous cases,
$$\begin{aligned}\|g(x) - g(y)\| &\leq \|g(x) - g(b_n)\| + \|f(b_n) - f(a_m)\| + \|g(a_m) - g(y)\| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$
Therefore g is continuous. ■

2.4.4. COROLLARY. *There are Peano curves with range equal to the unit square in* $\mathbb{R}^2$*, the unit ball in* $\mathbb{R}^3$ *and the Hilbert cube.*

2.5. Compact sets in *C(X)*

In this section, we will be concerned with $C(X)$ where (X, d) is a compact metric space. By the Extreme Value Theorem 2.3.2, every function in $C(X)$ attains its maximum modulus. So the supremum norm makes sense without assumimg boundedness.

We are interested in compact subsets of $C(X)$. By Proposition 2.1.3, a compact subset $K \subset C(X)$ must be closed and bounded. The following examples show that this is not sufficient, and we need to look for another condition.

2.5.1. EXAMPLES.
(1) Let $K = \{f_n(x) = x^n, n \geq 1\} \subset C[0, 1]$. In Example 1.8.3(1), we observed that f_n converges pointwise to the discontinuous function $\chi_{\{1\}}$. Any subsequence will also converge pointwise to this function. Hence no subsequence converges uniformly to a continuous limit. It follows that K is closed, and clearly K is bounded by 1. However this also shows that K is not compact, again because no subsequence of $(f_n)_{n\geq 1}$ converges uniformly.

(2) For $n \geq 2$, let $g_n(x) = \begin{cases} 1 & \text{if } x = \frac{1}{n} \\ 0 & \text{if } x = 0, \frac{1}{n+1}, \frac{1}{n-1}, 1 \\ \text{piecewise linear} & \text{in between.} \end{cases}$
Then $\|g_m - g_n\|_\infty = 1$ when $m \neq n$. So $K = \{g_n : n \geq 2\}$ has the discrete metric. In particular, it is closed and bounded but not compact.

Here is the key new notion that we need.

2.5.2. DEFINITION. A subset $\mathcal{F} \subset C(X)$ is *equicontinuous at* $x \in X$ if for every $\varepsilon > 0$, there is a $\delta > 0$ so that whenever $d(x', x) < \delta$, then $|f(x') - f(x)| < \varepsilon$ for all $f \in \mathcal{F}$. In symbols, $\forall_{\varepsilon>0}\ \exists_{\delta(x)>0}\ \forall_{f\in\mathcal{F}}\ \forall_{x'\in b_{\delta(x)}(x)}\ |f(x') - f(x)| < \varepsilon$.

Say that $\mathcal{F} \subset C(X)$ is *equicontinuous* if it is equicontinuous at every $x \in X$.

Say that $\mathcal{F}$ is *uniformly equicontinuous* if δ does not depend on x; i.e., for every $\varepsilon > 0$, there is a $\delta > 0$ so that if $f \in \mathcal{F}$ and $d(x_1, x_2) < \delta$, then $|f(x_1) - f(x_2)| < \varepsilon$. In symbols, $\forall_{\varepsilon>0}\ \exists_{\delta>0}\ \forall_{f\in\mathcal{F}}\ \forall_{d(x_1,x_2)<\delta}\ |f(x_1) - f(x_2)| < \varepsilon$.

2.5.3. LEMMA. *Let (X, d) be a compact metric space. Suppose that $K \subset C(X)$ is compact. Then K is uniformly equicontinuous.*

PROOF. Let $\varepsilon > 0$ be given. Since K is compact, it has a finite $\frac{\varepsilon}{3}$-net, say $f_1, \dots, f_n$. Each f_i is continuous on X, and hence is uniformly continuous by Theorem 2.3.3. Therefore there is a $\delta_i > 0$ so that $d(x_1, x_2) < \delta_i$ implies that $|f(x_1) - f(x_2)| < \frac{\varepsilon}{3}$. Define $\delta = \min\{\delta_1, \dots, \delta_n\}$. Suppose that $d(x_1, x_2) < \delta$ and $f \in \mathcal{F}$. Select i so that $\|f - f_i\|_\infty < \frac{\varepsilon}{3}$. Then

$$\begin{aligned} |f(x_1) - f(x_2)| &\le |f(x_1) - f_i(x_1)| + |f_i(x_1) - f_i(x_2)| + |f_i(x_2) - f(x_2)| \\ &< \|f - f_i\|_\infty + \frac{\varepsilon}{3} + \|f_i - f\|_\infty < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Therefore K is uniformly equicontinuous. ∎

The following is proved much like the proof of Theorem 2.3.3.

2.5.4. LEMMA. *Let (X, d) be a compact metric space. Suppose that $\mathcal{F} \subset C(X)$ is equicontinuous. Then $\mathcal{F}$ is uniformly equicontinuous.*

PROOF. Let $\varepsilon > 0$ be given. For each $x \in X$, there is a $\delta_x > 0$ so that if $d(x', x) < \delta_x$, then $|f(x') - f(x)| < \frac{\varepsilon}{2}$ for all $f \in \mathcal{F}$. The collection

$$\{b_{\delta_x/2}(x) : x \in X\}$$

is an open cover of X. By compactness of X, there is a finite subcover, say $b_{\delta_{x_i}/2}(x_i)$ for $1 \le i \le n$. Define $\delta = \min\{\frac{1}{2}\delta_{x_i} : 1 \le i \le n\}$.

Suppose that $y_1, y_2 \in X$ with $d(y_1, y_2) < \delta$. Select i so that $d(y_1, x_i) < \frac{1}{2}\delta_{x_i}$. Then

$$d(y_2, x_i) \le d(y_2, y_1) + d(y_1, x_i) < \delta + \frac{1}{2}\delta_{x_i} \le \delta_{x_i}.$$

Then if $f \in \mathcal{F}$,

$$|f(y_1) - f(y_2)| \le |f(y_1) - f(x_i)| + |f(x_i) - f(y_2)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore $\mathcal{F}$ is uniformly equicontinuous. ∎

2.5.5. Arzela-Ascoli Theorem. *Let (X, d) be a compact metric space. A subset $K \subset C(X)$ is compact if and only if it is closed, bounded and equicontinuous.*

Proof. If K is compact, then it is closed and bounded by Proposition 2.1.3, and uniformly equicontinuous by Lemma 2.5.3.

Conversely, suppose that K is closed, bounded and equicontinuous. Now $C(X)$ is complete by Theorem 1.8.4. Since K is closed, it is also complete by Proposition 1.6.4. We will show that K is totally bounded. Then the Borel-Lebesgue Theorem 2.1.5 will show that K is compact.

Fix an $\varepsilon > 0$. By Lemma 2.5.4, K is uniformly equicontinuous. Hence there is a $\delta > 0$ so that whenever $f \in K$ and $x_1, x_2 \in X$ with $d(x_1, x_2) < \delta$, then $|f(x_1) - f(x_2)| < \frac{\varepsilon}{4}$. Since X is compact, it has a finite δ-net, say $x_1, \dots, x_n$. Define a linear map $T : C(X) \to (\mathbb{F}^n, \|\cdot\|_\infty)$ by

$$Tf = (f(x_1), \dots, f(x_n)).$$

Note that $\|Tf\|_\infty = \max\{|f(x_i)| : 1 \leq i \leq n\} \leq \|f\|_\infty$. Therefore TK is bounded in $\mathbb{F}^n$. Hence $\overline{TK}$ is compact, and so TK is totally bounded. Therefore it has a finite $\frac{\varepsilon}{4}$-net, say $Tf_1, \dots, Tf_m$ for $f_j \in K$.

We claim that $f_1, \dots, f_m$ is an ε-net for K. Let $f \in K$. Select j so that $\|Tf - Tf_j\|_\infty < \frac{\varepsilon}{4}$. If $y \in X$, pick i so that $d(y, x_i) < \delta$. Then

$$\begin{aligned} |f(y) - f_j(y)| &\leq |f(y) - f(x_i)| + |f(x_i) - f_j(x_i)| + |f_j(x_i) - f_j(y)| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{3\varepsilon}{4}. \end{aligned}$$

Therefore $\|f - f_j\|_\infty \leq \frac{3\varepsilon}{4} < \varepsilon$. Thus K is totally bounded, and so is compact. ■

2.6. Connectedness

The notion of connectedness is introduced to generalize the ideas underlying the Intermediate Value Theorem.

2.6.1. Definition. A subset A of a metric space is *disconnected* if there are disjoint open sets U, V such that $A \subset U \cup V$ and $A \cap U \neq \emptyset \neq A \cap V$. A subset A of a metric space is *connected* if it is not disconnected; i.e. if U, V are disjoint open sets such that $A \subset U \cup V$, then either $A \subset U$ or $A \subset V$.

2.6.2. Examples.

(1) $[0, 1] \cup [2, 3]$ is disconnected. Take $U = (-1, 1.5)$ and $V = (1.5, 4)$.

(2) $\mathbb{Q}$ is disconnected. Take $U = (-\infty, \pi)$ and $V = (\pi, \infty)$.

2.6.3. REMARK. If X is a metric space which is not connected, then X is the union of non-empty disjoint open sets U and V. Thus $V = U^c$ is closed. Therefore U and V are clopen sets in X.

2.6.4. THEOREM. *$[a, b]$ is connected.*

PROOF. Suppose that U, V are disjoint open subsets of $\mathbb{R}$ such that $[a, b] \subset U \cup V$ and $[a, b] \cap U \neq \emptyset \neq [a, b] \cap V$, Without loss of generality, $a \in U$. Define $c = \sup\{x : a \leq x \leq b,\ [a, x] \subset U\}$. Choose x_n increasing to c so that $[a, x_n] \subset U$. Taking their union, we see that $[a, c) \subset U$. If $c \in U$, then there is an $r > 0$ so that $(c - r, c + r) \subset U$; and then $[a, c + r) \subset U$. This contradicts the definition of c unless $c = b$. So either $[a, c) \subset U$ and $c \notin U$ for some $c \leq b$ or $[a, b] \subset U$.

Suppose that $c \in V$. Then there is an $r > 0$ so that $(c-r, c+r) \subset V$. Therefore $U \cap V \supset (c - r, c)$, contradicting the fact that they are disjoint. Consequently, $[a, b] \subset U$. Thus, $[a, b]$ is connected. ■

2.6.5. THEOREM. *If A is connected and $f : A \to Y$ is continuous, then $f(A)$ is connected.*

PROOF. Suppose that U, V are disjoint open subsets of Y such that $f(A) \subset U \cup V$ and $f(A) \cap U \neq \emptyset \neq f(A) \cap V$, By continuity, $f^{-1}(U)$ and $f^{-1}(V)$ are open in A; and they are disjoint. Also $A \subset f^{-1}(U) \cup f^{-1}(V)$ and $A \cap f^{-1}(U) \neq \emptyset \neq A \cap f^{-1}(V)$. So A is disconnected. This is a contradiction. ■

2.6.6. INTERMEDIATE VALUE THEOREM. *If X is a connected metric space, and $f : X \to \mathbb{R}$ is continuous, then $f(X)$ is an interval (possibly infinite).*

PROOF. By Theorem 2.6.5, $f(X)$ is a connected subset of $\mathbb{R}$. Let

$$a = \inf f(X) \in \mathbb{R} \cup \{-\infty\} \quad \text{and} \quad b = \sup f(X) \in \mathbb{R} \cup \{\infty\}.$$

If $a < c < b$, we must have $c \in f(X)$, for otherwise

$$f(X) = \big(f(X) \cap (-\infty, c)\big) \cup \big(f(X) \cap (c, \infty)\big)$$

and both of these intersections must be non-empty. This shows that $f(X)$ is disconnected. Therefore $(a, b) \subset f(X) \subset [a, b]$, and hence $f(X)$ is an interval. ■

2.6.7. LEMMA. *If $X_\lambda \subset Y$ are connected sets for $\lambda \in \Lambda$ and $x_0 \in \bigcap X_\lambda$, then $X := \bigcup_{\lambda \in \Lambda} X_\lambda$ is connected.*

PROOF. Suppose that U and V are disjoint open sets, and $X \subset U \cup V$. We may assume that $x_0 \in U$. Then since X_λ is connected, $X_\lambda \subset U$. Thus $X \subset U$; whence X is connected. ■

2.6.8. LEMMA. *If $A \subset X$ is connected, then $\overline{A}$ is connected.*

PROOF. Suppose that U and V are disjoint open sets, and $\overline{A} \subset U \cup V$. Then $A \subset U \cup V$, so by connectedness it is contained in one of the open sets, say $A \subset U$. Then $\overline{A} \subset \overline{U} \subset V^c$, the last because V^c is closed and contains U. Thus $\overline{A} \cap V = \emptyset$, whence $\overline{A} \subset U$. So $\overline{A}$ is connected. ■

2.6.9. DEFINITION. If $x_0 \in X$, then the *connected component of* x_0 is the largest connected set containing x_0.

2.6.10. EXAMPLES.

(1) Let C be the Cantor set, and let $x \in C$. I will show that the connected component of x in C is just $\{x\}$. There is a unique sequence $a_1, a_2, \ldots$ in $\{0, 2\}$ so that $x \in C_{a_1,\ldots a_n}$ for each $n \geq 1$. Now $C_{a_1,\ldots a_n}$ is clopen in C, and thus $C = C_{a_1,\ldots a_n} \dot\cup (C \setminus C_{a_1,\ldots a_n})$ is a union of disjoint open sets. Therefore the connected component of x is contained in $C_{a_1,\ldots a_n}$; and thus it is contained in $\bigcap_{n\geq 1} C_{a_1,\ldots a_n} = \{x\}$.

(2) $\mathbb{Q}$ is totally disconnected. The connected component of $r \in \mathbb{Q}$ is contained in $(r - \frac{\pi}{n}, r + \frac{\pi}{n})$, because $\mathbb{Q}$ is contained in the disjoint union of the open sets $(r - \frac{\pi}{n}, r + \frac{\pi}{n})$ and $(-\infty, r - \frac{\pi}{n}) \cup (r + \frac{\pi}{n}, \infty)$. This holds for all $n \geq 1$; so the component is just their intersection, $\{r\}$.

2.6.11. PROPOSITION. *The connected component exists and is a closed set.*

PROOF. The connected component exists, since the union of all connected sets containing x_0 is connected by Lemma 2.6.7. This is clearly contains all others, so it is the largest. Moreover this connected component is closed by Lemma 2.6.8. ■

An easy way to show that a set X is connected is to construct a path between any two points in X. For example, any convex subset A of $\mathbb{R}^n$ is connected because if $a, b \in A$, then the line segment $[a, b]$ lies in A. This is connected by Theorem 2.6.4, and hence A consists of a single connected component. We formalize this idea.

2.6.12. DEFINITION. X is *path connected* if for every $x, y \in X$, there is a path from x to y in X, i.e., there is a continuous function $f : [0, 1] \to X$ such that $f(0) = x$ and $f(1) = y$.

2.6.13. PROPOSITION. *Path connected sets are connected.*

PROOF. Fix $x_0 \in X$. For each $y \in X$, find a continuous map $f : [0,1] \to X$ such that $f(0) = x_0$ and $f(1) = y$. By Theorems 2.6.4 and 2.6.5, $f([0,1])$ is connected. Therefore y belongs to the connected component of x_0. Hence X is connected. ∎

2.6.14. EXAMPLE. Let $f(x) = \begin{cases} 0 & \text{if } x \le 0 \\ \sin\frac{1}{x} & \text{if } x > 0 \end{cases}$.
Let $X = \mathcal{G}(f) = \{(x, f(x) : x \in \mathbb{R}\}$. Then $\overline{X} = X \cup L$ where $L = \{0\} \times [-1,1]$. See Figure 2.1. We show that both X and $\overline{X}$ are connected but not path connected.

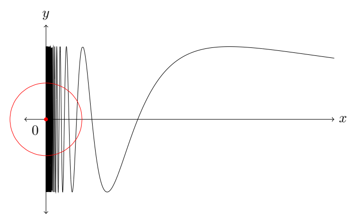

FIGURE 2.1. The topologists's sine curve.

Suppose that $X \subset U \cup V$ where U and V are disjoint open sets. One set, say U, contains $(0,0)$. So the connected component contains the left x-axis $(-\infty, 0] \times \{0\}$. It also contains a neighbourhood about $(0,0)$, and so contains $(\frac{1}{n\pi}, 0)$ for large n. The curve $\{(x, \sin\frac{1}{x}) : x > 0\}$ is path connected and hence connected, so it is also contained in U. Therefore X is connected. So $\overline{X}$ is also connected.

Now we show that neither X nor $\overline{X}$ is path connected. In fact there is no path from $(0,0)$ to $(\frac{1}{\pi}, 0)$ in $\overline{X}$, which establishes both claims. Suppose $g : [0,1] \to \overline{X}$ is continuous such that $g(0) = (0,0)$ and $g(1) = (\frac{1}{\pi}, 0)$. Let $c = \sup\{t : g(t) \in L\}$. Say $g(c) = (0, y)$. By continuity, there is a $\delta_1 > 0$ so that if $t < c + \delta_1$, then $\|g(t) - (0,y)\| < \frac{1}{2\pi}$. Let $t_1 = c + \delta_1/2$ and $g(t_1) = \big(x_1, \sin\frac{1}{x_1}\big)$. Now find $\delta_2 > 0$ so that if $t < c + \delta_2$, then $\|g(t) - (0,y)\| < \frac{x_1}{2}$. Let $t_2 = c + \delta_2/2$ and $g(t_2) = \big(x_2, \sin\frac{1}{x_2}\big)$. Since $g([t_2, t_1])$ is connected, it must contain $\{(x, \sin\frac{1}{x}) : x_2 \le x \le x_1\}$. Now $x_1 < \frac{1}{2\pi}$, and $x_2 < \frac{x_1}{2}$, so that $\frac{1}{x_2} - \frac{1}{x_1} \ge \frac{1}{x_1} > 2\pi$. So the function $f(x)$ takes all values in $[-1,1]$ on this interval. But not all of these values are within $\frac{1}{2\pi}$ of y, contradicting continuity of g.

2.7. The Cantor Set, Part II

Next we consider spaces which are very disconnected like C and $\mathbb{Q}$.

2.7.1. DEFINITION. X is *totally disconnected* if every connected component is a singleton.

The last result of this chapter is an abstract characterization of the Cantor set.

2.7.2. THEOREM. *If X is a non-empty compact metric space which is totally disconnected and perfect, than X is homeomorphic to the Cantor set.*

Before starting the main proof, we need the following key lemma.

2.7.3. LEMMA. *Let (X, d) be a compact, totally disconnected metric space, and let $\varepsilon > 0$. Then X has a finite cover consisting of disjoint non-empty clopen sets of diameter at most ε. If X is perfect, then the cardinality of this partition can be increased to any larger (finite) number.*

PROOF. For $x \in X$, let $A = \{a \in X : d(a, x) \geq \varepsilon/2\}$. Each $a \in A$ is not in the connected component of x, namely $\{x\}$. Hence there are clopen sets U_a and $V_a = U_a^c$ such that $x \in U_a$ and $a \in V_a$. Then $\{V_a : a \in A\}$ covers A. Since A is a closed subset of X, it is compact. Therefore there is a finite subcover $V_{a_1}, \dots, V_{a_n}$. Define $W_x = \bigcap_{i=1}^{n} U_{a_i}$. This is a clopen neighbourhood of x, and is the complement of $V = \bigcup_{i=1}^{n} V_{a_i}$, which is a clopen set containing A. Hence $W_x \subset b_{\varepsilon/2}(x)$. Thus the diameter of W_x is at most ε.

Now $\{W_x : x \in X\}$ is an open cover of X consisting of clopen sets of diameter at most ε. Select a finite subcover $W_{x_1}, \dots, W_{x_m}$. Define $W_k' = W_{x_k} \setminus \bigcup_{j=1}^{k-1} W_{x_j}$ for $1 \leq k \leq m$. After discarding any empty sets, this is the desired partition.

If X is perfect, then each non-empty clopen subset W is also perfect, and thus is not finite. So given two points $x, y \in W$, we can find a clopen set $U \ni x$ such that $y \in V := W \setminus U$. Replacing W by U, V increases the size of the partition by one. Repeat as often as required. ■

PROOF OF THEOREM 2.7.2. Using Lemma 2.7.3, partition X into $n_1 \geq 2$ non-empty clopen subsets $U_1, \dots, U_{n_1}$ of diameter at most 2^{-1}. Then partition C into the same number of clopen subsets $V_1, \dots, V_{n_1}$ of diameter at most 2^{-1}. In this case, we can do this using the standard intervals, as C_0 and C_2 are complementary clopen subsets of diameter $1/3$, and they can be further dissected as necessary to get n_1 sets.

Next we can partition each U_i into finitely many disjoint clopen sets of diameter at most 2^{-2}. By adding further divisions if required, we may assume that each has $n_2 \geq 2$ sets, enumerated $U_{i_1 i_2}$ for $1 \leq i_j \leq n_j$. Partition each V_i into n_2 disjoint non-empty clopen sets of diameter at most 2^{-2} called $V_{i_1 i_2}$ for $1 \leq i_j \leq n_j$. Recursively repeat this procedure, so that for each $k \geq 1$, X is partitioned into non-empty clopen sets $U_{i_1 \dots i_k}$ of diameter at most 2^{-k} where $1 \leq i_j \leq n_j$, and

$\bigcup_{i=1}^{n_k} U_{i_1\dots i_{k-1}i} = U_{i_1\dots i_{k-1}}$. And likewise partition C into non-empty clopen sets $V_{i_1\dots i_k}$ of diameter at most 2^{-k} where $1 \le i_j \le n_j$, and $\bigcup_{i=1}^{n_k} V_{i_1\dots i_{k-1}i} = V_{i_1\dots i_{k-1}}$.

The idea now is to use these partitions as a decision tree to identify individual points which are the intersection of a decreasing sequence of clopen sets. For each point $x \in X$, there is a unique choice of a sequence $U_{i_1}, U_{i_1 i_2}, \dots, U_{i_1\dots i_k}, \dots$ such that $x \in U_{i_1\dots i_k}$ for every $k \ge 1$ because at each level, X is partitioned into disjoint sets, so exactly one contains x. The intersection $\bigcap_{k\ge1} U_{i_1\dots i_k} = \{x\}$ is a single point because it has diameter 0. Conversely every choice of a decreasing sequence of these sets, $\bigcap_{k\ge1} U_{i_1\dots i_k}$ is non-empty since it is a decreasing sequence of closed sets with FIP, and X is compact. Again this intersection has diameter 0, so is a single point. Let's call it $x_{i_1 i_2 i_3\dots}$. Exactly the same is true for C for the same reasons. Denote the corresponding point in C by $c_{i_1 i_2 i_3\dots}$.

Define a function $f : X \to C$ by $f(x_{i_1 i_2 j_3\dots}) = c_{i_1 i_2 j_3\dots}$. By the discussion in the previous paragraph, this map is a bijection. We need to show that f and f^{-1} are continuous. First note that by construction, $f(U_{i_1\dots i_k}) = V_{i_1\dots i_k}$. So we also have $f^{-1}(V_{i_1\dots i_k}) = U_{i_1\dots i_k}$.

To finish the proof, we show that every open set $V \subset C$ is given by

$$V = \bigcup \{V_{i_1\dots i_k} : V_{i_1\dots i_k} \subset V\}. \tag{2.7.4}$$

Suppose that $x \in V$. Since V is open, there is an $r > 0$ so that $b_r(x) \subset V$. Choose k so that $2^{-k} < r$. There is a set $V_{i_1\dots i_k}$ which contains x, and has diameter at most 2^{-k}. Therefore $V_{i_1\dots i_k} \subset b_r(x) \subset V$. It follows that the right hand side of (2.7.4) contains every point in V, establishing the identity. Now we see that

$$f^{-1}(V) = \bigcup \{f^{-1}(V_{i_1\dots i_k}) : V_{i_1\dots i_k} \subset V\} = \bigcup \{U_{i_1\dots i_k} : V_{i_1\dots i_k} \subset V\}$$

is a union of open sets, and thus is open. Hence f is continuous.

The continuity of f^{-1} follows similarly, or we can apply Proposition 2.3.4. Hence X is homeomorphic to C. ■

Exercises

1. Consider the Cantor set as $C = \bigcap_{n\ge1} C_n$ as in § 2.4. The set C_n is the disjoint union of 2^n intervals, say $I_{k,n}$ for $0 \le k < 2^n$ in increasing order. Define a continuous function $f_n : C_n \to [0,1]$ by $f_n(x) = k2^{-n}$ for $x \in I_{k,n}$. Prove that $(f_n)_{n\ge1}$ converges uniformly on C to a continuous monotone function f which maps C onto $[0,1]$.

2. Prove that $C + C = \{x + y : x, y \in C\} = [0,2]$. HINT: consider $C_n + C_n$.

3. Let $\mathcal{F} = \{F \in C[0,1] : F(x) = \int_0^x f(t)\,dt,\ f \in C[0,1],\ \|f\|_\infty \le 1\}$.
(a) Show that $\mathcal{F}$ is bounded and equicontinuous, but not closed.
(b) Show that $\overline{\mathcal{F}} = \{f \in C[0,1] : f(0) = 0 \text{ and } \mathrm{Lip}(f) \le 1\}$, and this is compact.
HINT: find $F_n \in \mathcal{F}$ such that $F_n(\frac{k}{2^n}) = (1 - \frac{1}{n}) f(\frac{k}{2^n})$.

4. Show that the closed unit ball of $C[0,1]$ is not compact.

5. A continuous curve $\gamma : [0,1] \to (V, \|\cdot\|)$ is *rectifiable* if
$$L(\gamma) = \sup\Big\{ \sum_{i=0}^{n-1} \|\gamma(t_i) - \gamma(t_{i+1})\| : 0 = t_0 < t_i < t_{i+1} < t_n = 1 \Big\}$$
is finite. Prove that a space filling curve onto the unit square in $\mathbb{R}^2$ cannot be rectifiable.
HINT: use a fine grid in the square, and find a lower bound for the length of a curve that passes near every lattice point.

6. Let $f_n \in C_{\mathbb{R}}[a,b]$ for $n \ge 1$. Suppose that they are all Lipschitz with Lipschitz constant at most 5 and $|f_n(a)| \le 7$. Prove that there is a subsequence of $(f_n)_{n\ge1}$ which converges uniformly.

7. Let (X,d) be a compact metric space and let $\mathcal{F}$ be an equicontinuous family in $C(X)$ such that $M_x := \sup\{|f(x)| : f \in \mathcal{F}\} < \infty$ for $x \in X$. Prove that $\mathcal{F}$ is bounded.

8. Let $\mathcal{F}$ be an equicontinuous family of continuous functions on $\mathbb{R}$ such that $M_x := \sup\{|f(x)| : f \in \mathcal{F}\} < \infty$ for every $x \in \mathbb{R}$. Prove that every sequence $(f_n)_{n\ge1}$ in $\mathcal{F}$ has a subsequence which converges uniformly on every compact subset of $\mathbb{R}$.

9. Let U be an open subset of a normed vector space V. Prove that U is connected if and only if it is path connected.
HINT: fix $x \in U$, and show that the set of points in U which are path connected to x is open, and that the set of points in U which are not path connected to x is also open.

10. Show that the unit interval $[0,1]$ and the unit circle $\mathbb{T}$ are not homeomorphic.
HINT: $[0,\frac12) \cup (\frac12, 1]$ is not connected.

11. Show that $[0,1]$ and the unit square $[0,1]\times[0,1]$ are not homeomorphic.

12. Let $C_n = \{z \in \mathbb{C} : |z - \frac1n| = 1 - \frac1n\}$ for $n \ge 1$. Define $X = \bigcup_{n\ge0} C_n$. Show that X and $\overline{X}$ are path connected. Is there a continuous function from $[0,1]$ onto $\overline{X}$?

13. Consider the curve $\gamma : (0,1] \to \mathbb{C}$ by $\gamma(t) = (1+t)e^{i/t}$. Clearly $\operatorname{Ran}\gamma$ is path connected. Show that $\overline{\operatorname{Ran}\gamma}$ is connected but not path connected.

14. Let $X = \{(x,0) : 0 \le x \le 1\} \cup \{(\frac1n, y) : n \in \mathbb{N},\ 0 \le y \le 1\} \cup \{(0,1)\} \subset \mathbb{R}^2$. Prove that X is connected but not path connected; however $\overline{X}$ is path connected.

15. (a) Let $A_1 \supset A_2 \supset A_3 \dots$ be a decreasing sequence of connected compact subsets of (X,d). Prove that $\bigcap_{i\ge1} A_i$ is connected.
(b) Find an example of a decreasing sequence of connected closed subsets of $\mathbb{R}^2$ such that the intersection is not connected.

16. Show that if $x, y \in C$, the Cantor set, then there is a homeomorphism f of C onto itself such that $f(x) = y$. *Note:* this means that being the endpoint of an interval is a property of the *imbedding* of C into $[0,1]$, not a topological property of C.

17. Fix a real number $0 < t < 9$. If $x \in [0,1)$, choose the decimal expansion for x, $x = 0.x_1x_2\dots$, which does not end in all 9's. Let $a_n(x) = \frac1n \sum_{i=1}^n x_i$. Define $B = \{x \in [0,1) : a_n(x) \le t \text{ for all } n \ge 1\}$. Show that B is closed, nowhere dense and perfect.

CHAPTER 3

Completeness Revisited

3.1. The Baire Category Theorem

In this section, we prove a result about complete metric spaces which is not particularly difficult, but has many surprising consequences. We will see some of them in this course, and you will see several others if you take functional analysis.

3.1.1. DEFINITION. A subset A of a metric space X is *nowhere dense* if $\overline{A}$ has no interior. A set A is *first category in* X if $A = \bigcup_{n\geq 1} A_n$ and each A_n is nowhere dense; i.e., A is a *countable* union of nowhere dense sets.

Say that B is a *residual set in* X if B^c is first category.

If A is nowhere dense, then $\overline{A}^c$ is a dense open set. Thus A is small and its complement is pervasive within X. One should think of sets of first category sets as being small as well.

3.1.2. EXAMPLES.
(1) Single points in $\mathbb{R}$ are closed and nowhere dense. Thus $\mathbb{Q}$ is a countable union of $\{r\}$ for $r \in \mathbb{Q}$. Hence $\mathbb{Q}$ is first category and $\mathbb{R} \setminus \mathbb{Q}$ is a residual set.

(2) The Cantor set is closed and has no interior in $\mathbb{R}$. So it is nowhere dense even though it has the same cardinality as $\mathbb{R}$.

3.1.3. BAIRE CATEGORY THEOREM. *A non-empty complete metric space X is not first category; i.e., X is not a countable union of nowhere dense sets. Indeed, if A_n are nowhere dense subsets of X, then $\bigcap_{n\geq 1} \overline{A_n}^c$ is dense in X.*

PROOF. Let $x \in X$ and $r > 0$. We will find a point in $b_r(x) \setminus \bigcup_{n\geq 1} \overline{A_n}$. This will show that $\bigcap_{n\geq 1} \overline{A_n}^c$ is dense in X.

Since $\overline{A_1}$ has no interior, $V_1 := b_r(x) \cap \overline{A_1}^c$ is non-empty and open. Thus there is a point $x_1 \in X$ and an $0 < r_1 < r/2$ so that $\bar{b}_{r_1}(x_1) \subset V_1$. Proceed recursively. At stage n, we will have $x_1, \dots, x_n$ and $r_1, \dots, r_n$ so that $r_i < r/2^i$ and $\bar{b}_{r_i}(x_i) \subset V_i = b_{r_{i-1}}(x_{i-1}) \setminus \overline{A_i}$ for $1 \leq i \leq n$. Set $V_{n+1} = b_{r_n}(x_n) \setminus \overline{A_{n+1}}$. Since A_{n+1} is nowhere dense, this is a non-empty open set. So we may find a point x_{n+1} and an $r_{n+1} < r/2^{n+1}$ so that $\bar{b}_{r_{n+1}}(x_{n+1}) \subset V_{n+1}$. This completes the inductive step.

The balls $\bar{b}_{r_n}(x_n)$ form a decreasing nested sequence of closed sets. We claim that the sequence $(x_n)_{n\geq 1}$ is Cauchy. Indeed, if $N \leq m < n$, then $x_n, x_m \in \bar{b}_{r_N}(x_N)$ and hence $d(x_n, x_m) \leq 2r_N < 2^{1-N}r$. So if $\varepsilon > 0$ is given, choose N so that $2^{1-N}r < \varepsilon$. Since X is complete, this sequence has a limit, say $x_0 = \lim_{n\to\infty} x_n$. Hence x_0 belongs to $\bigcap_{n\geq 1} \bar{b}_{r_n}(x_n)$. (Alternatively, $\bar{b}_{r_n}(x_n)$ is a decreasing nested sequence of closed sets with diameter tending to 0, and thus they have non-empty intersection $\{x_0\}$ by the completeness of X. See Assignment 3.) Since $\bar{b}_{r_n}(x_n)$ is disjoint from $\overline{A_n}$, we have $x_0 \in \bigcap_{n\geq 1} \overline{A_n}^c$. Moreover, $x_0 \in V_1 \subset b_r(x)$, so that $d(x, x_0) < r$. Thus $\bigcap_{n\geq 1} \overline{A_n}^c$ is dense in X. ■

3.1.4. Definition. If X is a metric space, a set $A \subset X$ is a G_δ *set* if there are countably many open sets U_n, $n \geq 1$, so that $A = \bigcap_{n\geq 1} U_n$. A set $A \subset X$ is an F_σ *set* if there are countably many closed sets C_n, $n \geq 1$, so that $A = \bigcup_{n\geq 1} C_n$.

Since the complement of a closed nowhere dense set is a dense open set, the following corollary is immediate.

3.1.5. Corollary. *If X is a complete metric space and U_n are dense open sets for $n \geq 1$, then $\bigcap_{n\geq 1} U_n$ is a dense G_δ set.*

The Baire Category Theorem is often used using the contrapositive, which can be formulated as follows. Again the proof is immediate.

3.1.6. Corollary. *Let X be a complete metric space. Suppose that C_n are closed sets such that $X = \bigcup_{n\geq 1} C_n$. Then there is some n_0 so that C_{n_0} has non-empty interior.*

3.1.1. Pointwise Limits of Continuous Functions. We have seen in Example 1.8.3 that the pointwise limit of continuous functions need not be continuous. Functions which are pointwise limits of continuous functions are called *Baire one* functions. In this short section, we will show that Baire one functions retain some good properties.

3.1.7. Definition. Let (X, d) be a metric space, and let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$. If $f : X \to \mathbb{F}$ is a function, the *oscillation of* f at x, $\omega_f(x)$, is defined in two stages:

$$\omega_f(x, \delta) = \sup\big\{|f(y) - f(z)| : y, z \in b_\delta(x)\big\} \quad \text{for} \quad \delta > 0$$
$$\omega_f(x) = \inf_{\delta>0} \omega_f(x, \delta).$$

The following easy lemma is left as an exercise.

3.1.8. LEMMA. *Let $f : X \to \mathbb{F}$. Then f is continuous at x if and only if $\omega_f(x) = 0$.*

We need another easy lemma.

3.1.9. LEMMA. *Let $f : X \to \mathbb{F}$ and let $\varepsilon > 0$. Then $\{x : \omega_f(x) < \varepsilon\}$ is open.*

PROOF. Suppose that $\omega_f(x) < \varepsilon$. Then for some $\delta > 0$, $\omega_f(x, \delta) < \varepsilon$. If $d(x, y) = r < \delta$, then $b_{\delta-r}(y) \subset b_\delta(x)$. Therefore, $\omega_f(y, \delta - r) \le \omega_f(x, \delta) < \varepsilon$. Hence $\omega_f(y) < \varepsilon$. That is, $b_\delta(x) \subset \{x : \omega_f(x) < \varepsilon\}$. So this is an open set. ∎

The main result of this subsection is the following.

3.1.10. THEOREM. *Suppose that $f_i \in C[a, b]$ converge pointwise to a function f. Then f is continuous on a residual G_δ set.*

PROOF. Observe that the points of continuity of f are

$$\{x : \omega_f(x) = 0\} = \bigcap_{n\ge 1} \{x : \omega_f(x) < \tfrac{1}{n}\} = \Big(\bigcup_{n\ge 1} \{x : \omega_f(x) \ge \tfrac{1}{n}\}\Big)^c.$$

Since $U_n = \{x : \omega_f(x) < \frac{1}{n}\}$ is open by Lemma 3.1.9, the points of continuity form a G_δ set. The plan is to show that the closed sets $A_n = \{x : \omega_f(x) \ge \frac{1}{n}\} = U_n^c$ are nowhere dense. Let I be a (small) open interval in $[a, b]$. We will show that I contains a point with $\omega_f(x) < \frac{1}{n}$. Hence $I \not\subset A_n$. As I is arbitrary, A_n has no interior.

Set $\varepsilon < \frac{1}{3n}$. For all $i, j \ge 1$, set $X_{i,j} = \{x \in \overline{I} : |f_i(x) - f_j(x)| \le \varepsilon\}$. By the continuity of $f_i - f_j$, these are closed sets. Define closed sets for $n \ge 1$ by $E_n = \bigcap_{i,j\ge n} X_{i,j}$. Since $f_i(x)$ converges to $f(x)$, there is some $N \in \mathbb{N}$ so that $|f_i(x) - f(x)| < \frac{\varepsilon}{2}$ for all $i \ge N$. Hence if $i, j \ge N$, we have $|f_i(x) - f_j(x)| < \varepsilon$; and thus $x \in E_N$. It follows that $\overline{I} = \bigcup_{n\ge 1} E_n$.

Now $\overline{I}$ is complete, so Corollary 3.1.6 shows that there is some n_0 so that E_{n_0} has interior, say $E_{n_0} \supset J$ where J is an open interval. Hence $|f_i(x) - f_{n_0}(x)| \le \varepsilon$ for all $i > n_0$ and $x \in J$. Take the limit as $i \to \infty$ to see that $|f(x) - f_{n_0}(x)| \le \varepsilon$. Since f_{n_0} is uniformly continuous, there is a $\delta_0 > 0$ so that $|x - y| < \delta_0$ implies that $|f_{n_0}(y) - f_{n_0}(x)| < \varepsilon$. If $x \in J$, let $\delta_x = \min\{\delta_0/2, d(x, J^c)\}$. Then if $y, z \in b_{\delta_x}(x) \subset J$, we have $d(y, z) < \delta_0$; so

$$\begin{aligned}|f(y) - f(z)| &\le |f(y) - f_{n_0}(y)| + |f_{n_0}(y) - f_{n_0}(z)| + |f_{n_0}(z) - f(z)| \\ &\le \varepsilon + \varepsilon + \varepsilon = 3\varepsilon < \frac{1}{n}.\end{aligned}$$

Therefore $\omega_f(x) \leq \omega_f(x, \delta_x) \leq 3\varepsilon < \frac{1}{n}$ for all $x \in J \subset I$. This shows that A_n has no interior; so it is nowhere dense.

Hence $\bigcup_{n \geq 1} A_n = \{x : \omega_f(x) > 0\}$ is first category. By the Baire Category Theorem 3.1.3, $\{x : \omega_f(x) = 0\} = \bigcap_{n \geq 1} A_n^c$ is a dense G_δ set. By Lemma 3.1.8, this is the set of points of continuity of f. ■

3.2. Nowhere Differentiable Functions

Our major application of the Baire Category Theorem will be to show that most continuous functions on an interval are not differentiable even at a single point. Such functions are called *nowhere differentiable*.

We need the following local variant on Lipschitz functions. The reason for discussing a Lipschitz condition is that these functions are better behaved under limits than differentiable functions. The easy lemma is a local version of Example 1.4.7(3).

3.2.1. Definition. A function $f \in C(X)$ is *Lipschitz at* x for $x \in X$ is there is a constant C so that

$$|f(x) - f(y)| \leq C d(x, y) \quad \text{for all} \quad y \in X.$$

3.2.2. Lemma. *If $f \in C[a, b]$ is differentiable at x, then it is Lipschitz at x.*

Proof. We are given that $f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x}$ exists. Hence there is a $\delta > 0$ so that when $0 < |y - x| < \delta$,

$$\left| \frac{f(y) - f(x)}{y - x} - f'(x) \right| < 1.$$

If follows that $|f(y) - f(x)| \leq \big(|f'(x)| + 1\big)\,|y - x|$ if $|y - x| < \delta$. Now if $|y - x| \geq \delta$,

$$|f(y) - f(x)| \leq 2\|f\|_\infty \leq \big(2\|f\|_\infty \delta^{-1}\big)\,|y - x|.$$

Hence f is Lipschitz at x with constant $C = \max\big\{|f'(x)| + 1,\ 2\|f\|_\infty \delta^{-1})\big\}$. ■

3.2.3. Theorem. *The set of functions $f \in C[a, b]$ which are differentiable at one or more points is a set of first category. So the set of nowhere differentiable functions on $[a, b]$ is a residual set, and in particular is dense in $C[a, b]$.*

Proof. For $k \geq 1$, let

$$A_k = \{f \in C[a, b] : \exists\, x \in [a, b] \text{ s.t. } f \text{ is Lipschitz at } x \text{ with constant } k\}.$$

Our goal is to show that A_k is closed and nowhere dense.

First suppose that $f_n \in A_k$ and $f_n \to f$ uniformly on $[a,b]$. For each f_n, there is a point $x_n \in [a,b]$ so that $|f_n(y) - f_n(x_n)| \le k|y - x_n|$ for $y \in [a,b]$. The bounded sequence $(x_n)_{n\ge1}$ has a convergent subsequence by the Bolzano-Weierstrass Theorem, say $x_0 = \lim_{i\to\infty} x_{n_i}$. Then

$$\begin{aligned}|f(y) - f(x_0)| &\le |f(y) - f_{n_i}(y)| + |f_{n_i}(y) - f_{n_i}(x_{n_i})| + \\ &\qquad + |f_{n_i}(x_{n_i}) - f_{n_i}(x_0)| + |f_{n_i}(x_0) - f(x_0)| \\ &\le \|f - f_{n_i}\|_\infty + k|y - x_{n_i}| + k|x_{n_i} - x_0| + \|f_{n_i} - f\|_\infty \\ &= 2\|f - f_{n_i}\|_\infty + k\big(|y - x_{n_i}| + |x_{n_i} - x_0|\big).\end{aligned}$$

Now take the limit as $i \to \infty$ to obtain that

$$|f(y) - f(x_0)| \le k|y - x_0|.$$

Thus $f \in A_k$. So A_k is closed.

Next we show that A_k has no interior. Take $f \in A_k$ and let $\varepsilon > 0$ be given. The idea is to first find a nice (in this case, piecewise linear) function close to f. Then we will add to this function a very wild function to obtain a function that does not have a small local Lipschitz constant anywhere.

Since f is uniformly continuous, there is a $\delta > 0$ so that $|x-y| < \delta$ implies that $|f(x) - f(y)| < \varepsilon/4$. Choose a finite set of points $a = x_0 < x_1 < \cdots < x_n = b$ so that $x_{i+1} - x_i < \delta$ for $0 \le i < n$. Define h to be the piecewise linear function determined by $h(x_i) = f(x_i)$ for $1 \le i \le n$. Then if $x_i < x < x_{i+1}$,

$$\begin{aligned}|h(x) - f(x)| &\le |h(x) - h(x_i)| + |h(x_i) - f(x_i)| + |f(x_i) - f(x)| \\ &< |h(x_{i+1}) - h(x_i)| + 0 + \frac{\varepsilon}{4} \\ &= |f(x_{i+1}) - f(x_i)| + \frac{\varepsilon}{4} < \frac{\varepsilon}{2}.\end{aligned}$$

Thus $\|h - f\|_\infty < \frac{\varepsilon}{2}$. (Strict inequality follows from the Extreme Value Theorem.)

Since h is piecewise linear, it is Lipschitz with constant L equal to the maximim absolute value of the slope on each segment. Let $M > 4\pi\varepsilon^{-1}(L + k)$. Define $g = h + \frac{\varepsilon}{2}\sin Mx$. The small function $\frac{\varepsilon}{2}\sin Mx$ has a big derivative at many points, and this will ensure that g is not in A_k. Note that

$$\|g - f\|_\infty \le \|g - h\|_\infty + \|h - f\|_\infty < \frac{\varepsilon}{2}\|\sin Mx\|_\infty + \frac{\varepsilon}{2} \le \varepsilon.$$

For any $x_0 \in [a,b]$, we will show that g is not k-Lipschitz at x_0. In any interval of length $2\pi/M$, the function $\sin Mx$ will take all values in $[-1,1]$. So choose a point $x \in [a,b]$ so that

$$|x - x_0| < \frac{2\pi}{M} \quad \text{and} \quad \sin Mx = \begin{cases} +1 & \text{if } \sin Mx_0 < 0 \\ -1 & \text{if } \sin Mx_0 \ge 0. \end{cases}$$

Then

$$|g(x) - g(x_0)| = \Big|h(x) - h(x_0) + \frac{\varepsilon}{2}(\sin Mx - \sin Mx_0)\Big|$$

$$\begin{aligned}
&\geq \frac{\varepsilon}{2}\big|\sin Mx - \sin Mx_0\big| - \big|h(x) - h(x_0)\big| \\
&\geq \frac{\varepsilon}{2} - L|x - x_0| \\
&\geq \frac{\varepsilon}{2}\,\frac{M}{2\pi}\,|x - x_0| - L|x - x_0| \\
&= \Big(\frac{\varepsilon M}{4\pi} - L\Big)|x - x_0| > k|x - x_0|.
\end{aligned}$$

Hence g is not Lipschitz at x_0 with constant k. As x_0 was arbitrary, $g \notin A_k$. Hence A_k has no interior.

We have shown that each A_k is nowhere dense. So $\bigcup_{k\geq 1} A_k$ is first category. The complement consists of all functions which are not locally Lipschitz at any point. By Lemma 3.2.2, this implies in particular that they are nowhere differentiable. Hence the set of nowhere differentiable functions is also a residual set. By the Baire Category Theorem 3.1.3, the set of nowhere differentiable functions is dense in $C[a, b]$. ■

3.2.1. Weierstrass's Nowhere Differentiable Function. This Baire Category argument is not how people first discovered nowhere differentiable functions. Weierstrass constructed a whole family of such functions as sums of infinite series. We will provide one of his examples here.

Define

$$f(x) = \sum_{k\geq 1} 2^{-k}\cos(10^k \pi x) = \sum_{k\geq 1} f_k(x) \quad \text{for} \quad x \in \mathbb{R}.$$

Since $\|f_k\|_\infty = 2^{-k}$, the Weierstrass M-test 1.8.5 shows that this series converges uniformly to a continuous function on $\mathbb{R}$. Moreover each f_k is 1-periodic so f has period 1. Thus we need only consider $x \in [0, 1]$.

Let $x = 0.x_1x_2x_3\cdots \in [0, 1]$. For each $n \geq 1$, let $a_n = 0.x_1x_2x_3\ldots x_n$ and $b_n = a_n + 10^{-n}$. Notice that $10^n a_n$ is an integer and $10^n b_n = 10^n a_n + 1$; so

$$\begin{aligned}
f_n(a_n) &= 2^{-n}\cos(10^n \pi a_n) = 2^{-n}(-1)^{10^n a_n} \\
f_n(b_n) &= 2^{-n}\cos(10^n \pi b_n) = 2^{-n}(-1)^{10^n a_n + 1}.
\end{aligned}$$

Therefore $|f_n(a_n) - f_n(b_n)| = 2^{1-n}$.

If $k > n$, $10^k a_n$ and $10^k b_n$ are both even integers, so that $f_k(a_n) = f_k(b_n)$. If $1 \leq k < n$, the Mean Value Theorem shows that

$$|f_k(a_n) - f_k(b_n)| \leq \|f_k'\|_\infty\,(b_n - a_n) = (2^{-k}10^k\pi)10^{-n} = 2^{-n}5^{k-n}\pi.$$

Therefore

$$\begin{aligned}
|f(a_n) - f(b_n)| &= \Big|\sum_{k=1}^{\infty} f_k(a_n) - f_k(b_n)\Big| \\
&\geq |f_n(a_n) - f_n(b_n)| - \sum_{k=1}^{n-1} |f_k(a_n) - f_k(b_n)|
\end{aligned}$$

$$\geq 2^{1-n} - 2^{-n}\pi \sum_{k=1}^{n-1} 5^{k-n}$$
$$> 2^{-n}(2 - \frac{\pi}{4}) > 2^{-n}.$$

It follows that choosing the endpoint $y_n \in \{a_n, b_n\}$ judiciously, we can arrange that $|f(y_n) - f(x)| > 2^{-n-1}$. However $|y_n - x| \leq 10^{-n}$. Therefore

$$\left|\frac{f(y_n) - f(x)}{y_n - x}\right| > \frac{2^{-n-1}}{10^{-n}} = \frac{5^n}{2}.$$

This tends to ∞, from which we deduce that f is not differentiable at x.

3.3. The Contraction Mapping Principle

This section provides another easy but powerful consequence of completeness.

3.3.1. Definition. Let (X, d) be a metric space. A map $T : X \to X$ is a *contraction mapping* if it is Lipschitz with a Lipschitz constant $c < 1$.

A *fixed point* of a map $T : X \to X$ is a point $x \in X$ such that $Tx = x$.

3.3.2. Contraction Mapping Principle. *Let (X, d) be a complete metric space, and let $T : X \to X$ be a contraction mapping with Lipschitz constant $c < 1$. Then T has a unique fixed point x_*. Moreover, for any $x_0 \in X$, the sequence $x_n := T^n x$ converges to x_*, and*

$$d(x_n, x_*) \leq c^n d(x_0, x_*) \leq \frac{c^n}{1-c} d(x_0, Tx_0).$$

Proof. Start with any point $x_0 \in X$ and define $x_{n+1} = Tx_n$ for $n \geq 0$. Then for $n \geq 1$,

$$d(x_{n+1}, x_n) = d(Tx_n, Tx_{n-1}) \leq c d(x_n, x_{n-1}).$$

By induction, we see that $d(x_{n+1}, x_n) \leq c^n d(x_1, x_0)$.

We claim that $(x_n)_{n\geq 0}$ is a Cauchy sequence. Indeed, if $N \leq m < n$, then by the triangle inequality,

$$d(x_n, x_m) \leq \sum_{i=m}^{n-1} d(x_{i+1}, x_i) \leq \sum_{i=m}^{n-1} c^i d(x_1, x_0)$$
$$< \sum_{i\geq N} c^i d(x_1, x_0) = \Big(\frac{c^N}{1-c}\Big) d(x_1, x_0).$$

Given $\varepsilon > 0$, we may choose N so large that $c^N \frac{d(x_1,x_0)}{1-c} < \varepsilon$. Then $d(x_n, x_m) < \varepsilon$ for all $N \leq m < n$, so the sequence is Cauchy.

Let $x_* = \lim_{n\to\infty} x_n$, which exists since X is complete. Then

$$Tx_* = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} x_{n+1} = x_*.$$

Thus x_* is a fixed point for T. Moreover,

$$d(x_{n+1}, x_*) = d(Tx_n, Tx_*) \le cd(x_n, x_*).$$

By induction, we again show that $d(x_n, x_*) \le c^n d(x_0, x_*)$. From the previous paragraph with $N = 0$, we have that $d(x_n, x_0) \le \frac{1}{1-c} d(x_1, x_0)$. Letting $n \to \infty$ yields $d(x_*, x_0) \le \frac{1}{1-c} d(x_1, x_0)$. Hence

$$d(x_n, x_*) \le c^n d(x_0, x_*) \le \frac{c^n}{1-c} d(Tx_0, x_0).$$

Suppose that y_* is a fixed point of T. Then

$$d(x_*, y_*) = d(Tx_*, Ty_*) \le cd(x_*, y_*).$$

Since $c < 1$, this shows that $d(x_*, y_*) = 0$; that is, $y_* = x_*$. So x_* is the unique fixed point. ∎

3.3.3. EXAMPLES.

(1) The condition $c < 1$ is required to ensure a fixed point. If $S : \mathbb{R} \to \mathbb{R}$ by $Sx = x + 1$, then this is an isometry and has Lipschitz constant $c = 1$. Clearly it has no fixed points.

A somewhat more subtle example is $T : [1, \infty) \to [1, \infty)$ by $Tx = x + \frac{1}{x}$. Then

$$|Tx - Ty| = \left|x + \frac{1}{x} - y - \frac{1}{y}\right| = |x - y|\left(1 - \tfrac{1}{xy}\right) < |x - y|.$$

Thus T shrinks the distance between any two pairs of points. However for x, y very large, this ratio gets arbitrarily close to 1, so the Lipschitz constant is 1. This map has no fixed point because $Tx > x$ for all x.

(2) Put your calculator in radian mode, and enter an arbitrary number. Repeatedly compute the cosine by hitting the cos button many times. You will see that fairly quickly, you get the answer 0.739085133 and if you use a computer with higher precision, you will get $x_* = 0.73908513321516064\ldots$.

This corresponds to the map $T : \mathbb{R} \to \mathbb{R}$ by $Tx = \cos x$. Whatever the starting point x_0 is, $x_1 \in [-1, 1]$ and $x_2 \in [\cos 1, 1]$. The Mean Value Theorem shows that if $x, y \in [-1, 1]$, then there is some $\theta \in (x, y)$ so that

$$\frac{\cos x - \cos y}{x - y} = \sin\theta \le \sin 1.$$

Hence $|Tx - Ty| \le (\sin 1)|x - y|$ is a contraction mapping with $c = \sin 1$ once we restrict T to its range, $[-1, 1]$. Thus by the Contraction Mapping Principle, there is a unique fixed point. This fixed point is the unique solution to the equation $\cos x_* = x_*$. Graph the curves $y = \cos x$ and $y = x$ and find the intersection point, (x_*, x_*).

(3) Let $T : [-1,1] \to [-1,1]$ by $Tx = 1.8(x - x^3)$. Note that $T(\pm 1) = 0$. Compute $T'x = 1.8(1 - 3x^2)$. Then $T'x = 0$ when $x = \pm 1/\sqrt{3}$; and $T(\pm 1/\sqrt{3}) = \frac{\pm 2\sqrt{3}}{5}$. These are the local max and local min respectively, and $2\sqrt{3} < 5$; so T maps $[-1,1]$ into itself. You can also see by inspection that the derivative $T'(0) = 1.8 > 1$ and $|T'(\pm 1)| = 3.6 > 1$. So T is not a contraction mapping.

Solve for fixed points: $Tx = x$ if and only if $x = 1.8x - 1.8x^3$, which holds if $x(x^2 - \frac{4}{9}) = 0$. This has *three* solutions, $x = 0, \pm\frac{2}{3}$. Note that $T'(\pm\frac{2}{3}) = -0.6$ while $T'(0) = 1.8$.

We will show that: *if $Tx_* = x_*$ and T is C^1 with $|T'(x_*)| < 1$, then there is a (small) interval $I = (x_* - \delta, x_* + \delta)$ so that $T : I \to I$ is a contraction mapping.* This is called an attracting fixed point. Use the continuity of T' to select an interval containing x_* on which $\max |T'x| = c < 1$. By the Mean Value Theorem, if $x \in I$, then there is a point ξ in (x, x_*) so that

$$\frac{|Tx - Tx_*|}{|x - x_*|} = |T'\xi| \leq c.$$

Hence $|Tx - Tx_*| \leq c|x - x_*|$. So T maps I into itself, and is a contraction mapping. Thus $T^n x$ converges to x_* if we start with $x \in I$. This is what occurs in our example near $x_* = \pm\frac{2}{3}$.

Also *if if $Tx_* = x_*$ and T is C^1 with $|T'(x_*)| > 1$, then there is a (small) interval $I = (x_* - \delta, x_* + \delta)$ and $c > 1$ so that for $x \in I$, $|Tx - Tx_*| \geq c|x - x_*|$.* This is called a repelling fixed point. You can similarly bound $\min |T'x| = c > 1$ on some interval I. Another application of MVT shows that $|Tx - Tx_*| \geq c|x - x_*|$. So the points are being pushed away from x_*. This is the case for $x_* = 0$.

A useful extension of the contraction Mapping Principle is the following variant.

3.3.4. COROLLARY. *Let (X, d) be a complete metric space, and let $T : X \to X$. Suppose that there is a positive integer k so that T^k is a contraction mapping. Then T has a unique fixed point x_*. Given any $x_0 \in X$, $x_* = \lim_{n\to\infty} T^n x_0$.*

PROOF. Since T^k is a contraction, it has a unique fixed point x_*. Observe that

$$T^k(Tx_*) = T(T^k x_*) = Tx_*.$$

Thus Tx_* is a fixed point of T^k. By uniqueness, $Tx_* = x_*$. So x_* is fixed for T. Given x_0 and $0 \leq i < k$, starting with $T^i x_0$, repeated application of T^k yields x_*. Thus

$$\lim_{n\to\infty} T^{nk+i} x_0 = \lim_{n\to\infty} T^{nk}(T^i x_0) = x_*.$$

Therefore $T^n x_0$ converges to x_*. Conversely, if y_* is any fixed point of T, then $T^k y_* = y_*$ as well. So the fixed point for T is unique. ■

3.3.1. Fractals. Suppose that X is a closed subset of $\mathbb{R}^n$ and $T_1, \dots, T_n$ are affine invertible contraction mappings of X with Lipschitz constants $c_i < 1$ for $1 \le i \le k$. (An affine map is a translation of a linear map.) Look for a closed subset $A \subset X$ such that

$$A = T_1 A \cup \cdots \cup T_k A.$$

Since we are using invertible affine mappings, each $T_i A$ is similar to A geometrically. If the $T_i A$ are almost disjoint (say except for a finite number of points), then this self-similarity property will be more apparent. A will look like the union of k smaller copies of A, each of which is a union of k even smaller copies, etc. Such a figure is called a *fractal*.

3.3.5. EXAMPLE. Let $X = \mathbb{R}^2$ and let

$$T_1 x = \frac{x}{2}, \quad T_2 x = (1,0) + \frac{x}{2} \quad \text{and} \quad T_3 x = (\tfrac{1}{2}, \tfrac{\sqrt{3}}{2}) + \frac{x}{2}.$$

Let A_0 be the solid equilateral triangle with vertices $(0,0)$, $(2,0)$ and $(1,\sqrt{3})$. Define $A_n = T_1 A_{n-1} \cup T_2 A_{n-1} \cup T_3 A_{n-1}$ for $n \ge 1$ Then $T_i A_0$ is an equilateral triangle of half the size. So A_1 is the union of three triangles, and looks like A_0 was divided into four equal triangles and the centre was removed. Likewise, A_2 looks like each solid triangle in A_1 had the middle triangle removed from each. In the limit, it converges to a figure known as Sierpinski's triangle.

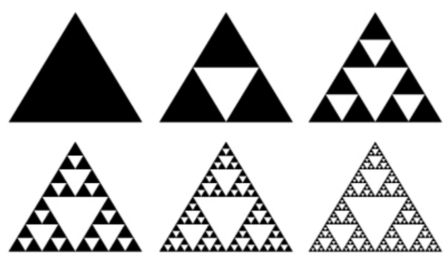

FIGURE 3.1. $A_0, A_1, A_2, A_3, A_4, A_5.$

The fact that T_i are affine and invertible turns out to be unimportant in terms of existence of a unique solution. It does, however, produce more symmetrical and pleasing results.

3.3.6. THEOREM. *Let (X,d) be a complete metric space, and let $T_i : X \to X$ be contraction mappings with Lipschitz constants $c_i < 1$ for $1 \le i \le k$. Let $\mathcal{H}(X)$ denote the space of non-empty closed bounded subsets of X with the Hausdorff metric. Define $T : \mathcal{H}(X) \to \mathcal{H}(X)$ by $TA = T_1 A \cup \cdots \cup T_k A$. Then T is a contraction mapping. It has a unique fixed point A_*, and it satisfies*

$$A_* = T_1 A_* \cup \cdots \cup T_k A_*.$$

PROOF. First we show that if $A_i, B_i \in \mathcal{H}(X)$ for $1 \le i \le n$, then

$$d_H(A_1 \cup \cdots \cup A_n, B_1 \cup \cdots \cup B_n) \le \max_{1 \le i \le n} d_H(A_i, B_i).$$

Let r be the right hand side. Then $A_i \subset (B_i)_r := \{x : d(x, B_i) \le r\}$ and similarly $B_i \subset (A_i)_r$. Therefore

$$A_1 \cup \cdots \cup A_n \subset (B_1 \cup \cdots \cup B_n)_r \quad \text{and} \quad (B_1 \cup \cdots \cup B_n) \subset (A_1 \cup \cdots \cup A_n)_r.$$

That proves the claim.

If $A, B \in \mathcal{H}(X)$, then

$$\begin{aligned} d_H(T_iA, T_iB) &= \max\{\sup_{a \in A} d(T_ia, T_iB), \sup_{b \in B} d(T_ib, T_iA)\} \\ &\le c_i \max\{\sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A)\} \\ &= c_i d_H(A, B). \end{aligned}$$

Let $c = \max\{c_i : 1 \le i \le n\}$. Then if $A, B \in \mathcal{H}(X)$, we have

$$\begin{aligned} d_H(TA, TB) &= d_H(T_1A \cup \cdots \cup T_kA, T_1B \cup \cdots \cup T_kB) \\ &\le \max\{c_i d_H(A, B)\} = c d_H(A, B). \end{aligned}$$

Therefore T is a contraction mapping with Lipschitz constant c.

By the Contraction Mapping Principle 3.3.2, T has a unique fixed point A_*. ■

3.4. Newton's Method

Newton's method is an iterative algorithm for finding zeros of nice functions that you have probably seen in your calculus class. It is frequently implemented for computer computation because it converges very quickly. In fact, it converges quadratically. This means that once you get sufficiently close to the solution, each iteration essentially doubles the number of significant digits. In fact the precision of the calculation may become the more serious problem.

3.4.1. DEFINITION. An algorithm for approximating a solution x_* by a sequence $(x_n)_{n \ge 0}$ *converges quadratically* if there is a constant C so that

$$|x_{n+1} - x_*| \le C|x_n - x_*|^2.$$

Start with a function $f \in C^2[a, b]$. Suppose that there is a point x_* such that $f(x_*) = 0$ and $f'(x_*) \ne 0$. You need to start with a point x_0 *sufficiently close* to x_*. Exactly how close depends on the function, but in some cases, there is a lot of leeway.

The idea is to take the tangent line through $(x_0, f(x_0))$ and solve for its root, x_1, which is the first step of the algorithm. Repeat, generating a sequence of approximations. As with any numerical method, it is important to have good error estimates.

The line through $(x_n, f(x_n))$ with slope $f'(x_n)$ is
$$y = f(x_n) + f'(x_n)(x - x_n).$$
The solution to $y = 0$ is x_{n+1} given by
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$
Define the mapping $Tx = x - \frac{f(x)}{f'(x)}$. Observe that $Tx_* = x_*$ precisely when $f(x_*) = 0$. We won't have any problem with the denominator being 0 if we are close to x_* because f' is continuous, and is non-zero at x_*. Compute
$$T'x = 1 - \frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2}.$$
Notice that $T'(x_*) = 0$, and hence T is very contractive near x_*.

3.4.2. Newton's Method. *Suppose that $f \in C^2$ and there is a point x_* such that $f(x_*) = 0$ and $f'(x_*) \neq 0$. Then there is an $r > 0$ so that $Tx = x - \frac{f(x)}{f'(x)}$ is a contraction mapping on $I = [x_* - r, x_* + r]$. Moreover there is a constant M so that $|x_{n+1} - x_*| \leq M|x_n - x_*|^2$.*

Proof. Based on the calculations preceding the proof, we can choose $r > 0$ so that $|T'x| \leq \frac{1}{2}$ on some interval $I = [x_* - r, x_* + r]$. This implies that $f'(x) \neq 0$ on I. By the Mean Value Theorem, if $x, y \in I$, there is a point $\xi \in (x, y)$ so that
$$|Tx - Ty| = |f'(\xi)(x - y)| \leq \tfrac{1}{2}|x - y|.$$
Thus T is a contraction mapping with Lipschitz constant $\frac{1}{2}$. In particular, if $x \in I$, $|Tx - x_*| = |Tx - Tx_*| \leq \frac{1}{2}|x - x_*|$. So Tx belongs to I; whence $TI \subset I$. The Contraction Mapping Principle shows that $x_n = T^n x_0$ converges to x_* and satisfies $|x_n - x_*| \leq 2^{-n}|x_0 - x_*|$. This is good, but it isn't quadratic convergence.

Let $A = \sup_{x \in I} |f''(x)|$ and $B = \inf_{x \in I} |f'(x)|$. We need to apply the MVT twice. First there is a point $\xi \in (x_*, x_n)$ so that
$$\frac{f(x_n) - f(x_*)}{x_n - x_*} = f'(\xi).$$
So $f(x_n) = f'(\xi)(x_n - x_*)$ because $f(x_*) = 0$. Therefore
$$\begin{aligned} x_{n+1} - x_* = (x_n - x_*) + (x_{n+1} - x_n) &= \frac{f(x_n)}{f'(\xi)} - \frac{f(x_n)}{f'(x_n)} \\ &= \frac{f(x_n)(f'(x_n) - f'(\xi))}{f'(\xi)f'(x_n)} = \frac{(x_n - x_*)(f'(x_n) - f'(\xi))}{f'(x_n)}. \end{aligned}$$

Now apply the MVT a second time to find $\zeta \in (\xi, x_n)$ so that

$$\frac{f'(x_n) - f'(\xi)}{x_n - \xi} = f''(\zeta).$$

Plugging this back in yields

$$|x_{n+1} - x_*| = \left|\frac{(x_n - x_*)(x_n - \xi)f''(\zeta)}{f'(x_n)}\right| \le \frac{A}{B}|x_n - x_*|^2.$$

This is quadratic convergence with $M = A/B$. ∎

3.4.3. EXAMPLE. The favourite example of this algorithm is the computation of square roots. Let $f(x) = x^2 - 153$. Note that f is smooth and $f'(x) = 2x$ is not zero on $I = [12, 13]$. Indeed, $B = \inf\{|f'(x)| : x \in I\} = 24$ and $A = \sup|f''(x)| = 2$. The map T is $Tx = x - \frac{x^2-153}{2x} = \frac{1}{2}(x + \frac{153}{x})$. Also $T'x = \frac{f(x)f''(x)}{f'(x)^2} = \frac{x^2-153}{2x^2}$. On I, this is bounded by $\frac{169-153}{2(144)} = \frac{1}{18}$. So T is a contraction. The constant $M = \frac{A}{B} = \frac{1}{12}$. Thus

$$\left|x_{n+1} - \sqrt{153}\right| \le \frac{|x_n - \sqrt{153}|^2}{12}.$$

Start with $x_0 = 12.5$. we can estimate

$$12.5 - \sqrt{153} = \frac{12.5^2 - 153}{12.5 + \sqrt{153}} < \frac{3.25}{12.5 + 12} < 0.14.$$

Hence

$$\left|x_1 - \sqrt{153}\right| < \frac{1}{12}(.14)^2 < .0017$$

and

$$\left|x_2 - \sqrt{153}\right| < \frac{1}{12}(1.7\,10^{-3})^2 < 2.5\,10^{-7}.$$

This is very rapid convergence. The real computational problem is computing $\frac{153}{x_n}$ to sufficient accuracy.

Exercises

1. Let (X, d) be a complete metric space. If $A \subset X$ is a G_δ set, prove that $\overline{A} \setminus A$ is first category.

2. Let (X, d) be a countable complete metric space. Prove that X has isolated points.

3. (a) Show that $\mathbb{R}^2$ is not the union of countably many lines.
   (b) More generally, show that a complete normed vector space is not the union of countably many translates of proper closed subspaces.
   (c) Show that there is no norm on the vector space $\mathbb{C}[x]$ of polynomials in which it is complete.

4. (a) Let (X,d) be a metric space, and let $f : X \to \mathbb{C}$ be a function. Show that the set of points of continuity of f is a G_δ.
(b) Prove that $\mathbb{Q}$ is not a G_δ subset of $\mathbb{R}$. Hence deduce that there is no function on $\mathbb{R}$ which is continuous precisely on $\mathbb{Q}$.

5. Let (X,d) be a compact metric space. Let V be a closed subspace of $C_{\mathbb{R}}(X)$ such that every $f \in V$ is Lipschitz. Prove that V is finite dimensional.
HINT: show that $A_n = \{f \in V : |f(x) - f(y)| \leq nd(x,y)\}$ has interior for some n. Hence show that the closed unit ball of V is equicontinuous.

6. Show that the unit cube $C = \{\mathbf{x} \in \mathbb{R}^d : 0 \leq x_i \leq 1,\ 1 \leq i \leq d\}$ is not the union of countably many *disjoint* non-empty closed sets by the following plan:
Suppose that A_n are disjoint non-empty closed sets such that $C = \bigcup_{n\geq 1} A_n$. Define the boundary of A_n to be $B_n = A_n \setminus \text{int}(A_n)$, (interior w.r.t. to C). Set $X = C \setminus \bigcup_{n\geq 1} \text{int}(A_n) = \bigcup_{n\geq 1} B_n$. Show that $b_r(\mathbf{x})$ for $\mathbf{x} \in B_n$ must intersect some B_m for $m \neq n$. What does the Baire Category Theorem say about $X = \bigcup_{n\geq 1} B_n$?

7. A *nowhere monotonic function* on $[0,1]$ is not monotonic on any interval. Show that these functions are a residual subset of $C[0,1]$.
HINT: Let $A_n = \big\{\pm f \in C[0,1] : \exists_{x\in[0,1]} \text{ s.t. } \big(f(t)-f(x)\big)(t-x) \geq 0 \text{ if } |t-x| < \frac{1}{n}\big\}$.

8. Let $f : [1,\infty) \to \mathbb{R}$ be continuous such that for every $x \geq 1$, $\lim_{n\to\infty} f(nx) = 0$. Prove that $\lim_{x\to\infty} f(x) = 0$.

9. (a) Show that $|\sin x - \sin y| < |x-y|$ for all $x \neq y \in \mathbb{R}$.
(b) Show that $Tx = \sin x$ is not a contraction mapping on $\mathbb{R}$.
(c) Show that if $x_0 \in \mathbb{R}$, then $x_n = T^n x_0$ converges .
(d) Use the Taylor expansion about $x = 0$ to show that if $x_0 = 1$, then convergence of x_n is much slower than geometric (i.e. for any $c < 1$, $\lim_{n\to\infty} \frac{|x_n|}{c^n} = +\infty$).

10. Let (X,d) be a complete metric space.
(a) Suppose that S and T are both contraction mappings with Lipschitz constant $c < 1$; with fixed points be x_S and x_T respectively. Define $d_\infty(S,T) = \sup_{x\in X} d(Sx,Tx)$. Prove that $d(x_S,x_T) \leq (1-c)^{-1} d_\infty(S,T)$.
(b) Show that if $t \to T_t$ for $t \in [0,1]$ is a continuous path of contraction mappings on X (with respect to d_∞) and they have a uniform Lipschitz constant $c < 1$, then the map $t \to x_t$ to the fixed points x_t of T_t is a continuous path in X.
(c) Construct a continuous path of maps $T_t : [0,1] \to [0,1]$ consisting of contractions, but the supremum of the Lipschitz constants is 1.

11. Let $T : C[0,1] \to C[0,1]$ be defined by $Tf(x) = 1 + \int_0^x f(t)\,dt$.
(a) Show that T is not a contraction mapping, but that T^2 is.
(b) Find the fixed point of T.

12. Let $A = [a_{ij}]$ be a linear transformation from $l_1^{(n)}$ to itself.
(a) Show that $\|A\| = \max\{\sum_{i=1}^n |a_{ij}| : 1 \leq j \leq n\} = \max_{1\leq j\leq n} \|Te_j\|_1$ is a norm on the vector space M_n of $n \times n$ matrices.
(b) Show that A is a contraction mapping if and only if $\|A\| < 1$.
(c) Suppose $\|I-A\| < 1$. Solve $Ax = b$ by finding the fixed point of $Tx = x - Ax + b$.

3.5. Metric Completion

In this section, we will show that every metric space sits inside a unique smallest complete metric space. This will be used soon to discuss the construction of the real numbers.

3.5.1. DEFINITION. If (X, d) is a metric space, a *completion* of X is a complete metric space (Y, ρ) together with a map $J : X \to Y$ which is isometric, i.e., $\rho(Jx_1, Jx_2) = d(x_1, x_2)$ for all $x_1, x_2 \in X$, and has dense range, i.e., $\overline{JX} = Y$.

We provide two proofs of our main result. The first is slick, but the second is more informative.

3.5.2. THEOREM. *Every metric space has a completion.*

FIRST PROOF. Recall from Theorem 1.8.4 that $C^b(X)$ is complete. Fix a point $x_0 \in X$, and for $x \in X$, define a function

$$f_x(y) = d(y, x) - d(y, x_0) \quad \text{for} \quad y \in X.$$

Note that this function is continuous because $d(y, x)$ and $d(y, x_0)$ are Lipschitz functions of y. By the triangle inequality,

$$-d(x, x_0) \leq d(y, x) - d(y, x_0) \leq d(x, x_0);$$

so that $\|f_x\|_\infty \leq d(x, x_0)$. This is sharp because $f_x(x_0) = d(x, x_0)$. Hence f_x belongs to $C^b(X)$. Define $J : X \to C^b(X)$ by $Jx = f_x$ for $x \in X$.

Now if $x_1, x_2 \in X$,

$$f_{x_1}(y) - f_{x_2}(y) = d(y, x_1) - d(y, x_0) - d(y, x_2) + d(y, x_0) = d(y, x_1) - d(y, x_2).$$

As above, the triangle inequality shows that $|f_{x_1}(y) - f_{x_2}(y)| \leq d(x_1, x_2)$. Taking $y = x_2$ shows that $\|f_{x_1} - f_{x_2}\|_\infty = d(x_1, x_2)$. Therefore J is an isometry,

Define $Y = \overline{JX}$. This is a closed subset of the complete space $C^b(X)$. Hence it is complete. By construction JX is dense in Y. Thus Y is a completion of X. ■

SECOND PROOF. Let $\mathcal{C} = \{(x_n)_{n \geq 1} : \text{Cauchy sequences in } X\}$ be the set of all Cauchy sequences in X. Let (x_n) and (y_n) belong to $\mathcal{C}$. By the triangle inequality

$$\big|d(x_n, y_n) - d(x_m, y_m)\big| \leq d(x_n, x_m) + d(y_n, y_m).$$

The right hand side is small if m, n are big enough. So $\big(d(x_n, y_n)\big)_{n \geq 1}$ is a Cauchy sequence. Therefore we may define a function $f : \mathcal{C} \times \mathcal{C} \to [0, \infty)$ by

$$R\big((x_n), (y_n)\big) = \lim_{n \to \infty} d(x_n, y_n).$$

This is a pseudo-metric: clearly it is symmetric and the triangle inequality is:

$$\begin{aligned} R\big((x_n),(z_n)\big) &= \lim_{n\to\infty} d(x_n,z_n) \\ &\le \lim_{n\to\infty} d(x_n,y_n) + d(y_n,z_n) \\ &= R\big((x_n),(y_n)\big) + R\big((y_n),(z_n)\big). \end{aligned}$$

By Assignment 1, A3, we obtain a metric space as follows. Put an equivalence relation on $\mathcal{C}$ by setting $(x_n) \sim (y_n)$ if $R\big((x_n),(y_n)\big) = 0$. It is easy to see that this is reflexive: $(x_n) \sim (x_n)$, symmetric: $(x_n) \sim (y_n)$ implies that $(y_n) \sim (x_n)$, and transitive: $(x_n) \sim (y_n)$ and $(y_n) \sim (z_n)$ imply that $(x_n) \sim (z_n)$. These are the requirements of an equivalence relation.

Let $Y = \mathcal{C}/\sim$ denote the set of equivalence classes of $\mathcal{C}$. Put a metric on Y by

$$\rho([(x_n)],[(y_n)]) = R\big((x_n),(y_n)\big).$$

First we show that this is well-defined, meaning that it is independent of the choice of representatives for the equivalence classes. So suppose that $(x'_n) \sim (x_n)$ and $(y'_n) \sim (y_n)$. Then

$$\begin{aligned} R\big((x'_n),(y'_n)\big) &= \lim_{n\to\infty} d(x'_n,y'_n) \\ &\le \lim_{n\to\infty} d(x'_n,x_n) + d(x_n,y_n) + d(y_n,y'_n) \\ &= R\big((x_n),(y_n)\big). \end{aligned}$$

Reversing the roles of the two representatives shows that

$$R\big((x'_n),(y'_n)\big) = R\big((x_n),(y_n)\big).$$

Thus ρ is well defined.

Clearly $\rho([(x_n)],[(y_n)]) = 0$ if and only if $\lim_{n\to\infty} d(x_n,y_n) = 0$ if and only if $(x_n) \sim (y_n)$ if and only if $[(x_n)] = [(y_n)]$. It is also clear that ρ is symmetric. For the triangle inequality, take $[(x_n)],[(y_n)],[(z_n)] \in Y$. Then

$$\begin{aligned} \rho\big([(x_n)],[(z_n)]\big) &= \lim_{n\to\infty} d(x_n,z_n) \\ &\le \lim_{n\to\infty} d(x_n,y_n) + d(y_n,z_n) \\ &= \rho\big([(x_n)],[(y_n)]\big) + \rho\big([(y_n)],[(z_n)]\big). \end{aligned}$$

Therefore ρ is a metric on Y.

Imbed X into Y by $Jx = (x,x,x,\dots)$. Then

$$\rho(Jx_1,Jx_2) = \lim_{n\to\infty} d(x,y) = d(x,y).$$

Thus J is an isometry. To see that JX is dense, let $[(x_n)] \in Y$ and let $\varepsilon > 0$. Since (x_n) is Cauchy, choose N so that for all $m,n \ge N$, $d(x_m,x_n) < \varepsilon/2$. Then Jx_N satisfies

$$\rho([(x_n)],Jx_N) = \lim_{n\to\infty} d(x_n,x_N) \le \varepsilon/2 < \varepsilon.$$

Therefore JX is dense in Y.

Finally we show that Y is complete. Let $(y_k)_{k\geq 1}$ be a Cauchy sequence in Y. For each k, choose $x_k \in X$ so that $\rho(y_k, Jx_k) < 2^{-k}$. Let $y_0 = [(x_k)]$. We claim that $\lim_{k\to\infty} y_k = y_0$. So let $\varepsilon > 0$. By the Cauchy property, there is an integer N so that if $N \leq m \leq n$, then $\rho(y_m, y_n) < \varepsilon/2$. Make N bigger if necessary so that $2^{-N} < \varepsilon/4$. Then for $N \leq m \leq n$

$$\begin{aligned} d(x_m, x_n) = \rho(Jx_m, Jx_n) &\leq \rho(Jx_m, y_m) + \rho(y_m, y_n) + \rho(y_n, Jx_n) \\ &< 2^{-m} + \frac{\varepsilon}{2} + 2^{-n} < \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

Thus (x_n) is a Cauchy sequence in X, so $y_0 = [(x_n)]$ is a point in Y. Moreover, if $m \geq N$,

$$\rho(Jx_m, y_0) = \lim_{n\to\infty} d(x_m, x_n) \leq \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, this shows that $\lim_{n\to\infty} Jx_n = y_0$. Finally,

$$\lim_{k\to\infty} \rho(y_k, y_0) \leq \lim_{k\to\infty} \rho(y_k, Jx_k) + \rho(Jx_k, y_0) = 0.$$

So $\lim_{k\to\infty} y_k = y_0$. Thus Y is complete. ■

It turns out that this completion is unique. To establish this, we first need a result of independent interest.

3.5.3. THE EXTENSION THEOREM. *Let (X, d) be a metric space with completion (Y, ρ), and let (Z, σ) be another complete metric space. If $f : X \to Z$ is a uniformly continuous function, then there is a unique (uniformly) continuous function $\tilde{f} : Y \to Z$ such that $\tilde{f}(Jx) = f(x)$ for $x \in X$.*

PROOF. We first show that if (x_n) is a Cauchy sequence in X, then $(f(x_n))$ is a Cauchy sequence in Z. Let $\varepsilon > 0$. By uniform continuity, there is a $\delta > 0$ so that $d(x, x') < \delta$ implies that $\sigma(f(x), f(x')) < \varepsilon$. Since (x_n) is Cauchy, there is an integer N so that if $N \leq m \leq n$, then $d(x_m, x_n) < \delta$; thus $\sigma(f(x_m), f(x_n)) < \varepsilon$. This just says that $(f(x_n))$ is Cauchy in Z.

Since Y is a completion of X, each point of Y is a limit of points in JX. So for $y \in Y$, choose a sequence (x_n) in X so that $y = \lim_{n\to\infty} Jx_n$. As (Jx_n) converges, it is a Cauchy sequence. And since J is an isometry, (x_n) is Cauchy in X. By the previous paragraph, we can define

$$\tilde{f}(y) = \lim_{n\to\infty} f(x_n).$$

We need to show that $\tilde{f}$ is well defined. That is, if (x'_n) is another sequence in X so that $y = \lim_{n\to\infty} Jx'_n$, we need to show that we assign the same value to $\tilde{f}(y)$. We see that the sequence $(Jx_1, Jx'_1, Jx_2, Jx'_2, \dots)$ converges to y and thus is Cauchy. So $(x_1, x'_1, x_2, x'_2, \dots)$ is a Cauchy sequence in X. Thus by the first paragraph, $(f(x_1), f(x'_1), f(x_2), f(x'_2), \dots)$ is a Cauchy sequence in Z. Hence $\lim_{n\to\infty} f(x'_n) = \lim_{n\to\infty} f(x_n)$. So $\tilde{f}$ is well defined.

Next, since (Jx) converges to Jx, we have

$$\tilde{f}(Jx) = \lim_{n\to\infty} f(x) = f(x) \quad \text{for} \quad x \in X.$$

So $\tilde{f}$ extends f. Finally we show $\tilde{f}$ is uniformly continuous. Let $\varepsilon > 0$. Again, there is a $\delta > 0$ so that $d(x, x') < \delta$ implies that $\sigma(f(x), f(x')) < \varepsilon$. Let $y, y' \in Y$ with $\rho(y_1, y_2) < \delta$. Then there are Cauchy sequences (x_n) and (x'_n) so that $y = \lim_{n\to\infty} Jx_n$ and $y' = \lim_{n\to\infty} Jx'_n$. So

$$\lim_{n\to\infty} d(x_n, x'_n) = \lim_{n\to\infty} \rho(Jx_n, Jx'_n) = \rho(y_1, y_2) < \delta.$$

Hence there is some integer M so that $d(x_n, x'_n) < \delta$ for $n \geq M$. Therefore $\sigma(f(x_n), f(x'_n)) < \varepsilon$. Taking limits yields $\sigma(\tilde{f}(y), \tilde{f}(y')) \leq \varepsilon$. So $\tilde{f}$ is uniformly continuous. Finally $\tilde{f}$ is unique because it is defined on a dense subset by $\tilde{f}(Jx) = f(x)$, and thus there is at most one way to extend it to be continuous on Y. ■

An important consequence of this result is the uniqueness of the metric completion,

3.5.4. COROLLARY. *The metric completion of (X, d) is unique in the sense that if $J_i : X \to (Y_i, \rho_i)$, $i = 1, 2$, are two metric completions of X, then there is a unique isometry κ of Y_1 onto Y_2 such that $J_2 = \kappa J_1$.*

PROOF. Define $\kappa_0 = J_2 : X \to Y_2$. Then κ_0 is an isometry, and hence is uniformly continuous. By the Extension Theorem 3.5.3, there is a unique continuous function $\kappa : Y_1 \to Y_2$ such that $\kappa J_1 = J_2$. If $y, y' \in Y_1$, choose sequences (x_n) and (x'_n) in X so that $y = \lim_{n\to\infty} J_1 x_n$ and $y' = \lim_{n\to\infty} J_1 x'_n$. Then

$$\begin{aligned} \rho_2(\kappa y, \kappa y') &= \lim_{n\to\infty} \rho_2(\kappa J_1 x_n, \kappa J_1 x'_n) \\ &= \lim_{n\to\infty} \rho_2(J_2 x_n, J_2 x'_n) = \lim_{n\to\infty} d(x_n, x'_n) \\ &= \lim_{n\to\infty} \rho_1(J_1 x_n, J_1 x'_n) = \rho_1(y, y'). \end{aligned}$$

Therefore κ is an isometry.

Since κ is an isometry, it takes Y onto a complete subset of Z, and thus it is closed. Also κY contains the dense set $J_2 X$. Therefore κ is onto. ■

3.6. The p-adic Numbers

Let p be a fixed prime number. For $x \in \mathbb{Q} \setminus \{0\}$, factor $x = p^a \frac{r}{s}$ where r, s, p are all relatively prime; and define $|x|_p = p^{-a}$ and $|0|_p = 0$. We showed in Example 1.2.2(5) that $d_p(x, y) = |x - y|_p$ is a metric on $\mathbb{Q}$. Let $\mathbb{Q}_p$ denote the completion of $(\mathbb{Q}, d_p)$. We will use the Cauchy sequence construction to describe elements of $\mathbb{Q}_p$. For $x \in \mathbb{Q}_p$, we define $|x|_p := d_p(x, 0)$, extending the definition of the norm.

3.6.1. PROPOSITION. *Let $x \in \mathbb{Q}_p \setminus \{0\}$, then $|x|_p \in \{p^a : a \in \mathbb{Z}\}$. If $x_n \in \mathbb{Q}$ such that $x = \lim_{n\to\infty} x_n$, then $|x|_p = \lim_{n\to\infty} |x_n|_p$ and $|x_n|_p$ is eventually constant.*

PROOF. Suppose that $x_n \in \mathbb{Q}$ such that $x = \lim_{n\to\infty} x_n$. Then

$$|x|_p = d_p(0,x) = \lim_{n\to\infty} d_p(x_n, 0) = \lim_{n\to\infty} |x_n|_p.$$

For n large, $x_n \neq 0$ so $d_p(x_n, 0) \geq \delta > 0$. Now $|x_n|_p \in \{p^a : a \in \mathbb{Z}\}$ and $\overline{\{p^a : a \in \mathbb{Z}\}} = \{p^a : a \in \mathbb{Z}\} \cup \{0\}$. So $|x|_p \in \{p^a : a \in \mathbb{Z}\} \cup \{0\}$, and is not 0 because $x \neq 0$. So $|x|_p = p^{a_0}$ is an isolated point of $\{p^a : a \in \mathbb{Z}\} \cup \{0\}$, and so for n sufficiently large, $|x_n|_p = |x|_p$. ∎

3.6.2. PROPOSITION. *Let $x = [(x_n)]$ and $y = [(y_n)]$ belong to $\mathbb{Q}_p$. Define $x \pm y = [(x_n \pm y_n)]$ and $xy = [(x_n y_n))]$. This makes $\mathbb{Q}_p$ into a commutative ring, and*

$$|xy|_p = |x|_p\, |y|_p \quad \textit{and} \quad |x \pm y|_p \leq \max\{|x|_p, |y|_p\}.$$

PROOF. Notice that

$$d_p(x_n + y_n - x_m - y_m) \leq \max\{d_p(x_n - x_m), d_p(y_n - y_m)\}.$$

Since (x_n) and (y_n) are Cauchy, so is their sum. Moreover it is easy to see that if $(x'_n) \sim (x_n)$ and $(y'_n) \sim (y_n)$ in the sense that

$$\lim_n d_p(x_n, x'_n) = 0 = \lim_n d_p(y_n, y'_n),$$

then $(x'_n + y'_n) \sim (x_n + y_n)$. Therefore addition is well defined. Moreover

$$|x \pm y|_p = \lim |x_n \pm y_n|_p \leq \lim \max\{|x_n|_p, |y_n|_p\} = \max\{|x|_p, |y|_p\}.$$

Similarly, if $x \neq 0 \neq y$, there is an N so that if $N \leq m \leq n$, then

$$\begin{aligned} d_p(x_n y_n, x_m y_m) &\leq \big|(x_n - x_m)y_y + x_m(y_n - y_m)\big|_p \\ &\leq \max\{|x_n - x_m|_p\, |y_n|_p, |x_m|_p\, |y_n - y_m|_p\} \\ &\leq \max\{|x_n - x_m|_p\, |y|_p, |x|_p\, |y_n - y_m|_p\}. \end{aligned}$$

It follows that $(x_n y_n)$ is a Cauchy sequence. Again it is easy to check that this multiplication is well defined. For $n \geq M$, we have that $|x_n|_p = |x|_p$ and $|y_n|_p = |y|_p$. Hence we have that $|x_n y_n| = |x_n|_p\, |y_n|_p = |x|_p\, |y|_p$. It follows that $|xy|_p = |x|_p\, |y|_p$. It is easy to check that $0x = 0$.

Now taking limits of the various ring axioms shows that $\mathbb{Q}_p$ is a commutative ring. ∎

Now we are ready for the main result of this section.

3.6.3. THEOREM. *$\mathbb{Q}_p$ is a topologically complete field containing $\mathbb{Q}$ as a dense subfield.*

PROOF. By construction $\mathbb{Q}_p$ is a complete metric space, and $\mathbb{Q}$ is a dense subset. The definition of addition and multiplication extend the operations on $\mathbb{Q}$. We need to show that non-zero elements are invertible.

Let $0 \neq x = [(x_n)] \in \mathbb{Q}_p$. Then there is an N so that $|x_n|_p = |x|_p \neq 0$ for $n \geq N$, Moreover

$$\Big|\frac{1}{x_n} - \frac{1}{x_m}\Big|_p = \frac{|x_n - x_m|_p}{|x_n|_p\,|x_m|_p} = \frac{|x_n - x_m|_p}{|x|_p^2}.$$

It follows that $\left(\frac{1}{x_n}\right)$ is a Cauchy sequence, and hence $y = \left[\left(\frac{1}{x_n}\right)\right] \in \mathbb{Q}_p$. Moreover $xy = [(1)] = 1$. Hence $\mathbb{Q}_p$ is a field. ∎

3.6.4. PROPOSITION. $\mathbb{Z}_p := \overline{\mathbb{Z}} = \{x \in \mathbb{Q}_p : |x|_p \leq 1\}$ *is a subring of* $\mathbb{Q}_p$.

PROOF. Note that by the strong triangle inequality, the set $\{x \in \mathbb{Q}_p : |x|_p \leq 1\}$ is closed and contains $\mathbb{Z}$; and hence contains $\mathbb{Z}_p$. Conversely, suppose that $x \in \mathbb{Q}_p$ with $|x|_p \leq 1$. Select $x_n \in \mathbb{Q}$ so that $d_p(x, x_n) \leq p^{-n} < 1$ for $n \geq 1$; so that $x = [(x_n)]$. Then $|x_n|_p \leq \max\{|x|_p, |x_n - x|_p\} \leq 1$. Write $x_n = p^{a_n}\frac{r_n}{s_n}$ where p, r_n, s_n are relatively prime, and $a_n \geq 0$ (because $1 \geq |x_n|_p = p^{-a_n}$). Solve the modular equation $r_n + p^n b_n \equiv 0 \pmod{s_n}$ for an integer b_n, which is possible because $\gcd(p^n, s_n) = 1$. Define $c_n = (r_n + p^n b_n)/s_n$. Define

$$x'_n = x_n + p^{a_n+n}\frac{b_n}{s_n} = p^{a_n}\frac{r_n + p^n b_n}{s_n} = p^{a_n}c_n$$

Then $x'_n \in \mathbb{Z}$ and $d_p(x'_n, x_n) = p^{-a_n-n}$ converges to 0 as $n \to \infty$; so $[(x'_n)] = [(x_n)] = x$. Hence $x = \lim_{n\to\infty} x'_n$ belongs to $\overline{\mathbb{Z}}$.

Since $\mathbb{Z}$ is closed under addition and multiplication, both of which are continuous operations, it follows that $\mathbb{Z}_p$ is also closed under addition and multiplication. Thus it is a subring of $\mathbb{Q}_p$. ∎

We now show that every element of $\mathbb{Q}_d$ has a p-adic expansion which we now describe,

3.6.5. LEMMA. *Let* $x \in \mathbb{Z}_p$*. Then there is a unique integer* $\alpha_0 \in \{0, 1, \dots, p–1\}$ *so that* $|x - \alpha_0|_p \leq \frac{1}{p}$*. For each* $i \geq 0$*, there is a unique integer* $\alpha_i \in \{0, 1, \dots, p–1\}$ *so that* $\left|x - \sum_{i=0}^{n} \alpha_i p^i\right|_p \leq \frac{1}{p^{n+1}}$.

PROOF. Since $\mathbb{Z}$ is dense in $\mathbb{Z}_p$, choose an integer k so that $|x - k|_p < 1$. Choose $\alpha_0 \in \{0, 1, \dots, p-1\}$ so that $k \equiv \alpha_0 \pmod{p}$. So $p | k - \alpha_0$, and hence $|k - \alpha_0|_p \leq \frac{1}{p}$. Then

$$|x - \alpha_0|_p \leq \max\{|x - k|_p, |k - \alpha_0|_p\} \leq \frac{1}{p}.$$

On the other hand, if $\beta \not\equiv k \pmod p$, then $k-\beta$ is not a multiple of p, so $|k-\beta|_p = 1$. Thus

$$1 = |k-\beta|_p \le \max\{|k-x|_p, |x-\beta|_p\} \le \max\{\frac{1}{p}, |x-\beta|_p\}.$$

Hence $|x-\beta|_p \ge 1$. So α_0 is unique.

Suppose that I have found $\alpha_i \in \{0, 1, \dots, p-1\}$ so that $\big|x - \sum_{i=0}^{n-1} \alpha_i p^i\big|_p \le \frac{1}{p^n}$. Let $y = p^{-n}\big(x - \sum_{i=0}^{n-1} \alpha_i p^i\big)$, and note that

$$|y|_p = p^n \Big|x - \sum_{i=0}^{n-1} \alpha_i p^i\Big|_p \le 1.$$

By the first paragraph, there is a unique $\alpha_n \in \{0, 1, \dots, p-1\}$ so that $|y-\alpha_n|_p \le \frac{1}{p}$. Therefore

$$\Big|x - \sum_{i=0}^{n} \alpha_i p^i\Big|_p = |p^n(y-\alpha_n)|_p \le \frac{1}{p^n}\,\frac{1}{p} = \frac{1}{p^{n+1}}.$$ ■

This leads to our p-adic expansion for elements of $\mathbb{Q}_p$.

3.6.6. THEOREM. *If $x \in \mathbb{Q}_p$ and $|x|_p = p^k$, then x has a unique expansion as an infinite series of the form $x = \sum_{i=-k}^{\infty} \alpha_i p^i$ where $\alpha_i \in \{0, 1, \dots, p-1\}$.*

PROOF. Let $y = p^k x$, so that $|y|_p = 1$. By Lemma 3.6.5, there is a unique sequence of integers $\beta_i \in \{0, 1, \dots, p-1\}$ so that $\big|y - \sum_{i=0}^{n} \beta_i p^i\big|_p \le \frac{1}{p^{n+1}}$. Multiply by p^{-k} and rename $\beta_i = \alpha_{i-k}$ to get

$$\Big|x - \sum_{i=-k}^{n-k} \alpha_i p^i\Big|_p \le \frac{p^k}{p^{n+1}}.$$

Letting $n \to \infty$ yields the desired convergent series. ■

Another consequence of this approach is the compactness of $\mathbb{Z}_p$.

3.6.7. PROPOSITION. *$\mathbb{Z}_p$ is compact.*

PROOF. $\mathbb{Z}_p$ is a closed subset of a complete space, and hence is complete. We claim that it is also totally bounded. Indeed we will show that $\{0, 1, \dots, p^n-1\}$ is a p^{-n}-net for $\mathbb{Z}_p$. Indeed, if $x \in \mathbb{Z}_p$, it follows from Lemma 3.6.5 that

$$\Big|x - \sum_{i=0}^{n-1} \alpha_i p^i\Big|_p \le \frac{1}{p^n}.$$

Moreover $k = \sum_{i=0}^{n-1} \alpha_i p^i < p^n$, so we have a p^{-n}-net. The Borel-Lebesgue Theorem 2.1.5 shows that $\mathbb{Z}_p$ is compact. ■

3.7. The Real Numbers

We have not talked about the construction of the real numbers. One approach, which fits well with our course, is to complete the rationals. Of course, we can't use our first proof that embeds $\mathbb{Q}$ into $C^b(\mathbb{Q})$ because this presumes the existence of $\mathbb{R}$. However the Cauchy sequence approach basically works, though again the way we define the metric presumes the existence of $\mathbb{R}$. Nevertheless we will see that it does work. There are several other constructions of $\mathbb{R}$ which we will also outline. The more serious issue is whether these various constructions produce the 'same' real numbers. This requires more subtlety.

3.7.1. DEFINITION. An *ordered field* is a field $\mathbb{F}$ which contains a subset $\mathbb{P}$ of *positive* elements satisfying

(i) $\mathbb{F} = \mathbb{P} \,\dot\cup\, \{0\} \,\dot\cup\, -\mathbb{P}$ is the disjoint union of the three sets $\mathbb{P}$, $\{0\}$ and $-\mathbb{P}$.

(ii) If $x, y \in \mathbb{P}$, then $x + y$ and xy are in $\mathbb{P}$.

We say that $x < y$ if $y - x \in \mathbb{P}$.

An ordered field has the *Least upper bound property* (LUBP) if whenever $\emptyset \neq S \subset \mathbb{F}$ is a nonempty subset which is bounded above (i.e., there is some $x \in \mathbb{F}$ so that $s \leq x$ for every $s \in S$), then there is a $y := \sup S$ such that y is an upper bound, and whenever x is another upper bound, then $y \leq x$. An ordered field with the LUBP is called *complete*.

An ordered field is *Archimedean* if whenever $x \in \mathbb{P}$, there is an $n \in \mathbb{N}$ so that $\frac{1}{n} < x$.

Notice that a complete ordered field has nothing to do with Cauchy sequences. One can define Cauchy sequences for any ordered field using arbitrary $\varepsilon \in \mathbb{P}$, but it turns out that it is a different property from the LUBP.

3.7.2. PROPOSITION. *Let $\mathbb{F}$ be an ordered field.*

(1) *Then $\mathbb{Q} \subset \mathbb{F}$ and $\mathbb{Q} \cap \mathbb{P} = \{r \in \mathbb{Q} : r > 0\}$.*

(2) *If $\mathbb{F}$ has the LUBP, then $\mathbb{F}$ is Archimedean.*

(3) *If $\mathbb{F}$ is Archimedean and $x < y$, then there is an $r \in \mathbb{Q}$ so that $x < r < y$. $($i.e. $\mathbb{Q}$ is order dense in $\mathbb{F}.)$*

PROOF. (1). $1 \in \mathbb{F}$ and $1 \neq 0$. So either $1 \in \mathbb{P}$ or $-1 \in \mathbb{P}$; but in either case, $1 = 1^2 = (-1)^2$ belongs to $\mathbb{P}$ by (ii) (and thus $-1 \in -\mathbb{P}$). For $n \in \mathbb{N}$, $n = 1 + \cdots + 1$ is the sum of n ones. This is positive by repeated application of (ii). In particular, n never equals 0; so in fact, they are all distinct because $n < n+1$ for all $n \in \mathbb{N}$. Thus $-\mathbb{N} \subset -\mathbb{P}$. If $n \in \mathbb{N}$, then $\frac{1}{n} \in \mathbb{F} \setminus \{0\}$. It must be positive, because if $\frac{1}{n} \in -\mathbb{P}$, then $-\frac{1}{n} \in \mathbb{P}$; so that $n(-\frac{1}{n}) = -1$ would be in $\mathbb{P}$. Therefore if $m, n \in \mathbb{N}$, then $\frac{m}{n} \in \mathbb{P}$ by (ii). So $\mathbb{Q} \subset \mathbb{F}$ and $\mathbb{Q} \cap \mathbb{P} = \mathbb{Q}^+$.

(2). Note that $\mathbb{F}$ is Archimedean if and only if

$$J := \{x \in \mathbb{P} : nx < 1 \text{ for all } n \in \mathbb{N}\}$$

is empty. If it is non-empty, say $x_0 \in J$, then let $y = \sup J$. Take any $x \in J$ and note that $x + x_0 \in J$ since if $n \in \mathbb{N}$, we have $2nx < 1$ and $2nx_0 < 1$, so that $2n(x + x_0) < 2$. Divide by 2 to get $n(x + x_0) < 1$ for all $n \in \mathbb{N}$. Therefore $y - x_0$ must be a smaller upper bound because $x + x_0 \le y$ implies that $x \le y - x_0$ for all $x \in J$. This shows that J does not have a least upper bound, which is a contradiction. So J is empty and $\mathbb{F}$ is Archimedean.

(3). If $\mathbb{F}$ is Archimedean and $x \in \mathbb{P}$, then there is some $m \in \mathbb{N}$ such that $0 \le x \le m$. If not, then $0 < \frac{1}{x} < \frac{1}{n}$ for all $n \in \mathbb{N}$, which contradicts the Archimedean property. Suppose that $x < y$. Then by the Archimedean property, there is an $n \in \mathbb{N}$ so that $\frac{1}{n} < y - x$. If $x \in \mathbb{P}$, then for some $m \in \mathbb{N}$, $0 \le x \le m$; and if $x \in -\mathbb{P}$, then for some $m \in \mathbb{N}$, we have $-m \le x \le 0$. In either case, among the finite set of numbers $\frac{k}{n}$ with $|k| \le mn$, there is a smallest one larger than x, say

$$\frac{k-1}{n} \le x < \frac{k}{n} \le x + \frac{1}{n} < y.$$

Thus $\mathbb{Q}$ is order dense in $\mathbb{F}$. ■

3.7.3. DEFINITION. An *embedding* of ordered fields $\mathbb{F}$ and $\mathbb{K}$ is an order preserving homomorphism $\gamma : \mathbb{F} \to \mathbb{K}$.

In the following result, we sketch the ideas, but some details are left to the reader to complete.

3.7.4. PROPOSITION. *Let $\mathbb{F}$ be an Archimedean ordered field, and let $\mathbb{K}$ be a complete ordered field. Then there is an embedding $\gamma : \mathbb{F} \to \mathbb{K}$.*

PROOF. Both fields contain a copy of $\mathbb{Q}$ as the subfield generated by the identity element, which we denote by $\mathbb{Q}_\mathbb{F}$ and $\mathbb{Q}_\mathbb{K}$. Let $\gamma_0 : \mathbb{Q}_\mathbb{F} \to \mathbb{Q}_\mathbb{K}$ be the identity homomorphism. For each $x \in \mathbb{F}$, define $S_x = \{r \in \mathbb{Q}_\mathbb{F} : r < x\}$. Define

$$\gamma(x) = \sup \gamma_0(S_x) \in \mathbb{K}.$$

Observe that if $x, y \in \mathbb{F}$, then

$$S_x + S_y := \{r + s : r \in S_x,\ s \in S_y\} = S_{x+y}.$$

It follows that $\gamma(x) + \gamma(y) = \gamma(x+y)$. If $x > 0$, then $S_x = \big(\mathbb{Q} \cap (-\mathbb{P} \cup \{0\})\big) \cup S_x^+$ where $S_x^+ = \{r \in \mathbb{Q}_\mathbb{F} : 0 < r < x\}$. Multiplication is a bit more delicate, but one can check that if $x, y \in \mathbb{P}$, then

$$S_{xy} = \big(\mathbb{Q} \cap (-\mathbb{P} \cup \{0\})\big) \cup S_x^+ S_y^+ = \big(\mathbb{Q} \cap (-\mathbb{P} \cup \{0\})\big) \cup \{rs : r \in S_x^+, s \in S_y^+\}.$$

From this, we deduce that $\gamma(x)\gamma(y) = \gamma(xy)$ when $x, y > 0$. With a bit of work, one can verify that γ is a homomorphism. ■

3.7.5. THEOREM. *There is a unique complete ordered field up to order preserving isomorphism.*

PROOF. Let $\mathbb{K}$ and $\mathbb{L}$ be two complete ordered fields. By Proposition 3.7.2(2), they are both Archimedean. By Proposition 3.7.4, there is an embedding $\gamma : \mathbb{K} \to \mathbb{L}$ and an embedding $\gamma' : \mathbb{L} \to \mathbb{K}$, Then $\gamma'\gamma : \mathbb{K} \to \mathbb{K}$ is an order preserving homomorphism. Since it carries 1 to 1, it must be the identity map on $\mathbb{Q}$. Now for each $x \in \mathbb{K}$, $x = \sup S_x$. Since $\gamma'\gamma$ preserves order,

$$\gamma'\gamma(x) = \sup \gamma'\gamma(S_x) = \sup S_x = x.$$

Thus $\gamma'\gamma = \mathrm{id}_{\mathbb{K}}$. Likewise $\gamma\gamma' = \mathrm{id}_{\mathbb{L}}$. Therefore γ is an order preserving isomorphism; i.e., $\mathbb{K}$ is unique up to order preserving isomorphism. ∎

3.7.6. DEFINITION. The unique complete ordered field is called $\mathbb{R}$, the field of real numbers.

Now we describe a few methods for constructing $\mathbb{R}$. There are many. We will be a bit sketchy on some of the details.

3.7.1. Cauchy sequences. We modify the second proof of Theorem 3.5.2 to complete $\mathbb{Q}$. Start with the set $\mathcal{C}$ of all Cauchy sequences in $\mathbb{Q}$ and define the equivalence relation $(x_n) \sim (y_n)$ if $\lim_{n\to\infty} x_n - y_n = 0$. Let $R = \mathcal{C}/\sim$ be the collection of equivalence classes. We embed $\mathbb{Q}$ into R by $\gamma(r) = [(r, r, r, \dots)]$. Note that these constructions do not require the real numbers.

We make R into a commutative ring by defining

$$\mathbf{0} = \gamma(0) \quad \text{and} \quad \mathbf{1} = \gamma(1)$$
$$[(x_n)] \pm [(y_n)] = [(x_n \pm y_n)]$$
$$[(x_n)]\,[(y_n)] = [(x_n y_n)].$$

The details to check that this is indeed a commutative ring $(R, 0, 1, +, \cdot)$ is left to the reader. This includes associativity of addition and multiplication, and the distributive law. They are all easy to deduce from the corresponding property of $\mathbb{Q}$. We leave the issue of inverses until later.

What we can't do is define a metric on R by taking limits, since these limits generally do not exist in $\mathbb{Q}$. However we can define the order. If $x = [(x_n)] \neq \mathbf{0}$, say that $x > 0$ (i.e. $x \in \mathbb{P}$) if there is an integer N so that $x_n > 0$ for all $n \geq N$. Note that this extends the notion of positivity for $\mathbb{Q}$. We need to check that every non-zero element x is either positive or negative; and that this is a well-defined notion.

So let $x = [(x_n)] \neq \mathbf{0}$. That means that (x_n) does not converge to 0. Thus there is some $\varepsilon > 0$ in $\mathbb{Q}$ and $n_i \to \infty$ so that $|x_{n_i}| \geq \varepsilon$. Use the Cauchy property to find N so that if $N \leq m \leq n$, then $|x_n - x_m| < \varepsilon/2$. Now choose $n = n_i > N$. If $x_{n_i} > \varepsilon$, then $x_m > \varepsilon/2$ for all $m \geq N$; while if $x_{n_i} < -\varepsilon$, we have $x_m < -\varepsilon/2$

for all $m \geq N$. In the first case, $x > \mathbf{0}$ and in the second case, $x < \mathbf{0}$. If $(x'_n) \sim (x_n)$, then since $\lim_{n\to\infty} x'_n - x_n = 0$, we have $|x'_n - x_n| < \varepsilon/4$ for $n \geq N'$. Thus for $n \geq \max\{N, N'\}$, x'_n and x_n have the same sign. So positivity is well defined. It is now easy to see that if $x, y \in \mathbb{P}$, then $x+y$ and xy also belong to P.

Finally consider inverses. By the argument in the previous paragraph, if $x = [(x_n)] \neq \mathbf{0}$, then for some N, $|x_n| > \varepsilon/2$ for all $n \geq N$. So we can define $y = [(y_n)]$ where $y_n = 0$ if $n < N$ and $y_n = x_n^{-1}$ for $n \geq N$. It is routine to check that (y_n) is Cauchy. Moreover $xy = [(x_n y_n)] = \mathbf{1}$ since $x_n y_n = 1$ for $n \geq N$. Thus every non-zero element of R has an inverse. So R is an ordered field.

Next we show that R is Archimedean. Suppose that $x \in \mathbb{P}$. We showed above that there is some $\varepsilon > 0$ in $\mathbb{Q}$ so that $x_n \geq \varepsilon/2$ for all $n \geq N$. Choose $n \in \mathbb{N}$ so that $\frac{1}{n} < \frac{\varepsilon}{2}$. Then it is easy to see that $\gamma(\frac{1}{n}) < x$. Therefore by Proposition 3.7.2(3), $\gamma(\mathbb{Q})$ is order dense in R.

Finally we need to verify the LUBP for R. Let $S \subset R$ be a nonempty set which is bounded above by $z \in R$, and let $s \in S$. Since R is Archimedean, we can find integers $a < s \leq z < b$. Recursively define sequences x_n and y_n of rational numbers as follows. Let $x_1 = a$ and $y_1 = b$. Suppose that x_i and y_i have been defined in $\mathbb{Q}$ for $1 \leq i < n$ so that $\gamma(x_i)$ is not an upper bound for S and $\gamma(y_i)$ is an upper bound for S and $y_i - x_i = 2^{1-i}(b-a)$. Let $c_n = \frac{1}{2}(x_{n-1} + y_{n-1})$. If c_n is an upper bound for S, then let $x_n = x_{n-1}$ and $y_n = c_n$; while if c_n is not an upper bound for S, then let $x_n = c_n$ and $y_n = y_{n-1}$. Let $x = [(x_n)]$. Then $x = [(y_n)]$ because $\lim_{n\to\infty} y_n - x_n = 0$. We claim that $\sup S = x$.

Let $s = [(s_n)] \in S$. If $s > x$, then by the Archimedean property, $s > x + \gamma(\frac{1}{d})$ for some $d \in \mathbb{N}$. So there is an integer N so that $s_n > y_n + \frac{1}{2d}$ for all $n \geq N$. Choose $M \geq N$ so that $2^{1-M}(b-a) < \frac{1}{4d}$. Then for $n \geq M$

$$y_n = y_M + \sum_{i=M+1}^{m} (y_i - y_{i-1}) > y_M - \sum_{i=M+1}^{m} 2^{1-i}(b-a) > y_M - \frac{1}{4d}.$$

Therefore for $n \geq M$, we have $s_n > y_M + \frac{1}{4d}$. This contradicts the fact that $\gamma(y_M)$ is an upper bound. So no such s exists, and x is an upper bound for S. A similar argument shows that if $w < x$, then w is not an upper bound.

The result is a construction of a complete ordered field, which we call $\mathbb{R}$.

3.7.2. Dedekind cuts.

A clever construction of $\mathbb{R}$ due the Richard Dedekind uses the following notion. It is perhaps the easiest construction.

3.7.7. DEFINITION. A *cut* is a proper subset C of $\mathbb{Q}$ with no largest element such that it is downward directed: if $r \in C$ and $s < r$, then $s \in C$. Let $\mathcal{R}$ denote the collection of all cuts.

We embed $\mathbb{Q}$ into $\mathcal{R}$ by $\gamma(r) = \{s \in \mathbb{Q} : s < r\}$. However there are other cuts, such as $C = \{r : r < 0 \text{ or } r^2 < 2\}$.

Order is easily defined: $C < D$ if $C \subsetneq D$ and $C > D$ if $C \supsetneq D$. In particular, $C > 0$ if and only if $0 \in C$.

This makes the LUBP very easy. Suppose that $S \subset \mathcal{R}$ has an upper bound D, let $E = \bigcup_{C \in S} C$. It is clear that E is proper since it is contained in D, and it is downward directed since it is the union of such sets. E cannot contain a largest element r_0 because then there would be a $C \in S$ with $r_0 \in C$ and r_0 would also be the largest element of C. So E is a cut. By construction, $C \leq E$ for all $C \in S$. But if F is another upper bound, then $C \subseteq F$ for all $C \in S$, and hence $E \subseteq F$; so $E \leq F$. Therefore $E = \sup S$.

It remains to define the field operations. Addition is defined by

$$C + D = \{r + s : r \in C,\ s \in D\}.$$

Multiplication of two *nonnegative* numbers is defined by

$$CD = \{r \in \mathbb{Q} : r < 0\} \cup \{rs : r \in C,\ r \geq 0,\ s \in D, s \geq 0\}.$$

Now you have to extend the definition to the rest, and verify all of the field laws. It is a bit tedious, but isn't difficult.

3.7.3. Infinite decimals. This is an intuitive and familiar construction, but it is actually more difficult to carry out than the others. However because of its familiarity, it is a good 'working version' of the real numbers.

We consider the set of all infinite decimal expansions $x = a_0.a_1a_2a_3\dots$ where where $a_0 \in \mathbb{Z}$ is an integer and each $a_i \in \{0, 1, 2, \dots, 9\}$ for $i \geq 1$. While this is familiar, there are a number of problems. First it seems to depend on being in base 10. Secondly some numbers have two names. Thirdly, defining addition is hard and multiplication is harder. We will only discuss the bare bones of this approach.

We will interpret this number $x = a_0.a_1a_2a_3\dots$ as a real number lying in the interval $[a_0, a_0 + 1]$. Note that this is a bit different from the standard practice of writing negative real numbers as the additive inverse of a positive infinite decimal expansion. So that what we normally call $-1.73000\dots$ will be written as $(-2).27000\dots$. At this point, we think of the infinite decimal expansion as a *name* for the real number x, and it does not imply that there is an infinite convergent series in the background. Nevertheless, this idea leads us to one important issue: some numbers have two names. For example, $x = 1.000\dots$ and $y = 0.999\dots$ should both represent the number 1.

We put an equivalence relation on the infinite decimals: let $x = a_0.a_1a_2a_3\dots$ and $y = b_0.b_1b_2b_3\dots$; say $x \sim y$ if

(1) $a_i = b_i$ for all $i \geq 0$, or

(2) there is some $i_0 \geq 0$ so that $a_i = b_i$ for $0 \leq i < i_0$, $b_{i_0} = a_{i_0} + 1$ and $a_i = 9$ and $b_i = 0$ for $i > i_0$, or

(3) there is some $i_0 \geq 0$ so that $a_i = b_i$ for $0 \leq i < i_0$, $a_{i_0} = b_{i_0} + 1$ and $a_i = 0$ and $b_i = 9$ for $i > i_0$.

When $x \sim y$, we will write $x = y$. Equivalence relations require three properties: reflexivity $x \sim x$, symmetry $x \sim y$ means that $y \sim x$ and transitivity $x \sim y$ and $y \sim z$ mean that $x \sim z$. In this case, it isn't hard to see that the equivalence class contains either a single infinite decimal which does not end in an infinite sequence of 0's or 9's, or it contains two infinite decimals, one (usually called a finite decimal) ending in an infinite string of 0's and a second ending in an infinite string of 9's. So it is easy to verify that this is an equivalence relation.

We can define the order by saying that $x = a_0.a_1a_2a_3 \cdots < y = b_0.b_1b_2b_3 \ldots$ if $x \not\sim y$ and there is an integer $i_0 \geq 0$ so that $a_i = b_i$ for $i < i_0$ and $a_{i_0} < b_{i_0}$. We establish the LUBP using the proof of Theorem 1.7.1.

The tricky bit is to define addition and multiplication. However one can use the order to help. Using x and y as before, we have that for any $n \in \mathbb{N}$

$$a_0.a_1a_2 \ldots a_n \leq x \leq a_0.a_1a_2 \ldots a_n + 10^{-n}$$
$$b_0.b_1b_2 \ldots b_n \leq y \leq b_0.b_1b_2 \ldots b_n + 10^{-n}.$$

We can add the left and right hand sides because these are rational numbers to get

$$z = c_0.c_1c_2 \ldots c_n \leq x + y \leq z + 2 \cdot 10^{-n}.$$

A similar argument works for multiplication when x, y are nonnegative. Details are omitted.

A similar construction can be obtained using any other base, for example binary numbers. Theorem 3.7.5 shows that the field that we obtain is independent of the construction, and in this case, of the base used. If we use one of the other contructions of $\mathbb{R}$, it is also straightforward to show that every number has a decimal expansion and that our equivalence relation describes the occasions when two different decimal expansions represent the same number.

Exercises

1. Show that the completion of a normed vector space is a normed vector space.

2. Show that the completion of (X, d) is compact if and only if X is totally bounded.

3. Consider $(C([0,1], \|\cdot\|_1)$. Let $L^1(0,1)$ denote its completion in this norm.
(a) Show $L^1(0,1)$ is a Banach space.
(b) Show that the Riemann integral $J(f) = \int_0^1 f(x)\,dx$ extends to a continuous function on $L^1(0,1)$. We will write $\int f$ for this extended function.
(c) (i) Show that $\int sf + tg = s\int f + t\int g$ for all $f, g \in L^1(0,1)$ and $s, t \in \mathbb{C}$.
(ii) Prove that $\left|\int f\right| \leq \int |f| = \|f\|_1$.
(d) Say that $f \in L^1(0,1)$ satisfies $f \geq 0$ if it is a limit in $L^1(0,1)$ of positive functions in $C[0,1]$. Thus $f \leq g$ if $g - f \geq 0$. Suppose that $f_n \in L^1(0,1)$, $n \geq 1$, is monotone increasing and $\sup_{n\geq 1} \|f_n\|_1 < \infty$. Prove that $\lim f_n = f$ exists in $L^1(0,1)$ and that $\int \lim f_n = \lim \int f_n$.

4. Find the p-adic expansion of -1.

5. Show that if $p > 2$ and $x^2 \equiv 2 \pmod p$ has a solution, then 2 has a square root in $\mathbb{Z}_p$.

6. Show that an ordered field $\mathbb{F}$ is Archimedean if and only if for every $x > 0$, there is an $n \in \mathbb{N}$ so that $x < n$.

7. Show that $\mathbb{C}$ can't be ordered to be an ordered field. HINT: what about $\pm i$?

8. Let $\mathbb{F} = \left\{ \frac{p}{q} : p, q \in \mathbb{R}[x], q \neq 0 \right\}$, where $q \neq 0$ means that q is not the 0 polynomial.
(a) Show that $\mathbb{F}$ is a field.
(b) Say that $\frac{p}{q} > 0$ if p and q are non-zero with leading terms $a_n x^n$ and $b_m x^m$ and $a_n b_m > 0$. Show that $\mathbb{F}$ is an ordered field.
(c) Show that $0 < \frac{1}{x} < \frac{1}{n}$ for all $n \in \mathbb{N}$, so that $\mathbb{F}$ is not Archimedean.
(d) Show that $\mathbb{N}$ is bounded above but does not have a least upper bound.

CHAPTER 4

Approximation Theory

4.1. Polynomial Approximation

Problem: Given $f \in C[a,b]$ and $\varepsilon > 0$, find a polynomial p so that

$$\|f-p\|_\infty = \sup_{a\le x\le b} |f(x)-p(x)| < \varepsilon.$$

First attempt. Let $x_i = a + \frac{i(b-a)}{n}$ for $0 \le i \le n$. There is a unique polynomial p of degree at most n such that $p(x_i) = f(x_i)$ for $0 \le i \le n$, obtained by Lagrange interpolation. Define

$$q_i(x) = \prod_{\substack{0\le j\le n\\ j\ne i}} \frac{x-x_i}{x_j-x_i}$$

and observe that $q_i(x_i) = 1$ and $q_i(x_j) = 0$ if $j \ne i$. Then the desired polynomial is $p(x) = \sum_{i=0}^n f(x_i)q_i(x)$. It is unique because if p_1 is another polynomial of degree at most n such that $p_1(x_i) = f(x_i)$ for $0 \le i \le n$, then $r(x) = p_1(x) - p(x)$ has degree at most n and $r(x_i) = 0$ for $0 \le i \le n$. If a polynomial of degree at most n has $n+1$ zeros, then it is the zero polynomial; thus $p_1 = p$.

This seems like a reasonable approach, but it doesn't work. A counterexample was constructed by Carl Runge in 1901.

Second attempt. Taylor polynomials. These work very well for certain very nice functions like $\sin x$ and e^x. However in general, a continuous function may have no derivative.

Even if f is C^∞, the radius of convergence of the power series may be too small. A typical example from calculus is

$$\frac{1}{1+x^2} = \sum_{n\ge 0}(-x^2)^n \quad \text{for} \quad |x|<1.$$

However this function is defined on the whole real line, but the series diverges if $|x| \ge 1$. (The power series actually makes sense in $\mathbb{C}$, but has a pole at $\pm i$ on the circle of radius 1. That is what causes the problems.)

Another even more troublesome example from calculus is

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if} \quad x \ne 0 \\ 0 & \text{if} \quad x = 0. \end{cases}$$

It turns out that this function is C^∞ and $f^{(k)}(0) = 0$ for all $n \geq 0$. So the Taylor series about $x = 0$ is the zero series. This converges uniformly on $\mathbb{R}$ to 0, but only agrees with f at one point!

4.1.1. WEIERSTRASS APPROXIMATION THEOREM. *The polynomials are dense in $C[a, b]$.*

Bernstein's proof. We will prove the theorem for the unit interval $[0, 1]$. Consider the terms arising in the binomial theorem:

$$1 = \big(x + (1-x)\big)^n = \sum_{i=0}^{n} \binom{n}{i} x^i (1-x)^{n-i} =: \sum_{i=0}^{n} P_i^n(x)$$

where $P_i^n(x) = \binom{n}{i} x^i (1-x)^{n-i}$ for $0 \leq i \leq n$. Note that $P_i^n(x) \geq 0$ on $[0, 1]$.

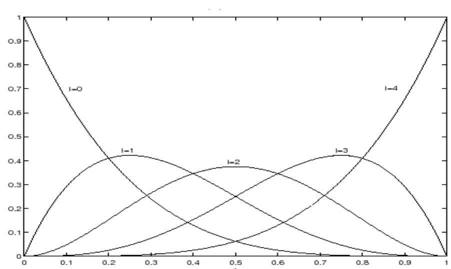

FIGURE 4.1. P_i^4 for $0 \leq i \leq 4$.

Moreover it is easy to check that $(P_i^n)'(x) = 0$ at $x = \frac{i}{n}$. So P_i^n has a maximum at $\frac{i}{n}$. Bernstein defined a *linear map* $B_n : C_{\mathbb{R}}[0, 1] \to \mathbb{R}[x]$ by

$$(B_n f)(x) = \sum_{i=0}^{n} f\big(\tfrac{i}{n}\big) P_i^n(x) = \sum_{i=0}^{n} f\big(\tfrac{i}{n}\big) \binom{n}{i} x^i (1-x)^{n-i}.$$

Since $B_n f$ is a linear combination of polynomials of degree n, it follows that this is a polynomial of degree at most n.

4.1.2. LEMMA. *B_n is a positive linear map, i.e.*

(1) *$B_n(sf + tg) = sB_n f + tB_n g$ for $f, g \in C_{\mathbb{R}}[0, 1]$ and $s, t \in \mathbb{R}$. (linear)*

(2) *$f \geq 0$ implies that $B_n f \geq 0$. (positive)*

(3) *$f \geq g$ implies $B_n f \geq B_n g$. (monotone)*

(4) *$|f| \leq g$ implies $|B_n f| \leq B_n g$.*

PROOF. (1) is easy. (2) follows because each $P_i^n \geq 0$, so if $f \geq 0$, then $B_n f$ is a sum of positive functions.

(3) If $f \le g$, then $0 \le g - f$, so that $0 \le B_n(g-f) = B_n g - B_n f$. Thus $B_n f \le B_n g$.

(4) $|f| \le g$ means $-g \le f \le g$, and hence $-B_n g \le B_n f \le B_n g$. Since $B_n g \ge 0$, we have $|B_n f| \le B_n g$. ∎

4.1.3. LEMMA.

(1) $B_n 1 = 1$.

(2) $B_n x = x$.

(3) $B_n x^2 = \frac{n-1}{n}x^2 + \frac{1}{n}x = x^2 + \frac{x-x^2}{n}$. *Hence* $B_n x^2$ *converges uniformly to* x^2 *as* $n \to \infty$.

PROOF. (1) follows from the binomial theorem:

$$B_n 1 = \sum_{k=0}^{n} 1 \binom{n}{k} x^k (1-x)^{n-k} = 1.$$

For (2), compute $\frac{\partial}{\partial x}(x+y)^n$ in two ways to get

$$\sum_{k=0}^{n} \binom{n}{k} k x^{k-1} y^{n-k} = n(x+y)^{n-1}. \tag{4.1.4}$$

Multiply equation (4.1.4) by $\frac{x}{n}$ and substitute $y = 1-x$ to get

$$B_n x = \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k} = x(x + (1-x))^{n-1} = x. \tag{4.1.5}$$

Now for (3), take $\frac{\partial}{\partial x}$ of equation (4.1.4) to get

$$\sum_{k=0}^{n} \binom{n}{k} k(k-1) x^{k-2} y^{n-k} = n(n-1)(x+y)^{n-2}. \tag{4.1.6}$$

Multiply equation (4.1.6) by $\frac{x^2}{n^2}$ and substitute $y = 1-x$ to get

$$\sum_{k=0}^{n} \frac{k^2-k}{n^2} \binom{n}{k} x^k (1-x)^{n-k} = \frac{n-1}{n} x^2 (x + (1-x))^{n-2} = \frac{n-1}{n} x^2. \tag{4.1.7}$$

Now add $\frac{1}{n}$ times equation(4.1.5) to equation(4.1.7) to get

$$\begin{aligned} B_n x^2 &= \sum_{k=0}^{n} \frac{k^2}{n^2} \binom{n}{k} x^k (1-x)^{n-k} \\ &= \frac{n-1}{n} x^2 + \frac{1}{n} x = x^2 + \frac{x-x^2}{n}. \end{aligned}$$

Then $\|x^2 - B_n x^2\|_\infty = \frac{1}{n}\|x - x^2\|_\infty = \frac{1}{4n}$. So $B_n x^2$ converges to x^2 uniformly. ∎

4.1.8. Corollary. $\left|\big(B_n(x-a)^2\big)(a)\right| \le \frac{1}{4n}$ *for* $a \in [0,1]$.

Proof. $B_n(x-a)^2 = B_n x^2 - 2aB_n x + a^2 B_n 1 = (x-a)^2 + \frac{1}{n}(x-x^2)$. Substitute $x = a$ to get $\big(B_n(x-a)^2\big)(a) = \frac{1}{n}(a-a^2) \in [0, \frac{1}{4n}]$. ∎

Proof of Weierstrass's Theorem. Fix a function f in $C_{\mathbb{R}}[0,1]$, and let $\varepsilon > 0$. Since f is uniformly continuous on $[0,1]$, there is $\delta > 0$ so that

$$|f(x) - f(a)| \le \varepsilon \quad \text{if} \quad |x-a| \le \delta, \;\; x, a \in [0,1].$$

Also, if $|x - a| \ge \delta$ for $x, a \in [0,1]$,

$$|f(x) - f(a)| \le 2\|f\|_\infty \le \frac{2\|f\|_\infty}{\delta^2}(x-a)^2.$$

Therefore for all $x, a \in [0,1]$,

$$|f(x) - f(a)| \le \varepsilon + \frac{2\|f\|_\infty}{\delta^2}(x-a)^2.$$

Hence if we make a a constant and x the variable, Lemma 4.1.2(4) shows

$$|B_n f(x) - f(a)| \le \varepsilon B_n 1 + \frac{2\|f\|_\infty}{\delta^2} B_n(x-a)^2.$$

Plug in $x = a$ and use Corollary 4.1.8 to get

$$|B_n f(a) - f(a)| \le \varepsilon + \frac{2\|f\|_\infty}{\delta^2}\frac{1}{4n}.$$

Now if $n \ge N(\varepsilon) := \left\lceil \frac{2\|f\|_\infty}{4\delta^2\varepsilon} \right\rceil$, we get

$$\|B_n f - f\|_\infty \le 2\varepsilon.$$

Therefore $B_n f$ converges uniformly to f.

If f is a complex valued continuous function, decompose $f = g + ih$ where $g(x) = \operatorname{Re} f(x)$ and $h(x) = \operatorname{Im} f(x)$. Find real polynomials p_n and q_n converging uniformly to g and h, respectively. Then $p_n + iq_n$ does the job.

Now consider an arbitrary interval $[a,b]$. Make a linear change of variables: if $f \in C[a,b]$, define $g(t) = f\big(a + (b-a)t\big)$ for $t \in [0,1]$. Since $x = a + (b-a)t$, we have $t = \frac{x-a}{b-a}$. Find polynomials p_n converging uniformly to g on $[0,1]$. Let $q_n(x) = p_n\big(\frac{x-a}{b-a}\big)$. Then q_n converges uniformly to $g\big(\frac{x-a}{b-a}\big) = f(x)$ on $[a,b]$. ∎

4.2. Best Approximation

The question we address in this section is whether there is a best approximation of degree n, i.e., a closest polynomial in the subspace $\mathbb{P}_n[a,b] \subset C[a,b]$ of all polynomials of degree at most n. Here $\mathbb{P}_n[a,b]$ means the vector space of polynomials of degree at most n with norm $\|p\|_\infty = \sup\{|p(x)| : a \le x \le b\}$. In general, when finding approximations in an infinite dimensional space like $C[a,b]$, there need not be a closest point. However in this case, $\mathbb{P}_n[a,b]$ is finite dimensional, so it turns out that there is a closest one. Also in norms with flat spots on the unit ball, which happens with the supremum norm, there can sometimes be many closest points. So we want to know if the closest point is unique.

4.2.1. DEFINITION. The *error of approximation* for $f \in C[a,b]$ is

$$E_n(f) = \operatorname{dist}(f, \mathbb{P}_n[a,b]) := \inf\{\|f - p\|_\infty : \deg p \leq n\}.$$

4.2.2. PROPOSITION. *If $f \in C[a,b]$ and $n \geq 0$, there exists a polynomial p of degree at most n so that*

$$\|f - p\|_\infty = E_n(f).$$

PROOF. The closest polynomial in $\mathbb{P}_n[a,b]$ is no further from f than 0, namely $\|f\|_\infty$. Hence it must lie in $X_n := \overline{B_{\|f\|_\infty}(f)} \cap \mathbb{P}_n[a,b]$. The subspace $\mathbb{P}_n[a,b]$ has dimension $n+1$. Thus by Corollary 1.7.7, $\mathbb{P}_n[a,b]$ is complete and hence closed in $C[a,b]$. Therefore X_n is a closed and bounded subset of a finite dimensional space. Hence it is compact by the Heine-Borel Theorem.

Define a function on X_n by $D(p) = \|f - p\|_\infty$. This function is Lipschitz and thus continuous. By the Extreme Value Theorem, D attains its minimum value. ∎

4.2.3. EXAMPLES.

(1) Let $\mathcal{S} = \{f \in C[0,1] : f(0) = 0\}$. This is a closed infinite dimensional subspace of $C[0,1]$. Let $\mathbf{1}$ denote the constant function. What is $\operatorname{dist}(\mathbf{1}, \mathcal{S})$? If $f \in \mathcal{S}$, then

$$\|\mathbf{1} - f\|_\infty \geq |(\mathbf{1} - f)(0)| = 1.$$

On the other hand, if $0 \leq f \leq 2$ and $f(0) = 0$, then $f \in \mathcal{S}$ and $\|\mathbf{1} - f\|_\infty = 1$. So there are infinitely many closest points, such as x, $3x^2 - x^4$, $2\sin^2 6\pi x$, $xe^x/2$, etc.

(2) Let $\mathcal{T} = \{f \in C[0,1] : f(0) = 0 \text{ and } \int_0^1 f(x)\,dx = 0\}$. Again this is a closed subspace. Let $g(x) = x$. What is $\operatorname{dist}(g, \mathcal{T})$? If $f \in \mathcal{T}$, then

$$\|g - f\|_\infty = \int_0^1 \|g - f\|_\infty \,dx \geq \Big|\int_0^1 (g - f)(x)\,dx\Big| = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}.$$

If this were an equality for some f, it is necessary that $g(x) - f(x) = \frac{1}{2}$ for all x. But then $f(x) = x - \frac{1}{2}$ does not vanish at 0, and thus $f \notin \mathcal{T}$. So the distance $\frac{1}{2}$ is not attained. Nevertheless, this is the distance to $\mathcal{T}$. Let

$$h_n(x) = \begin{cases} -\frac{(n-2)^2}{8n}x & \text{for} \quad 0 \leq x \leq \frac{4}{n+2} \\[2ex] x - \frac{1}{2} - \frac{1}{n} & \text{for} \quad \frac{4}{n+2} \leq x \leq 1. \end{cases}$$

Then you can check that $h_n \in \mathcal{T}$ and $\|g - h_n\|_\infty = \frac{1}{2} + \frac{1}{n}$.

4.2.4. CHEBYCHEV APPROXIMATION THEOREM. *If $f \in C_{\mathbb{R}}[a,b]$, then there is a unique closed polynomial of degree at most n.*

The key notion recognizes a geometric property of the smallest difference.

4.2.5. DEFINITION. A function $g \in C_{\mathbb{R}}[a,b]$ satisfies *equioscillation of degree n* if there are $n+2$ points $a \leq x_1 < x_2 < \cdots < x_{n+2} \leq b$ so that

$$g(x_i) = (-1)^i \|g\|_\infty \text{ or } g(x_i) = (-1)^{i+1} \|g\|_\infty \text{ for } 1 \leq i \leq n+2.$$

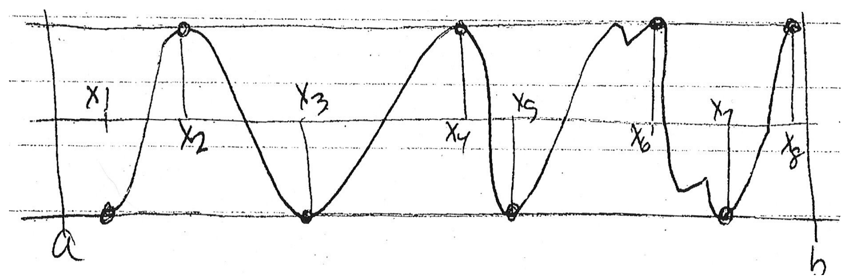

FIGURE 4.2. Equioscillation for $n = 6$.

4.2.6. LEMMA. *Suppose that $f \in C_{\mathbb{R}}[a,b]$ and $p \in \mathbb{P}_n[a,b]$ such that $r = f - p$ satisfies equioscillation of degree n. Then $\|f - p\|_\infty = E_n(f)$.*

PROOF. Let $a \leq x_1 < x_2 < \cdots < x_{n+2} \leq b$ so that

$$r(x_i) = (-1)^i \|r\|_\infty \text{ or } r(x_i) = (-1)^{i+1} \|g\|_\infty \text{ for } 1 \leq i \leq n+2.$$

Suppose that $q \in \mathbb{P}_n[a,b]$ so that $f - p - q$ has smaller norm, i.e.

$$\|f - p - q\|_\infty = \|r - q\|_\infty < \|r\|_\infty.$$

Then $|r(x_i) - q(x_i)| = \big| \pm (-1)^i \|r\|_\infty - q(x_i) \big| < \|r\|_\infty$. Therefore $\operatorname{sign}(q(x_i)) = \operatorname{sign}(r(x_i))$ for $1 \leq i \leq n+2$. Since r changes sign between x_i and x_{i+1}, so does q. By the Intermediate Value Theorem, there are points $y_i \in (x_i, x_{i+1})$ for $1 \leq i \leq n+1$ so that $q(y_i) = 0$. Since q has degree at most n and $n+1$ roots, $q = 0$. This is a contradiction. So p is a closest point. ∎

The converse is trickier.

4.2.7. LEMMA. *Suppose that $f \in C_{\mathbb{R}}[a,b]$ and that $p \in \mathbb{P}_n[a,b]$ satisfies $\|f - p\|_\infty = E_n(f)$. Then $r = f - p$ satisfies equioscillation of degree n.*

PROOF. If f is a polynomial of degree at most n, then $p = f$ is clearly the unique closest polynomial. So we may suppose that $r \neq 0$.

Since r is uniformly continuous, there is a $\delta > 0$ so that $|x - y| < \delta$ implies that $|r(x) - r(y)| < \frac{1}{2}\|r\|_\infty$. Partition $[a, b]$ into intervals of length less than δ. Let $I_1,\dots,I_s$ be those intervals in this partition (in order) on which r attains one of the values $\pm\|r\|_\infty$ on $\overline{I_j}$ (possibly an endpoint). Pick a point $x_j \in \overline{I_j}$ so that $r(x_j) = \pm\|r\|_\infty$. Set $\varepsilon_j = \text{sign}(r(x_j)) \in \{\pm 1\}$. Then $|f(y)| \geq \frac{1}{2}\|r\|_\infty$ on $\overline{I_j}$, and in particular does not change sign. We need to show that $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$ changes sign at least $n + 1$ times, for then we have equioscillation of degree n by choosing $n + 2$ of the x_j's in order with alternating signs.

If there are at most n sign changes, we will construct a closer element of $\mathbb{P}_n$. Group together all adjacent intervals I_j of the same sign into groupings J_1, J_2, $\dots$, J_t (still in order) where $t \leq n + 1$. Pick a point c_k between J_k and J_{k+1} for $1 \leq k \leq n$. Then define $q(x) = \prod_{k=1}^{t-1}(x - c_k) \in \mathbb{P}_n$. Then q changes its sign at each c_k, and in particular has constant sign on each J_k and alternates sign. Multiply q by -1 if neccesary so that $\text{sign}(q)$ agrees with $\text{sign}(r)$ on each J_k. We will show that subtracting a small multiple of q from r will reduce the norm.

Let $L = \bigcup_{j=1}^{s} \overline{I_j}$ and $M = \overline{[a, b] \setminus L}$. Let

$$m = \min\{|q(x)| : x \in L\} \quad \text{and} \quad \sup\{|r(x)| : x \in M\} = \|r\|_\infty - d.$$

Since q only vanishes on the points c_k, which are not in the compact set L, we have $m > 0$. Also $d > 0$ because M is the union of those closed intervals in our partition on which r does not attain its norm. Let $\varepsilon = \frac{d}{2\|q\|_\infty}$ and consider $p_1 = p + \varepsilon q$. Then

$$\begin{aligned}\|f - p_1\|_\infty = \|r - \varepsilon q\|_\infty &= \max\big\{\sup_{x\in L}|r(x) - \varepsilon q(x)|,\ \sup_{x\in M}|r(x) - \varepsilon q(x)|\big\}\\ &\leq \max\big\{\|r\|_\infty - \varepsilon m,\ \|r\|_\infty - d + \varepsilon\|q\|_\infty\big\}\\ &= \max\big\{\|r\|_\infty - \varepsilon m,\ \|r\|_\infty - \tfrac{d}{2}\big\} < \|r\|_\infty.\end{aligned}$$

This contradicts p being a closest polynomial, and thus there must have been at least $n + 1$ sign changes, and so r satisfies equioscillation of degree n. ■

PROOF OF CHEBYCHEV'S APPROXIMATION THEOREM. By Proposition 4.2.2, there is a polynomial $p \in \mathbb{P}[a, b]$ so that $\|f - p\|_\infty = E_n(f) =: d$. Suppose that q also satisfies $\|f - q\|_\infty = d$. Then

$$\big\|f - \tfrac{p+q}{2}\|_\infty = \big\|\tfrac{f-p}{2} + \tfrac{f-q}{2}\|_\infty \leq \tfrac{1}{2}\|f - p\|_\infty + \tfrac{1}{2}\|f - q\|_\infty = d.$$

So $\frac{p+q}{2}$ is also a closest polynomial. By Lemma 4.2.7, $r := f - \frac{p+q}{2}$ satisfies equioscillation of degree n. Let $a \leq x_1 < x_2 < \cdots < x_{n+2} \leq b$ so that

$$r(x_i) = (-1)^i d \quad \text{or} \quad r(x_i) = (-1)^{i+1} d \quad \text{for } 1 \leq i \leq n + 2.$$

Therefore

$$d = \Big|f(x_i) - \tfrac{p(x_i)+q(x_i)}{2}\Big|$$

$$\le \tfrac{1}{2}\big|f(x_i) - p(x_i)\big| + \tfrac{1}{2}\big|f(x_i) - q(x_i)\big|$$
$$\le \tfrac{d}{2} + \tfrac{d}{2} = d.$$

This is an equality, and therefore

$$f(x_i) - p(x_i) = f(x_i) - q(x_i) = \pm d \quad \text{for} \quad 1 \le i \le n+2.$$

Hence $(p-q)(x_i) = 0$ for $1 \le i \le n+2$. Since $p-q$ has degree at most n and has $n+2$ roots, this means that $p = q$. Hence p is the unique polynomial in $\mathbb{P}_n[a,b]$ which is closest to f. ■

4.3. The Stone-Weierstrass Theorems

In this section, we establish a very general approximation result which explains when an algebra of continuous functions is dense in $C_{\mathbb{R}}(X)$ or $C(X)$.

4.3.1. DEFINITION. Let (X,d) be a compact metric space. A subset $\mathcal{A}$ of $C(X)$ or $C_{\mathbb{R}}(X)$ is an *algebra* if it is a subspace such that if $f,g \in \mathcal{A}$, then the product fg is in $\mathcal{A}$.

A subset $\mathcal{A}$ of $C_{\mathbb{R}}(X)$ is a *vector lattice* if it is a subspace such that if $f,g \in \mathcal{A}$, then $f \vee g := \max\{f,g\}$ and $f \wedge g := \min\{f,g\}$ belong to $\mathcal{A}$.

A subset $\mathcal{A}$ of $C(X)$ or $C_{\mathbb{R}}(X)$ *separates points* if for all $x,y \in X$, $x \ne y$, there is an $f \in \mathcal{A}$ such that $f(x) \ne f(y)$.

A subset $\mathcal{A}$ of $C(X)$ or $C_{\mathbb{R}}(X)$ *vanishes at* x_0 if $f(x_0) = 0$ for all $f \in \mathcal{A}$.

4.3.2. STONE-WEIERSTRASS THEOREM. *Let (X,d) be a compact metric space. Suppose that $\mathcal{A} \subset C_{\mathbb{R}}(X)$ is an algebra which separates points and does not vanish at any point of X. Then $\mathcal{A}$ is dense in $C_{\mathbb{R}}(X)$.*

A key innovation by Stone was the recognition that the vector lattice property was very useful.

4.3.3. LEMMA. *If $\mathcal{A}$ is a subalgebra of $C_{\mathbb{R}}(X)$, then $\overline{A}$ is a closed subalgebra and a vector lattice.*

PROOF. If $f_n, g_n \in \mathcal{A}$ and $\lim f_n = f$ and $\lim g_n = g$, then for $r,s \in \mathbb{R}$, $rf_n + sg_n$ and $f_n g_n$ belong to $\mathcal{A}$ and converge to $rf + sg$ and fg, respectively. Therefore $\overline{A}$ is a subspace and an algebra.

To show that it is a lattice, it is enough to show that if $f \in \overline{A}$, then $|f|$ is also in $\overline{A}$. This is because

$$f \vee g = \frac{f+g}{2} + \frac{|f-g|}{2} \quad \text{and} \quad f \wedge g = \frac{f+g}{2} - \frac{|f-g|}{2}.$$

Let $f \in \overline{A}$. By Weierstrass's Theorem, there are polynomials $p_n(t)$ which converge uniformly to $|t|$ on $[-\|f\|_\infty, \|f\|_\infty]$. Say $p_n(t) = \sum_{i=0}^{k_n} a_{ni} t^i$. In particular, $p_n(0) = a_{n0} \to 0$. Hence $q_n(t) = p_n(t) - a_{n0} = \sum_{i=1}^{k_n} a_{ni} t^i$ also converges uniformly to $|t|$ on $[-\|f\|_\infty, \|f\|_\infty]$. Note that $q_n(f) = \sum_{i=1}^{k_n} a_{ni} f^i$ belongs to $\overline{A}$. We use q_n rather than p_n because we do not know that the constants belong to $\mathcal{A}$. Moreover

$$\begin{aligned}\big\|q_n(f) - |f|\big\|_\infty &= \sup_{x \in X} \big|q_n(f(x)) - |f(x)|\big| \\ &\le \sup_{|t| \le \|f\|_\infty} \big|q_n(t) - |t|\big| = \big\|q_n - |t|\big\|_\infty \to 0.\end{aligned}$$

Therefore $|f|$ belongs to $\overline{\mathcal{A}}$. So $\overline{\mathcal{A}}$ is a vector lattice. ■

4.3.4. LEMMA. *Suppose that $\mathcal{A}$ is a subalgebra of $\mathcal{C}_{\mathbb{R}}(X)$ which separates points and doesn't vanish at any $x \in X$. If x, y are distinct points in X and $r, s \in \mathbb{R}$, then there is an function $h \in \mathcal{A}$ so that $h(x) = r$ and $h(y) = s$.*

PROOF. As $\mathcal{A}$ separates points, there is an $f \in \mathcal{A}$ so that $a = f(x) \neq f(y) = b$. At least one of a, b is non-zero, so we may suppose that $b \neq 0$.

Case 1. $a \neq 0$. We look for h in span$\{f, f^2\}$. Note that

$$\left| \begin{bmatrix} a & a^2 \\ b & b^2 \end{bmatrix} \right| = ab^2 - ba^2 = ab(b-a) =: \Delta \neq 0.$$

Hence the matrix $T = \begin{bmatrix} a & a^2 \\ b & b^2 \end{bmatrix}$ is invertible. Hence we can solve the linear system of equations

$$\begin{bmatrix} a & a^2 \\ b & b^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} r \\ s \end{bmatrix}.$$

Indeed, the solution is

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a & a^2 \\ b & b^2 \end{bmatrix}^{-1} \begin{bmatrix} r \\ s \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} b^2 & -a^2 \\ -b & a \end{bmatrix} \begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} \frac{b^2 r - a^2 s}{ab(b-a)} \\ \frac{-br + as}{ab(b-a)} \end{bmatrix}.$$

Therefore if we set $h = uf + vf^2$,

$$\begin{aligned} h(x) &= (uf + vf^2)(x) = ua + va^2 = r \\ h(y) &= (uf + vf^2)(y) = ub + vb^2 = s. \end{aligned}$$

Case 2. $a = 0$. Then since $\mathcal{A}$ does not vanish at x, there is some $g \in \mathcal{A}$ such that $g(x) = c \neq 0$. Let $g(y) = d$; and set

$$h(z) = \frac{r}{c} g(z) + \frac{cs - rd}{bc} f(z) \quad \text{for} \quad z \in X.$$

Then $h \in \mathcal{A}$ and

$$h(x) = \frac{r}{c}c + 0 = r \quad \text{and} \quad h(y) = \frac{r}{c}d + \frac{cs - rd}{bc}b = s. \qquad \blacksquare$$

Proof of the Stone-Weierstrass Theorem. By Lemma 4.3.3, $\overline{\mathcal{A}}$ is a vector lattice. Fix $f \in C_{\mathbb{R}}(X)$ and $\varepsilon > 0$. Let $a \in X$. For each $x \in X \setminus \{a\}$, use Lemma 4.3.4 to find functions $h_x \in \mathcal{A}$ so that $h_x(a) = f(a)$ and $h_x(x) = f(x)$. Let

$$U_x = \{y \in X : h_x(y) > f(y) - \varepsilon\} = (h_x - f)^{-1}(-\varepsilon, \infty).$$

Then U_x is open and contains both a and x. So $\{U_x : x \neq a\}$ is an open cover of X. Thus there is a finite subcover $U_{x_1}, \dots, U_{x_n}$. Define $g_a = h_{x_1} \vee \dots \vee h_{x_n}$. This belongs to $\overline{\mathcal{A}}$, $g_a(a) = f(a)$ and

$$g_a(x) \geq h_{x_i}(x) > f(x) - \varepsilon \quad \text{for} \quad x \in U_{x_i},\ 1 \leq i \leq n.$$

Therefore $g_a > f - \varepsilon$.

Let $V_a = \{y \in X : g_a(y) < f(y) + \varepsilon\} = (g_a - f)^{-1}(-\infty, \varepsilon)$. This is an open set containing a. Hence $\{V_a : a \in X\}$ is an open cover of X. By compactness, there is a finite subcover $V_{a_1}, \dots, V_{a_m}$. Let $g = g_{a_1} \wedge \dots \wedge g_{a_m}$. This belongs to $\overline{\mathcal{A}}$. Then $g(x) > f(x) - \varepsilon$ since this is true for every g_a. Also

$$g(x) \leq g_{a_j}(x) < f(x) + \varepsilon \quad \text{for all} \quad x \in V_{a_j},\ 1 \leq j \leq m.$$

Hence $g < f + \varepsilon$,

Consequently, $|g(x) - f(x)| < \varepsilon$ for all $x \in X$, and therefore $\|g - f\| < \varepsilon$. Since $\overline{\mathcal{A}}$ is closed, it must equal $C_{\mathbb{R}}(X)$. $\blacksquare$

We mention a few of the many applications of this result.

4.3.5. Corollary. *Let X be a compact subset of $\mathbb{R}^n$. Then the algebra of polynomials in the coordinates $x_1, \dots, x_n$ is dense in $C(X)$.*

Proof. First consider the algebra $\mathcal{A}$ of polynomials with real coefficients $\mathcal{A} = \mathbb{R}[x_1, \dots, x_n]$. Then $\mathcal{A}$ is an algebra, it contains the constant function 1, so it does not vanish at any point. Also $x_1, \dots, x_n$ separate points in X. Hence by the Stone-Weierstrass Theorem, $\overline{\mathcal{A}} = C_{\mathbb{R}}(X)$.

In the complex case, we can write $f = g + ih$ where $g, h \in C_{\mathbb{R}}(X)$. Since both g, h are uniform limits of polynomials, f is also a uniform limit of polynomials with complex coefficients. $\blacksquare$

4.3.6. Corollary. *Let X, Y be two compact metric spaces. Then*

$$\mathcal{A} = \Big\{h(x,y) = \sum_{i=1}^{n} f_i(x) g_i(y) : f_i \in C(X),\ g_i \in C(Y)\Big\}$$

is dense in $C(X \times Y)$.

PROOF. First consider the real version $\mathcal{A}_{\mathbb{R}}$ which consists of finite sums of products of functions in $C_{\mathbb{R}}(X)$ and $C_{\mathbb{R}}(Y)$. This is a real algebra. It contains 1, so does not vanish anywhere. It separates points, because $C_{\mathbb{R}}(X)$ separates the X-coordinate and $C_{\mathbb{R}}(Y)$ separates the Y-coordinate. Thus the Stone-Weierstrass Theorem shows that $\overline{\mathcal{A}_{\mathbb{R}}} = C_{\mathbb{R}}(X \times Y)$. By taking real and imaginary parts, we obtain the complex version. ■

4.3.1. The complex case. Both corollaries were applicable to the complex case by taking real and complex parts. However this is not possible to do within all subalgebras.

4.3.7. EXAMPLE. Let $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ be the open unit disc in the complex plane. Let $A(\mathbb{D}) = \overline{\{p(z) = \sum_{k=0}^{n} a_k z^k\}} = \mathbb{C}[z]$ considered as a subalgebra of $C(\overline{\mathbb{D}})$. $1 \in A(\mathbb{D})$ so $A(\mathbb{D})$ does not vanish at any point. The function z separates points. However $A(\mathbb{D}) \neq C(\overline{\mathbb{D}})$.

To see this, observe that

$$\tfrac{1}{2\pi}\int_0^{2\pi} p(e^{it})\,dt = \sum_{k=0}^{n} a_k \,\tfrac{1}{2\pi}\int_0^{2\pi} e^{int}\,dt = a_0 = p(0).$$

By taking uniform limits, this identity extends to all $f \in A(\mathbb{D})$. But in $C(\overline{\mathbb{D}})$, there is no relationship between $f(0)$ and the restriction of f to the unit circle $\mathbb{T}$. For example, $f(z) = |z|$ satisfies $f(0) = 0$ and $f(e^{it}) = 1$, so that

$$\tfrac{1}{2\pi}\int_0^{2\pi} f(e^{it})\,dt = 1 \neq f(0).$$

We can also consider $A(\mathbb{D})$ as a subalgebra of $C(\mathbb{T})$. This follows from the maximum modulus principle which shows that each polynomial attains its supremum on the boundary. This property extends to limits. So the supremum norm over $\mathbb{T}$ is the same norm as the supremum norm over $\overline{\mathbb{D}}$. Again $1 \in A(\mathbb{D})$ and z separates points, but still $A(\mathbb{D}) \neq C(\mathbb{T})$. Using the integral above, we see that

$$\tfrac{1}{2\pi}\int_0^{2\pi} f(e^{it})e^{it}\,dt = 0 \quad \text{for all} \quad f \in A(\mathbb{D}).$$

However the function $\bar{z}(e^{it}) = e^{-it}$ does not belong to $A(\mathbb{D})$ because

$$\tfrac{1}{2\pi}\int_0^{2\pi} \bar{z}(e^{it})e^{it}\,dt = \tfrac{1}{2\pi}\int_0^{2\pi} dt = 1.$$

4.3.8. DEFINITION. A subalgebra $\mathcal{A} \subset C(X)$ is *self-adjoint* if $f \in \mathcal{A}$ implies that $\bar{f} \in \mathcal{A}$.

Since $\operatorname{Re} f = \frac{1}{2}(f + \bar{f})$ and $\operatorname{Im} f = \frac{1}{2i}(f - \bar{f})$, a subalgebra is self-adjoint if and only if it contains the real and imaginary parts of it elements. This was key in our two corollaries. This generalizes.

4.3.9. THEOREM. *Let X be a compact metric space. Let $\mathcal{A}$ be a subalgebra of $C(X)$ which is self-adjoint, separates points, and does not vanish at any point. Then $\mathcal{A}$ is dense in $C(X)$.*

PROOF. Let $\mathcal{A}_{\mathbb{R}} = \{\operatorname{Re} f : f \in \mathcal{A}\}$. Since $\mathcal{A}$ is self-adjoint, $\mathcal{A}_{\mathbb{R}}$ is a real algebra contained in $\mathcal{A}$. It separates points, and does not vanish at any point. By the real Stone-Weierstrass Theorem, $\overline{\mathcal{A}_{\mathbb{R}}} = C_{\mathbb{R}}(X)$. Every function $f = g + ih$ where $g = \operatorname{Re} f$ and $h = \operatorname{Im} f$. Since $g, h \in \overline{\mathcal{A}_{\mathbb{R}}}$, we have $f \in \overline{\mathcal{A}}$. ■

Exercises

1. Let $C^{(n)}[a,b]$ denote the space of functions on $[a,b]$ with n continuous derivatives.

(a) Define $\|f\|_{C^{(n)}} = \sum_{k=0}^{n} \frac{1}{k!}\|f^{(k)}\|_\infty$. Show that this norm satisfies
$$\|fg\|_{C^{(n)}} \le \|f\|_{C^{(n)}}\|g\|_{C^{(n)}}.$$
HINT: Leibnitz rule for $(fg)^{(k)}$.

(b) Show that the polynomials are dense in $C^{(n)}[a,b]$. HINT: approximate $f^{(n)}$ first.

2. Let p be the best polynomial approximation of degree n to $\sqrt{x}$ on $[0,1]$. Show that $q(x) = p(x^2)$ is the best polynomial approximation of degree $2n+1$ to $|x|$ on $[-1,1]$.

3. For $\mathcal{S} \subset C[a,b]$, let $E_n(\mathcal{S}) = \sup_{f\in\mathcal{S}} E_n(f)$. Let $\mathcal{S}$ be the functions in $C[0,1]$ with Lipschitz constant 1. Show that $E_n(\mathcal{S}) \ge \frac{1}{2n+2}$. HINT: consider a piecewise linear function such that $f\big(\frac{k}{n+1}\big) = (-1)^k$ for $0 \le k \le n+1$.

4. Show that the real linear span A of $\{1, \sin nx, \cos nx : n \ge 1\}$ is a dense *subalgebra* of $B = \{f \in C_{\mathbb{R}}[-\pi,\pi] : f(-\pi) = f(\pi)\}$.

5. Let X be a compact subset of $\mathbb{R}^n$, and let $F = \{\mathbf{y}_1, \dots, \mathbf{y}_p\}$ be a finite subset of X. Show that the set of polynomials in the coordinates $x_1, \dots, x_n$ which vanish on F is dense on the ideal $I(F) = \{f \in C_{\mathbb{R}}(X) : f|_F = 0\}$. HINT: build polynomials q_i such that $q_i(\mathbf{y}_j) = \delta_{ij}$.

6. Let (X,d) be a compact metric space. Let J be an *ideal* of $C(X)$. Define
$$Z = Z(J) = \textstyle\bigcap_{f\in J} f^{-1}(0).$$

(a) Let $\varepsilon > 0$ and let $Y_\varepsilon = \{x : d(x,Z) \ge \varepsilon\}$. Show that J contains a function f such that $f|_{Y_\varepsilon} = 1$. HINT: Find finitely many $f_i \in J$ so that $\sum_{i=1}^n |f_i(x)|^2 \ge 1$ for all $x \in Y_\varepsilon$.

(b) Hence show that J contains the ideal $I_0(Z) = \{f \in C(X) : Z \subset \operatorname{int}(f^{-1}(0))\}$ of all functions that vanish on a neighbourhood of Z.

(c) Show that $\overline{J} = I(Z) = \{f \in C(X) : f|_Z = 0\}$.

CHAPTER 5

Differential Equations

Ordinary differential equations or ODEs are equations that relate a function of one variable to one or more of its derivatives. *Partial differential equations* or PDEs relate functions of several variables to their various partial derivatives. They arise in many contexts: in physics, chemistry and engineering in modelling various physical phenomena. In the life sciences modelling various global properties among populations, such as predator-prey cycles, are governed by differential equations. In economics and mathematical finance, many processes are governed by differential equations. In differential geometry and mathematical physics, differential equations underlie most phenomena that are studied.

In this chapter, we study ODEs not with the idea of learning solution techniques, but rather to understand when we can solve them and what we can say in general about the behaviour of solutions.

Here is the basic idea:

$$y'(x) = f(x, y) \quad \text{and} \quad y(0) = y_0 \quad \text{for} \quad x, y \in \mathbb{R}.$$

In geometry and physics, one can consider a *vector field* which puts an arrow at every point (x, y) with slope $f(x, y)$. A common demonstration of this is to use iron filings scattered around a magnet. The filings align with the direction of the magnetic flow. Conversely, looking at the filings (which are the arrows), one can "see" the flow lines. These flow lines are solutions to the ODE. There is an intuitive sense that generally there is exactly one solution through each point, except for a few exceptional points at the boundary of the magnet. This intuition is excellent as long as the ODE is nice enough.

5.0.1. EXAMPLES.

(1) Consider $y' = f(x)$ and $y(a) = y_0$. This is easily solved by integration:

$$y(x) = y_0 + \int_a^x f(t)\, dt.$$

If f is continuous on $\mathbb{R}$, then this has a solution on the whole line. But if f is defined on a smaller set, say $(-1, 1)$ by $f(x) = \frac{1}{1-x^2}$ and $a = 0$, then the solution will also 'blow up' as x tends to ± 1,

(2) Consider the ODE $y' = xy$. This can be solved by the technique of *separation of variables* which involes rearranging things to get all the y's on one side and all the x's on the other:

$$\frac{y'}{y} = x.$$

Integrate to get $\log y = \frac{1}{2}x^2 + c$. Thus $y = Ce^{x^2/2}$. The constant C is a parameter determining a whole family of solutions. If we know some point on the curve, say $y(1) = e$, then we can solve for $C = \sqrt{e}$. If we change the initial value a little, this changes C a little, and the new solution is close to the original near the starting point. However it moves away dramatically as we get further from the initial position.

(3) Level lines. The equation $xy = \log y + c$, where c is a constant determines a family of curves, though they are difficult to express in closed form. These curves are the solutions to a DE obtained by differentiation:

$$y + xy' = \frac{y'}{y} \quad \text{or} \quad y' = \frac{y^2}{1 - xy}.$$

If c is known, we can see that $(0, e^{-c})$ lies on the curve. So we could add the *initial value condition*: $y(0) = e^{-c}$.

(4) Some ODEs have no solutions. For example, $(y')^2 + 1 = 0$.

The question addressed in this chapter is: under what conditions does an ODE have a solution? When is it unique? How does it change if we vary a parameter?

The following is a key example which illustrates some of the important ideas.

5.0.2. EXAMPLE. We will solve the initial value problem

$$\begin{aligned} y' &= 1 + x - y \quad \text{for} \quad -\tfrac{1}{2} \le x \le \tfrac{1}{2} \\ y(0) &= 1. \end{aligned}$$

Integrate to get

$$\begin{aligned} y(x) &= y(0) + \int_0^x y'(t)\,dt \\ &= 1 + \int_0^x 1 + t - y(t)\,dt \\ &= 1 + x + \tfrac{1}{2}x^2 - \int_0^x y(t)\,dt. \end{aligned}$$

This changes the differential equation into an *integral equation*.

Define a map T on $C[-\frac{1}{2}, \frac{1}{2}]$ by sending f to the function Tf given by

$$Tf(x) = 1 + x + \tfrac{1}{2}x^2 - \int_0^x f(t)\,dt.$$

The solution of the integral equation is a fixed point of T. Conversely, if

$$f(x) = Tf(x) = 1 + x + \tfrac{1}{2}x^2 - \int_0^x f(t)\,dt,$$

then $f(0) = 1$ and $f'(x) = 1 + x - f(x)$ by the Fundamental Theorem of Calculus.

We compute for $x \in [-\frac{1}{2}, \frac{1}{2}]$,

$$\begin{aligned} |Tf(x) - Tg(x)| &= \Big| \int_0^x f(t) - g(t)\,dt \Big| \le \Big| \int_0^x |f(t) - g(t)|\,dt \Big| \\ &\le |x|\,\|f - g\|_\infty \le \frac{1}{2}\|f - g\|_\infty. \end{aligned}$$

Therefore

$$\|Tf - Tg\|_\infty \le \frac{1}{2}\|f - g\|_\infty.$$

Hence T is a contraction mapping!

By the Contraction Mapping Principle, T has a unique fixed point f_∞ that will solve our DE. Moreover, we can compute the solution be setting $f_0(x) = 1$, defining $f_{n+1}(x) = Tf_n(x)$ for $n \ge 0$, and taking a limit. Then

$$f_1(x) = Tf_0(x) = 1 + x + \tfrac{1}{2}x^2 - \int_0^x 1\,dt = 1 + \tfrac{1}{2}x^2.$$

Similarly,

$$\begin{aligned} f_2(x) &= Tf_1(x) = 1 + x + \tfrac{1}{2}x^2 - \int_0^x 1 + \tfrac{1}{2}t^2\,dt \\ &= 1 + \tfrac{1}{2}x^2 - \tfrac{1}{6}x^3. \end{aligned}$$

We can show by induction that

$$f_n(x) = 1 + \tfrac{1}{2}x^2 - \tfrac{1}{3!}x^3 + \tfrac{1}{4!}x^4 - \tfrac{1}{5!}x^5 + \cdots + \tfrac{1}{(n+1)!}(-x)^{n+1}.$$

We obtain the solution

$$f_\infty(x) = \lim_{n\to\infty} f_n(x) = x + \sum_{k=0}^{\infty} \frac{1}{k!}(-x)^k = x + e^{-x}.$$

Note that this solution actually makes sense on the whole real line. Why that happens is part of the story.

5.1. Reduction to first order

Our plan is to start with a DE, and convert it to the problem of finding a fixed point of an associated integral operator. Generally this will need to be vector valued.

5.1.1. DEFINITION. The *order of a DE* is the highest derivative occuring in the equation. It is said to be in *standard form* if the highest derivative can be expressed as a function of x and the derivatives of lower order. An *initial value problem* contains the DE and specifies values at some point for all derivatives of lower order.

Initial value problem of order n. Let $c \in [a, b]$. Look for a functions $f \in C^n[a, b]$ satisfying:

$$f^{(n)}(x) = \varphi(x, f(x), f'(x), \dots, f^{(n-1)}(x)) \qquad n\text{th order ODE} \tag{5.1.2}$$

$$\left.\begin{aligned} f(c) &= \gamma_0 \\ f'(c) &= \gamma_1 \\ &\vdots \\ f^{(n-1)}(c) &= \gamma_{n-1}, \end{aligned}\right\} \qquad \text{initial data.}$$

where φ is a real-valued continuous function on $[a, b] \times \mathbb{R}^n$.

Reduce this to a first order ODE with values in $\mathbb{R}^n$. In order to proceed as in our example, we require a first order DE. This can be arranged if we allow functions with values in $\mathbb{R}^n$. Let $F : [a, b] \to \mathbb{R}^n$ by

$$F(x) = (f(x), f'(x), \dots, f^{(n-1)}(x)).$$

This belongs to $C([a, b], \mathbb{R}^n)$ and each coordinate is differentiable. Hence

$$F'(x) = \Big(f'(x), \dots, f^{(n-1)}(x), \varphi\big(x, f(x), \dots, f^{(n-1)}(x)\big)\Big),$$

and the initial data become

$$F(c) = (\gamma_0, \gamma_1, \dots, \gamma_{n-1}) =: \Gamma.$$

Define a continuous function Φ from $[a, b] \times \mathbb{R}^n$ to $\mathbb{R}^n$ by

$$\Phi(x, y_0, \dots, y_{n-1}) = \big(y_1, y_2, \dots, y_{n-1}, \varphi(x, y_0, \dots, y_{n-1})\big).$$

Then (5.1.2) becomes the first-order vector-valued initial value problem :

$$\begin{aligned} F'(x) &= \Phi(x, F(x)) \\ F(c) &= \Gamma. \end{aligned} \tag{5.1.3}$$

It is easy to see that a solution of (5.1.2) yields a solution of (5.1.3). Conversely, suppose (5.1.3) has a solution

$$F(x) = (f_0(x), f_1(x), \dots, f_{n-1}(x)).$$

Then

$$\begin{aligned} F'(x) &= (f_0'(x), f_1'(x), \dots, f_{n-2}'(x), f_{n-1}'(x)) \\ &= \Phi(x, f_0(x), f_1(x), \dots, f_{n-1}(x)) \\ &= \Big(f_1(x), \dots, f_{n-1}(x), \varphi\big(x, f(x), \dots, f^{(n-1)}(x)\big)\Big). \end{aligned}$$

By identifying each coordinate, we get $f_i'(x) = f_{i+1}(x)$ for $0 \le i \le n-2$ and

$$f_{n-1}'(x) = \varphi(x, f_0(x), f_1(x), \dots, f_{n-1}(x)).$$

Thus $f_1 = f_0'$, $f_2 = f_1' = f_0''$, $f_{i+1} = f_i' = f_0^{(i+1)}$ for $0 \le i \le n-2$ and

$$f_0^{(n)} = f_{n-1}' = \varphi(x, f_0(x), f_0'(x), \dots, f_0^{(n-1)}(x)).$$

So f_0 is a solution to the ODE. Finally from $F(c) = \Gamma$, we get that $f_0^{(i)}(c) = \gamma_i$ for $0 \le i \le n-1$. So the initial value data is satisfied. Therefore f_0 is a solution to (5.1.2).

Convert to an integral equation. We now integrate (5.1.3) to get

$$F(x) = F(c) + \int_c^x F'(t)\,dt = \Gamma + \int_c^x \Phi(t, F(t))\,dt.$$

Define a map on $C([a,b], \mathbb{R}^n)$ by

$$TF(x) = \Gamma + \int_c^x \Phi(t, F(t))\,dt.$$

A solution of (5.1.3) is clearly a fixed point of T. Conversely, by the Fundamental Theorem of Calculus, a fixed point of T is a solution of (5.1.3). So the problem (5.1.2) is equivalent to finding the fixed point(s) of the mapping T.

5.1.4. Example. We will express the unknown function as y, instead of $f(x)$. Consider the differential equation

$$\begin{gathered}\big(1 + (y')^2\big)y^{(3)} = y'' - xy'y + \sin x \quad \text{for} \quad -1 \le x \le 1 \\ y(0) = 1, \quad y'(0) = 0, \quad y''(0) = 2.\end{gathered}$$

Rewrite it in standard form as

$$y^{(3)} = \frac{y'' - xy'y + \sin x}{1 + (y')^2}.$$

Define $\Gamma = (1, 0, 2)$ and

$$\Phi\big(x, y_0, y_1, y_2\big) = \Big(y_1, y_2, \frac{y_2 - xy_1y_0 + \sin x}{1 + y_1^2}\Big).$$

Then define a mapping T from $C([-1,1], \mathbb{R}^3)$ into itself by sending a function $F(x) = \big(f_0(x), f_1(x), f_2(x)\big)$ to

$$\begin{aligned} TF(x) &= \Gamma + \int_0^x \Phi(t, F(t))\,dt \\ &= \bigg(1 + \int_0^x f_1(t)\,dt\,,\ \int_0^x f_2(t)\,dt\,,\ 2 + \int_0^x \frac{f_2(t) - tf_1(t)f_0(t) + \sin t}{1 + f_1(t)^2}\,dt\bigg). \end{aligned}$$

This converts the differential equation into the integral equation $TF = F$.

5.2. Global Solutions of ODEs

In this section, we use a modification of the Contraction Mapping Principle to establish the existence and uniqueness of solutions to a large class of ODEs. To obtain this, we need a strong condition on the function Φ.

5.2.1. DEFINITION. Let $\Omega \subset \mathbb{R}^n$. A continuous function $\Phi = (\varphi_1, \dots, \varphi_n) : [a, b] \times \Omega \to \mathbb{R}^n$ is *Lipschitz in* $y = (y_1, \dots, y_n)$ if there is a constant L so that

$$\|\Phi(x, y) - \Phi(x, z)\| = \Big(\sum_{i=1}^{n} \big|\varphi_i(x, y) - \varphi_i(x, z)\big|^2 \Big)^{1/2} \le L\|y - z\|$$

for all $x \in [a, b]$ and $y, z \in \Omega$.

5.2.2. EXAMPLES.
(1) Let $\Phi = (\varphi_1, \dots, \varphi_n)$ be defined on a convex set Ω. Suppose that Φ has continuous partial derivatives in the y-variables. Recall the gradient $\nabla_y \varphi_i = \Big(\frac{\partial \varphi_i}{\partial y_1}, \dots, \frac{\partial \varphi_i}{\partial y_n}\Big)$. Suppose that the derivative is bounded in Ω:

$$\max_{1 \le i \le n} \sup_{\substack{x \in [a,b] \\ y \in \Omega}} \|\nabla_y \varphi_i(x, y)\| = M < \infty.$$

Then by the Mean Value Theorem, there is a point ξ_i on $[y, z]$ so that

$$\|\varphi_i(x, y) - \varphi_i(x, z)\| = \|\nabla_y \varphi_i(x, \xi_i) \bullet (y - z)\| \le M\|y - z\|.$$

Therefore $\|\Phi(x, y) - \Phi(x, z)\| \le \sum_{i=1}^{n} \|\varphi_i(x, y) - \Phi(x, z)\| \le Mn\|y - z\|$. If Ω is compact, then $M < \infty$ by the Extreme Value Theorem. However if $\Omega = \mathbb{R}^n$, this is quite a stringent condition, since even very nice functions like y^2 have unbounded derivative on the whole line.

(2) One important example where this condition does hold globally on $\mathbb{R}^n$ is the case of *linear ODEs* in which the function Φ is linear in the y-variables. The function φ has the form

$$y^{(n)}(x) = \varphi(x, y) = \sum_{i=0}^{n-1} a_i(x) y^{(i)}(x) + b(x).$$

Writing $F(x) = (f_0(x), f_1(x), \dots, f_{n-1}(x))$, the DE becomes

$$F'(x) = \begin{bmatrix} f_0'(x) \\ f_1'(x) \\ \vdots \\ f_{n-1}'(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_0(x) & a_1(x) & a_2(x) & \dots & a_{n-1}(x) \end{bmatrix} \begin{bmatrix} f_0(x) \\ f_1(x) \\ \vdots \\ f_{n-2}(x) \\ f_{n-1}(x) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(x) \end{bmatrix}$$

$$= A(x)F(x) + B(x).$$

Here $A(x)$ is an $n \times n$ matrix function of x and $B(x)$ is a vector valued function of x. The dependence on the y variables is linear. That is, $\Phi(x, y) = A(x)y + B(x)$.

Let $M = \sup_{a \le x \le b} \big| \sum_{i=0}^{n-1} a_i(x)^2 \big|^{1/2}$. Since this is linear in y, we get (we write vectors horizontally to save space)

$$\begin{aligned}\|\Phi(x,y) - \Phi(x,z)\| &= \|A(x)(y-z)\| \\ &= \Big\| \big(y_1 - z_1, \dots, y_{n-1} - z_{n-1}, \sum_{i=0}^{n-1} a_i(x)(y_i - z_i) \big) \Big\} \\ &\le \|y - z\| + M\|y - z\| = (M+1)\|y - z\|.\end{aligned}$$

where the Cauchy-Schwarz inequality shows that

$$\Big| \sum_{i=0}^{n-1} a_i(x)(y_i - z_i) \Big| \le \Big| \sum_{i=0}^{n-1} a_i(x)^2 \Big|^{1/2} \Big| \sum_{i=0}^{n-1} (y_i - z_i)^2 \Big|^{1/2} \le M\|y - z\|.$$

Hence Φ is Lipschitz in y.

The following technical estimate is key. The point of the estimates is that they improve dramatically with repeated application.

5.2.3. LEMMA. *Let Φ be a continuous function from $[a, b] \times \mathbb{R}^n$ into $\mathbb{R}^n$ which is Lipschitz in y with constant L. Let $c \in [a, b]$, and*

$$TF(x) = \Gamma + \int_c^x \Phi(t, F(t))\, dt.$$

If $F, G \in C([a, b], \mathbb{R}^n)$ satisfy $\|F(x) - G(x)\| \le \frac{M|x-c|^k}{k!}$ for some $k \ge 0$, then

$$\|TF(x) - TG(x)\| \le \frac{LM|x - c|^{k+1}}{(k+1)!}.$$

Thus, T is Lipschitz; and there is an integer k_0 so that T^{k_0} is a contraction map.

PROOF. Compute

$$\begin{aligned}\|TF(x) - TG(x)\| &= \Big\| \Gamma + \int_c^x \Phi(t, F(t))\, dt - \Gamma - \int_c^x \Phi(t, G(t))\, dt \Big\| \\ &= \Big\| \int_c^x \Phi(t, F(t)) - \Phi(t, G(t))\, dt \Big\| \\ &\le \int_c^x \big\| \Phi(t, F(t)) - \Phi(t, G(t)) \big\|\, dt \le \int_c^x L \big\| F(t) - G(t) \big\|\, dt \\ &\le \frac{LM}{k!} \int_c^x |t - c|^k\, dt = \frac{LM}{(k+1)!}\, |x - c|^{k+1}.\end{aligned}$$

Now $\|F(x)-G(x)\| \leq \|F-G\|_\infty = \|F-G\|_\infty \frac{|x-c|^0}{0!}$. It follows by induction that

$$\|T^k F(x) - T^k G(x)\| \leq \frac{\|F-G\|_\infty L^k |x-c|^k}{k!} \leq \frac{\|F-G\|_\infty L^k (b-a)^k}{k!}.$$

Thus $\|T^k F - T^k G\|_\infty \leq \frac{L^k(b-a)^k}{k!}|F-G\|_\infty$. This shows that T^k is Lipschitz with constant $C_k = \frac{L^k(b-a)^k}{k!}$. Observe that

$$\lim_{k\to\infty} \frac{C_{k+1}}{C_k} = \lim_{k\to\infty} \frac{L(b-a)}{k+1} = 0.$$

By the ratio test, $\lim_{k\to\infty} C_k = 0$. Hence one can choose k_0 so that $C_{k_0} < 1$. Therefore T^{k_0} is a contraction mapping. ■

5.2.4. Global Picard Theorem. *If $\Phi : [a,b] \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous and Lipschitz in y, and $c \in [a,b]$, then the ODE*

$$F'(x) = \Phi(x, F(x)), \quad F(c) = \Gamma$$

has a unique solution on $[a,b]$.

Proof. We first convert the problem to a fixed point problem with the same solutions. Let T map $C([a,b],\mathbb{R}^n)$ into itself by

$$TF(x) = \Gamma + \int_c^x \Phi(t, F(t))\, dt \quad \text{for} \quad F \in C([a,b],\mathbb{R}^n).$$

By Lemma 5.2.3, there is an integer k_0 so that T^{k_0} is a contraction mapping. Therefore by Corollary 3.3.4, T has a unique fixed point F_*. Thus this is the unique solution to the ODE. ■

Starting with any initial function $F_0 \in C([a,b],\mathbb{R}^n)$, the sequence $F_n = T^k F_0$ converges uniformly to the solution. A convenient choice for F_0 is the constant function $F_0(x) = \Gamma$. In the next section, we will need some specific estimates using this starting point.

5.3. Local Solutions

The stipulation that Φ has to be Lipschitz over all of $\mathbb{R}^n$ is quite restrictive. However, many functions satisfy a Lipschitz condition in y on a bounded set of the form $[a,b] \times \overline{b_R(\Gamma)}$ for $R < \infty$. When this is the case, we can run the proof of the Global Picard Theorem for a while, until the solution escapes the set.

5.3.1. DEFINITION. Let $\Omega \subset \mathbb{R}^n$. A continuous function $\Phi : [a,b] \times \Omega \to \mathbb{R}^n$ is *locally Lipschitz in* y if for every point $(x,y) \in [a,b] \times \Omega$, there is a neighbourhood $\big([x-\varepsilon, x+\varepsilon] \cap [a,b]\big) \times b_\varepsilon(y)$ on which Φ is Lipschitz in y.

5.3.2. LEMMA. *Suppose that a continuous function* $\Phi : [a,b] \times \Omega \to \mathbb{R}^n$ *is locally Lipschitz in* y *and that* $K \subset \Omega$ *is compact and convex. Then* Φ *is Lipschitz in* y *on* $[a,b] \times K$.

PROOF. Each point $(x,y) \in [a,b] \times K$ has a convex open neighbourhood $U_{x,y}$ on which Φ is Lipschitz in y with constant $L_{x,y}$. Since $[a,b] \times K$ is compact and $\{U_{x,y} : (x,y) \in [a,b] \times K\}$ is an open cover, there is a finite subcover, say U_{x_i,y_i} for $1 \le i \le N$. Let $L = \max\{L_{x_i,y_i} : 1 \le i \le N\}$. For any $x \in [a,b]$ and $y, z \in K$, the line segment $[y,z] \subset K$ by convexity. Since $\{x\} \times [y,z]$ is covered by our finite subcover, there is a sequence $y = y_0, y_1, \dots, y_p = z$ on $[y,z]$ so that $\{x\} \times [y_i, y_{i+1}]$ is covered by a single set from the subcover. Therefore

$$\begin{aligned} \|\Phi(x,y) - \Phi(x,z)\| &\le \sum_{i=0}^{p-1} \|\Phi(x,y_i) - \Phi(x,y_{i+1})\| \\ &\le \sum_{i=0}^{p-1} L\|y_i - y_{i+1}\| = L\|y-z\|. \end{aligned}$$

Thus Φ is Lipschitz in y on $[a,b] \times K$. ■

The main example is the case in which Φ has continuous partial first derivatives in the y variables. This is Example 5.2.2(1).

In the following local version of the Picard theorem, one requires a local Lipschitz condition. However it is important to note that the interval on which a solution is obtained is not dependent on the Lipschitz constant. For convenience, we start at the left endpoint a, rather than some arbitrary point c. But this makes no real difference. Also the argument works in the negative direction equally well.

5.3.3. LOCAL PICARD THEOREM. *Suppose that* $\Phi : [a,b] \times \overline{b_R(\Gamma)} \to \mathbb{R}^n$ *is continuous and locally Lipschitz in* y*. Then the DE*

$$F'(x) = \Phi(x, F(x)), \quad F(a) = \Gamma$$

has a unique solution on the interval $[a, a+h]$*, where* $h = \min\{b-a, R/\|\Phi\|_\infty\}$.

PROOF. As for the Global Picard Theorem, we define a map T. However it will only be defined on $C([a,a+h], \overline{b_R(\Gamma)})$, those vector valued continuous functions on $[a, a+h]$ with values in $\overline{b_R(\Gamma)}$, because it is only in this range that Φ is defined. This is a closed subset of the complete normed space $C([a,a+h], \mathbb{R}^n)$, and thus it is complete. Set

$$TF(x) = \Gamma + \int_a^x \Phi(t, F(t))\,dt.$$

Observe that

$$\|TF(x) - \Gamma\| \le \int_a^x \|\Phi(t, F(t))\|\,dt \le \|\Phi\|_\infty\,|x - a| \le \|\Phi\|_\infty\,h \le R.$$

Therefore $TF \in C([a, a+h], \overline{b_R(\Gamma)})$.

$C([a, a+h], \overline{b_R(\Gamma)})$ is a complete metric space and so the Contraction Mapping Principle is applicable. Now the proof of Lemma 5.2.3 works as before. It follows that there is some k_0 so that T^{k_0} is a contraction mapping. Hence by Corollary 3.3.4, T has a unique fixed point F_*, and that this is a solution of the ODE.

Uniqueness of a local solution to the ODE requires a bit more care. Suppose that there is a solution G on a smaller interval $[a, a+k]$ for $0 < k \le h$. We can then restrict the mapping T to the smaller domain $C([a, a+k], \overline{b_R(\Gamma)})$. Again it has a unique fixed point, and this must be $F_*|_{[a,a+k]}$. Hence $G = F_*|_{[a,a+k]}$. ■

5.3.4. EXAMPLE. Consider the DE

$$y' = y^2, \qquad y(0) = 1, \qquad 0 \le x \le 2.$$

The function $\Phi(x, y) = y^2$ is not Lipschitz globally because $\varphi'(y) = 2y$ is unbounded. However since it is C^1, the derivative is bounded on $[0, 2] \times [1-R, 1+R]$ with constant $\sup_{|y-1|\le R} |\Phi'(y)| = 2R + 2$. Also $\sup_{|y-1|\le R} |\Phi(y)| = (R+1)^2$. By the Local Picard Theorem, this has a unique solution on $[0, h]$ where

$$h = \min\left\{2, \frac{R}{(R+1)^2}\right\} = \frac{R}{(R+1)^2}.$$

If we pick $R = 1$ (which is optimal), then we get a solution on $[0, \frac{1}{4}]$.

We can solve this DE by separation of variables. Rewrite it as $\dfrac{y'}{y^2} = 1$. Then by integration, we get

$$x = \int_0^x 1\,dt = \int_0^x \frac{y'(t)\,dt}{y(t)^2} = -\frac{1}{y(t)}\bigg|_0^x = 1 - \frac{1}{y(x)}.$$

Therefore, $y(x) = \frac{1}{1-x}$.

This is a solution on the interval $[0, 1)$. However it ‘blows up’ at $x = 1$. So the solution does not extend to all of $[0, 2]$. But this is definitely better than our solution on $[0, \frac{1}{4}]$. This can be improved by starting at $x = \frac{1}{4}$ and applying the Local Picard Theorem again. The next result explains why a repeated application of this idea will succeed.

Again we consider the initial condition at $x = a$ for convenience. However it works just as well with any initial point c in both directions with appropriate rewording.

5.3.5. Continuation Theorem. *Suppose that* $\Phi : [a,b] \times \mathbb{R}^n \to \mathbb{R}^n$ *is continuous and locally Lipschitz in* y. *Consider the DE*

$$(\ddagger) \qquad F'(x) = \Phi(x, F(x)), \quad F(a) = \Gamma.$$

Then either

(1) *the DE has a unique solution* $F_*(x)$ *on* $[a,b]$*; or*

(2) *there is a* $c \in (a,b]$ *so that the DE has a unique solution* $F_*(x)$ *on* $[a,c)$ *and* $\lim_{x\to c^-} \|F_*(x)\| = +\infty$.

Proof. Let $c := \sup\{d \in [a,b] : (\ddagger)$ has a solution on $[a,d]\}$. The Local Picard Theorem shows that there is a unique solution on $[a, a+h]$ for some $h > 0$, so $c > a$. Also no solution F_0 can extend to include c if $c < b$, for in that case, we can apply the Local Picard Theorem with initial data $F(c) = F_0(c)$ and extend the solution to an interval $[c, c+h]$.

Next we show that if $F_i(x)$ is a solution on $[a, d_i]$ with $d_1 < d_2$, then $F_2|_{[a,d_1]} = F_1$. If this were false, then

$$\inf\{x \in [a, d_1] : F_2(x) \neq F_1(x)\} = d < d_1.$$

Let $\Gamma_1 = F_1(d) = F_2(d)$. Apply the Local Picard Theorem with initial data $F(d) = \Gamma_1$ to obtain a unique solution on $[d, d+h_1]$. By uniqueness, we have $F_2(x) = F_1(x)$ on $[a, \min\{d+h_1, d_1\}]$, contradicting the definition of d. Thus $F_2|_{[a,d_1]} = F_1$. Therefore, there is a unique solution $F_*(x)$ of $(\ddagger)$ defined on $[a,c)$ by combining the solutions for $d < c$.

If $\lim_{x\to c^-} \|F_*(x)\| = +\infty$, then we satisfy part (2) of the theorem; and if the solution extends to include b, we have part (1). The remaining possibility is that the solution does not extend to include c, but

$$\liminf_{x\to c^-} \|F_*(x)\| < \infty.$$

In this last case, choose $x_n \to c^-$ such that $\|F_*(x_n)\| \leq K$ for $n \geq 1$. Since Φ is locally Lipschitz in y, it is Lipschitz in y on $D = [a,b] \times \overline{b_{2K+1}}(0)$ by Lemma 5.3.2. Let

$$M = \sup_{(x,y)\in D} \|\Phi(x,y)\| \quad \text{and} \quad \delta = \begin{cases} \min\{\frac{b-c}{2}, \frac{K+1}{2M}\} & \text{if } c < b \\ \frac{K+1}{2M} & \text{if } c = b. \end{cases}$$

Choose N so that $x_N > c - \delta$. Let $\Gamma_N = F_*(x_N)$. Apply the Local Picard Theorem to the DE $(\dagger)$ $F'(x) = \Phi(x, F(x))$ and $F(x_N) = \Gamma_N$. Since $\overline{b_{K+1}}(\Gamma_N) \subset$

$\overline{b_{2K+1}}(0)$, we obtain a unique solution on $[x_N, x_N + h]$ where

$$h = \min\left\{b - x_N, \frac{K+1}{M}\right\} \quad \text{and} \quad \frac{K+1}{M} \geq 2\delta.$$

Thus either the solution extends to include b or it extends to $[x_N, x_N + 2\delta] \supset [x_N, c + \delta]$. In either case, the solution extends beyond $[a, c)$, contrary to our hypothesis. Hence it must be the case that $\lim_{x \to c^-} \|F_*(x)\| = +\infty$. ■

5.3.6. REMARKS.

(1) The solution obtained in the Continuation Theorem is called the *maximal continuation* of the solution to the DE.

(2) Like the Local Picard Theorem, the Continuation Theorem works in both directions by symmetry.

(3) If the DE is defined on $\mathbb{R} \times \mathbb{R}^n$, one may have to restrict the x-domain as well as the y domain to get a local Lipschitz condition. One can restrict to $[-R, R] \times \overline{b_R}(0)$ to apply the Local Picard Theorem and the Continuation Theorem, and then piece the unique solutions together as in the proof above.

5.3.7. EXAMPLE. Not all solutions of differential equations blow up in the manner of the previous theorem. Consider this DE which was cooked up to have $f(x) = \sin \frac{1}{x}$ as a solution.

$$x^4 y'' + 2x^3 y' + y = 0 \qquad \text{and} \qquad y(\tfrac{2}{\pi}) = 1, \quad y'(\tfrac{2}{\pi}) = 0 \quad \text{for} \quad x \in \mathbb{R}.$$

This looks like a reasonably nice linear DE. However it is not in standard form, and the leading coefficient x^4 vanishes at $x = 0$. In standard form, it becomes

$$y'' = -\frac{2y'}{x} - \frac{y}{x^4} \quad \text{and} \quad \Phi(x, y_0, y_1) = \big(y_1, -\frac{2y_1}{x} - \frac{y_0}{x^4}\big).$$

The function Φ is discontinuous at $x = 0$, so we cannot expect any solution to include 0 in the domain. On the other hand, Φ satisfies a global Lipschitz condition on $[\varepsilon, \infty) \times \mathbb{R}^2$ because

$$\nabla_y \Phi = \begin{bmatrix} 0 & 1 \\ -\frac{1}{x^4} & -\frac{2}{x} \end{bmatrix}$$

is clearly bounded for $x \geq \varepsilon$. Therefore the Global Picard Theorem shows that this has a unique solution on $[\varepsilon, \infty)$. The uniqueness ensures that these solutions agree on the intersection, and so we can continue the solution to $(0, \infty)$.

The reader can check that $f(x) = \sin \frac{1}{x}$ is a solution. Clearly this does not extend beyond $(0, \infty)$. Since $f'(x) = \frac{-1}{x^2} \cos \frac{1}{x}$, the solution to the first order vector valued DE is

$$F(x) = \big(f(x), f'(x)\big) = \big(\sin \frac{1}{x}, -\frac{1}{x^2} \cos \frac{1}{x}\big).$$

Look at the points $x_n = \frac{2}{(2n+1)\pi}$. Then $F(x_n) = ((-1)^n, 0)$. Thus

$$\liminf_{x\to 0^+} \|F(x)\| \leq 1 < \infty.$$

This does not contradict the Continuation Theorem because Φ is not defined at $x = 0$.

5.4. Existence without Uniqueness

5.4.1. EXAMPLE. Consider the differential equation

$$y' = y^{2/3}, \qquad y(0) = 0, \qquad x \in \mathbb{R}.$$

The function $\Phi(x, y) = y^{2/3}$ is not Lipschitz because $\frac{\partial\Phi}{\partial y} = \frac{2}{3}y^{-1/3}$ blows up at $y = 0$. Hence the Picard Theorems do not apply.

We can try to solve it using separation of variables and integration to get:

$$3y^{1/3} = x + c.$$

Then $c = y(0) = 0$, and hence $y = x^3/27$. By inspection, one can see that $f(x) = 0$ is a solution. This is a second solution valid on the whole real line. Hence the solution is not unique.

In fact, there are infinitely many solutions. For any $b \leq 0 \leq a$, the function

$$f(x) = \begin{cases} \frac{(x-b)^3}{27} & \text{if} \quad x \leq b \\ 0 & \text{if} \quad b \leq x \leq a \\ \frac{(x-a)^3}{27} & \text{if} \quad x \geq a. \end{cases}$$

is a solution.

Note that as long as y is bounded away from 0, there is a Lipschitz condition on Φ. So any non-zero initial value for $y(c)$ will yield a solution by the Local Picard Theorem and the Continuation Theorem. What will happen is that this solution will be a cubic that will continue in both directions. But in one direction, it will eventually hit the x-axis tangentially. At that point, the solution may run along the axis for a while before continuing as a cubic on the other side of the axis. So the uniqueness of the continuation works until we hit the axis, where Φ is fails to be locally Lipschitz.

It turns out that only continuity of the function Φ is necessary to ensure the existence of a solution on a small interval. However the proof is more sophisticated, and relies on the Arzela–Ascoli Theorem 2.5.5.

5.4.2. PEANO'S THEOREM. *Let* $\Gamma \in \mathbb{R}^n$ *and* $\Phi : D = [a, b] \times \overline{b_R(\Gamma)} \to \mathbb{R}^n$ *be continuous. Then the DE*

$$F'(x) = \Phi(x, F(x)), \qquad f(a) = \Gamma \quad \textit{for} \quad a \leq x \leq b$$

has a solution on $[a, a+h]$*, where* $h = \min\{b-a, R/\|\Phi\|_\infty\}$.

PROOF. As in Picard's proof, we convert the problem to finding a fixed point for the map T, where

$$TF(x) = \Gamma + \int_a^x \Phi(t, F(t))\,dt.$$

Pick N so that $\frac{1}{N} < h$. For each $n \geq N$, define a function $F_n(x)$ on $[a, a+h]$ by:

$$F_n(x) = \begin{cases} \Gamma & \text{for} \quad a \leq x \leq a + \frac{1}{n} \\ \Gamma + \displaystyle\int_a^{x-1/n} \Phi(t, F_n(t))\,dt & \text{for} \quad a + \frac{1}{n} \leq x \leq a + h. \end{cases}$$

It appears that this definition depends on knowing F_n to define F_n, but actually, it is defined as the constant on $[a, a+\frac{1}{n}]$. Then the integral makes sense for x in $[a+\frac{1}{n}, a+\frac{2}{n}]$. Once F_n is defined there, the integral will make sense on $[a+\frac{2}{n}, a+\frac{3}{n}]$ as long as $F_n(x) \in \overline{b_R(\Gamma)}$. This follows from

$$\|F_n(x) - \Gamma\| \leq \int_a^{x-1/n} \|\Phi(t, F_n(t))\|\,dt \leq \|\Phi\|_\infty(h - \tfrac{1}{n}) \leq R.$$

Therefore $\Phi(t, F_n(t))$ is defined for $a \leq t \leq a+h$. So one proceeds in this manner, step by step, to define F_n on $[a, a+h]$.

Next we calculate for $a \leq x \leq a + \frac{1}{n}$,

$$\begin{aligned} \|TF_n(x) - F_n(x)\| &= \Big\| \Gamma + \int_a^x \Phi(t, F_n(t))\,dt - \Gamma \Big\| \\ &\leq \int_a^x \|\Phi(t, F_n(t))\|\,dt \leq \|\Phi\|_\infty(x-a) \leq \tfrac{1}{n}\|\Phi\|_\infty. \end{aligned}$$

Then for $a + \frac{1}{n} \leq x \leq a + h$,

$$\begin{aligned} \|TF_n(x) - F_n(x)\| &= \Big\| \int_{x-\frac{1}{n}}^x \Phi(t, F_n(t))\,dt \Big\| \\ &\leq \int_{x-\frac{1}{n}}^x \|\Phi(t, F_n(t))\|\,dt \leq \tfrac{1}{n}\|\Phi\|_\infty. \end{aligned}$$

Therefore $\|TF_n - F_n\|_\infty \leq \frac{1}{n}\|\Phi\|_\infty$.

Next we show that $\{F_n : n \geq N\}$ is equicontinuous. Given $\varepsilon > 0$, take $\delta = \varepsilon/\|\Phi\|_\infty$. If $a \leq x_1 < x_2 \leq a+h$ and $|x_2 - x_1| < \delta$, then

$$\begin{aligned} \|F_n(x_2) - F_n(x_1)\| &\leq \int_{x_1-\frac{1}{n}}^{x_2-\frac{1}{n}} \|\Phi(t, F_n(t))\|\,dt \\ &\leq \|\Phi\|_\infty |(x_1 - \tfrac{1}{n}) - (x_2 - \tfrac{1}{n})| < \|\Phi\|_\infty \delta = \varepsilon. \end{aligned}$$

The family of functions $\{F_n : n \geq 1\}$ is bounded by $\|\Gamma\| + R$ and equicontinuous. Therefore we may apply the Arzela–Ascoli Theorem to deduce that its closure is compact. Here we have a family of vector valued functions. However each of the n coordinate functiona are closed, bounded and equicontinuous. So the set is itself compact.

Hence we can extract a subsequence F_{n_i} which converges uniformly on $[a, a+h]$ to a function $F(x)$. We will show that F is a fixed point of T, and hence a solution of the DE. Compute

$$\begin{aligned}\|F(x)-TF(x)\| &\le \|F(x)-F_{n_i}(x)\|+\|F_{n_i}(x)-TF_{n_i}(x)\|+\|TF_{n_i}(x)-TF(x)\| \\ &\le \|F-F_{n_i}\|_\infty+\frac{\|\Phi\|_\infty}{n_i}+\int_a^{a+h}\big\|\Phi(t,F(t))-\Phi(t,F_{n_i}(t))\big\|\,dt.\end{aligned}$$

Since D is compact, Φ is uniformly continuous on D. For $\varepsilon > 0$, there is a $\delta > 0$ so that $\|y-z\| < \delta$ implies that $\|\Phi(x,y) - \Phi(x,z)\| < \varepsilon/h$ for all $x \in [a, a+h]$. Choose n_i so large that

$$\|F - F_{n_i}\|_\infty < \min\{\delta, \varepsilon\} \quad \text{and} \quad \|\Phi\|_\infty < n_i\varepsilon.$$

Then

$$\|F(x) - TF(x)\| < \varepsilon + \varepsilon + \int_a^{a+h} \varepsilon/h\,dt = 3\varepsilon.$$

The left side is constant, and the right side can be made arbitrarily small. Hence $TF = F$ is a solution. ■

5.5. Stability of DEs

Many DEs are designed to model behaviour in a perfect environment. In real life, there are often things which interfere with the process. For example, some physical actions are hypothesized to happen in a vacuum. When they occur in air, there is a component of friction from passing by air particles. In economics, there is often a significant noise component.

It is important to understand how a solution changes under such perturbations. The famous quote about a butterfly in the Amazon jungle affecting weather in Canada is a very long-term effect. Even under the ideal conditions, small changes in the initial conditions can have a dramatic effect far in the future. What we are concerned with here are local issues: how does the solution change nearby? Does a small change in the DE or in the initial conditions result in a small change in the solution? The short answer is yes if the initial DE is locally Lipschitz.

5.5.1. PERTURBATION THEOREM. *Let $\Phi : D = [a,b] \times \overline{b_R(\Gamma)} \to \mathbb{R}^n$ satisfy a Lipschitz condition in y with constant L. Suppose that $\Psi : D \to \mathbb{R}^n$ is another continuous function on D such that*

$$\|\Psi - \Phi\|_\infty = \sup_{(x,y)\in D} \|\Psi(x,y) - \Phi(x,y)\| \le \varepsilon.$$

(The function Ψ is not assumed to be Lipschitz.) Suppose that F and G are solutions of the differential equations

$$F'(x) = \Phi(x, F(x)), \qquad F(a) = \Gamma$$

and

$$G(x)' = \Psi(x, G(x)), \qquad G(a) = \Delta$$

such that $(x, F(x))$ and $(x, G(x))$ belong to D for $a \le x \le b$. Define $\delta := \|\Delta - \Gamma\|$. Then, for all $x \in [a, b]$,

$$\|G(x) - F(x)\| \le \delta e^{L|x-a|} + \frac{\varepsilon}{L}(e^{L|x-a|} - 1).$$

Thus

$$\|G - F\|_\infty \le \delta e^{L|b-a|} + \frac{\varepsilon}{L}(e^{L|b-a|} - 1).$$

PROOF. Define

$$\tau(x) = \|G(x) - F(x)\| = \Big(\sum_{i=0}^{n-1} \big(g_i(x) - f_i(x)\big)^2\Big)^{1/2}.$$

In particular, $\tau(c) = \|\Delta - \Gamma\| = \delta$. Then by the Cauchy–Schwarz inequality,

$$\begin{aligned} 2\tau(x)\tau'(x) = (\tau(x)^2)' &= \sum_{i=0}^{n-1} 2\big(g_i(x) - f_i(x)\big)\big(g_i'(x) - f_i'(x)\big) \\ &\le 2\Big(\sum_{i=0}^{n-1}\big(g_i(x) - f_i(x)\big)^2\Big)^{1/2}\Big(\sum_{i=0}^{n-1}\big(g_i'(x) - f_i'(x)\big)^2\Big)^{1/2} \\ &= 2\tau(x)\|G'(x) - F'(x)\|. \end{aligned}$$

Also

$$\begin{aligned} \|G'(x) - F'(x)\| &= \|\Psi(x, G(x)) - \Phi(x, F(x))\| \\ &\le \|\Psi(x, G(x)) - \Phi(x, G(x))\| + \|\Phi(x, G(x)) - \Phi(x, F(x))\| \\ &\le \varepsilon + L\|G(x) - F(x)\| = \varepsilon + L\tau(x). \end{aligned}$$

Therefore if $\tau(x) \ne 0$, we get

$$\tau'(x) \le \|G'(x) - F'(x)\| \le \varepsilon + L\tau(x).$$

Consider $x \in [a, b]$. In the inequality above, we needed to divide by $\tau(x)$. However it clearly is not a problem for us if $\tau(x) = 0$ because this will improve our estimates. Define

$$d = \sup\{a \le t \le x : \tau(t) = 0\} \cup \{a\}.$$

If $d = x$, then $\tau(x) = 0$. Otherwise

$$x - a \ge x - d = \int_d^x dt \ge \int_d^x \frac{\tau'(t)}{\varepsilon + L\tau(t)}\, dt$$

$$= \frac{1}{L}\log(\varepsilon + L\tau)\Big|_d^x = \frac{1}{L}\log\Big(\frac{L\tau(x)+\varepsilon}{L\tau(d)+\varepsilon}\Big).$$

Therefore

$$\frac{L\tau(x)+\varepsilon}{L\tau(d)+\varepsilon} \le e^{L|x-a|} \quad \text{or} \quad L\tau(x)+\varepsilon \le e^{L|x-a|}(L\tau(d)+\varepsilon).$$

Now $\tau(d) \le \tau(a) = \delta$. Thus by solving for $\tau(x)$, we get

$$\|G(x) - F(x)\| = \tau(x) \le \delta e^{L|x-a|} + \frac{\varepsilon}{L}(e^{L|x-a|} - 1).$$

The last statement is straightforward. ■

An important consequence of this result is that the solution of a DE with Lipschitz condition is a continuous function of the initial conditions. For simplicity, we take $c = a$, but it is readily modified for other points.

5.5.2. COROLLARY. *Suppose that Φ satisfies a local Lipschitz condition in y on $[a, b] \times \mathbb{R}^n$. Then the solution F_Γ of*

$$F'(x) = \Phi(x, F(x)), \qquad F(a) = \Gamma$$

is a continuous function of Γ.

PROOF. Fix $R > 0$ and let L be the Lipschitz constant in y on $D = [a, b] \times \overline{b_{R+1}(\Gamma)}$ and let $M = \sup_{(x,y)\in D} \|\Phi(x, y)\|$. Assume that $\|\Delta - \Gamma\| < 1$. Then $\overline{b_R(\Delta)} \subset \overline{b_{R+1}(\Gamma)}$. Then applying the Local Picard Theorem with either initial condition $f(a) = \Gamma$ or $F(a) = \Delta$, we obtain a unique solution F_Γ or F_Δ on $[a, a+h]$ where $h = \min\{b - a, \frac{R}{M}\}$ and the solution graphs remain in D.

Therefore we may apply the Perturbation theorem with $\Psi = \Phi$ in the Perturbation Theorem (so that $\varepsilon = 0$) to show that

$$\|F_\Delta(x) - F_\Gamma(x)\| \le \|\Delta - \Gamma\| e^{L|x-a|} \quad \text{and} \quad \|F_\Delta - F_\Gamma\|_\infty \le \|\Delta - \Gamma\| e^{Lh}.$$

Thus the solution depends continuously on the initial conditions. ■

The Perturbation Theorem can be interpreted as a stability result. If the DE and initial data are measured empirically, then this theorem assures us that the approximate solution based on the measurements remains reasonably accurate.

5.5.3. EXAMPLE. The DE $y'' + y = 0$, $y(0) = 0$, $y'(0) = 1$ for $x \in [-\pi, \pi]$ has unique solution $f(x) = \sin x$. The associated function $\Phi(x, y_0, y_1) = (y_1, -y_0)$ has global Lipschitz constant 1. Consider

$$y'' + y = e(x, y, y') \qquad y(0) = 0 \quad \text{and} \quad y'(0) = 1 \quad \text{for} \quad x \in [-\pi, \pi],$$

where $e(x, y, y')$ is a small function bounded by ε; and let $g(x)$ be the solution. By the Perturbation Theorem, since $\delta = 0$, we get

$$\big\|\big(g(x) - \sin x, g'(x) - \cos x\big)\big\| \le \varepsilon(e^{|x|} - 1).$$

Exercises

1. Consider the DE: $y' = 1 + xy$ and $y(0) = 0$ for $x \in [-b, b]$, where $b > 0$.
(a) Reduce this to finding the fixed point of a mapping T. Show that when $b = 1$, the map T is a contraction mapping.
(b) Prove that the DE has a unique solution on $[-b, b]$ for any $b > 0$. Hence deduce that there is a unique solution on the whole line $\mathbb{R}$.
(c) Let $f_0(x) = 1$ and compute $f_n(x) = T^n f_0$ by induction. Prove directly (rather than by quoting a theorem) that the sequence f_n converges uniformly on $[-b, b]$.

2. Consider the linear DE $y'' - x^{-1}y' + x^{-2}y = 0$ for $x \in [1, 3]$.
(a) Show that this DE has a unique solution for each choice of initial values $\Gamma = (y(1), y'(1)) \in \mathbb{R}^2$.
(b) Check that $y = x$ is a solution. Find a solution of the form $f(x) = xg(x)$ by showing that g' satisfies a 1st order DE, and solving it.
(c) Show that the set of solutions that you obtain is a 2-dimensional vector space that contains all possible solutions.

3. Consider the DE: $y'' = (1 + (y')^2)^{3/2}$ and $y(0) = 0$, $y'(0) = 0$.
(a) Convert this to a first order vector valued DE. Show that it satisfies a local Lipschitz condition, and find an $h > 0$ so that a solution exists on $[-h, h]$.
(b) Show that $f(x) = 1 - \sqrt{1 - x^2}$ is the unique solution on $(-1, 1)$.
(c) This solution does not continue further, yet $|f(x)| \leq 1$. Why does this not contradict the Continuation Theorem?

4. Suppose that Φ and Ψ are Lipschitz functions defined on $[a, b] \times \mathbb{R}$. Let f and g be solutions of $f' = \Phi(x, f(x))$ and $g' = \Psi(x, g(x))$, respectively. Suppose that $f(a) \leq g(a)$ and $\Phi(x, y) \leq \Psi(x, y)$ for all $(x, y) \in [a, b] \times \mathbb{R}$. Show that $f(x) \leq g(x)$ for all $x \in [a, b]$. HINT: If $f(x) = g(x)$, what about $f'(x)$ and $g'(x)$?

5. Consider the DE: $y' = x^2 + y^2$ and $y(0) = 0$.
(a) Show that this DE satisfies a local Lipschitz condition but not a global one.
(b) Integrate the inequality $y' \geq 1 + y^2$ for $x \geq 1$ to prove that the solution must go off to infinity in a finite time.

6. Show that the set of all solutions on $[a, a+h]$ to the DE of Peano's Theorem is closed, bounded, and equicontinuous; and thus is compact.

7. Let $\gamma \in \mathbb{R}$ and let Φ be a continuous real-valued function on $[a, b] \times [\gamma - R, \gamma + R]$. Consider the DE $y'(x) = \Phi(x, y)$ and $y(a) = \gamma$. Then Peano's Theorem guarantees a solution on $[a, a + h]$ for some $h > 0$.
(a) If f and g are both solutions on $[a, a + h]$, show that $f \vee g(x) = \max\{f(x), g(x)\}$ and $f \wedge g(x) = \min\{f(x), g(x)\}$ are also solutions.
HINT: Verify the DE in $U = \{x : f(x) > g(x)\}$, $V = \{x : f(x) < g(x)\}$, and $X = \{x : f(x) = g(x)\}$ separately.
(b) Prove that the set of all solutions on $[a, a + h]$ has a largest and smallest solution.
HINT: use Exercise (6). Find a countable dense subset $\{f_n\}$ of the set of solutions. Let $g_k = \max\{f_1, \ldots, f_k\}$ for $k \geq 1$. Show that g_k converges to the maximal solution $f_{\max}$.

8. Prove Gronwall's inequality: suppose that $u \in C[0,b]$ satisfies $u \geq 0$ and there are non-negative constants C and K so that

$$u(x) \leq C + K \int_0^x u(t)\,dt \quad \text{for} \quad 0 \leq x \leq b.$$

Prove $u(x) \leq Ce^{Kx}$ for $x \in [0,b]$.
HINT: let $v(x)$ denote the RHS, derive a differential inequality, and integrate.

INDEX

F_σ set, 51
G_δ set, 51
ε-net, 30
p-adic metric, 6

Stone-Weierstrass Theorem, 85

accumulation point, 11
algebra, 85
Arzela-Ascoli Theorem, 43

Baire Category Theorem, 50
Baire one, 51
Banach space, 19
Bolzano-Weierstrass Theorem, 24
Borel-Lebesgue Theorem, 31

Cantor set, 38
Cauchy sequence, 19
Chebychev Approximation Theorem, 83
clopen set, 10
closed ball, 9
closed set, 9
closure, 11
compact, 29
complete, 19
complete ordered field, 71
completion, 64
connected component, 45
connected set, 43
Continuation Theorem, 100
continuity, 13
continuous dependence on initial conditions, 106
contraction mapping, 56
Contraction Mapping Principle, 56
convergent sequence, 11

Dedekind cut, 74
dense, 34
discrete metric, 5

embedding, 72
equicontinuity, 42
equioscillation of degree n, 83
Extension Theorem, 66
Extreme Value Theorem, 35

finite subcover, 29
fixed point, 56
fractal, 59

geodesic distance, 5
Global Picard Theorem, 97

Hamming distance, 5
Hausdorff metric, 5
Heine-Borel Theorem, 33
homeomorphism, 15

induced metric, 7
initial value problem, 93
integral equation, 91
Intermediate Value Theorem, 44
isolated point, 11
isometry, 14

Least Upper Bound Principle, 23
Least upper bound property, 71
Lebesgue number, 37
limit point, 11
linear functionals, 21
linear ODE, 95
Lipschitz at x, 53
Lipschitz function, 15
Lipschitz in y, 95
Local Picard Theorem, 98
locally Lipschitz in y, 98

maximal continuation, 101
metric space, 4
Minkowski's inequality, 2

Newton's Method, 61
norm, 1
normed vector space, 1
nowhere dense, 38, 50
nowhere differentiable, 53

open ball, 9
open cover, 29
open set, 9
order of a DE, 93
ordered field, 71
oscillation of f at x, 51

path, 40
path connected, 45
Peano curve, 40
Peano's Theorem, 102
perfect, 38
Perturbation Theorem, 104
pointwise convergence, 26

rectifiable curve, 49
residual set, 50
reverse triangle inequality, 4

self-adjoint, 88
seminorm, 1
separable, 34
separates points, 85
separation of variables, 90
sequential continuity, 13
sequentially compact, 29
space filling curve, 40
standard form, 93
Stone-Weierstrass Theorem, 87
subcover, 29
subsequence, 11

totally bounded, 30
totally disconnected, 47
triangle inequality, 4

uniform continuity, 13
uniform convergence, 25
uniform equicontinuity, 42

vanishes at x, 85
vector lattice, 85

Weierstrass Approximation Theorem, 79
Weierstrass M-test, 27