Chapter 4

The Space of Continuous Functions

Two fundamental results concerning the space of continuous functions are present. In Section 1 we characterize precompact sets in the space of continuous functions, and, as an application, Cauchy-Peano Theorem on the existence of the initial value problem for differential equations is derived. In Section 2 the notions of first and second category are introduced and then Baire Category Theorem is proved. As an application, it is shown that there are many continuous, but nowhere differentiable functions.

4.1 Ascoli's Theorem

Recall that for any metric space (X, d), the space of all bounded, continuous functions $C_b(X)$ forms a complete metric space under the supnorm. In this section we will consider X being $\overline{G}$, where is G be a bounded, open set in $\mathbb{R}^n$. Note that every continuous function in $\overline{G}$ is uniformly continuous and, in particular, bounded. This space shares many common features with its one-dimensional special case $C[a, b]$.

Unlike the Euclidean space, $C(\overline{G})$ is an infinite dimensional vector space. A classical result is Bolzano-Weierstrass Theorem asserting that every sequence in a bounded set in the Euclidean space contains a convergent subsequence. This property no longer holds for our space. Examples are bountiful, for instance, consider the unit ball in $C[0, 1]$ which is clearly bounded. It contains the sequence $\{f_n\}$, $f_n(x) = x^n$, $x \in [0, 1]$ whose pointwise limit exists and is equal to the function $f(x) = 0$, $x \neq 1$ and $f(1) = 1$. Since f is not continuous, this sequence cannot have any convergent sequence in supnorm. Bolzano-Weierstrass Theorem does not hold in $C[0, 1]$.

To proceed further, let us introduce some basic definitions. A set E in a metric space is called a **precompact** set if its closure is compact. Thus every sequence in a precom-

pact set E contains a convergent subsequence whose limit may or may not be in E. A precompact set is compact iff it is closed. Bolzano-Weierstrass Theorem tells us that a set in $\mathbb{R}^n$ is precompact if and only if it is bounded, and a set is compact if and only if it is closed and bounded. On the other hand, as we have seen, only boundedness is not sufficient to ensure precompactness for sets in $C(\overline{G})$. In this section we will characterize these sets.

The crux is the notion of equicontinuity. A set $\mathcal{C}$ in $C(\overline{G})$ is **equicontinuous** if for every $\varepsilon > 0$, there exists some δ such that

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

Recall that a function is uniformly continuous in $\overline{G}$ if for each $\varepsilon > 0$, there exists some δ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$, $x, y \in \overline{G}$. Equicontinuity means that δ can further be chosen to fit all functions in $\mathcal{C}$. Applying to any single function in the set, it implies that each element in $\mathcal{C}$ is uniformly continuous.

There are various ways to establish equicontinuity of a given set. Recall that a function f defined in a subset $\overline{G}$ of $\mathbb{R}^n$ is called Hölder continuous if there exists some $\alpha \in (0, 1)$ such that

$$|f(x) - f(y)| \leq L|x - y|^\alpha, \quad \text{for all } x, y \in \overline{G}, \tag{4.1}$$

for some constant L. The number α is called the Hölder exponent. The function is called Lipschitz continuous if (4.1) holds for $\alpha = 1$. The set $\mathcal{C}$ in $C(\overline{G})$ is said to satisfy a **uniform Hölder** or **Lipschitz condition** if all f's are Hölder continuous with the same α and L or Lipschitz continuous and (4.1) holds for the same constant L. Clearly, such a set is equicontinuous. In fact, for any $\varepsilon > 0$, any δ satisfying $L\delta^\alpha < \varepsilon$ can do the job. 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 the functions.

Proposition 4.1. *Let $\mathcal{C}$ be in $C(\overline{G})$ where $\overline{G}$ is convex in $\mathbb{R}^n$. Suppose that each $f \in \mathcal{C}$ is differentiable and there is a uniform bound on their partial derivatives. Then $\mathcal{C}$ is equicontinuous.*

Proof. We repeat an old argument. For, x and y in X, $(1 - t)x + ty$, $t \in [0, 1]$, belongs to $\overline{G}$ by convexity. Let $\psi(t) \equiv f((1 - t)x + ty)$, $f \in \mathcal{C}$. By 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

$$
\begin{aligned}
f(y) - f(x) &= \psi(1) - \psi(0) \\
&= \int_0^1 \psi'(t)dt \\
&= \sum_{j=1}^n \int_0^1 \frac{\partial f}{\partial x_j}(x + t(y-x))(y_j - x_j).
\end{aligned}
$$

Therefore,

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

where $M = \sup\{|\partial f/\partial x_j(x)| : x \in X, \ j = 1,\dots,n,\}$ after using Cauchy-Schwarz Inequality. We conclude that $\mathcal{C}$ satisfies a uniform Lipschitz condition with Lipschitz constant $n^{1/2}M$.

$\square$

Theorem 4.2 (Ascoli's Theorem). *Consider $C(\overline{G})$ where G is bounded, open in $\mathbb{R}^n$. A set $\mathcal{E}$ in $C(\overline{G})$ is precompact if it is bounded and equicontinuous.*

We need the following useful lemma from elementary analysis.

Lemma 4.3. *Let A be a countable set and $\{f_n\}$ be a sequence of real-valued functions defined on A. Suppose that for each $z \in A$, there exists an M such that $|f_n(z)| \le M$ for all $n \ge 1$. There is a subsequence of $\{f_n\}$, $\{f_{n_k}\}$, such that $\{f_{n_k}(z)\}$ is convergent at each $z \in A$.*

Proof. Let $A = \{z_j\}, j \ge 1$. 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 $n \ge 1$, is a subsequence of $\{f_n\}$ which converges at every z_j. $\square$

The subsequence selected in this way is usually called a Cantor's diagonal sequence. It first came up in Cantor's study of infinite sets.

Proof of Ascoli's Theorem. Let $\{f_n\}$ be a sequence in $\mathcal{E}$. For each $k \ge 0$, consider the set

$$
E_k \equiv \left\{ x = (x_1, \cdots, x_n) \in \overline{G} : x_j = \frac{m}{2^k}, \ m \in \mathbb{Z}, \ j = 1, \cdots, n \right\},
$$

and

$$
E = \bigcup_k E_k .
$$

Each E_k is a finite set and hence E is a countable set $\{z_j\}$. By assumption, there is some M satisfying $|f_n(x)| \leq M$ for all $x \in \overline{G}$. Using the lemma above, we can pick a subsequence from $\{f_n\}$, $\{g_n\}$, such that $\{g_n(z_j)\}$ is convergent for each z_j's. We claim that $\{g_n\}$ is a Cauchy sequence in $C(\overline{G})$. For, by equicontinuity, for every $\varepsilon > 0$, there exists a δ such that $|g_n(x) - g_n(y)| < \varepsilon/3$, whenever $|x - y| < \delta$, $x, y \in \overline{G}$. We fix k so that $\sqrt{n}/2^{k+1} < \delta$. For each $x \in \overline{G}$, it must belong to one of the cubes formed by the lattice points. Hence there exists some lattice point z_j satisfying $|x - z_j| \leq \sqrt{n}/2^{k+1} < \delta$ (see below). We have

$$
\begin{aligned}
|g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(z_j)| + |g_n(z_j) - g_m(z_j)| + |g_m(z_j) - g_m(x)| \\
&< \frac{\varepsilon}{3} + |g_n(z_j) - g_m(z_j)| + \frac{\varepsilon}{3}.
\end{aligned}
$$

As $\{g_n(z_j)\}$ converges, there exists some $n_0(j)$ (depending on z_j) such that

$$
|g_n(z_j) - g_m(z_j)| < \frac{\varepsilon}{3}, \quad \text{for all } n, m \geq n_0(j). \tag{4.2}
$$

As E_k is a finite set, Letting $n_0 = \max_j\{n_0(j)\}$, (4.2) holds for all z_j and $n, m \geq n_0$. It follows that

$$
|g_n(x) - g_m(x)| < \varepsilon, \quad \text{for all } n, m \geq n_0,
$$

i.e., $\{g_n\}$ is a Cauchy sequence in $C(\overline{G})$. By the completeness of $C(\overline{G})$, $\{g_n\}$ converges uniformly to some function in $C(\overline{G})$. We have completed the proof of Ascoli's Theorem.

For a cube of side $1/2^k$ in $\mathbb{R}^n$, the length of its diagonal is $\sqrt{n}/2^k$ by Pythagoras' theorem. Now, the union of all lattice cubes with vertices in E_k covers $\overline{G}$. Each point x belongs to some of these cubes and the distance from x to the nearest vertices cannot be greater than half of $\sqrt{n}/2^k$. That is, $|x - z_j| \leq \sqrt{n}/2^{k+1}$.

We examine the assumptions in Ascoli's theorem in the following examples.

Example 4.1. First of all, the boundedness of the class of functions. This is easy to see. Fix some $f_0 \in C(\overline{G})$ and let $\mathcal{E}_1 = \{f_0 + n : n \in \mathbb{N}\}$. Then $\mathcal{E}_1$ consists of the sequence obtained by shifting f_0 upward to infinity. Clearly it is equicontinuous, but cannot have any convergent subsequence.

Example 4.2. Let $\mathcal{E}_2 = \{f_n(x) = x^n, x \in [0,1]\}$. This set is bounded in $C[0,1]$. $\{f_n\}$ has a pointwise limit given by $g(x) = 0, x \in [0,1)$ and $g(1) = 1$ which is not continuous at $x = 1$. Since uniform convergence implies pointwise convergence, this shows that any subsequence of $\{f_n\}$ must be divergent, since it does not have a limit in $C[0,1]$. The set $\mathcal{E}_2$ is therefore not equicontinuous in view of Ascoli's theorem. A direct proof runs as follows. Suppose on the contrary it is, then for $\varepsilon = 1/2$, there is δ such that $|x^n - y^n| < 1/2$ whenever $|x - y| < \delta$ for all n. Taking $y = 1$, we have $|x^n - 1| < 1/2$ for all n and $|x - 1| < \delta, x \in [0,1]$. Fixing some $x \in (0,1)$ and letting $n \to \infty$, we have $|-1| < 1/2$,

contradiction holds.

Example 4.3. This example shows that the underlying space G must be bounded. Let φ be a non-zero, C^1-function in $[0,1]$ which vanishes outside $[1/2, 3/4]$. Define a sequence $\{f_n\}$ in $C[0,\infty)$ by setting $f_n(x) = \varphi(x-n), x \in [n, n+1]$ and $= 0$ elsewhere. Then $\|f_n\|_\infty = \|\varphi\|_\infty > 0$ is bounded and clearly it is also equicontinuous. But it has no convergent subsequence. For, suppose some $\{f_{n_j}\}$ converges to f in supnorm (that is, uniformly convergent). It must also be pointwisely convergent, $f_{n_j}(x) \to f(x)$ for all x. As $f_n(x) = 0$ for all $n \geq x+1$, $f(x) \equiv 0$. But then, $\|\varphi\|_\infty = \|f_{n_j}\|_\infty = \|f_{n_j} - f\|_\infty \to 0$, which is impossible.

The following result provides a converse to Ascoli's theorem.

Theorem 4.4 (Arzela's Theorem). *Every precompact set in $C(\overline{G})$ must be bounded and equicontinuous.*

Proof. Let $\mathcal{E}$ be precompact. If $\mathcal{E}$ is unbounded, there exists some sequence $\{f_n\}$ in $\mathcal{E}$ satisfying $\sup_n \|f_n\| = \infty$. Pick a subsequence $\{f_{n_j}\}$ satisfying $\|f_{n_j}\| \to \infty$ as $n_j \to \infty$. Clearly this subsequence cannot contain any convergent subsequence.

Next, assume that $\mathcal{E}$ is bounded and precompact but not equicontinuous. We are going to draw a contradiction. First of all, there is some $\varepsilon_0 > 0$ such that, for each $n \geq 1$, we can find f_n, $x_n, y_n \in \overline{G}$ satisfying

$$|f_n(x_n) - f_n(y_n)| \geq \varepsilon_0, \quad |x_n - y_n| < \frac{1}{n} .$$

By precompactness, $\{f_n\}$ contains a convergent subsequence, let it be $\{f_{n_j}\}$. Assume that $f_{n_j} \rightrightarrows f \in C(\overline{G})$ as $n_j \to \infty$. The sequence $\{x_{n_j}\} \subset \overline{G}$ contains a convergent subsequence $x_{n_{j_k}} \to z \in \overline{G}$. To simplify notation, we let $g_k = f_{n_{j_k}}$ and $x'_k = x_{n_{j_k}}, y'_k = y_{n_{j_k}}$. For $\varepsilon > 0$, there is some n_0 such that

$$
\begin{aligned}
|g_k(x'_k) - g_k(y'_k)| &\leq\; |g_k(x'_k) - f(x'_k)| + |f(x'_k) - f(y'_k)| + |f(y'_k) - g_k(y'_k)| \\
&\leq\; \|g_k - f\|_\infty + |f(x'_k) - f(y'_k)| + \|g_k - f\|_\infty \\
&<\; \varepsilon + |f(x'_k) - f(y'_k)| + \varepsilon , \quad \forall n_{j_k} \geq n_0 .
\end{aligned}
$$

As $x'_k \to z$ and $|x'_k - y'_k| < 1/n_{j_k} \to 0$, $y'_k \to z$ too. Therefore, there exists some n_1 such that $|f(x'_k) - f(y'_k)| \leq |f(x'_k) - f(z)| + |f(z) - f(y'_k)| < \varepsilon$, for all $n \geq n_1$. It follows that

$$|g_k(x'_k) - g_k(y'_k)| < 3\varepsilon ,$$

for all $n_{j_k} \geq \max\{n_0, n_1\}$. We get a contradiction when $\varepsilon < \varepsilon_0/3$. Hence, $\mathcal{E}$ cannot be equicontinuous. $\qquad \square$

We present an application of Ascoli's Theorem to ordinary differential equations. Consider the initial value problem again,

$$\begin{cases} \dfrac{dx}{dt} = f(t,x), \\[2mm] x(t_0) = x_0. \end{cases} \tag{IVP}$$

where f is a continuous function defined in the rectangle $R = [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$. In Chapter 3 we proved that this Cauchy problem has a *unique* solution when f satisfies the Lipschitz condition in a smaller rectangle. Furthermore, Proposition 3.14 asserts the existence of a maximal solution defined on a maximal interval.

Proposition 4.5. *Consider the (IVP) where $f \in C(\overline{G}), G$ a bounded, open set in $\mathbb{R}^2$, and satisfies a Lipschitz condition, and $(t_0, x_0) \in G$. Let x be the maximal solution of this (IVP) defined on the maximal interval (α, β). Then there exists some $(\beta, z) \in \partial G$ such that $x(t) \to z$ as $t \to \beta^+$.*

Proof. The solution curve $\{(t, x(t)) : t \in [t_0, \beta)\}$ is contained in the bounded set G. Bolzano-Weierstrass theorem asserts that any bounded sequence has a convergent subsequence. Hence we can find a sequence $(t_n, x(t_n))$ and $(\beta, z) \in \overline{G}$ such that $(t_n, x(t_n)) \to (\beta, z)$, that is, $x(t_n) \to z$ as $t_n \to \beta^+$. We first claim: $x(t) \to z, \ t \to \beta^+$. For, from the formula

$$x(t) = x(t_n) + \int_{t_n}^{t} f(s, x(s)) \, ds \ ,$$

we have

$$|x(t) - x(t_n)| = |\int_{t_n}^{t} f(s, x(s)) \, ds| \le M|t - t_n| \ .$$

Letting $n \to \infty$, we arrive at $|x(t) - z| \le M|t - \beta|$.

Next, we claim that $(\beta, z) \in \partial G$. Assume on the contrary that (β, z) sits in the interior of G. We can find a small rectangle $R = [\beta - a_1, \beta + a_1, z_1 - b_1, z + b_1] \in G$. For those t_1 sufficiently close to β, the rectangle $R' = [t_1 - a_1/2, t_1 + a_1/2, x(t_1) - b_1/2, x(t_1) + b_1/2]$ is contained in R. By solving the (IVP) taking $(t_1, x(t_1))$ as initial data, the solution $x(t)$ can be extended up to $\gamma = t_0 + \min\{a_1/2, b_1/2M, 1/L\}$ which is independent of t_1. When $\beta - t_1 < \gamma$, it means that $x(t)$ exists beyond β, contradicting the maximality of $x(t)$. Hence (β, z) must lie on the boundary of G.

Alternatively, in the second step you may solve the (IVP) taking the initial condition to be $x(\beta) = z$. As (β, z) lies in the interior of G, it has a unique solution in $(\beta - \rho, \beta + \rho)$ for some $\rho > 0$. Thus the solution extends from (α, β) to $(\alpha, \beta + \rho)$, contradicting the maximality of β. $\qquad\square$

We will show that the (IVP) is solvable without the Lipschitz condition. Recall that Picard-Lindelöf Theorem ensures the existence of a unique solution on on the interval

$[t_0 - a', t_0 + a']$ where a' is any number satisfying

$$a' < \min\left\{a, \frac{b}{M}, \frac{1}{L}\right\} .$$

Now we remove $1/L$ in this dependence.

Proposition 4.6. *Under the setting of Picard-Lindelöf Theorem, there is a unique solution of the (IVP) on the interval $[t_0 - a^*, t_0 + a^*]$ where a^* is given by*

$$a^* = \min\left\{a, \frac{b}{M}\right\} .$$

Proof. Let x be the maximal solution on the maximal interval (α, β). We will show that $\beta \geq t_0 + a^*$. Similarly one can show that $\alpha \leq t_0 - a^*$.

First, let $y_n = (M + 1/n)(t - t_0) + x_0 + 1/n$ which solves the (IVP) $x' = M + 1/n, x(t_0) = x_0 + 1/n$. By the comparison principle, $x(t) < y_n(t)$ for $t \geq t_0$ as long as x exists. Letting $n \to \infty$, $x(t) \leq M(t - t_0) + x_0$ on $[t_0, \beta)$. Similarly, $-M(t - t_0) + x_0 \leq x(t)$ holds. Now, consider two cases: (a) $b/M \leq a$ and (b) $b/M > a$. In (a), we claim x exists in $[t_0, t_0 + b/M)$, that is, $t_0 + b/M \leq \beta$. Suppose on the contrary $\beta < t_0 + b/M$. Then the solution curve $(t, x(t))$ lies in the triangle T formed by the rays $x = M(t - t_0) + x_0, x = -M(t - t_0) + x_0$ and the vertical line $t = \beta$. This triangle T is contained in the interior of R. By Proposition 4.5 this is impossible. Hence $\beta \geq t_0 + b/M$. The solution exists on $[t_0, t_0 + b/M)$. In (b), we claim $x(t)$ exists on $[t_0, t_0 + a)$ so that $a^* = a$. The argument is similar to (a). If $\beta < t_0 + a$, the triangle formed by the rays $x = M(t - t_0) + x_0, x = -M(t - t_0) + x_0$ and the vertical line $t = \beta$ is contained in the interior of R, but this is impossible. Therefore x exists on $[t_0, t_0 + a)$. We can take $a^* = a$. Finally, by Q3 in Ex 8, the solution actually exists up to the right endpoint, that is, in $[t_0, t_0 + a^*]$.

□

Theorem 4.7 (Cauchy-Peano Theorem). *Consider the (IVP) where f is continuous on $R = [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$. There exists a C^1-function $x : [t_0 - a^*, t_0 + a^*] \to [x_0 - b, x_0 + b]$, where $a^* = \min\{a, b/M\}$, solving the (IVP).*

The theorem is also valid for systems.

Proof. By Weierstrass Approximation Theorem, there exists a sequence of polynomials $\{p_n\}$ approaching f in $C(R)$ uniformly. Then $c_n = \sup_G |f(t, x) - p_n(t, x)| \to 0$ as $n \to \infty$. If now we replace p_n by $p_n - c_n$, we still have $p_n - c_n \to 0$. Therefore, we may additionally assume $p_n \leq f$ so that $M_n \leq M$ and $M_n \to M$, where

$$M_n = \max\{|p_n(t, x)| : (t, x) \in R\}.$$

As each p_n satisfies the Lipschitz condition, by Proposition 4.6 there is a unique solution x_n defined on $I_n = [t_0, t_0 + a_n], a_n = \min\{a, b/M_n\}$ for the initial value problem

$$\frac{dx}{dt} = p_n(t, x), \quad x_0(t_0) = x_0.$$

As $M_n \leq M$, all x_n exist on $[t_0, t_0 + a^*]$.

The sequence $\{x_n\}$ is uniformly bounded as $|x_n(t) - x_0| \leq b$. Moreover, from $|dx_n/dt| \leq M_n \leq M$, it is equicontinuous on $[t_0, t_0 + a^*]$ By Ascoli's Theorem, it contains a subsequence $\{x_{n_j}\}$ converging uniformly to a continuous function x^* on $[t_0, t_0 + a^*]$. It remains to check that x^* solves (IVP) on this interval. Consider the formula

$$x_n(t) = x_n(t_0) + \int_{t_0}^{t} p_n(s, x_n(s)) \, ds , \quad t \in [t_0, t_0 + a^*] . \tag{4.3}$$

Letting $n \to \infty$, clearly its left hand side tends to $x^*(t)$. To treat its right hand side, we observe that, for $\varepsilon > 0$, there exists some δ such that

$$|f(s_2, x_2) - f(s_1, x_1)| < \varepsilon, \quad |s_2 - s_1|, |x_2 - x_1| < \delta . \tag{4.4}$$

This is because f is uniformly continuous on R. Next, there is some $n_2 \geq n_1$ such that $\|f - p_n\|_\infty < \varepsilon$, for all $n \geq n_2$ on R. It implies

$$|p_n(s, x) - f(s, x)| < \varepsilon, \quad \forall (s, x) \in R . \tag{4.5}$$

Putting (4.4) and (4.5) together, we have

$$\left| \int_{t_0}^{t} p_{n_j}(s, x_{n_j}(s)) ds - \int_{t_0}^{t} f(s, x^*(s)) ds \right|$$

$$\leq \left| \int_{t_0}^{t} |p_{n_j}(s, x_{n_j}(s)) - f(s, x_{n_j}(s))| ds \right| + \left| \int_{t_0}^{t} |f(s, x_{n_j}(s)) - f(s, x^*(s))| ds \right|$$

$$\leq 2a\varepsilon ,$$

which implies the right hand side of (4.3) tends to

$$x_0 + \int_{t_0}^{t} f(s, x(s)) ds ,$$

as $n_j \to \infty$. We conclude that x is a solution to (IVP) on $[t_0, t_0 + a^*]$.

Similarly one can show a solution exists on $[t_0 - a^*, t_0]$. By putting these two solutions together one obtains a solution on $[t_0 - a^*, t_0 + a^*]$.   □

In this proof we have used Weierstrass approximation theorem in two variables to produce a smooth approximation in the first step. Instead of using this theorem, a smooth approximation can be constructed by a general scheme called mollifiers, see exercise.

We conclude this section with an example of non-uniqueness.

Example 4.4. Consider the function $x(t) = t^2/4$. A direct check shows that it satisfies the differential equation $x' = x^{1/2}$ for $t \geq 0$. Therefore, the IVP

$$x' = |x|^{1/2} \, , \quad x(0) = 0 \, ,$$

has two solutions. First, $x_1(t) \equiv 0$ is the trivial solution. On the other hand, $x_2(t) = t^2/4, t \geq 0$, and $= 0, \ t < 0$ is also a solution. In fact, there are two more solutions, namely, $x_3(t) = 0, t \geq 0, \ = -t^2/4, t < 0$ and $x_4 = x_2 + x_3$. Note that here $f(t, x) = |x|^{1/2}$ does not satisfy the Lipschitz condition in any interval containing the origin.

4.2 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. Although this theorem applies to all complete metric spaces, it has wonderful applications when it comes to the space of continuous functions. In view of this, we include the theorem here.

We start with a definition. A set E in a metric space (X, d) is a **dense set** if for every $x \in X$, the metric ball $B_r(x) \cap E \neq \phi$ for all $r > 0$. For instance, the set of all rational numbers $\mathbb{Q}$ forms a dense set in $\mathbb{R}$, so is the set of all irrational numbers $\mathbb{I}$. On the other hand, the set of all polynomials, $\mathcal{P}$ forms a dense set in $C[0, 1]$ due to Weierstrass Approximation Theorem.

A dense set is considered to be a "large set" in a metric space. However, this classification is a little too rough. Consider the following sets in $\mathbb{R}$ under the usual Euclidean metric:

$$\mathbb{R}, \quad \mathbb{R} \setminus \{a_1, \cdots, a_n\}, \quad \mathbb{I}, \quad \mathbb{Q} \, .$$

Although all of them are dense, they are quite different. Apparently the first two sets, which are open and dense, are "more dense" than the other two. On the other hand, $\mathbb{I}$ is "more dense" than $\mathbb{Q}$. How can we make it precise? A more satisfactory description of the size of a set, as we will see, can be achieved when working on a complete space.

Let us follow most books to start by looking from the other direction . A set E in a metric space is called **nowhere dense** if its closure has empty interior. A closed, nowhere dense set is precisely the complement of an open dense set. From the definition we see that the closure of a nowhere dense set is nowhere dense, and so are its subsets. A nowhere dense set is a set whose elements are scattered sparsely in the metric space. You may say a nowhere dense set is a small set.

Proposition 4.8. *(a) The closure of a nowhere dense set is nowhere dense, so are its subsets.*

(b) The union of finitely many nowhere dense sets is nowhere dense.

(c) Every finite set is nowhere dense provided the metric space has no isolated points.

A point a in a metric space is an **isolated point** if $\{a\}$ is an open set. As it is obvious that a singleton set is closed, an isolated point forms an open and closed set. Every point in a discrete metric space is isolated. There are no isolated points in $\mathbb{R}$. However, if we consider $\mathbb{Z}$ as a subspace of $\mathbb{R}$, every point is an isolated point. In fact, the metric ball $B_{1/2}(n)$ is equal to the singleton set $\{n\}$ showing that the latter is open.

Proof. (a) trivially follows from the definition of a nowhere dense set.

(b) Let $F_j, 1 \leq j \leq n$, be nowhere dense. Without loss of generality we may assume they are closed. Let B_0 be any ball. We need to show that $B_0 \bigcap (X \setminus \bigcup_j F_j) \neq \phi$. As F_1 has empty interior, there exists some point $x_1 \in B_0$ lying outside F_1. Since F_1 is closed, we can find a ball $B_1 \subset B_0$ centering at x_1 such that $B_1 \cap F_1 = \phi$. Next, as F_2 has empty interior and closed, by the same reason there is a ball $B_2 \subset B_1$ centering at x_2 such that $B_2 \cap F_2 = \phi$. Repeating this process finitely many times, we obtain balls $B_j, 1 \leq j \leq n$, with center x_j satisfying (a) $B_{j+1} \subset B_j$, and (b) B_j is disjoint from $F_1, \cdots, F_j$. In particular, x_n belongs to B_0 and is disjoint from $F_1, \cdots, F_n$, that is, from $\bigcup_{j=1}^n F_j$.

(c). First of all, any finite set is closed. In view of (b), it suffices to show that the singleton set $\{a\}$ formed by an isolated point is nowhere dense, or, $X \setminus \{a\}$ is dense. But the latter is clearly true since every open set containing a must intersect $X \setminus \{a\}$.

$\square$

Example 4.5. Consider $\mathbb{R}$ under the Euclidean metric. As there are no isolated points, every finite set in $\mathbb{R}$ is nowhere dense. How about countably union of finite sets? The answer is sometimes yes and sometimes no. For instance the set $\{1, 2, 3, \cdots\}$ is closed and nowhere dense, but $\mathbb{Q}$, which is a countable set, is not nowhere dense.

Although countable unions of nowhere dense sets may not be nowhere dense, they are still considered by mathematicians to be small in size. We fix the notion by introducing the following definition. A set in a metric space is of **first category** or **meager** if it can be expressed as the countable union of nowhere dense sets. It is **of second category** if it is not meager. A set is called **residual** if its complement is of first category. As we will see, a residual set is most effective to describe a large set.

Proposition 4.9. *(a) Every subset of a set of first category is of first category.*

(b) The union of countably many sets of first category is of first category.

(c) Every countable set is of first category provided the metric space has no isolated points.

Proof. (a) Let $E = \bigcup_n E_n$ be a set of first category where E_n's are nowhere dense. For $F \subset E$, $F = \bigcup_n (F \cap E_n)$ is of first category since $F \cap E_n$ is nowhere dense.

(b) It is clear from definition.

(c) It follows from Proposition 4.8.

$\square$

Taking complement, Proposition 4.9 may be formulated in terms of residual sets.

Proposition 4.9'. (a) *Every set containing a residual set is residual.*

(b) *The intersection of countably many residual sets is a residual set.*

(c) *Every set whose complement is countable is a residual set provided the metric space has no isolated points.*

Example 4.6. Let $\mathbb{Q} = \{q_j\}, j \geq 1$ as a subset in $\mathbb{R}$. Since each singleton set is closed and nowhere dense, $\mathbb{Q}$, a countable union of rational numbers, is of first category. According to the definition, $\mathbb{I}$, the irrational numbers, is a residual set.

We agree that a nowhere dense set is regarded as a small set and so is its countable union. On the other hand, an open dense set is regarded as a large set and so is its countable intersection. This has been witnessed by the example $\mathbb{Q}$ and $\mathbb{I}$ in $\mathbb{R}$. Such observation can be justified when the ambient space is complete.

Theorem 4.10 (**Baire Category Theorem**). *In a complete metric space, the countable union of nowhere dense sets has empty interior.*

Proof. Let $F_n, n \geq 1$, be nowhere dense. Without loss of generality we may assume they are closed. Let B_0 be any ball. The theorem will be established if we can show that $B_0 \bigcap (X \setminus \bigcup_n F_n) \neq \phi$. As F_1 has empty interior, there exists some point $x_1 \in B_0$ lying outside F_1. Since F_1 is closed , we can find a closed ball $\overline{B}_1 \subset B_0$ centering at x_1 such that $\overline{B}_1 \cap F_1 = \phi$ and its diameter $d_1 \leq d_0/2$, where d_0 is the diameter of B_0. Next, as F_2 has empty interior and closed, by the same reason there is a closed ball $\overline{B}_2 \subset B_1$ centering at x_2 such that $\overline{B}_2 \cap F_2 = \phi$ and $d_2 \leq d_1/2$. Repeating this process, we obtain a sequence of closed balls $\overline{B}_n$ with center x_n satisfying (a) $\overline{B}_{n+1} \subset B_n$, (b) $d_n \leq d_0/2^n$, and (c) $\overline{B}_n$ is disjoint from $F_1, \cdots, F_n$. As the balls shrink to a point, $\{x_n\}$ is a Cauchy sequence. By the completeness of X, $\{x_n\}$ converges to some x^*. As each $\overline{B}_n$ is closed and $x_j \in \overline{B}_n$ for all $j \geq n$, $x^* \in \overline{B}_n$ for all n. In particular, it means that x^* does not belong to F_n for all n, so x^* is a point in B_0 not in $\bigcup_n F_n$.

$\square$

Using the fact that the complement of a closed, nowhere dense set is an open dense set, we have the following equivalent formulation of the Baire Category Theorem.

Theorem 4.10' (**Baire Category Theorem**). *Let (X, d) be a complete metric space and $\{G_n\}$ be a sequence of open, dense sets in X. The set $E = \bigcap_{n=1}^{\infty} G_n$ is dense.*

We have the following two immediate consequences of the theorem.

Corollary 4.11. *Let (X, d) be complete. Suppose*

$$X = \bigcup_{n=1}^{\infty} E_n \ ,$$

where E_n's are closed sets. Then at least one of these E_n's has non-empty interior.

In other words, it is impossible to decompose a complete metric space into a countable union of nowhere dense sets.

Proof. In case all these E_n's have empty interior, $X = \bigcup_{n=1}^{\infty} E_n$ has empty interior according to Baire Category Theorem. However, this is impossible since X is open. $\square$

This corollary is also called the Baire category theorem.

Corollary 4.12. *If a set in a complete metric space is of first category, it cannot be a residual set, and vice versa.*

Proof. Let E be of first category and let $E \subset \bigcup_{n=1}^{\infty} F_n$ where F_n's are closed with empty interior. If it is residaul, its complement is of first category. Thus, $X \setminus E \subset \bigcup_{n=1}^{\infty} E_n$ where E_n's are closed with empty interior. We put F_n's and E_n's together to form a sequence $\{H_n\}, H_{2n-1} = F_n, H_{2n} = E_n$. Then

$$X = E \bigcup (X \setminus E) \subset \bigcup_{n=1}^{\infty} H_n \ ,$$

which contradicts the corollary above. $\square$

Baire Category Theorem has many interesting applications. We end this section by giving two standard ones. 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.

To refresh your memory we state

Lemma 4.13. *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
$$|f(y) - f(x)| \leq L|y - x|, \quad \forall y, |y - x| < \delta_0,$$
where $L = |f'(x)| + 1$. For y lying outside $(x - \delta_0, x + \delta_0)$, $|y - x| \geq \delta_0$. Hence
$$\begin{aligned} |f(y) - f(x)| &\leq |f(x)| + |f(y)| \\ &\leq \frac{2M}{\delta_0}|y - x|, \quad \forall y \in [a,b] \setminus (x - \delta_0, x + \delta_0), \end{aligned}$$
where $M = \sup\{|f(x)| : x \in [a,b]\}$. It follows that f is Lipschitz continuous at x with an Lipschitz constant not exceeding $\max\{L, 2M/\delta_0\}$.
$\square$

Theorem 4.14. *The set of all continuous, nowhere differentiable functions forms a residual set in $C[a,b]$ and hence dense in $C[a,b]$.*

Proof. For each $L > 0$, define

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

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. We need to show $f \in S_L$. By passing to a subsequence if necessary, we may assume $\{x_n\}$ to some x^* in $[a,b]$. We have,

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

Letting $n \to \infty$, $|f(y) - f(x^*)| \leq L|y - x^*|$. 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 $2r$ satisfying $0 \leq \varphi \leq 1$ 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. We have $\|f - g\|_\infty \leq \|f - p\|_\infty + \varepsilon/2\|\varphi\|_\infty < \varepsilon$. On the other hand, 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 equal to $1/r$ or $-1/r$. We have

$$\begin{aligned} |g(y) - g(x)| &\geq \frac{\varepsilon}{2}|\varphi(y) - \varphi(x)| - |p(y) - p(x)| \\ &\geq \frac{\varepsilon}{2}|\varphi(y) - \varphi(x)| - L_1|y - x| \\ &= \frac{\varepsilon}{2}\frac{1}{r}|y - x| - L_1|y - x| \end{aligned}$$

Therefore, if we choose r so that

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

$|g(y) - g(x)| > L|y - x|$, 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 the lemma it must belong to S_N for some positive integer N. Therefore, $S \subset \cup_{N=1}^{\infty} S_N$ is of first category. Its complement, the collection of continuous, nowhere differentiable functions, is residual, and hence dense in $C[0, 1]$. The completeness of $C[0, 1]$ is implicitly used.   □

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.

Our second application is concerned with the basis of a vector space. Recall that a basis of a vector space is a set of linearly independent vectors such that every vector can be expressed as a linear combination of vectors from the basis. More precisely, Let V be a real vector space. A subset B of V is called a basis of V if (a) Let $w_1, w_2, \cdots, w_N$ be finitely many vectors from B. If $a_1 w_2 + a_2 w_2 + \cdots + a_N w_N = 0$, then $a_1 = a_2 = \cdots = a_n = 0$, that is, these w_j's are linearly independent; and (b) every vector from V can be written as $v = \sum_{j=1}^{M} a_j w_j$ for some finite number M, that is, every vector can be expressed as a linear combination of vectors from B. The construction of a basis in a finite dimensional vector space was done in MATH2040. However, in an infinite dimensional vector space the construction of a basis is not so easy. Nevertheless, using Zorn's lemma, a variant of the axiom of choice, one shows that a basis always exists. Some authors call a basis for an infinite dimensional basis a Hamel basis. The difficulty in writing down a Hamel basis is explained in the following result.

Theorem 4.15. *Every basis of an infinite dimensional Banach space consists of uncountably many vectors.*

Proof. Let V be an infinite dimensional Banach space and $\mathcal{B} = \{w_j\}$ be a countable basis. We aim for a contradiction. Indeed, let W_n be the subspace spanned by $\{w_1, \cdots, w_n\}$. We have the decomposition

$$V = \bigcup_n W_n.$$

If one can show that each W_n is closed and has empty interior, since V is complete, the corollary above tells us this decomposition is impossible. To see that W_n has empty interior, pick a unit vector v_0 outside W_n. For $w \in W_n$ and $\varepsilon > 0$, $w + \varepsilon v_0 \in B_\varepsilon(w) \cap (V \setminus W_n)$, so W_n cannot contain a ball. Next, letting v_j be a sequence in W_n and $v_j \to v_0$, we would like to show that $v \in W_n$. Indeed, every vector $v \in W_n$ can be uniquely expressed as $\sum_{j=1}^{n} a_j w_j$. The map $v \mapsto a \equiv (a_1, \cdots, a_n)$ sets up a linear bijection between W_n and

$\mathbb{R}^n$ and $\||a\|| \equiv \|v\|$ defines a norm on $\mathbb{R}^n$. Since any two norms in $\mathbb{R}^n$ are equivalent (see exercise), a convergent (resp. Cauchy) sequence in one norm is the same in the other norm. Since now $\{v_j\}$ is convergent in V, it is a Cauchy sequence in V. The corresponding sequence $\{a^j\}$, $a^j = (a_1^j, \cdots, a_n^j)$, is a Cauchy sequence in $\mathbb{R}^n$ with respect to $\||\cdot\||$ and hence in $\|\cdot\|_2$, the Euclidean norm. Using the completeness of $\mathbb{R}^n$ with respect to the Euclidean norm, $\{a^j\}$ converges to some $a^* = (a_1^*, \cdots, a_n^*)$. But then $\{v_j\}$ converges to $v^* = \sum_j a_j^* w_j$ in W_n. By the uniqueness of limit, we conclude that $v_0 = v^* \in W_n$, so W_n is closed.

$\square$

Comments on Chapter 4.

Arzela and Ascoli Theorems play 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. A more general version of Ascoli's Theorem asserts that Theorem 4.2 still holds when $C(\overline{G})$ is replaced by $C(X)$ where X is a compact metric space. Google for more.

Ascoli's Theorem (Theorem 4.2) 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

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

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

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

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.

There is an alternate proof of Cauchy-Peano Theorem without using Picard-Lindelöf Theorem. In this proof piecewise linear approximate solutions, called Euler's polygonal lines, to (IVP) are constructed and subconvergence to a solution is shown by Ascoli's theorem, see, Coddington-Levinson, Theory of Ordinary Differential Equations, 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 Hewitt-Stromberg, "Abstract Analysis". An simpler example can be found in Rudin's Principles.

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