Functional Analysis for Quantum Information

Jason Crann
David W. Kribs
Vern I. Paulsen

Contents

Chapter 1

Metric and Topological Spaces

1.1 Topics to be covered

In this chapter we cover basic topics in analysis and topology that provide a mathematical foundation for the topics from functional analysis presented in subsequent chapters. We also include an introductory discussion on a key notion from quantum information: the quantum bit or "qubit".

Readers with a background that includes a good course in real or complex analysis can move swiftly through this chapter, whereas those coming to the text from different fields should carefully work through the material.

- Review of Metric Spaces

  Definition of metric and examples

  Convergence of sequences

  $\epsilon - \delta$ definition of Continuity

  Open, closed and compact sets

  Sequential characterizations

  Connected and pathwise connected sets

  Equivalence and Uniform equivalence of metrics

  Cauchy sequences and completeness

  $C(\mathbb{R})$ as a metric space

  Baire's theorem

- Qubits and the Bloch Sphere

  Qubit metrics and fidelity

- General Topological Spaces and Nets

  Open and closed sets, continuity

  Nets and directed sets

  Closed sets, compact sets, and continuity in terms of nets

1.2 Metric Spaces

Many mathematical investigations require various notions of distance to be quantified. This is accomplished at a most basic level through metrics.

Definition 1.1. Given a set X a **metric** on X is a function, $d : X \times X \to \mathbb{R}$ that satisfies for all $x, y, z \in X$:

1. $d(x, y) \geq 0$,
2. $d(x, y) = 0$ if and only if $x = y$,
3. $d(x, y) = d(y, x)$,
4. $d(x, z) \leq d(x, y) + d(y, z)$.

The pair (X, d) is called a **metric space**. If in place of 2) we only have that $d(x, x) = 0$ for all $x \in X$, then d is called a **pseudometric**.

Example 1.2. Some standard examples of metrics on real spaces include the following (with complex versions defined similarly):

- $(\mathbb{R}, d)$ where $d(x, y) = |x - y|$.
- $(\mathbb{R}^n, d_p)$ where $d_p(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p}$, for any $1 \leq p < +\infty$. When $p = 2$ we call this the *Euclidean distance*.
- $(\mathbb{R}^n, d_\infty)$ where $d_\infty(x, y) = \max\{|x_i - y_i| : 1 \leq i \leq n\}$.
- X any non-empty set, define $d(x, y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$. This is called the *discrete metric*.
- Given a metric space (X, d), and a subset $W \subseteq X$, (W, d) is also a metric space. This is called a *metric subspace of X*.

Definition 1.3. Let $(X, d_1), (Y, d_2)$ be metric spaces, let $f : X \to Y$ be a function, and let $x_0 \in X$. We say that f is **continuous at** x_0 provided that for every $\epsilon > 0$, there exists $\delta > 0$ such that $d_1(x, x_0) < \delta$ implies $d_2(f(x), f(x_0)) < \epsilon$. We call f **continuous** if it is continuous at every $x_0 \in X$.

Example 1.4. Consider the subset $(-1, +1) \subseteq \mathbb{R}$ of $(\mathbb{R}, d)$. Observe that the function $f : \mathbb{R} \to (-1, +1)$ given by $f(x) = \frac{x}{1+|x|}$ is one-to-one, onto and continuous. The inverse function $g(t) = \frac{t}{1-|t|}$ is also continuous.

Definition 1.5. Let (X, d) be a metric space. A subset $\mathcal{O}$ is **open** provided that whenever $x \in \mathcal{O}$, then there exists $\delta > 0$ such that the set

$$B(x_0; \delta) := \{x \mid d(x_0, x) < \delta\} \subseteq \mathcal{O}.$$

The set $B(x_0; \delta)$ is called the **ball centered at** x_0 **of radius** δ. A set is called **closed** if its complement is open.

These concepts also characterize continuity.

Proposition 1.6. *Let* $(X, d_1), (Y, d_2)$ *be metric spaces. Let* $f : X \to Y$. *The following conditions are equivalent:*

1. f *is continuous,*
2. *for every open set* $U \subseteq Y$, *the set* $f^{-1}(U) := \{x \in X : f(x) \in U\}$ *is open,*
3. *for every closed set* $C \subseteq Y$, *the set* $f^{-1}(C)$ *is closed.*

Convergence of sequences also gives good characterizations of continuity.

Definition 1.7. Let (X, d) be a metric space. A sequence $\{x_n\}_{n \in \mathbb{N}}$ **converges to** x_0 provided that for every $\epsilon > 0$ there is N such that $n > N$ implies $d(x_n, x_0) < \epsilon$. We write $\lim_n x_n = x_0$ or $x_n \to x_0$.

Proposition 1.8. *Let* $(X, d_1), (Y, d_2)$ *be metric spaces and let* $f : X \to Y$. *Then*

1. f *is continuous at* $x_0 \in X$ *if and only if for every sequence* $\{x_n\}$ *such that* $\lim_n x_n = x_0$ *we have that* $\lim_n f(x_n) = f(x_0)$.
2. $C \subseteq X$ *is closed if and only if for every convergent sequence* $\{x_n\} \subseteq C$ *we have that* $\lim_n x_n \in C$.

1.2.1 Equivalent Metrics

Definition 1.9. Let X be a set and let d_1 and d_2 be metrics on X. We say that d_1 and d_2 are **equivalent** provided that a set is open in the d_1 metric if and only if it is open in the d_2 metric. We say that d_1 and d_2 are **uniformly equivalent** provided that there are constants, $A, B > 0$ such that $d_1(x, y) \leq A\, d_2(x, y)$ and $d_2(x, y) \leq B\, d_1(x, y)$ for all $x, y \in X$.

It is easy to see that if two metrics are uniformly equivalent then they are equivalent. Also, two metrics are equivalent if and only if the function $id : X \to X$ is continuous from (X, d_1) to (X, d_2) and from (X, d_2) to (X, d_1).

A map $f : X \to Y$ between metric spaces such that f is one-to-one, onto with both f and f^{-1} continuous is called a *homeomorphism.* Thus, two metrics are equivalent if and only if the identity map is a homeomorphism.

Example 1.10. Returning to the metrics d_1, d_2, d_∞ on $\mathbb{R}^n$, observe that for $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ we have:

- $d_\infty(x, y) \leq d_1(x, y)$ and $d_1(x, y) \leq n\, d_\infty(x, y)$,
- $d_\infty(x, y) \leq d_2(x, y)$ and $d_2(x, y) \leq \sqrt{n}\, d_\infty(x, y)$,
- $d_1(x, y) \leq \sqrt{n}\, d_2(x, y)$ (use Cauchy-Schwarz) and $d_2(x, y) \leq d_1(x, y)$.

It follows that these are all uniformly equivalent metrics.

Problem 1.11. Let (X, d) be a metric space. Define $r(x, y) = \frac{d(x,y)}{1+d(x,y)}$. Prove that r is a metric on X and that d and r are equivalent metrics.

1.2.2 Cauchy Sequences and Completeness

Definition 1.12. Let (X, d) be a metric space. A sequence $\{x_n\}$ is **Cauchy** provided that for every $\epsilon > 0$ there exists N so that when $n, m > N$ then $d(x_n, x_m) < \epsilon$. A metric space is called **complete** provided that every Cauchy sequence converges to a point in X.

Definition 1.13. Let (Y, d) be a metric space. A subset $X \subseteq Y$ is called **dense** provided that for every $y \in Y$ and every $\epsilon > 0$ there is a point $x \in X$ with $d(y, x) < \epsilon$.

The canonical example is that the rational numbers $\mathbb{Q}$ together with $d(x, y) = |x-y|$ is a metric space that is not complete. The method in which we construct $\mathbb{R}$ from $\mathbb{Q}$ by adding limits of Cauchy sequences, generalizes to any metric space.

Theorem 1.14. *Let (X, d) be a metric space. Then there is a metric space $(\hat{X}, \hat{d})$ so that*

1. *$X \subseteq \hat{X}$,*
2. *$\hat{d}(x, y) = d(x, y)$ for every pair $x, y \in X$,*
3. *X is a dense subset of $\hat{X}$.*

Moreover, if $(\tilde{X}, \tilde{d})$ is another metric space satisfying 1), 2), 3), then there is a homeomorphism $h : \hat{X} \to \tilde{X}$ such that $h(x) = x$ for every $x \in X$ and $\tilde{d}(h(\hat{x}), h(\hat{y})) = \hat{d}(\hat{x}, \hat{y})$ for every pair $\hat{x}, \hat{y} \in \hat{X}$.

Definition 1.15. The (unique) metric space $(\hat{X}, \hat{d})$ given in the above theorem is called the **completion** of (X, d).

The following problem shows that the property of being complete is NOT invariant under equivalence of metric.

Problem 1.16. On $\mathbb{R}$ define $d(x, y) = |x - y|$, $r(x, y) = \frac{d(x,y)}{1+d(x,y)}$ and $s(x, y) = |\frac{x}{1+|x|} - \frac{y}{1+|y|}|$. Prove that

1. s is a metric on $\mathbb{R}$,
2. d, r and s are all equivalent metrics,
3. $(\mathbb{R}, d)$ and $(\mathbb{R}, r)$ are complete metric spaces,
4. $(\mathbb{R}, s)$ is not complete.

1.2.3 Compact Sets

Definition 1.17. Let (X, d) be a metric space. Then a subset $K \subseteq X$ is called **compact** provided that whenever $\{U_a\}_{a \in A}$ is a collection of open sets such that $K \subseteq \cup_{a \in A} U_a$, then there is a finite subset $F \subseteq A$ such that $K \subseteq \cup_{a \in F} U_a$.

The following gives a nice characterization of this property in metric spaces.

Theorem 1.18. *Let (X, d) be a metric space. The following conditions are equivalent:*

1. *K is compact,*

2. *every sequence* $\{x_n\}_{n\in\mathbb{N}} \subseteq K$ *has a subsequence* $\{x_{n_k}\}$ *that converges to a point in* K*,*

3. (K,d) *is complete and for all* $\epsilon > 0$ *there is a finite subset* $\{x_1, ..., x_n\} \subseteq K$ *such that for each* $x \in K$ *there is an* x_i *with* $d(x, x_i) < \epsilon$*.*

Condition 2) is often called the *Bolzano-Weierstrass property*, and a subset as in 3) is called an ϵ*-net*. One consequence of this result is the important Heine-Borel theorem.

Corollary 1.19. *A subset* $K \subseteq \mathbb{R}^n$ *is compact if and only if it is closed and bounded.*

1.2.4 Uniform Convergence

Proposition 1.20. *Let* (K,d) *be a non-empty compact metric space and let* $f : K \to \mathbb{R}$ *be continuous. Then there are points* $x_M, x_m \in K$ *such that* $f(x_M) = \sup\{f(x) : x \in K\}$ *and* $f(x_m) = \inf\{f(x) : x \in K\}$*. In particular, both the supremum and infimum are finite.*

Proof. Choose a sequence $\{x_n\}$ so that $\lim_n f(x_n) = \sup\{f(x) : x \in K\}$. This then has a convergent subsequence with $\lim_k x_{n_k} = x_M \in K$. By continuity, $f(x_M) = \lim_k f(x_{n_k}) = \sup\{f(x) : x \in K\}$. The rest of the proof is similar. □

One motivation for studying metrics is that they can often be built to capture various kinds of convergence. We illustrate this with one example.

Definition 1.21. Let (X,d) be a metric space and let $K \subseteq X$. We say that a sequence of functions $\{f_n\} \subset C(X)$ **converges uniformly to** $f \in C(X)$ **on** K provided that

$$\lim_n \sup\{|f(x) - f_n(x)| : x \in K\} = 0.$$

We say that $\{f_n\}$ **converges uniformly on compact subsets to** f provided that $\{f_n\}$ converges uniformly to f on K for every compact subset K of X.

Example 1.22. We now construct a metric that for $(\mathbb{R}, d)$ captures uniform convergence on compact subsets. Let $C(\mathbb{R})$ denote the continuous real-valued functions on $\mathbb{R}$ and let $f, g \in C(\mathbb{R})$. For each n, set

$$d_n(f,g) = \sup\{|f(t) - g(t)| : -n \le t \le +n\},$$

and let $r_n(f,g) = \frac{d_n(f,g)}{1+d_n(f,g)}$. We define

$$r(f,g) = \sum_{n=1}^{\infty} \frac{r_n(f,g)}{2^n}.$$

Note that since $0 \leq r_n(f,g) \leq 1$ this series converges.

We shall make use of the following inequalities below.

Lemma 1.23. *Let $a, b, c \geq 0$. Then $a \leq b$ if and only if $\frac{a}{1+a} \leq \frac{b}{1+b}$. Further, if $a \leq b + c$, then $\frac{a}{1+a} \leq \frac{b}{1+b} + \frac{c}{1+c}$.*

Theorem 1.24. *Let $f_n, f \in C(\mathbb{R})$. Then $\lim_n r(f_n, f) = 0$ if and only if $\{f_n\}$ converges uniformly to f on compact subsets. Moreover, $(C(\mathbb{R}), r)$ is a complete metric space.*

Proof. First we show that r is a metric. It is clear that $r(f,g) = 0$ exactly when $f = g$ and that $r(f,g) = r(g,f)$. To see the triangle inequality, let $f, g, h \in C(\mathbb{R})$. It is clear that $d_n(f,g) \leq d_n(f,h) + d_n(h,g)$ so by the lemma, $r_n(f,g) \leq r_n(f,h) + r_n(h,g)$. Now it follows that $r(f,g) \leq r(f,h) + r(h,g)$.

Suppose that $\lim_n r(f_n, f) = 0$. Given a compact subset K and $\epsilon > 0$, pick an m so that $K \subseteq [-m, +m]$. Choose an N so that $n > N$ implies that $r(f_n, f) \leq 2^{-m} \frac{\epsilon}{1+\epsilon}$. Then for $n > N$ we have that

$$2^{-m} \frac{d_m(f_n, f)}{1 + d_m(f_n, f)} \leq r(f_n, f) < 2^{-m} \frac{\epsilon}{1+\epsilon},$$

and hence $d_m(f_n, f) < \epsilon$. But $\sup\{|f_n(t) - f(t)| : t \in K\} \leq d_m(f_n, f) < \epsilon$. Thus, $\{f_n\}$ converges uniformly to f on K.

Conversely, assume that $\{f_n\}$ converges uniformly to f on every K and that $\epsilon > 0$ is given. Pick M so that $\sum_{m=M+1}^{\infty} 2^{-m} < \epsilon/2$. Pick δ so that $\frac{\delta}{1+\delta} = \epsilon/2$, and finally pick an N so that for $n > N$, we have that $d_M(f_n, f) < \delta$.

Then for $n > N$ it follows that

$$\begin{aligned} r(f_n, f) &\leq \sum_{m=1}^{M} 2^{-m} r_m(f_n, f) + \sum_{m=M+1}^{\infty} 2^{-m} \\ &\leq \sum_{m=1}^{M} 2^{-m} r_M(f_n, f) + \epsilon/2 \\ &< \frac{d_M(f_n, f)}{1 + d_M(f_n, f)} + \epsilon/2 \\ &< \frac{\delta}{1+\delta} + \epsilon/2 = \epsilon. \end{aligned}$$

Hence, uniform convergence on compact sets implies convergence in the metric.

Finally, if a sequence $\{f_n\}$ is Cauchy in the metric r, then it is pointwise Cauchy and so there is a function, $f(x) = \lim_n f_n(x)$. Also it is easy to show that for each M and $\epsilon > 0$, there must be an N so that $n, m > N$ implies that $d_M(f_n, f_m) < \epsilon$. From this it follows that f_n converges uniformly to f on $[-M, +M]$ and so is continuous on $[-M, +M]$. Thus, f is continuous. More of the same shows that $r(f_n, f) \to 0$, and so the space is complete. □

1.2.5 Baire's Theorem

Some of the deepest results in functional analysis are consequences of Baire's theorem. One statement of the theorem is as follows.

Theorem 1.25. *Let (X, d) be a complete metric space and let $\{U_n\}_{n\in\mathbb{N}}$ be a countable collection of open, dense sets. Then $\cap_{n\in\mathbb{N}} U_n$ is dense in X.*

Proof. Given $\epsilon > 0$ and $x \in X$, since U_1 is dense, we can find $x_1 \in U_1$ so that $d(x, x_1) < \epsilon/2$. As U_1 is open we can pick $\delta_1 < \epsilon/2$ so that the closure of the ball of radius δ_1 around x_1 satisfies

$$B(x_1; \delta_1)^- \subseteq U_1 \cap B(x; \epsilon),$$

where the set on the left side of this inclusion denotes the closure of $B(x_1; \delta_1)$. Applying the same reasoning to U_2 we may pick $x_2 \in U_2$ so that $d(x_1, x_2) < \delta_1/2$, and a $\delta_2 < \delta_1/2$ so that $B(x_2, \delta_2)^- \subseteq U_2 \cap B(x_1; \delta_1)$.

Continuing inductively, we define x_n, δ_n such that $\delta_{n+1} < \delta_n/2 \le \epsilon/2^n$ and $B(x_{n+1}; \delta_{n+1})^- \subseteq U_{n+1} \cap B(x_n; \delta_n)$.

The fact that $d(x_n, x_{n+1}) < \epsilon/2^n$ is enough to show that $\{x_n\}$ is Cauchy, and so there is a point $x_0 = \lim_n x_n$. The containments imply the points $x_n, x_{n+1}, ...$ are all inside $B(x_n, \delta_n)^-$, and hence, $x_0 \in B(x_n; \delta_n)^- \subseteq U_n$. Thus, $x_0 \in \cap_{n\in\mathbb{N}} U_n$. Also, since each $B(x_n, \delta_n)^- \subseteq B(x, \epsilon)$ we have that $d(x, x_0) < \epsilon$. □

Definition 1.26. A set E is called **nowhere dense** provided that the set complement of the closure E^- is dense, i.e., $X \backslash E^-$ is dense.

Note that U_n is dense and open if and only if $E_n = X \backslash U_n$ is closed and nowhere dense.

A much weaker statement than saying that $\cap U_n$ is dense, is to say that it is non-empty. Taking complements, this is just the statement that $X \neq X \backslash \cap U_n = \cup (X \backslash U_n)$. Thus, the following result appears weaker than Baire's theorem, but nevertheless it is quite important and referred to as Baire's Category Theorem.

Theorem 1.27. *A complete metric space cannot be written as a countable union of nowhere dense sets.*

The name comes from the following. A set that can be written as a countable union of nowhere dense sets is called a set of *first category*. A set that cannot is called a set of *second category*. In this language, the above theorem says that a complete metric space is of second category. Another name for sets of first category is to call them *meagre sets*.

Example 1.28. An example that illustrates this theorem comes from consideration of $(\mathbb{Q}, d)$ with $d(x, y) = |x - y|$. Since $\mathbb{Q}$ is countable we may write it as $\mathbb{Q} = \{q_n : n \in \mathbb{N}\}$. Now set $E_n = \{q_n\}$. Then this is a countable collection of nowhere dense sets and $\mathbb{Q} = \cup E_n$. Thus, by the theorem, $(\mathbb{Q}, d)$ is not complete, something that we knew already.

Problem 1.29. What about $(\mathbb{N}, d)$ and $E_n = \{n\}$?

1.3 Qubits and the Bloch Sphere

The basic unit of information in quantum information theory is the quantum bit or *qubit*. We will give a more detailed treatment later; for the moment we can simply take a qubit to be a unit vector in $\mathbb{C}^2$. Thus, for the purposes of this initial discussion, a qubit is given by $\psi = \alpha|0\rangle + \beta|1\rangle$ where $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$ and

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

If $\alpha \neq 0 \neq \beta$, then ψ is said to be in a *superposition* of the (classical) states $|0\rangle$ and $|1\rangle$.

When qubits represent pure states on a 2-level quantum mechanical system (with $|0\rangle$ and $|1\rangle$ as the classical base states), then the vectors ψ and $\lambda\psi$ correspond to the same pure state, where λ is any complex number of modulus 1, i.e., a point on the unit circle in the complex plane. Thus, we are often interested in the set of equivalence classes of qubits where we define ψ and ϕ to be equivalent if there exists λ, $|\lambda| = 1$, such that $\psi = \lambda\phi$. If we let $[\psi]$ denote the equivalence class of the vector ψ, then the collection of all such equivalence classes is the complex projective space, denoted $\mathbb{CP}^1$.

Note that when ψ is a qubit whose first coordinate is non-zero, then ψ is equivalent to a unique vector of the form (a, b) with $0 < a \leq 1$. When $0 < a < 1$ then all equivalence classes are given by $(a, \lambda\sqrt{1-a^2})$, which is the product of an interval and a circle, i.e., a cylinder. However, when $a = 1$ there is only one equivalence class, $[(1, 0)]$ and similarly, when $a = 0$, there is only the equivalence class $[(0, 1)]$. Thus, the set of equivalence classes of qubits can be regarded as a cylinder where the circles at each endpoint are collapsed to single points, i.e., a two dimensional sphere. This object is known as the *Bloch sphere*.

Let us give one of the standard explicit descriptions of the Bloch sphere, which we will refer to later. Observe that we can write any unit vector $\psi \in \mathbb{C}^2$ as

$$|\psi\rangle = e^{i\gamma}(\cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle),$$

for some $\gamma \in \mathbb{R}$ and $0 \leq \theta, \phi < 2\pi$. As noted above, the "global phase factor" $e^{i\gamma}$ may be ignored, and thus the angles θ and ϕ describe a unique point on the surface of the unit sphere as depicted in Figure blochsphere 1.1.

Problem 1.30. In this parametrization, verify the following special cases of states identified with points on the Bloch sphere: the north pole ($\theta = 0$) and $|0\rangle$; the south pole ($\theta = \pi$) and $|1\rangle$; and important superposition states,

$$\theta = \frac{\pi}{2}, \phi = 0 \quad \text{and} \quad |+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle),$$
$$\theta = \frac{\pi}{2}, \phi = \pi \quad \text{and} \quad |-\rangle := \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle).$$

Later we will see another representation of equivalence classes of qubits in the context of the postulates of quantum mechanics, by identifying each qubit with the matrix of the projection onto the one dimensional space that it spans.

1.3.1 Qubit Metrics and Fidelity

There are several natural metrics on the Bloch sphere. The first comes from its identification as equivalence classes of vectors. We set

$$d([\phi], [\psi]) = \inf\{\|\phi - \lambda\psi\|_2 : |\lambda| = 1\},$$

where $\|(a, b)\|_2 = \sqrt{|a|^2 + |b|^2}$ denotes the usual Euclidean distance. The other metrics we shall use come from the concept of *fidelity*.

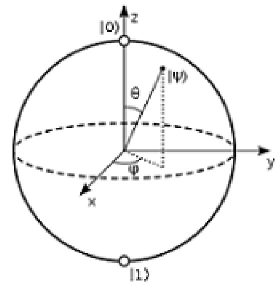

Figure 1.1: Unit vectors, also known as “pure states”, are represented as surface points on the sphere, and “mixed” (rank-2) states correspond to interior points. Moving inside, further from the surface means “more mixed” as we shall clarify further later, in particular the centre corresponds to the *maximally state* $\rho = \frac{1}{2}I$. (Image courtesy of Google Images) blochsphere

Definition 1.31. Given two qubits, $\phi = (a, b)$, $\psi = (c, d)$ their **fidelity** is the quantity

$$F(\phi, \psi) = |\langle\phi|\psi\rangle|^2 = |\bar{a}c + \bar{b}d|^2.$$

Note $0 \leq F(\phi, \psi) \leq 1$, also that fidelity is really a function on equivalence classes, and $F(\phi, \psi) = 1$ if and only if $[\phi] = [\psi]$.

Two other important metrics are given as follows.

Definition 1.32. The **Fubini-Study metric** on the Bloch sphere is given by

$$d_{FS}([\phi], [\psi]) = \arccos(F(\phi, \psi)).$$

The **Bures metric** is given by

$$d_B([\phi], [\psi]) = \sqrt{1 - F(\phi, \psi)}.$$

It can be shown that these three metrics are all uniformly equivalent and the Bloch sphere is a complete metric space in each of them.

1.4 Topological Spaces

Not all notions of convergence can be defined by a metric, this led to a generalization of metric spaces. Recall that in a metric space, arbitrary unions of open sets are open and finite intersections of open sets are open. This motivates the following definition.

Definition 1.33. Let X be a set $\mathcal{T}$ be a collection of subsets of X. We call $\mathcal{T}$ a **topology on X** and call $(X, \mathcal{T})$ **a topological space**, provided that:

1. $\emptyset \in \mathcal{T}$ and $X \in \mathcal{T}$,
2. whenever $U_a \in \mathcal{T}$, $\forall a \in A$, then $\big(\cup_{a \in A} U_a \big) \in \mathcal{T}$,
3. whenever $U_i \in \mathcal{T}$, $1 \leq i \leq n$, then $\big(\cap_{i=1}^n U_i \big) \in \mathcal{T}$.

We call a subset of X **open** if it is in $\mathcal{T}$ and we call a subset of X **closed** if its complement is in $\mathcal{T}$. Given a point $x \in X$, we let $\mathcal{N}_x = \{U \in \mathcal{T} : x \in U\}$ and call this collection the **neighborhoods of x**.

We base the notion of continuity on the open set description in metric spaces.

Definition 1.34. Given two topological spaces $(X, \mathcal{T})$ and $(Y, \mathcal{S})$ and a function $f : X \to Y$ we say that:

- f is **continuous at** x_0 provided that for every $V \in \mathcal{S}$ with $f(x_0) \in V$ there is $U \in \mathcal{T}$ with $x_0 \in U$ such that $f(U) \subseteq V$, i.e., every neighborhood of $f(x_0)$ contains $f(U)$ for some neighborhood of x_0;
- f is **continuous** provided that for every $V \in \mathcal{S}$ we have that $f^{-1}(V) := \{x \in X : f(x) \in V\} \in \mathcal{T}$.

It is not hard to see that when X and Y are metric spaces these reduce to the usual definitions of continuity.

Problem 1.35. Prove that $f : X \to Y$ is continuous if and only if it is continuous at every point in X.

1.4.1 Nets and Directed Sets

Many results about topological spaces have the same proofs as in the metric space case if one replaces sequences with nets.

Definition 1.36. A pair $(\Lambda, \leq)$ consisting of a set Λ together with a relation $\leq$ on Λ is called a **directed set** if the relation satisfies:

1. $x \leq y$ and $y \leq x$ implies that $x = y$ (**symmetric**),
2. $x \leq y$ and $y \leq z$ implies that $x \leq z$ (**transitive**),
3. for every $x_1, x_2 \in \Lambda$ there is $x_3 \in \Lambda$ such that $x_1 \leq x_3$ and $x_2 \leq x_3$.

Example 1.37. Below are some examples of directed sets.

- Let Λ be either $\mathbb{N}, \mathbb{Z}$, or $\mathbb{R}$ with the usual $\leq$.
- Let S be a set and let $\mathcal{F}$ be the collection of all finite subsets of S. Given $F_1, F_2 \in \mathcal{F}$ define $F_1 \leq F_2$ if and only if $F_1 \subseteq F_2$.
- Given a topological space $(X, \mathcal{T})$ and a point x, let $\Lambda = \mathcal{N}_x$ be the set of all open neighborhoods of x and define $U_1 \leq U_2$ if and only if $U_2 \subseteq U_1$. The intuition of this order is that "farther out" means smaller set and so "closer" to x.
- We find important examples in calculus as well. Given an interval $[a, b]$ recall that a *partition* is a set $P = \{x_0 = a < x_1 \cdots < x_n = b\}$. These are used to subdivide the interval. An *augmentation of* P is a collection, $P^* = \{x'_1 < \cdots < x'_n\}$ where $x_{i-1} \leq x'_i \leq x_i$. The pair $\mathcal{P} = (P, P^*)$ is called an *augmented partition*. We define $\mathcal{P}_1 = (P_1, P_1^*) \leq \mathcal{P}_2 = (P_2, P_2^*)$ provided that $P_1 \subseteq P_2$ and $P_1^* \subseteq P_2^*$. It is easy to see that $\leq$ satisfies the first two properties needed to be a directed set. The third property is trickier: Given any two augmented partitions $\mathcal{P}_1 = (P_1, P_1^*)$ and $\mathcal{P}_2 = (P_2, P_2^*)$, let $P_1 \cup P_1^* \cup P_2 \cup P_2^* = \{a = x_0 < ... < x_m = b\}$ add one extra point, c with $x_{m-1} < c < b$ and let $P_3 = \{x_0 < ... < x_{m-1} < c < x_m = b\}$. Now let $P_3^* = \{x_0 < ... < x_{m-1} < x_m\}$, then this is an augmentation of P_3 and $P_1 \cup P_2 \subseteq P_3$ and $P_1^* \cup P_2^* \subseteq P_3$, so that $\mathcal{P}_1 \leq \mathcal{P}_3$ and $\mathcal{P}_2 \leq \mathcal{P}_3$.

Definition 1.38. Let $(X, \mathcal{T})$ be a topological space. Then a **net in X** is a directed set $(\Lambda, \leq)$ together with a function $f : \Lambda \to X$. As with sequences we prefer to set $x_\lambda = f(\lambda)$ and write the net as $\{x_\lambda\}_{\lambda \in \Lambda}$. We say that the net $\{x_\lambda\}_{\lambda \in \Lambda}$ **converges to x** and write $\lim_\lambda x_\lambda = x$ provided that for each $U \in \mathcal{N}_x$ there is $\lambda_0 \in \Lambda$ such that when $\lambda_0 \leq \lambda$ then $x_\lambda \in U$.

Here is the careful statement of the Riemann sum theorem from calculus: Let $f : [a, b] \to \mathbb{R}$ be continuous. For each augmented partition $\mathcal{P} = (\{a = x_0 < \cdots < x_n = b\}, \{x'_1 < \cdots < x'_n\})$, the *Riemann sum* is given by

$$S_{\mathcal{P}} = \sum_{i=1}^{n} f(x'_i)(x_i - x_{i-1}).$$

The set of all Riemann sums forms a net of real numbers.

Theorem 1.39. *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then the net of Riemann sums converges, and we call this limit the "Riemann integral" of f.*

1.4.2 Unordered vs Ordered Sums

Given a set A and real numbers $r_a, a \in A$, we wish to define $\sum_{a\in A} r_a$. To do this consider the directed set $(\mathcal{F}, \leq)$ of all finite subsets of A. Given $F \in \mathcal{F}$ we set

$$s_F = \sum_{a\in F} r_a$$

and we call this number the *partial sum over* F. The collection $\{s_F\}_{F\in F}$ is a net of real numbers, called the *net of partial sums*.

Definition 1.40. We say that $\sum_{a\in A} r_a$ **converges to** s provided that s is the limit of the net of partial sums.

Thus, $s = \sum_{a\in A} r_a$ if and only if for each $\epsilon > 0$ there is a finite set $F_0 \subseteq A$ such that for every finite set F with $F_0 \subseteq F$ we have that

$$|s - s_F| < \epsilon.$$

Given $r_n, n \in \mathbb{N}$, it is interesting to compare the convergence of the unordered series, $\sum_{n\in\mathbb{N}} r_n$ with the convergence of the ordered series, $\sum_{n=1}^{+\infty} r_n$. For the latter case we only consider partial sums of the form $\sum_{n=1}^{K} r_n$, a very small collection of all finite subsets of $\mathbb{N}$.

Since we only need this smaller collection of partial sums to approach a value, it is easy to see that whenever $\sum_{n\in\mathbb{N}} r_n$ converges, then $\sum_{n=1}^{+\infty} r_n$ will converge.

The converse is not true. For instance, the series $\sum_{n=1}^{+\infty} \frac{(-1)^n}{n}$ converges, but $\sum_{n\in\mathbb{N}} \frac{(-1)^n}{n}$ does not converge.

Problem 1.41. Prove that $\sum_{n\in\mathbb{N}} r_n$ converges if and only if $\sum_{n=1}^{+\infty} |r_n|$ converges, i.e. if and only if the series *converges absolutely.*

Here are just some of the reasons that nets are convenient.

Proposition 1.42. *Let* $(X, \mathcal{T})$ *and* $(Y, \mathcal{S})$ *be topological spaces and let* $f : X \to Y$*. Then*

1. f *is continuous at* x_0 *if and only if for every net* $\{x_\lambda\}_{\lambda\in\Lambda}$ *such that* $\lim_\lambda x_\lambda = x_0$ *we have that* $\lim_\lambda f(x_\lambda) = f(x_0)$.

2. f *is continuous if and only if whenever a net* $\{x_\lambda\}_{\lambda\in\Lambda}$ *in* X *converges to a point* x*, then* $\lim_\lambda f(x_\lambda) = f(x)$.

3. *A set* $C \subseteq X$ *is closed if and only if whenever* $\{x_\lambda\}_{\lambda\in\Lambda}$ *is a net in* C *that converges to a point* $x \in X$*, then* $x \in C$.

4. *A set $K \subseteq X$ is compact if and only if every net in K has a subnet that converges to a point in K.*

Note that these results are the counterparts of many results for metric spaces. In general they are not true if you only use sequences instead of nets.

We have not yet defined subnets.

Definition 1.43. Let Λ and D be two directed sets. A function $g : D \to \Lambda$ is called **final** provided that $a \le b$ implies $g(a) \le g(b)$ and given any $\lambda_0 \in \Lambda$ there is $d_0 \in D$ such that $d_0 \le d$ implies $\lambda_0 \le g(d)$. Given a net $\{x_\lambda\}_{\lambda\in\Lambda}$ a **subnet** is any net of the form $\{x_{g(d)}\}_{d\in D}$ for some directed set D and final function $g : D \to \Lambda$.

Just as with subsequences, we will often write a subnet as $\{x_{\lambda_d}\}_{d\in D}$ where really, $\lambda_d = g(d)$.

Problem 1.44. When $\Lambda = D = \mathbb{N}$ then every subsequence is a subnet but a subnet need not be a subsequence. Explain why.

1.4.3 The Key Separation Axiom

Definition 1.45. A topological space $(X, \mathcal{T})$ is called **Hausdorff** provided that for any $x \ne y$ there are open set U, V such that $x \in U$, $y \in V$ and $U \cap V = \emptyset$.

This axiom guarantees that there are lots of continuous functions. The following result is called Urysohn's Lemma.

Theorem 1.46. *Let $(K, \mathcal{T})$ be a compact Hausdorff space and let A, B be closed subsets with $A \cap B = \emptyset$. Then there is a continuous function $f : K \to [0, 1]$ such that $f(a) = 0$, for all $a \in A$ and $f(b) = 1$, for all $b \in B$.*

The following result is Tietze's ExtensionTheorem.

Theorem 1.47. *Let $(K, \mathcal{T})$ be a compact, Hausdorff space and let $A \subseteq K$ be a closed subset and let $f : A \to [0, 1]$ be continuous. Then there is a continuous function $F : K \to [0, 1]$ such that $F(a) = f(a)$, for all $a \in A$.*

Bibliography

agmc [1] Jim Agler and John E. McCarthy, *Pick Interpolation and Hilbert Function Spaces*, Graduate Studies in Mathematics, vol. 44, American Mathematical Society, Providence, Rhode Island, 2002.

aron [2] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. **68** (1950), 337–404. MR0051437 (14,479c)

bass [3] R.F. Bass, *Stochastic Processes*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 33, Cambridge University Press, Cambridge, UK, 2011.

berg [4] S. Bergman, *Ueber die Entwicklung der harmonischen Funktionen der Ebene und des Raumes nach Orthogonalfunktionen*, Math. Ann. **86** (1922).

deBrRo [5] Louis de Branges and James Rovnyak, *Square summable power series*, Holt, Rinehart and Winston, New York, 1966. MR0215065 (35 #5909)

deBrRo [6] Louis de Branges, *A proof of the Bieberbach conjecture*, Acta Math. **154** (1985), no. 1-2, 137–152, DOI 10.1007/BF02392821. MR772434 (86h:30026)

deBr [7] ———, *Hilbert spaces of entire functions*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1968. MR0229011 (37 #4590)

conway [8] John B. Conway, *A course in functional analysis*, 2nd ed., Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1990. MR1070713 (91e:46001)

DoPa [9] Ronald G. Douglas and Vern I. Paulsen, *HIlbert Modules over Function Algebras*, Pitman Research Notes in Mathematics, vol. 217, Longman Scientific, 1989.

folland [10] Gerald B. Folland, *Real Analysis: Modern Techniques and Their Applications*, John Wiley & Sons, Inc., New York, New York, 1999.

moore [11] E. H. Moore, *General Analysis*. 2, Vol. 1, 1939.

ross [12] Sheldon Ross, *A First Course in Probability*, 9th ed., Pearson Education Limited, Harlow, Esex, England, 2012.

sarason [13] Donald Sarason, *Complex Function Theory*, American Mathematical Society, Providence, Rhode Island, 2007.

steinshakarchi [14] Elias M. Stein and Rami Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*, Princeton Lectures in Analysis III, Princeton University Press, Princeton, New Jersey, 2005.

tak [15] M Takesaki, *Theory of Operator Algebras I*, Encyclopedia of Mathematical Sciences, vol. 124, Springer-Verlag, Berlin–Heidelberg–New York, 2002.

rasswill [16] Christopher K. I. Williams and Carl Edward Rasmussen, *Gaussian Processes for Machine Learning*, MIT Press, 2006.