Chapter 3

The Space of Continuous Functions

仲尼適楚，出於林中，見痀僂者承蜩，猶掇之也。仲尼曰：子巧乎。 有道邪。曰：我有道也。五六月累丸二而不墜，則失者錙銖；累三而不墜，則失者十一； 累五而不墜，猶掇之也。吾處身也，若厥株拘；吾執臂也，若槁木之枝；雖天地之大，萬物之多，而唯蜩翼之知。吾不反不側，不以萬物易蜩之翼，何為而不得。 孔子顧謂弟子曰：用志不分，乃凝於神，其痀僂丈人之謂乎。 莊子 達生

In this chapter we study the space of continuous functions as a prototype of infinite dimensional normed spaces. In Section 1 we review these spaces. In Section 2 the notion of separability is introduced. A proof of Weierstrass approximation theorem different from the one given in Chapter 1 is present in Section 3, following by the general Stone-Weierstrass theorem. The latter is applied to establish the separability of the space of continuous functions when the underlying space is compact. Ascoli-Arezela theorem, which characterizes compact sets in the space of continuous functions, is established in Section 4. Finally in Section 5 we study complete metric spaces. Baire category theorem is proved and, as an application, it is shown that continuous, nowhere differentiable functions form a set of second category in the space of continuous functions.

3.1 Spaces of Continuous Functions

We studied continuous functions on an interval in MATH2050/60 and in a domain bounded by curves/surfaces in $\mathbb{R}^2$ or $\mathbb{R}^3$ in MATH2010/20. In this chapter we will focus on the space of continuous functions defined on a metric space.

Let $C(X)$ denote the vector space of all continuous functions defined on X where (X, d) is a metric space. Recall that in the exercise we showed that there are many continuous functions in X. In general, in a metric space such as the real line, a continuous function may not be bounded. In order to turn continuous functions into a normed space, we need to restrict to bounded functions. For this purpose let

$$C_b(X) = \{f : \ f \in C(X), \ |f(x)| \leq M, \ \forall x \in X \text{ for some } M\}.$$

It is readily checked that $C_b(X)$ is a normed space under the sup-norm. From now on, $C_b(X)$ is always regarded as a metric space under the metric induced by the sup-norm. In order words,

$$d_\infty(f, g) = \|f - g\|_\infty, \quad \forall f, g \in C_b(X).$$

Some basic properties of $C_b(X)$ are listed below:

Property 1. *$C_b(X)$ is a complete metric space.* Indeed, let $\{f_n\}$ be a Cauchy sequence in $C_b(X)$. For $\varepsilon > 0$, there exists some n_0 such that $\|f_n - f_m\|_\infty < \varepsilon/4$ for all $n \geq n_0$. In particular, it means for each x, $\{f_n(x)\}$ is a Cauchy sequence in $\mathbb{R}$. By the completeness of $\mathbb{R}$, the limit $\lim_{n\to\infty} f_n(x)$ exists and we define $f(x) \equiv \lim_{n\to\infty} f_n(x)$. Assuming that $f \in C_b(X)$, by taking $m \to \infty$ in the inequality above, we immediately obtain $\|f_n - f\|_\infty \leq \varepsilon/4 < \varepsilon$, hence $f_n \to f$ in $C_b(X)$. To show that $f \in C_b(X)$, we let $m \to \infty$ in $|f_n(x) - f_m(x)| < \varepsilon/4$ to get $|f_n(x) - f(x)| \leq \varepsilon/4$ for all x and $n \geq n_0$. Taking $n = n_0$ we get $|f(x)| \leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x)| \leq \varepsilon/4 + \|f_{n_0}\|_\infty$, hence f is bounded. On the other hand, as f_{n_0} is continuous, for each x we can find a δ such that $|f_{n_0}(y) - f_{n_0}(x)| < \varepsilon/4$ whenever $d(y, x) < \delta$. It follows that for all y, $d(y, x) < \delta$,

$$|f(y) - f(x)| \leq |f(y) - f_{n_0}(y)| + |f_{n_0}(y) - f_{n_0}(x)| + |f_{n_0}(x) - f(x)| \leq \frac{3\varepsilon}{4} < \varepsilon.$$

From this proof we see that the completeness of $C_b(X)$ is inherited from the completeness of $\mathbb{R}$, so the underlying space X does not play any role in this aspect.

Property 2. *$C_b(X) = C(X)$ when X is a compact metric space.* We need to show every continuous function on a compact set is bounded. Assume on the contrary that for some continuous f, there are points $\{x_k\}$ such that $|f(x_k)| \to \infty$. By compactness, there is a subsequence $\{x_{k_j}\}$ and $z \in X$ such that $\lim_{j\to\infty} x_{k_j} = z$. But, by continuity we would have $\lim_{j\to\infty} |f(x_{k_j})| = |f(z)| < \infty$, contradiction holds.

Property 3. *$C_b(X)$ forms an algebra under pointwise product.* Recall that an algebra is a vector space V in which a product satisfying the association law is well-defined between two points. The interaction between this product and the vector space structure is reflected in the rules $w(u + v) = (u + v)w = wu + wv$ and $(au)(bv) = ab(uv)$ for all $u, v, w \in V$ and $a, b \in \mathbb{R}$. It is clear the product of two bounded, continuous functions is again a bounded, continuous function, so $C_b(X)$ forms an algebra.

We will investigate various properties of the spaces of continuous functions. Recall that a consequence of Weierstrass approximation theorem tells that every continuous function on $[a, b]$ can be approximated by polynomials with rational coefficients. In general a set E in a metric space is called a **dense set** if its closure is equal to the space. Thus the collection of all polynomials with rational coefficients forms a dense set in $C[a, b]$. Since this set is countable, we know that every continuous function in $[a, b]$ can be approximated from continuous functions chosen from a countable set. Our question is, in a general space $C(X)$, when does this property still hold? To obtain a result in the positive direction, we need to establish a generalization of Weierstrass approximation theorem, namely, Stone-Weierstrass theorem. That is what we are going to do in the next section.

3.2 Stone-Weierstrass Theorem

In Chapter 1 we proved Weierstrass approximation theorem. Here we first present an alternate and optional proof of this theorem. Our proof in Chapter is via a short cut. Now the proof is a bit longer but inspiring.

Theorem 3.1 (**Weierstrass Approximation Theorem**). *Let $f \in C([a,b])$. For every $\varepsilon > 0$, there exists a polynomial p such that*

$$\|f - p\|_\infty < \varepsilon.$$

Proof. Assume the interval to be $[0,1]$ first. Let f be a continuous function on $[0,1]$. By subtracting it from a linear function (a polynomial of degree 0 or 1) which passes $(0, f(0))$ and $(1, f(1))$ we may assume $f(0) = f(1) = 0$. Extend it to be a continuous function in $\mathbb{R}$ which equals zero outside the unit interval and denote the extended function still by f. Now we approximate f by introducing

$$p_n(x) := \int_{-1}^{1} f(x+t) Q_n(t)\,dt, \quad x \in [0,1],$$

where $Q_n(x) := c_n(1-x^2)^n$. (In fact, as f vanishes outside $[0,1]$, the integration is in fact from $-x$ to $1-x$.) Q_n is a polynomial and the normalizing constant c_n is chosen so that $\int_{-1}^1 Q_n = 1$, in other words, c_n is given by

$$c_n^{-1} = \int_{-1}^{1} (1-x^2)^n dx.$$

We will need an upper estimate for c_n in a second,

$$\begin{aligned}
\int_{-1}^{1} (1-x^2)^n dx &= 2\int_0^1 (1-x^2)^n dx \\
&\ge 2\int_0^{1/\sqrt{n}} (1-x^2)^n dx \\
&\ge 2\int_0^{1/\sqrt{n}} (1-nx^2) dx \\
&= \frac{4}{3\sqrt{n}} > \frac{1}{\sqrt{n}},
\end{aligned}$$

after using Bernoulli's inequality in the form $(1-x^2)^n \geq 1-nx^2$ for all x. It follows that

$$c_n < \sqrt{n}.$$

From this estimate we obtain

$$Q_n(x) \leq \sqrt{n}(1-\delta^2)^n, \quad \forall x \in [\delta, 1],$$

so, in particular, $Q_n(x) \to 0$ uniformly on $[\delta, 1]$ for any fixed $\delta \in (0,1)$.

By a change of variables, it is clear that

$$p_n(x) = \int_{-x}^{1-x} f(x+t)Q_n(t)dt = \int_0^1 f(t)Q_n(t-x)dt$$

is a polynomial. We claim that

$$\|p_n - f\|_\infty \to 0.$$

For, given any $\varepsilon > 0$, there exists δ such that

$$|f(x)-f(y)| < \varepsilon, \quad |x-y| < \delta, \quad x, y \in [0,1].$$

We have, for all $x \in [0,1]$,

$$\begin{aligned}
|p_n(x) - f(x)| &= \Big|\int_{-1}^1 (f(x+t)-f(x))Q_n(t)dt\Big| \\
&\leq \int_{-1}^1 |f(x+t)-f(x)|Q_n(t)dt \\
&\leq 2M\int_{-1}^{-\delta} Q_n(t)dt + \frac{\varepsilon}{2}\int_{-\delta}^{\delta} Q_n(t)dt + 2M\int_\delta^1 Q_n(t)dt \\
&\leq 4M\sqrt{n}(1-\delta^2)^n + \frac{\varepsilon}{2} \\
&< \varepsilon, \quad \text{where } M := \sup|f|,
\end{aligned}$$

for all sufficiently large n.

For a continuous function f defined on $[a,b]$, the function $g(x) = f((b-a)x+a)$ belongs to $C[0,1]$. For $\varepsilon > 0$, there exists a polynomial p such that $\|g-p\|_\infty < \varepsilon$. Noting that the function $q(x) = p((x-a)/(b-a))$ is again a polynomial, we have $\|f-q\|_\infty = \|g-p\|_\infty < \varepsilon$. □

So far we have shown that trigonometric functions and polynomials are dense in the space of periodic, continuous functions and the space of continuous functions respectively. In this section we will establish a far-reaching generalization of these results in the space of continuous functions defined in a compact metric space. In such a space both trigonometric functions and polynomials are not available, so we need to seek a reasonable formulation. The answer relies on an extra algebraic structure we have exploited explicitly.

We have pointed out that the space $C_b(X)$ carries an extra property, namely, it is an algebra under pointwise product. A subspace $\mathcal{A}$ is called a subalgebra of $C_b(X)$ if it is closed under this product. It is readily checked that all polynomials on $[a,b]$ form a subalgebra of $C[a,b]$, so does the subalgebra consisting of all polynomials with rational coefficients. Similarly, the vector space consisting of all trigonometric functions and its subspace consisting of all trigonometric functions with rational coefficients are algebras in in the space of 2π-periodic continuous functions in $\mathcal{C}_{2\pi}$. Note that $\mathcal{C}_{2\pi}$ can be identified with $C(S^1)$ where $S^1 = \{(\cos t, \sin t) \in \mathbb{R}^2 : \ t \in [0, 2\pi]\}$ is the unit circle.

Before proceeding further, recall that aside from the constant ones, there are many continuous functions in a metric space. For instance, given any two distinct points x_1 and x_2 in X, it is possible to find some $f \in C(X)$ such that $f(x_1) \neq f(x_2)$. We simply take $f(x) = d(x, x_1)$ and $f(x_1) = 0 < f(x_2)$. It is even possible to find one in $C_b(X)$, e.g., $g(x) = f(x)/(1 + f(x))$ serves this purpose. We now consider what conditions a subalgebra must possess in order that it becomes dense in $C_b(X)$. A subalgebra is called to satisfy the **separating points property** if for any two points x_1 and x_2 in X, there exists some $f \in \mathcal{A}$ satisfying $f(x_1) \neq f(x_2)$. From the discussion above, it is clear that $\mathcal{A}$ must satisfy the separating points property if its closure is $C_b(X)$. Thus the separating points property is a necessary condition for a subalgebra to be dense. On the other hand, the polynomials of the form $\sum_{j=0}^{n} a_j x^{2j}$ form an algebra which does not have the separating point property, for it is clear that $p(-x) = p(x)$ for such p. Another condition is that, whenever $x \in X$, there must be some $g \in \mathcal{A}$ such that $g(x) \neq 0$. We will call this the **non-vanishing property.** The non-vanishing property fails to hold for the algebra consisting of all polynomials of the form $\sum_{j=1}^{n} a_j x^j$, for $p(0) = 0$ for all these p. The non-vanishing property is also a necessary condition for an algebra to be dense in $C(X)$. For, if for some particular $z \in X$, $f(z) = 0$ holds for all $f \in \mathcal{A}$, it is impossible to approximate the constant function 1 in the supnorm by functions in $\mathcal{A}$. Surprisingly, it turns out these two conditions are also sufficient when the underlying space is compact, and this is the content of the following theorem. This is for an optional reading.

Theorem 3.2 (**Stone-Weierstrass Theorem**). *Let $\mathcal{A}$ be a subalgebra of $C(X)$ where X is a compact metric space. Then $\mathcal{A}$ is dense in $C(X)$ if and only if it has the separating points and non-vanishing properties.*

Recall that $C_b(X) = C(X)$ when X is compact. For the proof of this theorem two lemmas are needed.

Lemma 3.3. *Let $\mathcal{A}$ be a subalgebra of $C_b(X)$. For every pair $f, g \in \mathcal{A}$, $|f|, f \vee g$ and $f \wedge g$ belong to the closure of $\mathcal{A}$.*

Proof. Observing the relations

$$f \vee g \equiv \max\{f, g\} = \frac{1}{2}(f+g) + \frac{1}{2}|f+g|,$$

and

$$f \wedge g \equiv \min\{f, g\} = \frac{1}{2}(f+g) - \frac{1}{2}|f+g|,$$

it suffices to show that $|f|$ belongs to the closure of $\mathcal{A}$. Indeed, given $\varepsilon > 0$, since $t \mapsto |t|$ is continuous on $[-M, M]$, $M = \sup\{|f(x)| : x \in X\}$, by Weierstrass approximation theorem, there exists a polynomial p such that $\big||t| - p(t)\big| < \varepsilon$ for all $t \in [-M, M]$. It follows that $\||f| - p(f)\|_\infty \le \varepsilon$. As $p(f) \in \mathcal{A}$, we conclude that $\mathcal{A}$ is dense in $C(X)$. □

Lemma 3.4. *Let $\mathcal{A}$ be a subalgebra in $C(X)$ which separates points and non-vanishing at all points. For $x_1, x_2 \in X$ and $\alpha, \beta \in \mathbb{R}$, there exists $\varphi \in \mathcal{A}$ such that $\varphi(x_1) = \alpha$ and $\varphi(x_2) = \beta$.*

Proof. Since $\mathcal{A}$ separates points, we can find some $\psi \in \mathcal{A}$ such that $\psi(x_1) \neq \psi(x_2)$. We claim that one can further choose ψ such that $\psi(x_1), \psi(x_2)$ are both non-zero. For, if, for instance, $\psi(x_1) = 0$, fix some $\xi \in \mathcal{A}$ satisfying $\xi(x_1) \neq 0$. This is possible due to the non-vanishing property. Consider a function $\psi_1 \in \mathcal{A}$ of the form $\psi + t\xi$. We would like to find $t \in \mathbb{R}$ such that (a) $\psi_1(x_1) \neq \psi_1(x_2)$, (b) $\psi_1(x_1) \neq 0$, and (c) $\psi_1(x_2) \neq 0$. There are two cases; when $\xi(x_2) \neq 0$, it suffices to choose t such that $t \neq 0, -\psi(x_2)/\xi(x_2)$ (if $\xi(x_2) \neq 0$). When $\xi(x_2) = 0$, we choose t such that $t \neq \psi(x_2)/\xi(x_1)$. Replacing ψ by ψ_1, we obtain our desired function which satisfies (a)–(c).

Now, we can find a and b such that the combination $\varphi = a\psi + b\psi^2 \in \mathcal{A}$ satisfies the requirement in the lemma. Indeed, what we need are the conditions $a\psi(x_1) + b\psi^2(x_1) = \alpha$ and $a\psi(x_2) + b\psi^2(x_2) = \beta$. As the determinant of this linear system (viewing a and b as the unknowns) is equal to $\psi(x_1)\psi(x_2)(\psi(x_1) - \psi(x_2))$ which is not equal to 0, a and b can always be found. □

Proof of Theorem 3.5. It remains to establish the necessary part of the theorem. Let $f \in C(X)$ be given. For each pair of x, y, there exists a function $\varphi_{x,y} \in \mathcal{A}$ satisfying $\varphi_{x,y}(x) = f(x)$ and $\varphi_{x,y}(y) = f(y)$. This is due to the previous lemma when x and y are distinct. When x is equal to y, such function still exists. Now, for each $\varepsilon > 0$, there exists an open set $U_{x,y}$ containing x and y such that

$$|f(t) - \varphi_{x,y}(t)| < \varepsilon, \quad \forall t \in U_{x,y}.$$

For fixed y, the sets $\{U_{x,y} : x \in X\}$ form an open cover of X. By the compactness of X, it admits a finite subcover $\{U_{x_j,y}\}_{j=1}^N$. The function $\varphi_y = \varphi_{x_1,y} \vee \cdots \vee \varphi_{x_N,y}$ belongs to $\overline{\mathcal{A}}$ according to Lemma 3.3. Furthermore, $\varphi_y > f - \varepsilon$ in X. For, let $x \in X$, there is some $U_{x_j,y}$ containing x. Therefore, $\varphi_y(x) \ge \varphi_{x_j,y}(x) > f(x) - \varepsilon$. Next, $G_y \equiv \cap_{j=1}^N U_{x_j,y}$ is an open set containing y and all these open sets together form an open cover of X when y runs over

X. Note that $\varphi_y < f+\varepsilon$ on X since $\varphi_{x_j,y} < f+\varepsilon$ in G_y for all $j = 1, \cdots, N$. By compactness, we can extract $y_1, \cdots, y_M$ such that $\{G_{y_k}\}_{k=1}^M$, cover X. Define $\varphi = \varphi_{y_1} \wedge \cdots \wedge \varphi_{y_M}$. By Lemma 3.3 it belongs to $\overline{\mathcal{A}}$ and $\varphi > f - \varepsilon$ in X. On the other hand, each x belongs to some G_{y_k}, so $\varphi(x) \le \varphi_{y_k}(x) < f(x)+\varepsilon$ holds. We conclude that $\|f-\varphi\|_\infty < \varepsilon, \varphi \in \overline{\mathcal{A}}$. □

3.3 Separabilty

Recall that a set E in a metric space (X, d) is dense if for every $x \in X$ and $\varepsilon > 0$, there exists some $y \in E$ such that $d(y, x) < \varepsilon$. Equivalently, E is dense if $\overline{E} = X$. The space X is called a **separable space** if it admits a countable dense subset. Equivalently, X is separable if there is a countable subset E satisfying $\overline{E} = X$. A set is **separable** if it is separable as a metric subspace. When a metric space is separable, every element can be approximated by elements from a countable set. Hence its structure is easier to study than the non-separable ones. Here are two basic properties of separable spaces.

Proposition 3.5. *Every subset of a separable space is separable.*

Proof. Let Y be a subset of the separable space (X, d) and $D = \{x_j\}$ a countable dense subset of X. For each n, pick a point z_j^n from $Y \cap B_{1/n}(x_j)$ if it is non-empty to form the countable set $E = \{z_j^n\}$. We claim that E is dense in Y. For, let $y \in Y$ and each $\varepsilon > 0$, there is some $x_j \in D$ such that $d(y, x_j) < \varepsilon/2$. Therefore, for $n > 2/\varepsilon$, $B_{1/n}(x_j) \cap Y$ is nonempty and we can find some $y_j^n \in B_{1/n}(x_j) \cap Y$. It follows that $d(y, y_j^n) \le d(y, x_j) + d(x_j, y_j^n) < \varepsilon/2 + 1/n < \varepsilon$.

□

Proposition 3.6. *Every compact metric space is separable.*

Proof. Every compact space is totally bounded. By Proposition 2.11, for each n, there exist finitely many points $x_1, \cdots, x_N$ such that the balls $B_{1/n}(x_j)$, $j = 1, \cdots N$, form an open cover of the space. It is clear that the countable set consisting of all centers of these balls when n runs from 1 to infinity forms a dense set of the space.

□

Example 3.1. Consider the Euclidean space $\mathbb{R}^n$. As it is well-known that the set of all rational numbers $\mathbb{Q}$ forms a countable dense subset of $\mathbb{R}$, $\mathbb{R}$ is a separable space. Similarly, $\mathbb{R}^n$ is separable for all $n \ge 1$ because it contains the dense subset $\mathbb{Q}^n$. According to Proposition 3.5, all sets in the Euclidean space are separable.

Example 3.2. $C[a, b]$ is a separable space. Although the proof was assigned as an exercise, we repeat it here. Without loss of generality we take $[a, b] = [0, 1]$. Denote by $\mathcal{P}$ the restriction of all polynomials to $[0,1]$. Let

$$\mathcal{S} = \{p \in \mathcal{P} : \text{ The coefficients of } p \text{ are rational numbers}\}.$$

It is clear that $\mathcal{S}$ is a countable set. Given any polynomial $p(x) = a_0 + a_1 x + \cdots + a_n x^n$, $a_j \in \mathbb{R}, j = 1, \cdots, n$. For every $\varepsilon > 0$, we can choose some $b_j \in \mathbb{Q}$ such that $|a_j - b_j| < \varepsilon/(n+1)$ for all j. It follows that for $q(x) = \sum_j b_j x^j \in S$, we have

$$\begin{aligned} |p(x) - q(x)| &\leq \sum_j |a_0 - b_0| + |a_1 - b_1||x| + \cdots + |a_n - b_n||x|^n \\ &< (n+1)\frac{\varepsilon}{2(n+1)} \\ &= \frac{\varepsilon}{2} \end{aligned}$$

for all x. We conclude that $\|p - q\|_\infty \leq \varepsilon/2$ Now, for any $f \in C[0,1]$ and $\varepsilon > 0$, we apply Weierstrass approximation theorem to obtain a polynomial p such that $\|f - p\|_\infty < \varepsilon/2$ and then find some $q \in \mathcal{S}$ such that $\|p - q\|_\infty \leq \varepsilon/2$. It follows that

$$\|f - q\|_\infty \leq \|f - p\|_\infty + \|p - q\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

that is, $\mathcal{S}$ is dense in $C[0,1]$.

Example 3.3. Let K be a closed, bounded set in $\mathbb{R}^n$. Then $C(K)$ is separable. This is a generalization of the previous example. The proof is again aided by considering the collection of all polynomials $\mathcal{P}$ in n-variables with rational coefficients. It is clear that this countable collection forms an algebra in $C(K)$. For the separating points property we observe that for every two distinct points x^0 and y^0 in $\mathbb{R}^n$ the polynomial $p(x) = \sum_{k=1}^n (x_k - x_k^0)^2$ satisfies $p(x^0) = 0$ and $p(y^0) \neq 0$. It is also nonvanishing because constant functions are polynomials. Henceforth, it is dense in $C(K)$ by Stone-Weierstrass theorem.

In fact, this example can be further generalized.

Proposition 3.7. *The space $C(X)$ is separable when X is a compact metric space.*

Proof. When X consists of a singleton, $C(X)$ is equal to $\mathbb{R}$ and so separable. We will always assume X has more than one points below. By Proposition 2.11, we can find a sequence of balls $\{B_j\}$ whose centers $\{z_j\}$ form a dense set in X. Define $f_j(x) = d(x, z_j)$ and let $\mathcal{M} \subset C(X)$ consist of functions which are finite product of f_j's. Then let $\mathcal{A}$ consist of functions of the form

$$f = \sum_{k=1}^N a_k h_k, \qquad h_k \in \mathcal{M}, \quad a_j \in \mathbb{Q}.$$

It is readily checked that $\mathcal{A}$ forms a subalgebra of $C(X)$. To verify separating points property let x_1 and x_2 be two distinct points in X. The function $f(x) = d(x, x_1)$ satisfies $f(x_1) = 0$ and $f(x_2) \neq 0$. By density, we can find some z_k close to x_1 so that the function $f_k(x) = d(x, z_k)$ separates x_1 and x_2. On the other hand, given any point x_0 we can fix another distinct point y_0 so that the function $d(x, y_0)$ is nonvanishing at x_0. By density again, there is some z_j close to y_0 such that $f_j(x_0) \neq 0$. By Stone-Weierstrass theorem, $\mathcal{A}$ is dense in $C(X)$. The theorem will be proved if we can show that $\mathcal{A}$ is countable. To see this, let $\mathcal{A}_n$ be the subset of $\mathcal{A}$ which only involves finitely many functions $f_1, \cdots, f_n$. We have $\mathcal{A} = \cup_n \mathcal{A}_n$, so it suffices to show each $\mathcal{A}_n$ is countable. Each function in $\mathcal{A}_n$ is composed of finitely many terms of the form $f_{n_1}^{a_1} \cdots f_{n_k}^{a_k},\ n_j \in \{1, \cdots, n\}$. Let $\mathcal{A}_n^m \subset \mathcal{A}_n$ consist of all those functions whose "degree" is less than or equal to m. It is clear that $\mathcal{A}_n^m$ is countable, so is $\mathcal{A}_n = \cup_m \mathcal{A}_n^m$. $\square$

To conclude this section, we note the existence of non-separable spaces. Here is one.

Example 3.4. Consider the space of all bounded functions on $[a, b]$ under the supnorm. It forms a metric space denoted by $B[a, b]$. We claim that it is not separable. For, let $f_z \in B[a, b]$ be given by $f_z(x) = 0$ for all $x \neq z$ and $f_z(z) = 1$. All f_z's form an uncountable set. Obviously the metric balls $B_{1/2}(f_z)$ are pairwise disjoint. If S is a dense subset of $B[a, b]$, $S \cap B_{1/2}(f_z)$ must be non-empty for each z. We pick $w_z \in S \cap B_{1/2}(f_z)$ to form an uncountable subset $\{w_z\}$ of S. We concldue that S must be uncountable, so there is no countable dense set of $B[a, b]$.

3.4 Compactness and Ascoli-Arzela Theorem

We pointed out that not every closed, bounded set in a metric space is compact. In Section 2.3 a bounded sequence without any convergent subsequence is explicitly displayed to show that a closed, bounded set in $C[a, b]$ needs not be compact. In view of numerous theoretic and practical applications, it is strongly desirable to give a characterization of compact sets in $C[a, b]$. The answer is given by the fundamental Arzela-Ascoli theorem. This theorem gives a necessary and sufficient condition when a closed and bounded set in $C[a, b]$ is compact. In order to have wider applications, we will work on a more general space $C(K)$, where K is a closed, bounded subset of $\mathbb{R}^n$, instead of $C[a, b]$. Recall that $C(K)$ is a complete, separable space under the sup-norm.

The crux for compactness for continuous functions lies on the notion of equicontinuity. Let X be a subset of $\mathbb{R}^n$. A subset $\mathcal{F}$ of $C(X)$ is **equicontinuous** if for every $\varepsilon > 0$, there exists some δ such that

$$|f(x) - f(y)| < \varepsilon, \quad \text{for all } f \in \mathcal{F}, \ \text{ and } \ |x - y| < \delta,\ x, y \in X.$$

Recall that a function is uniformly continuous in X if for each $\varepsilon > 0$, there exists some δ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta,\ x, y \in X$. So, equicontinuity means that δ can further be chosen independent of the functions in $\mathcal{F}$.

There are various ways to show that a family of functions is equicontinuous. Recall that a function f defined in a subset X of $\mathbb{R}^n$ is called Hölder continuous if there exists some $\alpha \in (0,1)$ such that

$$|f(x) - f(y)| \leq L|x-y|^\alpha, \quad \text{for all } x, y \in X, \tag{3.1}$$

for some constant L. The number α is called the Hölder exponent. The function is called Lipschitz continuous if (3.1) holds for α equals to 1. A family of functions $\mathcal{F}$ in $C(X)$ is said to satisfy a uniform Hölder or Lipschitz condition if all members in $\mathcal{F}$ are Hölder continuous with the same α and L or Lipschitz continuous and (3.1) holds for the same constant L. Clearly, such $\mathcal{F}$ is equicontinuous. The following situation is commonly encountered in the study of differential equations. The philosophy is that equicontinuity can be obtained if there is a good, uniform control on the derivatives of functions in $\mathcal{F}$.

Proposition 3.8. *Let $\mathcal{F}$ be a subset of $C(X)$ where X is a convex set in $\mathbb{R}^n$. Suppose that each function in $\mathcal{F}$ is differentiable and there is a uniform bound on the partial derivatives of these functions in $\mathcal{F}$. Then $\mathcal{F}$ is equicontinuous.*

Proof. For, x and y in X, $(1-t)x + ty$, $t \in [0,1]$, belongs to X by convexity. Let $\psi(t) \equiv f((1-t)x+ty)$. From the mean-value theorem

$$\psi(1) - \psi(0) = \psi'(t^*)(1-0),\ t^* \in [0,1],$$

for some mean value $t^* \in (0,1)$ and the chain rule

$$\psi'(t) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}((1-t)x+ty)(y_j - x_j),$$

we have

$$f(y) - f(x) = \psi(1) - \psi(0) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}((1-t^*)x+t^*y)(y_j - x_j).$$

Therefore,

$$|f(y) - f(x)| \leq \sqrt{n}M|y-x|,$$

where $M = \sup\{|\partial f/\partial x_j(x)| \ :\ x \in X,\ j = 1, \ldots, n,\ f \in \mathcal{F}\}$ after using Cauchy-Schwarz inequality. So $\mathcal{F}$ satisfies a uniform Lipschitz condition with Lipschitz constant $n^{1/2}M$. □

Theorem 3.9 (Arzela-Ascoli). *Let $\mathcal{F}$ be a closed set in $C(K)$ where K is a closed and bounded set in $\mathbb{R}^n$. Then $\mathcal{F}$ is compact if and only if it is bounded and equicontinuous.*

A set $\mathcal{E} \subset C(K)$ is bounded means it is contained in a ball, or, more specifically, there exists $M > 0$ such that

$$|f(x)| \leq M, \quad \text{for all } f \in \mathcal{E} \ \text{ and } x \in K.$$

We need the following lemma from elementary analysis.

Lemma 3.10. *Let $\{z_j,\ j \geq 1\}$ be a sequence in $\mathbb{R}^n$ and $\{f_n\}$ be a sequence of functions defined on $\{z_j,\ j \geq 1\}$. Suppose that for each j, there exists an M_j such that $|f_n(z_j)| \leq M_j$ for all $n \geq 1$. There is a subsequence of $\{f_n\}$, $\{g_n\}$, such that $\{g_n(z_j)\}$ is convergent for each j.*

Proof. Since $\{f_n(z_1)\}$ is a bounded sequence, we can extract a subsequence $\{f_n^1\}$ such that $\{f_n^1(z_1)\}$ is convergent. Next, as $\{f_n^1(z_2)\}$ is bounded, it has a subsequence $\{f_n^2\}$ such that $\{f_n^2(z_2)\}$ is convergent. Keep doing in this way, we obtain sequences $\{f_n^j\}$ satisfying (i) $\{f_n^{j+1}\}$ is a subsequence of $\{f_n^j\}$ and (ii) $\{f_n^j(z_1)\}, \{f_n^j(z_2)\}, \cdots, \{f_n^j(z_j)\}$ are convergent. Then the diagonal sequence $\{g_n\},\ g_n = f_n^n,$ for all $n \geq 1$, is a subsequence of $\{f_n\}$ which converges at every z_j. $\square$

The subsequence selected in this way is sometimes called Cantor's diagonal sequence.

Proof of Arzela-Ascoli Theorem. Assuming boundedness and equicontinuity of $\mathcal{F}$, we would like to show that $\mathcal{F}$ is compact.

Since K is compact in $\mathbb{R}^n$, it is totally bounded. By Proposition 2.11, for each $j \geq 1$, we can cover K by finitely many balls $B_{1/j}(x_1^j), \cdots, B_{1/j}(x_N^j)$ where the number N depends on j. All $\{x_k^j\},\ j \geq 1, 1 \leq k \leq N$, form a countable set. For any sequence $\{f_n\}$ in $\mathcal{F}$, by Lemma 3.11, we can pick a subsequence denoted by $\{g_n\}$ such that $\{g_n(x_k^j)\}$ is convergent for all x_k^j's. We claim that $\{g_n\}$ is a Cauchy sequence in $C(K)$. For, due to the equicontinuity of $\mathcal{F}$, for every $\varepsilon > 0$, there exists a δ such that $|g_n(x) - g_n(y)| < \varepsilon$, whenever $|x - y| < \delta$. Pick j_0, $1/j_0 < \delta$. Then for $x \in K$, there exists $x_k^{j_0}$ such that $|x - x_k^{j_0}| < 1/j_0 < \delta$,

$$\begin{aligned} |g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(x_k^{j_0})| + |g_n(x_k^{j_0}) - g_m(x_k^{j_0})| + |g_m(x_k^{j_0}) - g_m(x)| \\ &< \varepsilon + |g_n(x_k^{j_0}) - g_m(x_k^{j_0})| + \varepsilon. \end{aligned}$$

As $\{g_n(x_k^{j_0})\}$ converges, there exists n_0 such that

$$|g_n(x_k^{j_0}) - g_m(x_k^{j_0})| < \varepsilon, \quad \text{for all } n, m \geq n_0. \tag{3.2}$$

Here n_0 depends on $x_k^{j_0}$. As there are finitely many $x_k^{j_0}$'s, we can choose some N_0 such that (3.2) holds for all $x_k^{j_0}$ and $n, m \geq N_0$. It follows that

$$|g_n(x) - g_m(x)| < 3\varepsilon, \quad \text{for all } n, m \geq N_0,$$

i.e., $\{g_n\}$ is a Cauchy sequence in $C(K)$. By the completeness of $C(K)$ and the closedness of $\mathcal{F}$, $\{g_n\}$ converges to some function in $\mathcal{F}$.

Conversely, let $\mathcal{F}$ be compact. By Proposition 2.11, for each $\varepsilon > 0$, there exist $f_1, \cdots, f_N \in \mathcal{F}$ such that $\mathcal{F} \subset \bigcup_{j=1}^N B_\varepsilon(f_j)$ where N depends on ε. So for any $f \in \mathcal{F}$, there exists f_j such that

$$|f(x) - f_j(x)| < \varepsilon, \quad \text{for all } x \in K.$$

As each f_j is continuous, there exists δ_j such that $|f_j(x) - f_j(y)| < \varepsilon$ whenever $|x-y| < \delta_j$. Letting $\delta = \min\{\delta_1, \cdots, \delta_N\}$, then

$$|f(x) - f(y)| \leq |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_j(y) - f(y)| < 3\varepsilon,$$

for $|x - y| < \delta$, so S is equicontinuous. As $\mathcal{F}$ can be covered by finitely many balls of radius 1, it is also bounded. We have completed the proof of Arzela-Ascoli theorem.

□

The following special case of Arzela-Ascoli theorem, sometimes called Ascoli's theorem, is most useful in applications.

Proposition 3.11. *A sequence in $C(K)$ where K is a closed, bounded set in $\mathbb{R}^n$ has a convergent subsequence if it is uniformly bounded and equicontinuous.*

Proof. Let $\mathcal{F}$ be the closure of the sequence $\{f_n\}$. We would like to show that $\mathcal{F}$ is bounded and equicontinuous. First of all, by the uniform boundedness assumption, there is some M such that

$$|f_n(x)| \leq M, \quad \forall x \in K, \ n \geq 1.$$

As every function in $\mathcal{F}$ is either one of these f_n or the limit of its subsequence, it also satisfies this estimate, so $\mathcal{F}$ is bounded in $C(K)$. On the other hand, by equicontinuity, for every $\varepsilon > 0$, there exists some δ such that

$$|f_n(x) - f_n(y)| < \frac{\varepsilon}{2}, \quad \forall x, y \in K, \ |x - y| < \delta.$$

As every $f \in \mathcal{F}$ is the limit of a subsequence of $\{f_n\}$, f satisfies

$$|f(x) - f(y)| \leq \frac{\varepsilon}{2} < \varepsilon, \quad \forall x, y \in K, \ |x - y| < \delta,$$

so $\mathcal{F}$ is also equicontinuous. Now the conclusion follows from the Arzela-Ascoli theorem.

□

We present an application of Arzela-Ascoli theorem to ordinary differential equations. Consider the Cauchy problem (2.3) again,

$$\begin{cases} \dfrac{dy}{dx} = f(x,y), \\ y(x_0) = y_0. \end{cases} \tag{2.3}$$

where f is a continuous function defined in the rectangle $R = [x_0-a, x_0+a] \times [y_0-b, y_0+b]$. In Chapter 2 we proved that this Cauchy problem has a unique solution when f satisfies the Lipschitz condition. Now we show that the existence part of Picard-Lindelöf theoorem is still valid without the Lipschitiz condition.

Theorem 3.12 (**Cauchy-Peano Theorem**). *Consider (2.3) where f is continuous on $R = [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]$. There exist $a' \in (0, a)$ and a C^1-function $y(x)$: $[x_0 - a', x_0 + a'] \to [y_0 - b, y_0 + b]$, solving (2.3).*

From the proof we will see that a' can be taken to be $0 < a' < \min\{a, b/M\}$ where $M = \sup\{|f(x, y)| : (x, y) \in R\}$. The theorem is also valid for systems.

Proof. Recalling in the proof of Picard-Lindelöf theorem we showed that under the Lipschitz condition the unique solution exists on the interval $[x_0 - a', x_0 + a']$ where $0 < a' < \min\{a, b/M, 1/L^*\}$ where L^* is the Lipschitz constant. Let us first argue that the maximal solution in fact exists in the interval $[x_0 - a', x_0 + a']$ where $0 < a' < \min\{a, b/M\}$. In other words, the Lipschitz condition does not play any role in the range of existence. Although this was done in the exercise, we include it here for the sake of completeness.

Take $x_0 = y_0 = 0$ to simplify notations. The functions $w(x) = Mx$ and $z(x) = -Mx$ satisfy $y' = \pm M$, $y(0) = 0$, respectively. By comparing them with y, our maximal solution to (2.3), we have $z(x) \le y(x) \le w(x)$ as long as y exists. In case y exists on $[0, \alpha)$ for some $\alpha < \min\{a, b/M\}$, $(x, y(x))$ would be confined in the triangle bounded by $y = Mx, y = -Mx$, and $x = \alpha$. As this triangle is compactly contained in the interior of R, the Lipschitz constant ensures that the solution could be extended beyond α. Thus the solution exists up to $\min\{a, b/M\}$. Similarly, one can show that the solution exists in $(-\min\{a, b/M\}, 0]$.

With this improvement at our disposal, we prove the theorem. First, of all, by Weierstrass approximation theorem, there exists a sequence of polynomials $\{f_n\}$ approaching f in $C([-a, a] \times [-b, b])$ uniformly. In particular, it means that $M_n \to M$, where $M_n = \max\{|f_n(x, y)| : (x, y) \in [-a, a] \times [-b, b]$. As each f_n satisfies the Lipschitz condition (why?), there is a unique solution y_n defined on $I_n = (-a_n, a_n), a_n = \min\{a, b/M_n\}$ for the initial value problem

$$\frac{dy}{dx} = f_n(x, y), \quad y(0) = 0.$$

From $|dy_n/dx| \le M_n$ and $\lim_{n\to\infty} M_n = M$, we know from Proposition 3.8 that $\{y_n\}$ forms an equicontinuous family. Clearly, it is also bounded. By Ascoli's theorem, it contains a subsequence $\{y_{n_j}\}$ converging uniformly to a continuous function $y \in I$ on every subinterval $[\alpha, \beta]$ of I and $y(0) = 0$ holds. It remains to check that y solves the differential equation for f.

Indeed, each y_n satisfies the integral equation

$$y_n(x) = \int_0^x f(t, y_n(t))dt, \quad x \in I_n.$$

As $\{y_{n_j}\} \to y$ uniformly, $\{f(x, y_{n_j}(x))\}$ also tends to $f(x, y(x))$ uniformly. By passing to limit in the formula above, we conclude that

$$y(x) = \int_0^x f(t, y(t))dt, \quad x \in I$$

holds. By the fundamental theorem of calculus, y is differentiable and a solution to (2.3). □

3.5 Completeness and Baire Category Theorem

In this section we discuss Baire category theorem, a basic property of complete metric spaces. It is concerned with the decomposition of a metric space into a countable union of subsets. The motivation is somehow a bit strange at first glance. For instance, we can decompose the plane $\mathbb{R}^2$ as the union of strips $\mathbb{R}^2 = \bigcup_{k\in\mathbb{Z}} S_k$ where $S_k = (k, k+1] \times \mathbb{R}$. In this decomposition each S_k is not so sharply different from $\mathbb{R}^2$. Aside from the boundary, the interior of each S_k is just like the interior of $\mathbb{R}^2$. On the other hand, one can make the more extreme decomposition: $\mathbb{R}^2 = \bigcup_{\alpha\in\mathbb{R}} l_\alpha$ where $l_\alpha = \{\alpha\} \times \mathbb{R}$. Each l_α is a vertical straight line and is very different from $\mathbb{R}^2$. It is simpler in the sense that it is one-dimensional and has no area. The sacrifice is now we need an uncountable union. The question is: Can we represent $\mathbb{R}^2$ as a countable union of these sets (or sets with lower dimension)? It turns out that the answer is no. The obstruction comes from the completeness of the ambient space.

We need one definition. Let (X, d) be a metric space. A subset E of X is called **nowhere dense** if its closure does not contain any metric ball. Equivalently, E is nowhere dense if $X \setminus \overline{E}$ is dense in X. Note that a set is nowhere dense if and only if its closure is nowhere dense. Also every subset of a nowhere dense set is nowhere dense.

Theorem 3.13 (**Baire Category Theorem**). *Let $\{E_j\}_1^\infty$ be a sequence of nowhere dense subsets of (X, d) where (X, d) is complete. Then $\bigcup_{j=1}^\infty \overline{E_j}$ has empty interior.*

A set with empty interior means that it does not contain any ball. It is so if and only if its complement is a dense set.

Proof. Replacing E_j by its closure if necessary, we may assume all E_j's are closed sets. Let B_0 be any ball. The theorem will be established if we can show that $B_0 \bigcap (X \setminus \bigcup_j E_j) \neq \phi$. As E_1 is nowhere dense, there exists some point $x \in B_0$ lying outside E_1. Since E_1 is closed , we can find a closed ball $\overline{B}_1 \subset B_0$ centering at x such that $\overline{B}_1 \cap E_1 = \phi$ and its diameter $d_1 \leq d_0/2$, where d_0 is the diameter of B_0. Next, as E_2 is nowhere dense and closed, by the same reason there is a closed ball $\overline{B}_2 \subset B_1$ such that $\overline{B}_2 \cap E_2 = \phi$ and $d_2 \leq d_1/2$. Repeating this process, we obtain a sequence of closed balls $\overline{B}_j$ satisfying (a) $\overline{B}_{j+1} \subset B_j$, (b) $d_j \leq d_0/2^j$, and (c) $\overline{B}_j$ is disjoint from $E_1, \cdots, E_j$. Pick x_j from $\overline{B}_j$ to form a sequence $\{x_j\}$. As the diameters of the balls tend to zero, $\{x_j\}$ is a Cauchy sequence. By the completeness of X, $\{x_j\}$ converges to some x^*. Clearly x^* belongs to all $\overline{B}_j$. If x^* belongs to $\bigcup_j E_j$, x^* belongs to some E_{j_1}, but then $x^* \in \overline{B}_{j_1} \bigcap E_{j_1}$ which means that $\overline{B}_{j_1}$ is not disjoint from E_{j_1}, contradiction holds. We conclude that $B_0 \bigcap (X \setminus \bigcup_j E_j) \neq \phi$. $\square$

Some remarks are in order.

First, taking complement in the statement of the theorem, it asserts that the intersection of countably many open, dense sets is again a dense set. Be careful it may not be open. For example, let $\{q_j\}$ be the set of all rational numbers in $\mathbb{R}$ and $D_k = \mathbb{R} \setminus \{q_j\}_{j=1}^k$. Each D_k is an open, dense set. However, $\bigcap_k D_k = \mathbb{R} \setminus \mathbb{Q}$ is the set of all irrational numbers. Although it is dense, it is not open any more.

Second, that the set $\bigcup_j \overline{E}_j$ has no interior in particular implies $X \setminus \bigcup_j \overline{E}_j$ is nonempty, that is, it is impossible to decompose a complete metric space into a countable union of nowhere dense subsets.

Third, the above remark may be formulated as, if X is complete and $X = \bigcup_j A_j$ where A_j are closed, then one of the A_j's must contain a ball.

When we describe the size of a set in a metric space, we could use the notion of a dense set or a nowhere dense set. However, sometimes some description is too rough. For instance, consider $\mathbb{Q}$ and $\mathbb{I}$, the set of rational numbers and irrational numbers respectively. Both of them are dense in $\mathbb{R}$. However, everyone should agree that $\mathbb{Q}$ and $\mathcal{I}$ are very different. The former is countable but the latter is uncountable. From a measure-theoretic point of view, $\mathbb{Q}$ is a set of measure zero and yet $\mathbb{I}$ has infinite measure. Thus simply calling them dense sets is not precise enough. Baire category theorem enables us to make a more precise description of the size of a set in a complete metric space. A set in a metric space is called **of first category** if it can be expressed as a countable union of nonwhere dense sets. Note that by definition any subset of a set of first category is again of first category. A set is **of second category** if its complement is of first category. According to Baire category theorem, a set of second category is a dense set when the underlying space is complete.

Proposition 3.14. *If a set in a complete metric space is of first category, it cannot be of second category, and vice versa.*

Proof. Let E be of first category and let $E = \cup_{k=1}^{\infty} E_k$ where E_k are nowhere dense sets. If it is also of second category, its complement is of first category. Thus, $X \setminus E = \cup_{k=1}^{\infty} F_k$ where F_k are nowhere dense. It follows that $X = E \cup (X \setminus E) = \cup_k (E_k \cup F_k)$ so the entire space is a countable union on of nowhere dense sets, contradicting the completeness of the space and the Baire category theorem. $\square$

Applying to $\mathbb{R}$, we see that $\mathbb{Q}$ is of first category and $\mathbb{I}$ is of second category although they both are dense sets. Indeed, $\mathbb{Q} = \cup_k \{x_k\}$ where k runs through all rational numbers and $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$ is of second category.

Baire category theorem has many applications. We end this section by giving a standard one. It is concerned with the existence of continuous, but nowhere differentiable functions. We knew that Weierstrass is the first person who constructed such a function in 1896. His function is explicitly given in the form of an infinite series

$$W(x) = \sum_{n=1}^{\infty} \frac{\cos 3^n x}{2^n}.$$

Here we use an implicit argument to show there are far more such functions than continuously differentiable functions.

We begin with a lemma.

Lemma 3.15. *Let $f \in C[a,b]$ be differentiable at x. Then it is Lipschitz continuous at x.*

Proof. By differentiability, for $\varepsilon = 1$, there exists some δ_0 such that

$$\left| \frac{f(y) - f(x)}{y - x} - f'(x) \right| < 1, \quad \forall y \neq x, |y - x| < \delta_0.$$

We have

$$|f(y) - f(x)| \le L|y - x|, \quad \forall y, |y - x| < \delta_0,$$

where $L = |f'(x)| + 1$. For y lying outside $(x - \delta_0, x + \delta_0)$, $|y - x| \ge \delta_0$. Hence

$$\begin{aligned} |f(y) - f(x)| &\le |f(x)| + |f(y)| \\ &\le \frac{2M}{\delta_0}|y - x|, \quad \forall y \in [a,b] \setminus (x - \delta_0, x + \delta_0), \end{aligned}$$

where $M = \sup\{|f(x)| : x \in [a,b]\}$. It follows that f is Lipschitz continuous at x with an Lipschitz constant not exceeding $\max\{L, 2M/\delta_0\}$.

$\square$

Proposition 3.16. *The set of all continuous, nowhere differentiable functions forms a set of second category in $C[a,b]$ and hence dense in $C[a,b]$.*

Proof. For each $L > 0$, define

$S_L = \{ f \in C[a,b] : f$ is Lipschitz continuous at some x with the Lipschitz constant $\le L \}$.

We claim that S_L is a closed set. For, let $\{f_n\}$ be a sequence in S_L which is Lipschitz continuous at x_n and converges uniformly to f. By passing to a subsequence if necessary, we may assume $\{x_n\}$ to some x^* in $[a,b]$. We have, by letting $n \to \infty$,

$$\begin{aligned}
|f(y) - f(x^*)| &\le |f(y) - f_n(y)| + |f_n(y) - f(x^*)| \\
&\le |f(y) - f_n(y)| + |f_n(y) - f_n(x_n)| + |f_n(x_n) - f_n(x^*)| + |f_n(x^*) - f(x^*)| \\
&\le |f(y) - f_n(y)| + L|y - x_n| + L|x_n - x^*| + |f_n(x^*) - f(x^*)| \\
&\to L|y - x^*|
\end{aligned}$$

Next we show that each S_L is nowhere dense. Let $f \in S_L$. By Weierstrass approximation theorem, for every $\varepsilon > 0$, we can find some polynomial p such that $\|f - p\|_\infty < \varepsilon/2$. Since every polynomial is Lipschitz continuous, let the Lipschitz constant of p be L_1. Consider the function $g(x) = p(x) + (\varepsilon/2)\varphi(x)$ where φ is the jig-saw function of period r satisfying $0 \le \varphi \le 1$ and $\varphi(0) = 1$. The slope of this function is either $1/r$ or $-1/r$. Both will become large when r is chosen to be small. Clearly, we have $\|f - g\|_\infty < \varepsilon$. On the other hand,

$$\begin{aligned}
|g(y) - g(x)| &\ge \frac{\varepsilon}{2}|\varphi(y) - \varphi(x)| - |p(y) - p(x)| \\
&\ge \frac{\varepsilon}{2}|\varphi(y) - \varphi(x)| - L_1|y - x| \ .
\end{aligned}$$

For each x sitting in $[a,b]$, we can always find some y nearby so that the slope of φ over the line segment between x and y is greater than $1/r$ or less than $-1/r$. Therefore, if we choose r so that

$$\frac{\varepsilon}{2}\frac{1}{r} - L_1 > L,$$

we have $|g(y) - g(x)| > L|y - x|$, that is, g does not belong to S_L.

Denoting by S the set of functions in $C[a,b]$ which are differentiable at at least one point, by Lemma 3.15 it must belong to S_L for some L. Therefore, $S \subset \cup_{k=1}^\infty S_k$ is of first category. □

Though elegant, a drawback of this proof is that one cannot assert which particular function is nowhere differentiable. On the other hand, the example of Weierstrass is a concrete one.

Comments on Chapter 3. Three properties, namely, separability, compactness, and completeness of the space of continuous functions are studied in this chapter.

Separability is established by various approximation theorems. For the space $C[a, b]$, Weierstrass approximation theorem is applied. Weierstrass (1885) proved his approximation theorem based on the heat kernel, that is, replacing the kernel Q_n in our proof in Section 1 by the heat kernel. The argument is a bit more complicated but essentially the same. It is taken from Rudin, Principles of Mathematical Analysis. A proof by Fourier series is already presented in Chapter 1. Another standard proof is due to Bernstein, which is constructive and had initiated a branch of analysis called approximation theory. The Stone-Weierstrass theorem is due to M.H. Stone (1937, 1948). We use it to establish the separability of the space $C(X)$ where X is a compact metric space. You can find more approximation theorem by web-surfing.

Arzela-Ascoli theorem plays the role in the space of continuous functions the same as Bolzano-Weierstrass theorem does in the Euclidean space. A bounded sequence of real numbers always admits a convergent subsequence. Although this is no longer true for bounded sequences of continuous functions on $[a, b]$, it does hold when the sequence is also equicontinuous. Ascoli's theorem (Proposition 3.11) is widely applied in the theory of partial differential equations, the calculus of variations, complex analysis and differential geometry. Here is a taste of how it works for a minimization problem. Consider

$$\inf \left\{ J[u] : \ u(0) = 0, u(1) = 5, \ u \in C^1[0, 1] \right\},$$

where

$$J[u] = \int_0^1 \Big(u'^2(x) - \cos u(x) \Big) \, dx.$$

First of all, we observe that $J[u] \geq -1$. This is clear, for the cosine function is always bounded by ± 1. After knowing that this problem is bounded from -1, we see that $\inf J[u]$ must be a finite number, say, γ. Next we pick a minimizing sequence $\{u_n\}$, that is, every u_n is in $C^1[0, 1]$ and satisfies $u(0) = 0, u(1) = 5$, such that $J[u_n] \to \gamma$ as $n \to \infty$. By Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\big|u_n(x) - u_n(y)\big| &\leq \int_x^y \big|u_n'(x)\big| dx \\
&\leq \sqrt{\int_x^y 1^2 dx} \sqrt{\int_x^y u_n'^2(x) dx} \\
&\leq \sqrt{\int_x^y 1^2 dx} \sqrt{\int_0^1 u_n'^2(x) dx} \\
&\leq \sqrt{J[u_n] + 1} \sqrt{|y - x|} \\
&\leq \sqrt{\gamma + 2} \ |y - x|^{1/2}
\end{aligned}$$

for all large n. From this estimate we immediately see that $\{u_n\}$ is equicontinuous and bounded (because $u_n(0) = 0$). By Ascoli's theorem, it has a subsequence $\{u_{n_j}\}$ converging to some $u \in C[0, 1]$. Apparently, $u(0) = 0, u(1) = 5$. Using knowledge from functional analysis, one can further argue that $u \in C^1[0, 1]$ and is the minimum of this problem.

Arzela showed the necessity of equicontinuity and boundedness for compactness while Ascoli established the compactness under equicontinuity and boundedness. Google under Arzela-Ascoli theorem for details.

There are some fundamental results that require completeness. The contraction mapping principle is one and Baire category theorem is another. The latter was first introduced by Baire in his 1899 doctoral thesis. It has wide, and very often amazing applications in all branches of analysis. Some nice applications are available on the web. Google under applications of Baire category theorem for more.

Weierstrass' example is discussed in some detailed in Hewitt-Stromberg, "Abstract Analysis". An simpler example can be found in Rudin's Principles.

Being unable to locate a single reference containing these three topics, I decide not to name any reference but let you search through the internet.