MATH3060: Mathematical Analysis III

onenylxus/math-notes

December 7, 2020

Remarks

(1) Context of this document is based on university course *MATH3060: Mathematical Analysis III* from *Department of Mathematics, The Chinese University of Hong Kong (CUHK)*. The original source can be found at `https://www.math.cuhk.edu.hk/course`. The author does not own the source.

(2) This document is assumed unavailable for unauthorized parties that have not attended the university course. It is prohibited to share, including distributing or copying this document to unauthorized parties in any means for any non-academic purpose.

(3) Context of this doucment may not be completely accurate. The author assumes no responsibility or liability for any errors or omissions in the context of this document.

(4) This document is under license CC-BY-SA 4.0. It is allowed to make any editions on this document, as long as terms of the license is not violated.

Source

The latest version of this document can be found at `https://www.github.com/onenylxus/math-notes`.

Contents

1 Fourier Series

1.1 Introduction to Fourier Series

1.1.1 Trigonometric Series

Definition 1.1. A **trigonometric series** on $[-\pi, \pi]$ is a series of functions in the form

$$\sum_{n=0}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

where $a_n, b_n \in \mathbb{R}$. Furthermore, if $a_n = 0$ for all n, the series is called a **sine series**. Similarly, if $b_n = 0$ for all n, the series is called a **cosine series**.

Note that it is possible to pull out the zeroth term of the sum, making the series in the form

$$a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

hence it can be assumed that $b_0 = 0$.

Proposition 1.2. Let $(a_n), (b_n)$ be infinite series. If

$$|a_n|, |b_n| \leq \frac{C}{n^s}$$

for some $C > 0$ and $s > 1$, then their corresponding series $\sum_{n=0}^{\infty} |a_n|$ and $\sum_{n=0}^{\infty} |b_n|$ are convergent.

Proposition 1.3. If $\sum_{n=0}^{\infty} |a_n|$ and $\sum_{n=0}^{\infty} |b_n|$ are convergent, then by Weierstrass M-test,

$$\sum_{n=0}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

is uniformly and absolutely convergent.

Proposition 1.4. Let $\phi(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$ be a continuous function on $[-\pi, \pi]$. If $\sum |a_n|, \sum |b_n| < \infty$, then $\phi(x)$ is 2π-periodic.

1.1.2 Fourier Series

Definition 1.5. Let f be a 2π-periodic function on $\mathbb{R}$ which is Riemann integrable on $[-\pi, \pi]$, then the **Fourier series** (or **Fourier expansion**) of f is the trigonometric series

$$a_0 + \sum_{n=1}^{\infty}(a_n \cos(nx) + b_n \sin(nx))$$

with Fourier coefficients of f

$$\begin{cases} a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(y)\,\mathrm{d}y \\ a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\cos(ny)\,\mathrm{d}y \\ b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(y)\sin(ny)\,\mathrm{d}y \end{cases}$$

Note that a_0 is actually the average of f over $[-\pi, \pi]$. Fourier series depends on the global information of f on $[-\pi, \pi]$ instead of a point in f. Fourier series also depends only on $f|_{(-\pi,\pi)}$, which means the end points of the closed interval are independent.

Proposition 1.6. Let f_1, f_2 are Fourier series where $f_1 \equiv f_2$ almost everywhere on $[-\pi, \pi]$, then f_1 and f_2 are the same Fourier series.

1.1.3 Motivation of Fourier Series

Recall the form of a Fourier series as

$$f(x) = a_0 + \sum_{n=1}^{\infty}(a_n \cos(nx) + b_n \sin(nx))$$

for all $x \in \mathbb{R}$. If f is uniformly convergent,

$$\begin{aligned} &\int_{-\pi}^{\pi} f(x)\cos(mx)\,\mathrm{d}x \\ = {} & a_0 \int_{-\pi}^{\pi}\cos(mx)\,\mathrm{d}x + \sum_{n=1}^{\infty}\left(a_n \int_{-\pi}^{\pi}\cos(nx)\cos(mx)\,\mathrm{d}x + b_n \int_{-\pi}^{\pi}\sin(nx)\cos(mx)\,\mathrm{d}x\right) \end{aligned}$$

Note that

$$\int_{-\pi}^{\pi}\cos(mx)\,\mathrm{d}x = \begin{cases} 2\pi & \text{if } m = 0 \\ 0 & \text{if } m \neq 0 \end{cases}$$

$$\int_{-\pi}^{\pi}\cos(nx)\cos(mx)\,\mathrm{d}x = \begin{cases} \pi & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases}$$

$$\int_{-\pi}^{\pi} \sin(nx)\cos(mx)\,\mathrm{d}x = 0\ \forall m, n \geq 1$$

which will deduce

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,\mathrm{d}x,\ a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(mx)\,\mathrm{d}x$$

Using similar method but instead of $\cos(mx)$, $\sin(mx)$ will deduce

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(mx)\,\mathrm{d}x$$

1.2 Complex Fourier Series

1.2.1 Definition of Complex Fourier Series

Definition 1.7. Let f be a 2π-periodic function on $\mathbb{C}$ which is Riemann integrable on $[-\pi, \pi]$, then its **complex Fourier series** is a Fourier series of the form

$$\sum_{-\infty}^{\infty} c_n e^{inx}$$

where $\{c_n\}_{-\infty}^{\infty}$ is a **bisequence** of complex numbers defined by

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,\mathrm{d}x$$

for all integers n. Moreover, $\sum_{-\infty}^{\infty} c_n e^{inx}$ is said to be convergent at x if

$$\lim_{N\to+\infty}\sum_{-N}^{N} c_n e^{inx}$$

exists.

Note that for a complex-valued function $f = u + iv$,

$$\int_a^b f = \int_a^b u + i\int_a^b v$$

In other words, f is said to be integrable if both u and v are integrable.

1.2.2 Motivation of Complex Fourier Series

Recall the form of a complex Fourier series as

$$f(x) = \sum_{-\infty}^{\infty} c_n e^{inx}$$

for all $x \in \mathbb{C}$. If f converges nicely,

$$\int_{-\pi}^{\pi} e^{-imx}\,\mathrm{d}x = \sum_{-\infty}^{\infty} c_n \int_{-\pi}^{\pi} e^{i(n-m)x}\,\mathrm{d}x$$

Note that

$$\int_{-\pi}^{\pi} e^{i(n-m)x}\,\mathrm{d}x = \begin{cases} 2\pi & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

which will deduce

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,\mathrm{d}x$$

1.2.3 Relations between Real and Complex Fourier Series

In this section the relationship between (real) Fourier series and complex Fourier series for a real-valued function f is discussed. Note that

$$\begin{aligned} c_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,\mathrm{d}x \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)(\cos(nx) - i\sin(nx))\,\mathrm{d}x \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,\mathrm{d}x - \frac{i}{2\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,\mathrm{d}x \end{aligned}$$

which will deduce

$$c_n = \begin{cases} \frac{1}{2}(a_n - ib_n) & \text{if } n \geq 1 \\ 0 & \text{if } n = 0 \\ \frac{1}{2}(a_{-n} + ib_{-n}) & \text{if } n \leq -1 \end{cases}$$

Proposition 1.8. Let f be a real-valued function, then its complex Fourier coefficient $c_{-n} = \overline{c_n}$.

Proposition 1.9. Let f be a 2π-periodic real function which is differentiable on $[-\pi, \pi]$ with f' integrable on $[-\pi, \pi]$. Denote the Fourier coefficients of f and f' by $\{a_n(f), b_n(f); c_n(f)\}$ and $\{a_n(f'), b_n(f'); c_n(f')\}$ respectively, then

$$\begin{cases} a_n(f') = nb_n(f) \\ b_n(f') = -na_n(f) \end{cases} \quad \text{and } c_n(f') = inc_n(f)$$

Note that one of the advantages of using complex Fourier series is to compute derivatives with more convenience.

1.3 Fourier Series and Extensions

1.3.1 Extensions of Periodic Functions

For any Riemann integrable function f on $[-\pi, \pi]$, one can define the Fourier coefficients to form a Fourier series. On the other hand, we can restrict f to $(-\pi, \pi]$ and extend periodically to a 2π-periodic function $\tilde{f}$ on $\mathbb{R}$. The function f and its extension $\tilde{f}$ have the same Fourier series.

Here the $\sim$ symbol in

$$f(x) \sim a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

represents $f(x)$ has the Fourier series on the right hand side. The equal sign $=$ is not used since the series may not converge.

Example 1.10. Let $f(x) = x$ be a function in $[-\pi, \pi]$. Find the Fourier series of f.

Answer. The Fourier coefficients are

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} x \, \mathrm{d}x = 0$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x \cos(nx) \, \mathrm{d}x = 0$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin(nx) \, \mathrm{d}x = (-1)^{n+1} \frac{2}{n}$$

Therefore

$$f(x) \sim \sum_{n=1}^{\infty} (-1)^{n+1} \frac{2}{n} \sin(nx)$$

which is a sine series.

With the example above, the following proposition can be introduced by observation:

Proposition 1.11. Let f be a real function, then its Fourier series is a sine series if f is odd, and it is a cosine series if f is even.

Note that the Fourier series may not be the same as the original function, especially when there are discontinuous points like $\pm\pi$. Also, the convergence of Fourier series is not clear since some terms like $\sum(1/n)$ does not converge.

1.3.2 Big-O and Little-O Notations

Definition 1.12. Let $\{x_n\}$ be a sequence, then the **big-O notation**, denoted by O, paired with x_n is defined as

$$x_n = O(n^s) \Leftrightarrow |x_n| \leq Cn^s$$

for some constant $C > 0$, as $n \to \infty$. Similarly, the **little-O notation**, denoted by o, paired with x_n is defined as

$$x_n = o(n^s) \Leftrightarrow \frac{|x_n|}{n^s} \to 0$$

as $n \to \infty$.

Example 1.13. Find the correlation of

$$x_n = \frac{2(-1)^{n+1}}{n} \sin(nx)$$

using big-O notation.

Answer. Since $|x_n| \leq 2/n$, $x_n = O(1/n)$.

Example 1.14. Find the correlation of

$$x_n = \log(n)$$

using little-O notation.

Answer. Since $|\log(n)|/n \to 0$ as $n \to \infty$, $x_n = o(n)$.

1.3.3 Fourier Series of Unusual Periodic Functions

Let f be a $2T$-periodic function. Note that $g(x) = f(\frac{T}{\pi}x)$ is a 2π-periodic function, then

$$g(x) \sim a_0 + \sum_{n=1}^{\infty}(a_n \cos(nx) + b_n \sin(nx))$$

with

$$\begin{cases} a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(x)\,\mathrm{d}x \\ a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(x)\cos(nx)\,\mathrm{d}x \\ b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(x)\sin(nx)\,\mathrm{d}x \end{cases}$$

along with the substitution $y = \frac{T}{\pi}x$ implies

$$f(y) \sim a_0 + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{n\pi}{T}y\right) + b_n \sin\left(\frac{n\pi}{T}y\right)\right)$$

with

$$\begin{cases} a_0 = \frac{1}{2T}\int_{-T}^{T} f(y)\,\mathrm{d}y \\ a_n = \frac{1}{T}\int_{-T}^{T} f(y)\cos\left(\frac{n\pi}{T}y\right)\,\mathrm{d}y \\ b_n = \frac{1}{T}\int_{-T}^{T} f(y)\sin\left(\frac{n\pi}{T}y\right)\,\mathrm{d}y \end{cases}$$

Such Fourier series is called Fourier series of $2T$-periodic function f.

1.4 Convergence of Fourier Series

1.4.1 Riemann-Lebesgue Lemma

Recall the definition of a step function on $[-\pi, \pi]$ as a function of the form

$$s(x) = \sum_{j=0}^{N-1} s_j \chi_{I_j}$$

where $-\pi = a_0 < a_1 < \cdots < a_N = \pi$, $I_0 = [a_0, a_1]$ and $I_j = (a_j, a_{j+1}]$ for $1 \leq j \leq N-1$. The characteristic function (or indicator function)

$$\chi_E = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

Proposition 1.15. For every step function s integrable on $[-\pi, \pi]$, there exists a constant $C > 0$ depending on s such that

$$|a_n(s)|, |b_n(s)| \leq \frac{C}{n}$$

for all $n \geq 1$. $a_n(s)$ and $b_n(s)$ are Fourier coefficients of s.

Proof. Let

$$s(x) = \sum_{j=0}^{N-1} s_j \chi_{I_j}$$

be the step function, then for $n \geq 1$,

$$\begin{aligned}\pi a_n(s) &= \int_{-\pi}^{\pi} s(x)\cos(nx)\,\mathrm{d}x \\ &= \sum_{j=0}^{N-1} s_j \int_{a_j}^{a_{j+1}} \cos(nx)\,\mathrm{d}x \\ &= \sum_{j=0}^{N-1} s_j \frac{\sin(na_{j+1}) - \sin(na_j)}{n}\end{aligned}$$

which implies $|a_n(s)| \leq \frac{C}{n}$. The proof for $|b_n(s)|$ is similar.

Proposition 1.16. Let f be integrable on $[-\pi, \pi]$, then for all $\epsilon > 0$, there exists a step function s such that $s \leq f$ on $[-\pi, \pi]$ and

$$\int_{-\pi}^{\pi} (f - s) < \epsilon$$

Proof. Since f is Riemann integrable, the function can be approximated with Darboux lower sum. For all $\epsilon > 0$, there exists a partition $-\pi = a_0 < a_1 < \cdots < a_N = \pi$ such that

$$\int_{-\pi}^{\pi} f - \sum_{j=0}^{N-1} m_j(a_{j+1} - a_j) < \epsilon$$

where $m_j = \inf\{f(x) \mid x \in [a_j, a_{j+1}]\}$. Define the step function

$$s(x) = \sum_{j=0}^{N-1} m_j \chi_{I_j}$$

then $s \leq f$ and

$$\int_{-\pi}^{\pi} s(x)\,\mathrm{d}x = \sum_{j=0}^{N-1} m_j(a_{j+1} - a_j)$$

implies the result.

With the propositions above, the **Riemann-Lebesgue Lemma** can be introduced:

Theorem 1.17. The Fourier coefficients of any 2π-periodic function f integrable on $[-\pi, \pi]$ converge to 0 as $n \to \infty$.

Proof. By *Proposition 1.16*, for any $\epsilon > 0$, there exists a step function s such that $s \leq f$ and

$$\int_{-\pi}^{\pi}(f-s) < \frac{\epsilon}{2}$$

On the other hand, by *Proposition 1.15*, there exists $n_0 > 0$ such that

$$|a_n(s)| < \frac{\epsilon}{2}$$

for all $n \geq n_0$. For instance, $n_0 = [\frac{2C}{\epsilon}] + 1$ with the constant C in *Proposition 1.15*. Note that

$$\begin{aligned}
|a_n(f) - a_n(s)| &= \frac{1}{\pi}\left|\int_{-\pi}^{\pi}(f-s)(x)\cos(nx)\,\mathrm{d}x\right| \\
&\leq \frac{1}{\pi}\int_{-\pi}^{\pi}(f-s) \quad \text{as } f \geq s \\
&\leq \frac{\epsilon}{2\pi}
\end{aligned}$$

Hence,

$$\begin{aligned}
|a_n(f)| &\leq |a_n(s)| - |a_n(f) - a_n(s)| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2\pi} < \epsilon
\end{aligned}$$

for all $n \geq n_0$, which means $a_n(f) \to 0$ as $n \to \infty$. The proof for $b_n(f)$ is similar to that above.

1.4.2 Lipschitz Continuity at Points

Definition 1.18. Let $f \sim a_0 + \sum_{n=1}^{\infty}(a_n\cos(nx) + b_n\sin(nx))$ be a function which has a Fourier series, then the n**-th partial sum** of Fourier series of f, denoted by $(S_nf)(x)$, is given by

$$(S_nf)(x) = a_0 + \sum_{k=1}^{n}(a_k\cos(kx) + b_k\sin(kx))$$

Definition 1.19. Let f be a function on $[a,b]$, then f is called **Lipschitz continuous** at a point $x_0 \in [a,b]$ if there exists $L > 0$ and $\delta > 0$ such that

$$|f(x) - f(x_0)| \leq L\,|x - x_0|$$

for all $|x - x_0| < \delta$.

Note that both L and δ may depend on the point x_0. Below is a proposition on extending Lipschitz continuity from a point to an interval:

Proposition 1.20. If f is Lipschitz continuous at $x_0 \in [a, b]$ and f is bounded on $[a, b]$, then there exists $L' > 0$ which may depends on x_0 such that

$$|f(x) - f(x_0)| \leq L' |x - x_0|$$

for all $x \in [a, b]$.

Proof. By *Definition 1.19*, there exists $L, \delta > 0$ such that

$$|f(x) - f(x_0)| \leq L |x - x_0|$$

for all $|x - x_0| < \delta$. If $|x - x_0| \geq \delta$, then

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x)| + |f(x_0)| \\ &\leq 2M \leq \frac{2M |x - x_0|}{\delta} \end{aligned}$$

where $M = \sup_{[a,b]} |f| \geq 0$. Pick $L' = \max \{L, 2M/\delta\} > 0$, then

$$|f(x) - f(x_0)| \leq L' |x - x_0|$$

for all $x \in [a, b]$.

1.4.3 Dirichlet Kernels

Definition 1.21. The **Dirichlet kernel**, denoted by $D_n(z)$, is defined by

$$D_n(z) = \begin{cases} \frac{\sin(n+1/2)z}{2\pi \sin(1/2)z} & \text{if } z \neq 0 \\ \frac{2n+1}{2\pi} & \text{if } z = 0 \end{cases}$$

Proposition 1.22. Below are the properties of Dirichlet kernels:

(a) Integral of a Dirichlet kernel $\int_{-\pi}^{\pi} D_n(z) \, \mathrm{d}z = 1$.

(b) $D_n(z)$ is even, continuous, 2π-periodic on $[-\pi, \pi]$ and

$$D_n\left(\frac{2k\pi}{2n+1}\right) = 0$$

for all $k = -n, -n+1, \cdots, n$.

(c) The maximum

$$\max_{[-\pi,\pi]} D_n(z) = D_n(0) = \frac{2n+1}{2\pi}$$

(d) For all $0 < \delta < \pi/2$,

$$\int_0^\delta |D_n(z)| \, \mathrm{d}z \to +\infty$$

as $n \to +\infty$

Proof. Part (a) This can be achieved by integrating

$$\int_{-\pi}^{\pi} \left(\frac{1}{2} + \sum_{k=1}^{n} \cos(kz) \right) \mathrm{d}z$$

Part (d) Let $0 < \delta < \pi/2$, then for all $n \in \mathbb{N}$, there exists $N \in \mathbb{N}$ such that

$$N < \frac{n+1/2}{\pi}\delta \leq N+1$$

where $N \to \infty$ as $n \to \infty$. Note that

$$\begin{aligned}
\int_0^\delta |D_n(z)| \, \mathrm{d}z &= \int_0^\delta \frac{|\sin(n+1/2)z|}{2\pi\,|\sin(z/2)|} \, \mathrm{d}z \\
&= \int_0^{(n+1/2)\delta} \frac{|\sin(t)|}{2\pi\,|\sin(t/(2n+1))|} \left(\frac{2\,\mathrm{d}t}{2n+1} \right) \quad \text{where } t = \left(n + \frac{1}{2}\right) z \\
&= \frac{1}{\pi} \int_0^{(n+1/2)\delta} \frac{\sin(t)}{t} \frac{t/(2n+1)}{|\sin(t/(2n+1))|} \, \mathrm{d}t \\
&\geq \frac{1}{\pi} \int_0^{(n+1/2)\delta} \frac{\sin(t)}{t} \, \mathrm{d}t \quad \text{since } \frac{\sin(x)}{x} < 1 \text{ for } 0 < x \\
&\geq \frac{1}{\pi} \int_0^{N\pi} \frac{\sin(t)}{t} \, \mathrm{d}t \\
&= \frac{1}{\pi} \sum_{k=1}^{N} \int_{(k-1)\pi}^{k\pi} \frac{\sin(t)}{t} \, \mathrm{d}t \\
&= \frac{1}{\pi} \sum_{k=1}^{N} \int_0^{\pi} \frac{|\sin(s)|}{s + (k-1)\pi} \, \mathrm{d}s \quad \text{where } s = t - (k-1)\pi \\
&\geq \frac{1}{\pi} \sum_{k=1}^{N} \int_0^{\pi} \frac{|\sin(s)|}{k\pi} \, \mathrm{d}s \quad \text{since } t \leq k\pi \\
&= \frac{1}{\pi^2} \left(\int_0^{\pi} |\sin(s)| \, \mathrm{d}s \right) \sum_{k=1}^{N} \frac{1}{k} = \frac{2}{\pi^2} \sum_{k=1}^{N} \frac{1}{k}
\end{aligned}$$

But since the sum of harmonic series $\sum_{k=1}^{N}(1/k)$ diverges when $N \to \infty$ as $n \to \infty$,

$$\lim_{n\to\infty} \int_0^\delta |D_n(z)| \, \mathrm{d}z = +\infty$$

With the definition and properties of Dirichlet kernels, the following proposition can be introduced.

Proposition 1.23. Let f be a 2π-periodic function integrable on $[-\pi, \pi]$. Suppose that f is Lipschitz continuous at x, then the sequence $\{S_n f(x)\}$ converges to $f(x)$ as $n \to +\infty$.

Proof. Let f be a function that is Lipschitz continuous at a point $x_0 \in [-\pi, \pi]$. By splitting

$$(S_n(f))(x_0) - f(x_0) = I_1 + I_2$$

into integrals I_1 and I_2 concentrated in $[-\delta, \delta]$ and essentially, outside the interval, respectively. Note that by *Proposition 1.22*,

Example 1.24. Let $f(x) = x$ be a 2π-periodic function integrable on $[-\pi, \pi]$. Its Fourier series

$$x \sim 2\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx)$$

clearly implies that $f(x)$ is Lipschitz continuous at any $x \in (-\pi, \pi)$.

Proposition 1.25. Let f be a 2π-periodic function integrable on $[-\pi, \pi]$. Suppose that for $x_0 \in [-\pi, \pi]$, the following are satisfied:

(a) The left-hand limit and right-hand limit both exist, which is

$$f(x_0^-) = \lim_{x\to x_0^-} f(x), f(x_0^+) = \lim_{x\to x_0^+} f(x)$$

(b) There exists $L > 0$ and $\delta > 0$ such that

$$\begin{cases} \left|f(x) - f(x_0^+)\right| \leq L(x - x_0) & \text{where } 0 < x - x_0 < \delta \\ \left|f(x) - f(x_0^-)\right| \leq L(x_0 - x) & \text{where } 0 < x_0 - x < \delta \end{cases}$$

then

$$S_n f(x) \to \frac{f(x_0^+) + f(x_0^-)}{2}$$

as $n \to +\infty$.

Example 1.26. Let $f(x) = x$ be a 2π-periodic function integrable on $[-\pi, \pi]$, where f is discontinuous at $x = \pi$. Note that $f(\pi^-) = \pi$ and $f(\pi^+) = -\pi$.

Now assume $\delta = \frac{\pi}{2}$. For $0 < x - \pi < \delta$,

$$\begin{aligned} \left|f(x) - f(\pi^+)\right| &= |f(x - 2\pi) - (-\pi)| \\ &= |x - 2\pi + \pi| \\ &= x - \pi \leq L(x - \pi) \end{aligned}$$

where $L = 1$. The approach for $0 < \pi - x < \delta$ is similar. Therefore, by *Proposition 1.25*,

$$S_n f(x) \to \frac{f(\pi^+) + f(\pi^-)}{2} = 0$$

as $n \to +\infty$.

1.4.4 Lipschitz Condition and Uniform Convergence

Definition 1.27. Let f be a function on $[a, b]$, then it is said to satisfy **Lipschitz condition** if there exists $L > 0$ such that

$$|f(x) - f(y)| \leq L\,|x - y|$$

for all $x, y \in [a, b]$.

Note that Lipschitz condition is uniform since L is independent of any choice of x, y. Also, if f satisfies a Lipschitz condition, f is Lipschitz continuous at every point on $[a, b]$.

Proposition 1.28. Let f be a 2π-periodic function satisfying a Lipschitz condition, then its Fourier series converge uniformly to f itself.

1.5 Weierstrass Approximation Theorem

1.5.1 Piecewise Linear Functions

Recall that a continuous function is piecewise linear if there exists a partition such that the function is linear within each subinterval.

Proposition 1.29. Let f be a continuous function on $[a, b]$, then for all $\epsilon > 0$, there exists a continuous and piecewise linear function g with $g(a) = f(a)$, $g(b) = f(b)$ such that

$$\|f - g\|_\infty < \epsilon$$

where

$$\|f - g\|_\infty = \sup_{[a,b]} |f(x) - g(x)|$$

1.5.2 Trigonometric Polynomials

Definition 1.30. A **trigonometric polynomial** is of the form $P(\cos(x), \sin(x))$ where $P(x, y)$ is a polynomial of 2 variables.

Note that a trigonometric polynomial is a finite Fourier series, and vice versa.

Proposition 1.31. Let f be a continuous function on $[0, \pi]$, then for all $\epsilon > 0$, there exists a trigonometric polynomial h such that $\|f - h\|_\infty < \epsilon$.

1.5.3 General Theorem

Below is the **Weierstrass Approximation Theorem**:

Theorem 1.32. Let $f \in C[a, b]$, then for all $\epsilon > 0$, there exists a polynomial q such that $\|f - q\|_\infty < \epsilon$.

1.6 Mean Convergence of Fourier Series

1.6.1 Bracket Products

Definition 1.33. Let f, g be Riemann integrable functions on $[-\pi, \pi]$, then the **bracket product** (or L^2**-product**, L^2 **inner product**) of f and g is given by

$$\langle f, g\rangle_2 = \int_{-\pi}^{\pi} f(x)g(x)\,\mathrm{d}x$$

Note that for complex functions, the bracket product is defined by

$$\langle f, g\rangle_2 = \int_{-\pi}^{\pi} f\overline{g}$$

Definition 1.34. Let f, g be Riemann integrable functions on $[-\pi, \pi]$, then the L^2**-norm** of f is given by

$$\|f\|_2 = \sqrt{\langle f, f \rangle_2}$$

Also, the L^2**-distance** between f and g is given by $\|f - g\|_2$.

1.6.2 Mean Convergence

Definition 1.35. Let f, f_n be Riemann integrable functions on $[-\pi, \pi]$, then $f_n \to f$ in L^2**-sense** if $\|f_n - f\|_2 \to 0$ as $n \to \infty$.

This definition brings out why such idea is called mean convergence:

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} (f_n - f)^2 \to 0$$

is actually a variation of root mean square. Note that L^2-norm and L^2-distance are not norm and distance in a strict sense since

$$\begin{cases} \|f\|_2 = 0 & \not\Rightarrow f = 0 \\ \|f - g\|_2 = 0 & \not\Rightarrow f = g \end{cases}$$

in $R[-\pi, \pi]$. It is only true for almost everywhere. Also, although it is not hard to show that $f_n \to f$ uniformly implies $\|f_n - f\|_2 \to 0$, its converse does not hold. Consider the following counterexample:

Example 1.36. Let

$$f_n(x) = \begin{cases} 1 & \text{if } x \in [0, 1/n] \\ 0 & \text{otherwise} \end{cases}$$

be a function, then $\|f_n\|_2^2 = \int_{-\pi}^{\pi} f_n^2 = 1/n$, which tends to 0 as $n \to \infty$. In this way, $f_n \to 0$ in L^2-sense. However, $f_n \not\to 0$ uniformly or even pointwisely.

1.7 Applications to Fourier Series

1.7.1 Minimizers

Consider the functions on $[-\pi, \pi]$

$$\begin{cases} \varphi_0 = \frac{1}{\sqrt{2\pi}} \\ \varphi_n = \frac{1}{\sqrt{\pi}} \cos(nx) \\ \psi_n = \frac{1}{\sqrt{\pi}} \sin(nx) \end{cases}$$

Note that

$$\begin{cases} \langle \varphi_m, \varphi_n \rangle_2 = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} \\ \langle \psi_m, \psi_n \rangle_2 = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} \\ \langle \varphi_m, \psi_n \rangle_2 = 0 \quad \text{for all } m, n \end{cases}$$

therefore

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos(nx), \frac{1}{\sqrt{\pi}} \sin(nx) \right\}_{n=1}^{\infty}$$

can be regarded as an orthogonal basis in $R[-\pi, \pi]$.

Definition 1.37. The $(2n+1)$ dimensional vector subspace of $R[-\pi, \pi]$ spanned by the first $(2n+1)$ trigonometric functions, denoted by E_n, is defined by

$$E_n = \operatorname{span} \left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos(kx), \frac{1}{\sqrt{\pi}} \sin(kx) \right\}_{k=1}^{n}$$

In general, if there is an orthogonal set (or ortogonal family) $\{\phi_n\}_{n=1}^{\infty}$ in $R[-\pi, \pi]$, let

$$S_n = \operatorname{span} \langle \phi_1, \phi_2, \cdots, \phi_n \rangle$$

be an n-dimensional subspace spanned by the first n functions in the orthogonal set, then for any $f \in R[-\pi, \pi]$, the **minimization problem** is

$$\inf \left\{ \|f - g\|_2 \mid g \in S_n \right\}$$

Proposition 1.38. The unique minimizer of

$$\inf \left\{ \|f - g\|_2 \mid g \in S_n \right\}$$

is attained at the function

$$g = \sum_{k=1}^{n} \langle f, \phi_k \rangle_2 \, \phi_k \in S_n$$

Proof. Note that to minimize $\|f - g\|_2$ is equivalent to minimize $\|f - g\|_2^2$. For all $g \in S_n$,

$$g = \sum_{k=1}^{n} \beta_k \phi_k \Rightarrow \|f - g\|_2^2 = \int_{-\pi}^{\pi} \left| f - \sum_{k=1}^{n} \beta_k \phi_k \right|^2$$

Let $\Phi(\beta) = \|f - g\|_2^2$, then

$$\begin{aligned}
\Phi(\beta) &= \int_{-\pi}^{\pi} \left| f - \sum_{k=1}^{n} \beta_k \phi_k \right|^2 \\
&= \left(\int_{-\pi}^{\pi} f^2 \right) - 2 \sum_{k=1}^{\infty} \left(\frac{\beta_k}{\sqrt{2}} \right) \left(\sqrt{2} \left\langle f, \phi_k \right\rangle_2^2 \right) + \sum_{k=1}^{n} \beta_k^2 \\
&\geq \left(\int_{-\pi}^{\pi} f^2 \right) - \sum_{k=1}^{\infty} \left(\frac{\beta_k^2}{2} + 2 \left\langle f, \phi_k \right\rangle_2^2 \right) + \sum_{k=1}^{n} \beta_k^2 \quad \text{since } 2ab \leq a^2 + b^2 \\
&= \left(\int_{-\pi}^{\pi} f^2 \right) - 2 \sum_{k=1}^{n} \left\langle f, \phi_k \right\rangle_2^2 + \frac{1}{2} \sum_{k=1}^{n} \beta_k^2 \to \infty
\end{aligned}$$

as

$$\|\beta\| = \sqrt{\sum_{k=1}^{n} \beta_k^2} \to \infty$$

Hence $\Phi(\beta)$ attains its minimum at some finite point β. By some calculus, the minimum required is given by $\beta_k = \left\langle f, \phi_k \right\rangle_2$ for all $1 \leq k \leq n$.

Note that the minimizer g of $\|f - g\|_2$ over S_n is called the **orthogonal projection** of f on S_n, denoted by $P_n(f)$. With the notation of orthogonal projection,

$$\operatorname{dist}(f, S_n) = \|f, P_n(f)\|_2$$

Corollary. For a 2π-periodic function f integrable on $[-\pi, \pi]$ and $n \geq 1$, $\|f - S_n(f)\|_2 \leq \|f - g\|_2$ where $S_n(f)$ represents the n-th partial sum of the Fourier series of f, for all

$$g = \alpha_0 + \sum_{k=1}^{n} (\alpha_k \cos(kx) + \beta_k \sin(kx))$$

with real coefficients.

Proof. By the definition of Fourier coefficients $S_n(f) = P_n(f)$ of E_n,

$$\begin{cases}
a_0 = \frac{1}{\sqrt{2\pi}} \left\langle f, \frac{1}{\sqrt{2\pi}} \right\rangle_2 \\
a_n \cos(nx) = \frac{1}{\sqrt{\pi}} \left\langle f, \frac{1}{\sqrt{\pi}} \cos(nx) \right\rangle_2 \cos(nx) \\
b_n \sin(nx) = \frac{1}{\sqrt{\pi}} \left\langle f, \frac{1}{\sqrt{\pi}} \sin(nx) \right\rangle_2 \sin(nx)
\end{cases}$$

1.7.2 Measure Zeroes of Fourier Series

Theorem 1.39. Let f be 2π-periodic integrable function on $[-\pi,\pi]$, then the n-th partial sum of the Fourier series of f converges to f in L^2-sense. In other words,

$$\lim_{n\to\infty} \|S_n(f)-f\|_2 = 0$$

Proof. For any $\epsilon > 0$, there exists a 2π-periodic Lipschitz continuous function g such that $\|f-g\|_2 < \epsilon/2$. This can be achieved by finding a step function approximating f. By *Proposition 1.28*, there exists $N > 0$ such that

$$\|g - S_N(g)\|_\infty < \frac{\epsilon}{2\sqrt{2\pi}}$$

where $\|\cdot\|_\infty$ represents the uniform convergence. This induces

$$\|g-S_N(g)\|_2 = \sqrt{\int_{-\pi}^{\pi} (g-S_N(g))^2} \le \sqrt{2\pi \|g-S_n(g)\|_\infty^2} = \frac{\epsilon}{2}$$

By the corollary of *Proposition 1.38*,

$$\|f-S_N(f)\|_2 \le \|f-S_N(g)\|_2 \le \|f-g\|_2 + \|g-S_n(g)\|_2 < \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon$$

Finally, since $E_N \subset E_n$ for all $n \ge N$,

$$\|f-S_n(f)\|_2 \le \|f-S_N(f)\|_2 < \epsilon$$

for any $n \ge N$, thus

$$\lim_{n\to\infty} \|S_n(f)-f\|_2 = 0$$

Corollary. Let f_1 and f_2 be 2π-periodic integrable functions on $[-\pi,\pi]$ with the same Fourier series, then $f_1 = f_2$ almost everywhere, or $f_1 = f_2$ except a set of measure zero. Furthermore, if f_1 and f_2 are both continuous on $[-\pi,\pi]$, then $f_1 = f_2$.

Proof. Let $f = f_1 - f_2$, then $a_n(f) = b_n(f) = 0$ gives $S_n(f) = 0$ for any $n \ge 0$. Therefore

$$\lim_{n\to\infty} \|S_n(f)-f\|_2 = 0 \Rightarrow \|f\|_2 = 0$$

and by theory of Riemann integrals, $f = 0$ almost everywhere. If f_1, f_2 are continuous, $f^2 \text{cts} \ge 0 \Rightarrow f^2 \equiv 0$.

Note that a set E is said to be measure zero if for any $\epsilon > 0$, there exists countably many intervals $\{I_k\}$ such that $E \subset \bigcup_k I_k$ and $\sum_k |I_k| < \epsilon$.

1.7.3 Parserval's Identity

Below is the **Parserval's Identity**:

Proposition 1.40. Let f be a 2π-periodic function f integrable on $[-\pi, \pi]$, then

$$\|f\|_2^2 = 2\pi a_0^2 + \pi \sum_{n=1}^{\infty} (a_n^2 + b_n^2)$$

where a_0, a_n, b_n are Fourier coefficients of f.

Proof. Note that

$$\begin{cases} \sqrt{2\pi} a_0 = \left\langle f, \frac{1}{\sqrt{2\pi}} \right\rangle_2 \\ \sqrt{\pi} a_n = \left\langle f, \frac{1}{\sqrt{\pi}} \cos(nx) \right\rangle_2 & \text{for all } n \geq 1 \\ \sqrt{\pi} b_n = \left\langle f, \frac{1}{\sqrt{\pi}} \sin(nx) \right\rangle_2 & \text{for all } n \geq 1 \end{cases}$$

then by corollary of *Proposition 1.38*,

$$\begin{aligned} \langle f, S_N(f) \rangle_2 &= \langle (f - S_N(f)) + S_N(f), S_N(f) \rangle_2 \\ &= \langle S_N(f), S_N(f) \rangle_2 \quad \text{since } (f - S_N(f)) \text{ is orthogonal} \\ &= \int_{-\pi}^{\pi} \left(a_0 + \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \sin(kx) \right)^2 \mathrm{d}x \end{aligned}$$

Finally by *Theorem 1.39*,

$$\begin{aligned} 0 &= \lim_{N \to \infty} \|f - S_N(f)\|_2^2 \\ &= \lim_{N \to \infty} (\|f\|_2^2 - 2 \langle f, S_N(f) \rangle_2 + \|S_N(f)\|_2^2) \\ &= \lim_{N \to \infty} (\|f\|_2^2 - 2 \|S_N(f)\|_2^2 + \|S_N(f)\|_2^2) \\ &= \lim_{N \to \infty} (\|f\|_2^2 - \|S_N(f)\|_2^2) \end{aligned}$$

therefore

$$\|f\|_2^2 = \lim_{N \to \infty} \left(2\pi a_0^2 + \pi \sum_{n=1}^{N} (a_n^2 + b_n^2) \right)$$

2 Metric Spaces

2.1 Introduction of Metric Spaces

2.1.1 Definition of Metrics Spaces

Definition 2.1. Let X be a nonempty set, then a **metric** on X is a function

$$d : X \times X \to [0, +\infty)$$

such that the following properties is satisfied for all $x, y, z \in X$:

(a) Metric is nonnegative, or $d(x, y) \geq 0$. Equality holds if and only if $x = y$.

(b) Metric is symmetric, or $d(x, y) = d(y, x)$.

(c) Metric is subadditive (satisfies triangle inequality), or $d(x, y) \leq d(x, z) + d(z, y)$.

The pair (X, d) is called a **metric space**.

Definition 2.2. Let (X, d) be a metric space, then the **metric ball** of radius r centered at x, denoted by $B_r(x)$, is defined as

$$B_r(x) = \{y \in X \mid d(x, y) < r\}$$

2.1.2 Examples of Metric Spaces

Example 2.3. Below are some examples of metric spaces:

(a) $(\mathbb{R}, |x - y|)$ is a metric space.

(b) Let $X = \mathbb{R}^n$. Denote the metrics

$$\begin{cases} d_k(x, y) = \sqrt[k]{\sum |x_i - y_i|^k} \\ d_\infty(x, y) = \max |x_i - y_i| \end{cases}$$

where $1 \leq i \leq n$, then $(\mathbb{R}^n, d_1)$, $(\mathbb{R}^n, d_2)$, $(\mathbb{R}^n, d_\infty)$ are metric spaces.

(c) Let $C[a, b]$ be the set of all (real) continuous functions on $[a, b]$ and

$$\begin{cases} d_k(f, g) = \sqrt[k]{\int_a^b |f - g|^k} \\ d_\infty(f, g) = \max \{|f(x) - g(x)| \mid x \in [a, b]\} \end{cases}$$

for all $f, g \in C[a, b]$, then $(C[a, b], d_1)$, $(C[a, b], d_2)$, $(C[a, b], d_\infty)$ are metric spaces.

Example 2.4. Let $X = R[a, b]$ be the set of Riemann integrable functions on $[a, b]$ and

$$d_1(f, g) = \int_a^b |f - g|$$

However, part (a) of *Definition 2.1* is not satisfied since $d_1(f, g) = 0$ only implies $f = g$ almost everywhere, but not exactly $f = g$. d_1 is then not a suitable metric on $R[a, b]$.

In order to fix this problem, consider $X = R[a, b]/\sim$ where $\sim$ is an equivalent relation on $R[a, b]$ defined by

$$f \sim g \Leftrightarrow f = g \text{ almost everywhere}$$

Denote

$$\overline{f} = \{g \in R[a, b] \mid f \sim g\}$$

and its corresponding metric

$$\tilde{d_k}(\overline{f}, \overline{g}) = d_k(f, g)$$

then $(X, \tilde{d_1})$ and $(X, \tilde{d_2})$ are metric spaces.

Note that $\tilde{d_2}$ in the example above is in fact L^2-distance defined in the last section.

2.1.3 Normed Spaces

Definition 2.5. Let X be a nonempty set, then a **norm** on X is a function

$$\|\cdot\| : X \to [0, +\infty)$$

such that the following properties is satisfied for all $x, y \in X$ and $\alpha \in \mathbb{R}$:

(a) Norm is nonnegative, or $\|x\| \geq 0$. Equality holds if and only if $x = 0$.

(b) Norm is absolutely scalable, or $\|\alpha x\| = |\alpha| \, \|x\|$.

(c) Norm is subadditive, or $\|x + y\| \leq \|x\| + \|y\|$.

The pair $(X, \|\cdot\|)$ is called a **normed space**. Furthermore, a metric d is said to be **induced** by the norm $\|\cdot\|$ if $d(x,y) = \|x - y\|$.

Example 2.6. Below are some examples of norms:

(a) Let $\|x\|_k = \sqrt[k]{\sum |x_i|^k}$ and $\|x\|_\infty = \max\{x_i\}$, then $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$ are norms on $\mathbb{R}^n$.

(b) Let $\|f\|_k = \sqrt[k]{\int_a^b |f|^k}$ and $\|f\|_\infty = \max\{f(x)\}$, then $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$ are norms on $C[a,b]$.

Note that a norm can induce a metric, but not all metrics are induced from norm.

Example 2.7. Let X be a nonempty set and

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

be a metric on X. Note that X is not necessary a vector space, so d is not induced by a norm. Moreover, even X is a vector space,

$$d(\alpha x, \alpha y) \neq |\alpha|\, d(x,y)$$

when $|\alpha| \neq 1$ and $x \neq y$.

Such metric d in the above example is called a **discrete metric** on X.

2.1.4 Metric Subspaces

Definition 2.8. Let (X,d) be a metric space, then for any nonempty set $Y \subset X$, (Y,d) is called a **metric subspace** of (X,d).

Note that a metric subspace of a normed space may not be also a normed space, only if the subset is also a vector subspace.

2.2 Limits and Continuity

2.2.1 Limits and Convergence in Metric Spaces

With the understanding of metric spaces, one can extend the definition of limits and convergence to any metric space:

Definition 2.9. Let $\{x_n\}$ be a sequence in a metric space (X, d), then the sequence is said to be **converge** to $x \in X$, denoted by $x_n \to x$, if

$$\lim_{n\to\infty} d(x_n, x) = 0$$

Proposition 2.10. Let $\{x_n\}$ be a sequence in a metric space (X, d). If $x_n \to x$ and $x_n \to y$, then $x = y$.

Example 2.11. Below are some examples on convergence in metric spaces:

(a) Convergence in $(\mathbb{R}^n, d_2)$ is the usual convergence in advanced calculus.

(b) Convergence in $(C[a, b], d_\infty)$ is the uniform convergence of sequence of functions in $C[a, b]$.

2.2.2 Strength of Convergence

There are many metrics suitable for the same nonempty set X, so it is natural to think of comparing among those metrics.

Definition 2.12. Let d and ρ be different metrics defined on X, then ρ is said to be **stronger than** d (or d is **weaker than** ρ) if there exists a constant $C > 0$ such that

$$d(x, y) \leq C\rho(x, y)$$

for all $x, y \in X$. d and ρ are **equivalent** to each other if d is stronger and weaker than ρ at the same time. In other words, there exists $C_1, C_2 > 0$ such that

$$d(x, y) \leq C_1\rho(x, y) \leq C_2 d(x, y)$$

for all $x, y \in X$.

Note that the equivalence of metrics defined above is an equivalent relation.

Proposition 2.13. Let d and ρ be different metrics defined on X. If ρ is stronger than d and a sequence $\{x_n\}$ converges in (X, ρ), then the sequence also converges in (X, d) with the same limit. If ρ is equivalent to d, then $\{x_n\}$ converges in (X, ρ) if and only if it converges in (X, d) also.

Example 2.14. Recall the metrics d_1, d_2 and d_∞ on $\mathbb{R}^n$, then

$$\begin{cases} d_1(x,y) \le n d_\infty(x,y) \le n d_1(x,y) \\ d_2(x,y) \le \sqrt{n} d_\infty(x,y) \le \sqrt{n} d_2(x,y) \end{cases}$$

shows that d_1, d_2 and d_∞ are equivalent metrics.

Example 2.15. Recall the metrics d_1 and d_∞ on $C[a,b]$, then

$$d_1(f,g) \le (b-a) d_\infty(f,g)$$

shows that d_∞ is stronger than d_1. However, d_1 is not stronger than d_∞, so d_1 and d_∞ are not equivalent.

2.2.3 Continuity in Metric Spaces

Definition 2.16. Let $f : (X,d) \to (Y,\rho)$ be a mapping between two metric spaces, then f is **continuous** at a point $x \in X$ if $f(x_n) \to f(x)$ in (Y,ρ) whenever $x_n \to x$ in (X,d). f is continuous on a set $E \in X$ if it is continuous at every point in E.

Proposition 2.17. Let $f : (X,d) \to (Y,\rho)$ be a mapping between two metric spaces and $x_0 \in X$ be a point, then f is continuous at x_0 if and only if for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$\rho(f(x), f(x_0)) < \epsilon \quad \text{for all } \{x \in X \mid d(x,x_0) < \delta\}$$

Proposition 2.18. Let $f : (X,d) \to (Y,\rho)$ and $g : (Y,\rho) \to (Z,m)$ be mappings between metric spaces, then if f is continuous at x and g is continuous at $f(x)$, then $g \circ f$ is also continuous at x. Similarly, if f is continuous at X and g is continuous at Y, then $g \circ f$ is also continuous at X.

Example 2.19. Let (X,d) be a metric space and $A \subset X$ be a nonempty set. Further define $\rho_A : X \to \mathbb{R}$ by

$$\rho_A(x) = \inf_{y \in A} d(y,x)$$

which is the shortest distance from x to the subset A. Show that

$$|\rho_A(x) - \rho_A(y)| \le d(x,y)$$

for any $x, y \in X$.

Answer. For fixed $x, y \in X$, along with the definition of ρ_A, for all $\epsilon > 0$, there exists $z \in A$ such that $\rho_A(y) + \epsilon > d(z, y)$. Hence

$$\rho_A(x) \leq d(z, x) \leq d(z, y) + d(y, x) < d(y, x) + \rho_A(y) + \epsilon$$

rearranging the equation gives

$$\rho_A(x) - \rho(A)(y) < d(x, y) + \epsilon$$

Note that since x and y are interchangable, and ϵ is arbitrary,

$$|\rho_A(x) - \rho_A(y)| \leq d(x, y)$$

In fact the example above shows that ρ_A is continuous (and even Lipschitz continuous) since $d(x_n, x) \to 0$ implies $\rho_A(x_n) \to \rho_A(n)$. This actually mean there are many continuous functions on a metric space.

For simplicity, define

$$\begin{cases} d(x, F) = \inf\{d(x, y) \mid y \in F\} \\ d(E, F) = \inf\{d(x, y) \mid x \in E, y \in F\} \end{cases}$$

for subsets E and F.

2.3 Open and Closed Sets

2.3.1 Open Sets

Definition 2.20. Let (X, d) be a metric space, then a set $G \in X$ is called an **open set** if for any $x \in G$, there exists $\epsilon > 0$ such that

$$B_\epsilon(x) = \{y \mid d(x, y) < \epsilon\} \subset G$$

Note that ϵ may vary depending on the choice of x, and the empty set ϕ is considered an open set. Therefore, the proposition applies:

Proposition 2.21. Let (X, d) be a metric space and G_α be a collection of open sets, then the following are true:

(a) X and ϕ are open sets.

(b) Arbitrary union of open sets $\bigcup_\alpha G_\alpha$ is an open set.

(c) Finite intersection of open sets $\bigcap_{i=1}^{n} G_i$ is an open set.

2.3.2 Closed Sets

Definition 2.22. Let (X, d) be a metric space, then a set $F \in X$ is called an **closed set** if $X \setminus F$ is an open set.

Proposition 2.23. Let (X, d) be a metric space and F_α be a collection of closed sets, then the following are true:

(a) X and ϕ are closed sets.

(b) Finite union of closed sets $\bigcup_{j=1}^{n} F_j$ is an closed set.

(c) Arbitrary intersection of closed sets $\bigcap_\alpha F_\alpha$ is an closed set.

Corollary. Let (X, d) be a metric space, then X and ϕ are both open and closed.

2.3.3 Applications of Open and Closed Sets

Proposition 2.24. Let (X, d) be a metric space, then a sequence $\{x_n\}$ converges to x if and only if for all open set G containing x, there exists n_0 such that $x_n \in G$ for all $n \geq n_0$.

Proposition 2.25. Let (X, d) be a metric space, then a set $A \subset X$ is closed if and only if whenever $\{x_n\} \subset A$ and $x_n \to x$ as $n \to \infty$ implies that $x \in A$.

Proposition 2.26. Let $f : (X, d) \to (Y, \rho)$ be a mapping between metric spaces, then the following applies:

(a) f is continuous at x if and only if for all open set $G \subset Y$ containing $f(x)$, $f^{-1}(G)$ contains $B_\epsilon(x)$ for some $\epsilon > 0$.

(b) f is continuous at x if and only if for all open set $G \subset Y$, $f^{-1}(G)$ is open in X.

In this case, f is also continuous at x if and only if for all closed set $F \subset Y$, $f^{-1}(F)$ is closed in X.

2.4 Points in Metric Space

2.4.1 Boundary Points and Closures

Definition 2.27. Let E be a set in a metric space (X, d), then a point $x \in X$ (which is not necessary in E) is called a **boundary point** of E if for all open set $G \subset X$ containing x,

$$G \cap E \neq \phi \text{ and } G \setminus E \neq \phi$$

In other words, this is satisfied when $G \cap (X \setminus E) \neq \phi$. The **boundary** of E, denoted by ∂E, is the set of boundary points of E. The **closure** of E, denoted by $\overline{E}$, is defined as $\overline{E} = E \cup \partial E$.

For the conditions of a boundary point in the definition above, it suffices to check G of the form $B_\epsilon(x)$ for all small $\epsilon > 0$, or even $B_{1/n}(x)$ for all $n \geq 1$. Also note that X and $X \setminus E$ shares the same boundary no matter the choice of E, or

$$\partial E = \partial X \setminus E \quad \text{for all } E \subset X$$

2.4.2 Properties of Boundaries and Closures

Proposition 2.28. Below are some properties of boundaries and closures:

(a) The boundary of an empty set is an empty set, or $\partial\phi = \phi$.

(b) For all $E \subset X$, ∂E is a closed set.

(c) If E is a closed set, $\overline{E} = E$.

Proposition 2.29. Let E be a subset of a metric space (X, d), then the following applies:

(a) $x \in \overline{E}$ if and only if $B_r(x) \cap E \neq \phi$ for all $r > 0$.

(b) If $A \subset B$, $\overline{A} \subset \overline{B}$ for all $A, B \subset (X, d)$.

(c) $\overline{E}$ is closed.

(d) $\overline{E}$ is the smallest closed set containing E, or $\overline{E} = \cap\{G \subset E \mid G \text{ is closed}\}$.

2.4.3 Interior Points

Definition 2.30. Let E be a subset of a metric space (X, d), then a point x is called an **interior point** of E if there exists an open set G such that $x \in G$ and $G \subset E$. The **interior** of E, denoted by E^0, is the set of interior points of E.

Proposition 2.31. Below are the properties of interiors:

(a) Interior of E, E^0 is open.

(b) Interior of E is the set without boundary, or $E^0 = E \setminus \partial E$.

(c) Interior of E, $E^0 = X \setminus \overline{X \setminus E}$.

(d) Interior of E, $E^0 = \cup \{G \subset E \mid G \text{ is open}\}$.

2.5 Elementary Inequalities for Functions

2.5.1 Young's Inequality

Theorem 2.32. By **Young's Inequality**, for $a, b > 0$ and $p > 1$,

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \quad \text{with } \frac{1}{p} + \frac{1}{q} = 1$$

Equality holds when $a^p = b^q$.

Note that $q = \frac{p}{p-1} > 1$ is called the **conjugate** of p. Also specifically if $p = 2$, the inequality reduces to $2ab \leq a^2 + b^2$.

2.5.2 Holder's Inequality

For the following inequality, denote the norm

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{1/p}$$

Theorem 2.33. Let $f, g \in R[a, b]$ be Riemann integrable functions and $p > 1$, then by **Holder's Inequality**,

$$\int_a^b |f(x)g(x)| \, \mathrm{d}x \leq \left(\int_a^b |f(x)|^p \, \mathrm{d}x\right)^{1/p} \left(\int_a^b |f(x)|^q \, \mathrm{d}x\right)^{1/q}$$

where q is the conjugate of p. Equality holds when one of the following conditions is satisfied:

(a) f or g equals to 0 almost everywhere.

(b) There exists a constant $\lambda > 0$ such that $|g(x)|^q = \lambda |f(x)|^p$ almost everywhere.

Note that Holder's Inequality can be written in norm form $\|fg\|_1 = \|f\|_p \|g\|_q$. Holder's Inequality also holds for limiting cases $(p, q) \to (1, \infty)$ and $(p, q) \to (\infty, 1)$.

2.5.3 Minkowski's Inequality

Theorem 2.34. By **Minkowski's Inequality**, for any $f, g \in R[a, b]$ and $p > 1$,

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

Equality holds when one of the following conditions is satisfied:

(a) f or g equals to 0 almost everywhere.

(b) $\|f\|_p, \|g\|_p > 0$ and there exists a constant $\lambda > 0$ such that $g(x) = \lambda f(x)$ almost everywhere.

3 Contraction Mapping Principle

3.1 Complete Metric Space

3.1.1 Definition of Complete Metric Space

Definition 3.1. Let (X, d) be a metric space, then a sequence $\{x_n\}$ in (X, d) is a **Cauchy sequence** if for any $\epsilon > 0$, there exists n_0 such that $d(x_n, x_m) < \epsilon$ for all $n, m > n_0$.

Definition 3.2. Let (X, d) be a metric space, then the metric space is **complete** if every Cauchy sequence in the metric space converges. A subset E is complete if the induce metric subspace (E, d) with $d = d|_{E \times E}$ is complete. In other words, every Cauchy sequence in E converges with limit in E.

Note that convergent sequence is a Cauchy sequence.

Proposition 3.3. Let (X, d) be a metric space, then the following applies:

(a) Every complete set in X is closed.

(b) If X is complete, then every closed set in X is complete.

Example 3.4. Below are the examples of complete metric space:

(a) $(\mathbb{R}, \text{standard})$ is complete.

(b) $[a, b]$, $(-\infty, b]$ and $[a, \infty)$ are complete.

Below are the counterexamples of complete metric space:

(a) $[a, b)$ where b is finite, is not complete since $x_n = b - 1/n \to b \notin [a, b)$.

(b) $\mathbb{Q}$ is not complete.

3.1.2 Completion of Metric Spaces

Definition 3.5. A metric space (X, d) is said to be **isometrically embedded** in metric space (Y, ρ) if there exists a mapping $\Phi : X \to Y$ such that $d(x, y) = \rho(\Phi(x), \Phi(y))$. If such mapping exists, Φ is called an **isometric embedding** (or a **metric preserving map**) from (X, d) to (Y, ρ).

Note that Φ must be injective and continuous.

Definition 3.6. Let (X, d) and (Y, ρ) be metric spaces, then (Y, ρ) is called a completion of (X, d) if the following statements are satisfied:

(a) (Y, ρ) is complete.

(b) There exists an isometric embedding Φ such that the closure $\overline{\Phi(X)} = Y$.

Example 3.7. Let $(X, d) = (\mathbb{Q}, \text{induced metric})$ and $(Y, \rho) = (\mathbb{R}, \text{standard})$. Since $\mathbb{Q} \subset \mathbb{R}$, (Y, ρ) is complete. Further let $\Phi : (X, d) \to (Y, \rho)$ where $\Phi(q) = q$, since $\mathbb{Q}$ is dense in $\mathbb{R}$, $\overline{\Phi(\mathbb{Q})} = \overline{\mathbb{Q}} = \mathbb{R}$. Therefore, (Y, ρ) is a completion of (X, d).

Theorem 3.8. Every metric space has a completion.

Note that the definition of isometric embedding can be extended to bijection as below:

Definition 3.9. Let (X, d) and (Y, ρ) be metric spaces, then they are called **isometric** if there exists a bijective isometric embedding between (X, d) and (Y, ρ).

Note that the inverse of a bijective isometric embedding is also an isometric embedding. Also, the two metric spaces are regarded as the same if they are isometric.

Theorem 3.10. If metric spaces (Y, ρ) and (Y', ρ') are both completions of a metric space (X, d), then (Y, ρ) and (Y', ρ') are isometric. In other words, completion is unique up to isometry.

3.2 Introduction to Contraction Mapping Principle

3.2.1 General Theorem

Definition 3.11. Let (X, d) be a metric space, then a map $T : (X, d) \to (X, d)$ is called a **contraction** if there exists a constant $\gamma \in (0, 1)$ such that

$$d(Tx, Ty) \leq \gamma d(x, y)$$

for all $x, y \in X$. A point $x \in X$ is called a **fixed point** of T if $Tx = x$.

Note that Tx is the notation for $T(x)$ but not the multiplication. With the definition of a contraction, below is the **Contraction Mapping Principle** (or the **Banach Fixed Point Theorem**):

Theorem 3.12. Every contraction in a complete metric space admit a fixed point.

3.2.2 Perturbation of Identity

Definition 3.13. A normed space $(X, \|\cdot\|)$ is a **Banach space** if it is complete as a metric space with respect to the induced metric $d(x, y) = \|x - y\|$ for all $x, y \in X$.

Example 3.14. Below are some examples of Banach space:

(a) $(\mathbb{R}^n, \|\cdot\|_p)$ is a Banach space if $p > 1$.

(b) $(C[a,b], \|\cdot\|_\infty)$ is a Banach space.

Theorem 3.15. Let $(X, \|\cdot\|)$ be a Banach space, and $\Phi : \overline{B_r(x_0)} \to X$ satisfies $\Phi(x_0) = y_0$. Suppose that $\Phi = \mathrm{Id}_X + \Psi$ such that there exists a constant $\gamma \in (0,1)$ such that

$$\|\Psi(x_2) - \Psi(x_1)\| \leq \gamma \|x_2 - x_1\|$$

for all $x_1, x_2 \in \overline{B_r(x_0)}$, then by **Perturbation of Identity**, for all $y \in \overline{B_R(y_0)}$ where $R = (1-\gamma)r$, there exists unique $x \in \overline{B_r(x_0)}$ such that $\Phi(x) = y$.

Example 3.16. Show that $3x^4 - x^2 + x = -0.05$ has a real root.

Proof. Notice that $3x^4 - x^2 + x = 0$ has a root $x = 0$. Let $\Phi(x) = x + \Psi(x)$ where $\Psi(x) = 3x^4 - x^2$, then $\Phi(0) = 0$. For $x_1, x_2 \in \overline{B_r(0)}$,

$$\begin{aligned}
|\Psi(x_1) - \Psi(x_2)| &= \left|3x_1^4 - x_1^2 - 3x_2^4 + x_2^2\right| \\
&= \left|3(x_1^4 - x_2^4) - (x_1^2 - x_2^2)\right| \\
&= \left|3(x_1^3 + x_1^2 x_2 + x_2^2 x^1 + x_2^3) - (x_1 + x_2)\right| |x_1 - x_2| \\
&= \left|12r^3 + 2r\right| |x_1 - x_2|
\end{aligned}$$

Choose $r > 0$ such that $\gamma = 12r^3 + 2^r < 1$ and $R = (1-\gamma)r \geq 0.05$ so that $-0.05 \in \overline{B_R(0)}$. Pick $r = 1/4$, then $\gamma = 11/16$ and $R = 5/64$. By Perturbation of Identity, for all $y \in \overline{B_R(0)}$, there exists $x \in \overline{B_r(0)}$ such that $\Phi(x) = y$. Therefore, there exists a real root for $3x^4 - x^2 + x = -0.05$ since $-0.05 \in \overline{B_R(0)}$.

The example above can be summarized into the following proposition:

Proposition 3.17. Let $\Phi(x) = x + \Psi(x)$ where $\Psi(x) : U \to \mathbb{R}^n$ be a C^1-function on some open set $U \subset \mathbb{R}^n$ containing 0, such that

$$\Psi(0) = 0 \text{ and } \lim_{x \to 0} \frac{\partial \Psi_i}{\partial x_j}(x) = 0$$

for all i, j, then there exists $r > 0$ and $R > 0$ such that for all $y \in B_R(0)$, $\Phi(x) = y$ has a unique solution $x \in B_r(0)$.

3.3 Inverse Function Theorem

3.3.1 Introduction to Inverse Function Theorem

Recall the chain rule: let $G : U \subset \mathbb{R}^n \to \mathbb{R}^m$ and $F : V \subset \mathbb{R}^m \to \mathbb{R}^l$ be differentiable functions where U, V open in $\mathbb{R}^n, \mathbb{R}^m$ respectively, and $G(U) \subset V$. Then $H = F \circ G :$

$U \to \mathbb{R}^l$ differentiable and $DH(x) = DF(G(x))DG(x)$ where

$$DG(x) = \left(\frac{\partial G_i}{\partial x_j}(x)\right)_{i,j}$$

and similarly for DF and DH. Besides the proposition is required:

Proposition 3.18. Let $F : B \to \mathbb{R}^n$ be C^1 function, where B is a ball in $\mathbb{R}^n$, then for any $x_1, x_2 \in B$,

$$F(x_1) - F(x_2) = \left(\int_0^1 DF(x_2 + t(x_1 - x_2))\,\mathrm{d}t\right) \cdot (x_1 - x_2)$$

in component form $F = (F_1, \cdots, F_n)$. In other words,

$$F_i(x_1) - F_i(x_2) = \sum_{j=1}^{n} \left(\int_0^1 \frac{\partial F_i}{\partial x_j}(x_2 + t(x_1 - x_2))\,\mathrm{d}t\right)(x_1 - x_2)_j$$

Finally, recall that if $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at a point p in an open set U of $\mathbb{R}^n$, then

$$F(p + x) - F(p) = DF(p)x + o(|x|)$$

for all $x = (x_1, \cdots, x_n)$ sufficiently small (or $|x|$ small) where $o(|x|)$ is a remaining term such that

$$\frac{o(|x|)}{|x|} \to 0 \text{ as } |x| \to 0$$

Definition 3.19. The condition in Inverse Function Theorem that $DF(x_0)$ is invertible is called the **nondegeneracy condition**.

Note that nondegeneracy condition is necessary for the differentiability of local inverse.

Proposition 3.20. Let $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ be a C^1 function where U is an open set and $x_0 \in U$. Suppose there exists open V such that $x_0 \in V \subset U$ and $F|_V$ has a differentiable inverse, then $DF(x_0)$ is nonsingular (or invertible).

Below is the **Inverse Function Theorem**:

Theorem 3.21. Let $F : U \to \mathbb{R}^n$ be a C^1 map from an open set $U \to \mathbb{R}^n$. Suppose $x_0 \in U$ and $DF(x_0)$ is invertible (as a matrix or linear transformation), then there exists open sets V, W containing $x_0, F(x_0)$ respectively such that the restriction of F on V is a bijection onto W with a C^1 inverse.

Moreover, the inverse is C^k when F is C^k where $1 \leq k \leq \infty$, in V.

Proof. *Part (a)* Consider the special case where $x_0 = 0, y_0 = F(x_0) = F(0) = 0$, then $DF(0) = I$, which is the identity. Let $\Psi(x) = -x + F(x)$. As $0 \in U$ and U is open, there exists $r_0 > 0$ such that $\overline{B_{r_0}(0)} \subset U$. Then

$$\Psi(x_1) - \Psi(x_2) = -x_1 + F(x_1) + x_2 - F(x_2)$$

By *Proposition 3.18*,

$$\begin{aligned}\Psi(x_1) - \Psi(x_2) &= \left(\int_0^1 DF(x_2 + t(x_1 - x_2))\,\mathrm{d}t\right) \cdot (x_1 - x_2) - (x_1 - x_2) \\ &= \left(\int_0^1 DF(x_2 + t(x_1 - x_2))\,\mathrm{d}t - I\right) \cdot (x_1 - x_2) \\ &= \left(\int_0^1 DF(x_2 + t(x_1 - x_2)) - DF(0)\,\mathrm{d}t\right) \cdot (x_1 - x_2)\end{aligned}$$

Since F is C^1, for all $\epsilon > 0$, there exists $0 < r \leq r_0$ such that

$$\|DF(x) - DF(0)\| < \epsilon$$

for all $x \in \overline{B_r(0)}$, where

$$\|(b_{ij})\| = \sqrt{\sum_{i,j} b_{ij}^2}$$

for any $n \times n$ matrix (b_{ij}).

Since $\overline{B_r(0)}$ is convex, $x_1, x_2 \in \overline{B_r(0)}$ implies $x_2 + t(x_1 - x_2) \in \overline{B_r(0)}$. Hence for all $\epsilon > 0$, there exists $0 < r \leq r_0$ such that

$$\|DF(x_2 + t(x_1 - x_2)) - DF(0)\| < \epsilon$$

for all $x_1, x_2 \in \overline{B_r(0)}$ and $t \in (0, 1)$. Therefore choosing $\epsilon = 1/2$ gives

$$|\Psi(x_1) - \Psi(x_2)| \leq \frac{1}{2}|x_1 - x_2|$$

for all $x_1, x_2 \in \overline{B_r(0)}$.

Part (b) Choose $r > 0$ as in Part (a), then for all $y \in B_{r/2}(0)$, there exists $x \in B_r(0)$ such that $F(x) = y$. This is true because of Perturbation of Identity (*Theorem 3.15*) with $\epsilon = 1/2$. The local inverse G of F,

$$G : B_{r/2}(0) \to G(B_{r/2}(0)) \subset B_r(0)$$

satisfies

$$|G(y_1) - G(y_2)| \leq \frac{1}{1-\epsilon} |y_1 - y_2| = 2 |y_1 - y_2|$$

for all $y_1, y_2 \in B_{r/2}(0)$ with $G(B_{r/2}(0))$ open in $B_r(0)$.

Part (c) Since $DF(0) = I$, assume that $DF(x)$ is invertible for all $x \in B_r(0)$ for $r > 0$ given in Part (a). Further let $W = B_{r/2}(0) = B_R(0)$, and $V = G(W) \ni 0$, then $G : W \to V$ (and similarly $F : V \to W$). If G is differentiable, by chain rule $DF(G(y))DG(y) = I$ for all $y \in W$. Rewriting the equation gives $DG(y) = (DF)^{-1}(G(y))$.

For any $y_1 \in W$ such that $y_1 + y \in W$,

$$y = (y_1 + y) - y_1 = F(G(y_1 + y)) - F(G(y_1))$$

let $x_1 = G(y_1 + y)$ and $x_2 = G(y_1)$, then by *Proposition 3.18*,

$$\begin{aligned} y = F(x_1) - F(x_2) &= \left(\int_0^1 DF(x_2 + t(x_1 - x_2))\,\mathrm{d}t \right) \cdot (x_1 - x_2) \\ &= \left(\int_0^1 DF(x_2 + t(x_1 - x_2)) - DF(x_2)\,\mathrm{d}t \right) \cdot (x_1 - x_2) + DF(x_2)(x_1 - x_2) \end{aligned}$$

Hence

$$\begin{aligned} (DF)^{-1}(x_2)y = (DF)^{-1}(x_2) &\left(\int_0^1 DF(x_2 + t(x_1 - x_2)) - DF(x_2)\,\mathrm{d}t \right) \cdot (x_1 - x_2) \\ &+ (x_1 - x_2) \end{aligned}$$

In other words, $G(y_1 + y) - G(y_1) = (DF)^{-1}(G(y_1))y + R$ where

$$R = (DF)^{-1}(x_2) \left(\int_0^1 DF(x_2 + t(x_1 - x_2)) - DF(x_2)\,\mathrm{d}t \right) \cdot (x_1 - x_2)$$

By Part (b), $|x_1 - x_2| \leq 2|y|$, so $|x_1 - x_2| \to 0$ as $|y| \to 0$ and

$$\frac{|R|}{|y|} \leq 2 \left\| (DF)^{-1}(x_2) \right\| \int_0^1 \| DF(x_2) - DF(x_2 + t(x_1 - x_2)) \| \,\mathrm{d}t$$

With the assumption that F is C^1,

$$\lim_{|y| \to 0} \frac{|R|}{|y|} = 0$$

Therefore $G(y_1+y)-G(y) = (DF)^{-1}(G(y_1))y+o(|y|)$ which implies G is differentiable at $y_1 \in W$ and $DG(y_1) = (DF)^{-1}(G(y_1))$.

Finally, for the special case, it is assumed that DF is continuous and invertible on $B_r(0)$, then by linear algebra $(DF)^{-1}$ is also continuous. Then $DG(y) = (DF)^{-1}(G(y))$ is also continuous, and implies G is C^1. Using induction and differentiating the identity $DG(y) = (DF)^{-1}(G(y))$ will finish the fact that F is C^k implies G is C^k.

3.3.2 Diffeomorphisms

Definition 3.22. Let $F : V \to W$ be a C^k map where V and W are open sets in $\mathbb{R}^n$, then F is called a **C^k-diffeomorphism** if F^{-1} exists and is also C^k.

With the definition of diffeomorphisms, the Inverse Function Theorem can be rephrased as follows:

Theorem 3.23. Let $F : U \to \mathbb{R}^n$ be a C^k map from an open set $U \to \mathbb{R}^n$. Suppose $x_0 \in U$ and $DF(x_0)$ is invertible (as a matrix or linear transformation), then F is a C^k-diffeomorphism between some open sets V, W of $x_0, F(x_0)$ respectively.

Also, if $F : V \to W$ is a C^k-diffeomorphism, then for all function $\varphi : W \to \mathbb{R}$, there corresponds a function $\psi = \varphi \circ F : V \to \mathbb{R}$. Conversely, for all function $\psi = V \to \mathbb{R}$, there corresponds a function $\varphi = \psi \circ F^{-1} : W \to \mathbb{R}$. Moreover, φ is C^k if and only if ψ is C^k. Thus every C^k-diffeomorphism gives rise to a **local C^k-change of coordinates**.

3.3.3 Examples of Inverse Function Theorem

Below are some examples about the Inverse Function Theorem:

Example 3.24. Let $F : (0, \infty), (-\infty, \infty) \to \mathbb{R}^2$ such that $F(r, \theta) = (r\cos(\theta), r\sin(\theta))$, then

$$DF = \begin{bmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{bmatrix}$$

is invertible for all (r, θ). By the Inverse Function Theorem, F is locally invertible at every point $(r, \theta) \in (0, \infty) \times (-\infty, \infty)$. However, F is not globally invertible as $F(r, \theta + 2\pi) = F(r, \theta)$ implies it is not injective.

Example 3.25. Let U be an open interval $(a, b) \in \mathbb{R}$, then a C^1 function $f : (a, b) \to \mathbb{R}$ with $f' \neq 0$ implies f is strictly increasing or decreasing, so global inverse exists. Therefore 1-dimensional case has stronger result than higher dimensions.

3.3.4 Implicit Function Theorem

A theorem similar to Inverse Function Theorem is the **Implicit Function Theorem**:

Theorem 3.26. Let U be an open set in $\mathbb{R}^n \times \mathbb{R}^m$, and $F : U \to \mathbb{R}^m$ is a C^1 map. Suppose that $(x_0, y_0) \in U$ satisfies $F(x_0, y_0) = 0$ and $D_y F(x_0, y_0)$ is invertible in $\mathbb{R}^m$, then the following applies:

(a) There exists an open set of the form $V_1 \times V_2 \in U$ containing (x_0, y_0) and a C^1 map

$$\varphi : V_1 \subset \mathbb{R}^n \times V_2 \subset \mathbb{R}^m$$

with $\varphi(x_0) = y_0$ such that $F(x, \varphi(x)) = 0$ for all $x \in V_1$.

(b) $\varphi : V_1 \to V_2$ is C^k when F is C^k where $1 \le k \le \infty$.

(c) Assume DF_y is invertible in $V_1 \times V_2$, then if $\psi : V_1 \to V_2$ is another C^1 map satisfying $F(x, \psi(x)) = 0$, then $\psi \equiv \varphi$.

Note that if

$$F = \begin{bmatrix} F_1(x_1, \cdots, x_n, y_1, \cdots, y_m) \\ \vdots \\ F_m(x_1, \cdots, x_n, y_1, \cdots, y_m) \end{bmatrix}$$

then

$$D_y F = \begin{bmatrix} \partial F_1/\partial y_1 & \cdots & \partial F_1/\partial y_m \\ \vdots & \ddots & \vdots \\ \partial F_m/\partial y_1 & \cdots & \partial F_m/\partial y_m \end{bmatrix}$$

is an $m \times m$ matrix and can be regarded as a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^m$. In general, for a map F such that $DF(x_0, y_0)$ has rank m, then one can rearrange the independent variables to make the $m \times m$ submatrix corresponding to the last m columns of the Jacobian matrix invertible, which is the situation in the theorem. Hence the condition $DF_y(x_0, y_0)$ is invertible in the Implicit Function Theorem can be generalized to rank$DF(x_0, y_0) = m$.

3.4 Picard-Lindelof Theorem

3.4.1 Initial Value Problems

Definition 3.27. Let f be a function defined on

$$R = [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$$

where $(t_0, x_0) \in \mathbb{R}^2$ and $a, b > 0$. An **initial value problem** (or **Cauchy problem**) is of the form

$$\begin{cases} \mathrm{d}x/\,\mathrm{d}t = f(t, x) \\ x(t_0) = x_0 \end{cases}$$

This means one has to find $x(t)$ defined in an interval

$$x : [t_0 - a', t_0 + a'] \to [x_0 - b, x_0 + b]$$

for some $0 < a' \le a$ such that $x(t)$ is differentiable, $x(t_0) = x_0$ and

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) = f(t, x(t))$$

for all $t \in [t_0 - a', t_0 + a']$.

Example 3.28. Consider the initial value problem

$$\begin{cases} \mathrm{d}x/\,\mathrm{d}t = 1 + x^2 \\ x(0) = 0 \end{cases}$$

Note that $f(t, x) = 1 + x^2$ is smooth on $[-a, a] \times [-b, b]$ for any $a, b > 0$, but the solution $x(t) = \tan(t)$ is defined only on $(-\pi/2, \pi/2)$. Therefore, even for a nice function f, it is still possible that $a' < a$.

3.4.2 Picard-Lindelof Theorem for Differential Equations

Recall the defintion of Lipschitz condition:

Definition 3.29. Let f be a function defined in $R : [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$, then f satisfies the Lipschitz condition (uniform in t) if there exists a Lipschitz constant $L > 0$ such that for all $(t, x_1), (t, x_2) \in R$,

$$|f(t, x_1) - f(t, x_2)| \le L\,|x_1 - x_2|$$

Also recall the properties related to Lipschitz condition:

Proposition 3.30. The following statements are true:

(a) $f(t, \cdot)$ is Lipschitz continuous in x for all $t \in [t_0 - a, t_0 + a]$.

(b) If L is a Lipschitz constant for f, then any $L' > L$ is also a Lipschitz constant.

(c) Continuity does not imply Lipschitz continuity. For example, $f(t, x) = tx^{1/2}$ is continuous but not Lipschitz continuous near 0.

(d) If $R = [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$ and $f(t, x) : R \to \mathbb{R}$ is C^1, then $f(t, x)$ satisfies the Lipschitz condition. In fact, for some $y \in [x_0 - b, x_0 + b]$,

$$|f(t, x_1) - f(t, x_2)| = \left| \frac{\partial f}{\partial x}(t, y)(x_2 - x_1) \right|$$

Hence $|f(t, x_1) - f(t, x_2)| \leq L\,|x_1 - x_2|$ for

$$L = \max \left\{ \left| \frac{\partial f}{\partial x}(t, x) \right| \mid (t, x) \in R \right\}$$

Proposition 3.31. Under assumption of *Theorem 3.30*, every solution x of the initial value problem from $[t_0 - a', t_0 + a']$ to $[x_0 - b, x_0 + b]$ satisfies the equation

$$x(t) = x_0 + \int_{t_0}^{t} f(t, x(t))\,\mathrm{d}t$$

Conversely, every $x(t) \in C[t_0 - a', t_0 + a']$ satisfying the equation above is C^1 and solves the initial value problem.

Proof. This is a result of Fundamental Theorem of Calculus.

Below is the **Picard-Lindelof Theorem**:

Theorem 3.32. Let f be a continuous function on $R : [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$ where $(t_0, x_0) \in \mathbb{R}^2$ and $a, b > 0$ If f satisfies Lipschitz condition on R (uniform in t), then there exists $a' \in (0, a]$ and $x \in C^1[t_0 - a', t_0 + a']$ such that

$$x_0 - b \leq x(t) \leq x_0 + b$$

for all $t \in [t_0 - a', t_0 + a']$ and solving the initial value problem. Furthermore, x is the unique solution in $[t_0 - a', t_0 + a']$.

Proof. For $a' > 0$ to be chosen later, let

$$X = \{\varphi \in C[t_0 - a', t_0 + a'] \mid \varphi(t_0) = x_0, \varphi(t) \in [x_0 - b, x_0 + b]\}$$

with uniform metric d_∞ on X. Note that X is a closed subset in the complete metric space $(C[t_0 - a', t_0 + a'], d_\infty)$, so (X, d_∞) is complete.

Define T on X by

$$(T\varphi)(t) = x_0 + \int_{t_0}^{t} f(s, \varphi(s))\,\mathrm{d}s$$

Note that it is well-defined since $\varphi(s) \in [x_0 - b, x_0 + b]$. To show $T\varphi \in X$, one requires $(T\varphi)(t) \in [x_0 - b, x_0 + b]$. Let $M = \sup\{|f(t,x)| \mid (t,x) \in R\}$, then for all $t \in [t_0 - a', t_0 + a']$,

$$\begin{aligned} |(T\varphi)(t) - x_0| &= \left| \int_{t_0}^{t} f(s, \varphi(s)) \, \mathrm{d}s \right| \\ &\leq M\,|t - t_0| \leq Ma' \end{aligned}$$

Choose $0 < a' \leq b/M$ gives $|(T\varphi)(t) - x_0| \leq b$ and so $T\varphi \in X$. Notice that $T : X \to X$ is a mapping from (X, d_∞) to itself. For contraction,

$$\begin{aligned} |(T\varphi_1 - T\varphi_2)(t)| &= \left| (x_0 + \int_{t_0}^{t} f(s, \varphi_1(s)) \, \mathrm{d}s) - (x_0 + \int_{t_0}^{t} f(s, \varphi_2(s)) \, \mathrm{d}s) \right| \\ &\leq \int_{t_0}^{t} |f(s, \varphi_1(s)) - f(s, \varphi_2(s))| \, \mathrm{d}s \\ &\leq L \int_{t_0}^{t} |\varphi_1(s) - \varphi_2(s)| \, \mathrm{d}s \\ &\leq L\,|t - t_0| \sup_{[t_0 - a', t_0 + a']} \{\varphi_1(s) - \varphi_2(s)\} \\ &\leq La' d_\infty(\varphi_1, \varphi_2) \end{aligned}$$

Therefore if $La' = \gamma < 1$, T is a contraction since $d_\infty(T\varphi_1, T\varphi_2) \leq \gamma d_\infty(\varphi_1, \varphi_2)$.

In conclusion, if $0 < a' < \min\{a, b/M, 1/L\}$, then T is a contraction on a complete metric space. By Contraction Mapping Principle, T admits a unique fixed point $x(t) \in X$.

Note that the existence part of Picard-Lindelof Theorem still holds with $f(t,x)$ being continuous only. However, the solution may not be unique. Consider the following example:

Example 3.33. Let $f(t,x) = |x|^{1/2}$ on $\mathbb{R} \times \mathbb{R}$. Note that f is continuous but not Lipschitz continuous. Then the initial value problem

$$\begin{cases} \mathrm{d}x/\,\mathrm{d}t = |x|^{1/2} \\ x(0) = 0 \end{cases}$$

has solutions

$$x_1 = 0, x_2 = \frac{1}{4}\,|t|\,t$$

for all $t \in \mathbb{R}$.

On the other hand, uniqueness of Picard-Lindelof Theorem holds regardless of the size of the interval of existence.

3.4.3 Picard-Lindelof Theorem for Systems

Theorem 3.34. Consider the initial value problem

$$\begin{cases} \mathrm{d}\mathbf{x}/\,\mathrm{d}t = \mathbf{f}(t, \mathbf{x}) \\ \mathbf{x}(t_0) = \mathbf{x}_0 \end{cases}$$

where

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \in [x_1 - b, x_1 + b] \times \cdots \times [x_n - b, x_n + b]$$

$$\mathbf{x}_0 = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{f}(t, x) = \begin{bmatrix} f_1(t, x) \\ \vdots \\ f_n(t, x) \end{bmatrix} \in C^1(R)$$

with

$$R = [t_0 - a, t_0 + a] \times [x_1 - b, x_1 + b] \times \cdots \times [x_n - b, x_n + b]$$

satisfying the Lipschitz condition (uniform in t),

$$|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})| \leq L\, |\mathbf{x} - \mathbf{y}|$$

for all $(t, \mathbf{x}), (t, \mathbf{y}) \in R$ and some constant $L > 0$. There exists a unique solution $\mathbf{x} \in C^1[t_0 - a', t_0 + a']$ with

$$\mathbf{x}(t) \in [x_1 - b, x_1 + b] \times \cdots \times [x_n - b, x_n + b]$$

for all $t \in [t_0 - a', t_0 + a']$ to the initial value problem, where a' satisfies

$$0 < a' < \min \left\{ a, \frac{b}{M}, \frac{1}{L} \right\}$$

with

$$M = \max_{j=1,\cdots,n} \sup_R |f_j(t, \mathbf{x})|$$

Note that the Picard-Lindelof Theorem for systems can be applied to initial value problems for higher order ordinary differential equations:

$$\begin{cases} \mathrm{d}^m x/\,\mathrm{d}t^m = f(t, x,\, \mathrm{d}x/\,\mathrm{d}t, \cdots,\, \mathrm{d}^{m-1}x/\,\mathrm{d}t^{m-1}) \\ x(t_0) = x_0 \\ \mathrm{d}x/\,\mathrm{d}t(t_0) = x_1 \\ \vdots \\ \mathrm{d}^{m-1}x/\,\mathrm{d}t^{m-1} = x_{m-1} \end{cases}$$

by letting

$$\mathbf{x} = \begin{bmatrix} x \\ \mathrm{d}x/\,\mathrm{d}t \\ \vdots \\ \mathrm{d}^{m-1}x/\,\mathrm{d}t^{m-1} \end{bmatrix}$$

then

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \begin{bmatrix} \mathrm{d}x/\,\mathrm{d}t \\ \mathrm{d}^2x/\,\mathrm{d}t^2 \\ \vdots \\ \mathrm{d}^m x/\,\mathrm{d}t^m \end{bmatrix} = \mathbf{f}(t, \mathbf{x})$$

with

$$\mathbf{x}(t_0) = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{m-1} \end{bmatrix}$$

4 Space of Continuous Functions

4.1 Arzela-Ascoli Theorem

4.1.1 Compact Sets

Definition 4.1. Let (X, d) be a metric space, then the vector space of all bounded continuous functions is denoted by

$$C_b(X) = \{f \in C(X) \mid |f(x)| \leq M, \forall x \in X, \exists M\}$$

It is simple to see that $C_b(X) \subset C(X)$, where $C(X)$ is the set of continuous functions on X.

Example 4.2. If G is a nonempty bounded open set in $\mathbb{R}^n$, then $C_b(\overline{G}) = C(\overline{G})$ as $\overline{G}$ is closed and bounded, then $f \in C(\overline{G})$ has to be bounded.

Recall that a norm $\|\cdot\|$ on a real vector space X is defined by the following properties:

(a) $\|x\| \geq 0$ is nonnegative, and $\|x\| = 0$ if and only if $x = 0$.

(b) $\|\alpha x\| = |\alpha| \|x\|$ for all $\alpha \in \mathbb{R}$.

(c) Triangle inequality holds, or $\|x + y\| \leq \|x\| + \|y\|$

A vector space with norm $(X, \|\cdot\|)$ is called a norm space. Note that a norm space has a natural metric $d(x, y) = \|x - y\|$.

Definition 4.3. Let $C_b(X)$ be the vector space of all bounded continuous functions, then the **supnorm** is a norm on $C_b(X)$ defined by

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

It is always assumed $C_b(X)$ with metric $d_\infty(f, g) = \|f - g\|_\infty$ given by the supnorm.

Proposition 4.4. $(C_b(X), d_\infty)$ is a complete metric space, for any metric space (X, d).

Note that $(C_b(X), d_\infty)$ is a Banach space since it is a complete normed vector space. $C_b(X)$ is usually of infinite dimensional (for example $X = \mathbb{R}^n$ or a subset with nonempty interior in $\mathbb{R}^n$ like $X = [0, 1]$), but it also could be of finite dimensional (for example $X = \{p_1, \cdots, p_n\}$ as a finite set of discrete metrics, which gives $X \to \mathbb{R}^n$ a linear bijection).

A reason for studying $C_b(X)$ instead of $C(X)$ is the fact that $C(X)$ may contain unbounded function and the supnorm is not defined (for example $X = \mathbb{R}$). However, in some cases, it is still possible to define a metric on $C(X)$:

Example 4.5. Let $X = \mathbb{R}^n$ and $\overline{B_n(0)} = \{|x| \leq n\}$ for all positive integers n. For all $f \in C(\mathbb{R}^n)$, define

$$d(f,g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\|f-g\|_{\infty,\overline{B_n(0)}}}{1+\|f-g\|_{\infty,\overline{B_n(0)}}}$$

where $\|\cdot\|_{\infty,\overline{B_n(0)}}$ is the supnorm on the closed ball $\overline{B_n(0)}$, then d is a complete metric on $C(\mathbb{R}^n)$.

Finally, recall the Bolzano-Weierstrass Theorem in $\mathbb{R}^n$:

Theorem 4.6. Every bounded sequence has a convergent subsequence. Similarly, every bounded set contains a convergent sequence.

$C_b(X)$ may not have Bolzano-Weierstrass property. Consider the following example:

Example 4.7. Observe that $C_b([0,1]) = C[0,1]$. Let $f_n(x) = x^n$ where $x \in [0,1]$ for all n, then $\|f_n\|_\infty = 1$. The pointwise limit

$$f_n(x) \to \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

implies that no subsequence converges in $C_b[0,1]$.

Because of this, further condition are required to find convergent sequences in subsets of $C_b(X)$.

Definition 4.8. Let (X,d) be a metric space. A set $E \subset X$ is called a **precompact** set if every sequence in E contains a convergent subsequence with limit in X (which is not necessary in E). If the limit is further restricted within E, then E is called a **compact** set.

Proposition 4.9. A compact set is a closed precompact set.

Proof. Let (X,d) be a metric space and $\{x_n\} \subset E$ be a sequence in $E \subset X$. If E is precompact, there exists a subsequence $\left\{x_{n_j}\right\}$ with limit $z \in X$. If E is closed, the limit $z \in E$, which implies compactness.

Also recall that by Bolzano-Weierstrass Theorem, $E \subset \mathbb{R}^n$ is precompact implies E is bounded. Therefore E is compact implies E is closed and bounded.

4.1.2 Equicontinuity

Definition 4.10. Let (X, d) be a metric space. A subset C of $C(X)$ is said to be **equicontinuous** if for all $\epsilon > 0$, there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $f \in C$ and $x, y \in X$ where $d(x, y) < \delta$.

In fact, equicontinuity is based on uniform continuity, but at the same time extends δ to fulfill every function $f \in C$. Therefore, equicontinuity implies every function in C is uniformly continuous. Then, it is simple to see that if C is equicontinuous, any $C' \subset C$ is also equicontinuous.

There are other ways to show that a set is equicontinuous. Recall that a function f is Holder continuous if there exists a Holder exponent $\alpha \in (0, 1)$ such that

$$|f(x) - f(y)| \leq L\,|x - y|^{\alpha}$$

for some constant L. f is Lipschitz continuous if the equation holds for $\alpha = 1$. A set C is equicontinuous too if every function $f \in C$ is Holder continuous or Lipschitz continuous.

Another method for equicontinuity requires the following definition:

Definition 4.11. A set C is said to be **convex** (in $\mathbb{R}^n$) if $x + t(y - x) \in C$ for all $x, y \in C$ and $t \in [0, 1]$.

Proposition 4.12. Let C be a subset of $C(\overline{G})$ where $\overline{G}$ is a convex in $\mathbb{R}^n$. Suppose that each function in C is differentiable and there is a uniform bound on their partial derivatives:

$$\left\| \frac{\partial f}{\partial x_i} \right\|_{\infty} \leq M$$

for all $f \in C(\overline{G})$ and i, then C is equicontinuous.

Proposition 4.13. Let $A = \{z_j\}$ be a countable set and $f_n : A \to \mathbb{R}$ where $n = 1, 2, \cdots$ be a sequence of functions defined on A. Suppose for each $z_j \in A$, $\{f_n(z_j)\}$ is a bounded sequence in $\mathbb{R}$, then there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that for all $z_j \in A$, $\{f_{n_k}(z_j)\}$ is convergent.

4.1.3 Ascoli's Theorem

Below is the **Ascoli's Theorem**:

Theorem 4.14. Suppose G is a bounded nonempty open set in $\mathbb{R}^m$, then a set $\mathcal{E} \subset C(\overline{G}) = C_b(\overline{G})$ is precompact if $\mathcal{E}$ is bounded (in supnorm) and equicontinuous.

Note that Ascoli's Theorem remains valid for bounded and equicontinuous subsets of $C(G)$ where G is not necessary to take closure. This is because equicontinuity implies uniform continuity of G, which can be further extended to uniform continuity of $\overline{G}$. However, boundedness of the domain $\overline{G}$ cannot be removed:

Example 4.15. Let $\overline{G} = [0, \infty) \subset \mathbb{R}$, then take $\varphi \in C^1[0,1]$ such that $\varphi \not\equiv 0$ and $\varphi(x) = 0$ when $x \in [0,1] \setminus [1/2, 3/4]$. Further define

$$f_n(x) = \begin{cases} \varphi(x-n) & \text{if } x \in [n, n+1] \\ 0 & \text{otherwise} \end{cases}$$

It is easy to check that $f_n \in C(\overline{G})$ and

$$\|f_n\|_{\infty,\overline{G}} = \|\varphi\|_{\infty,[0,1]} > 0$$

Thus $\mathcal{E} = \{f_n\}$ is a bounded subset of $C(\overline{G})$. By chain rule,

$$\left\|\frac{\mathrm{d}f_n}{\mathrm{d}x}\right\|_{\infty,\overline{G}} = \left\|\frac{\mathrm{d}\varphi}{\mathrm{d}x}\right\|_{\infty,[0,1]} > 0$$

Then *Proposition 4.12* states that $\mathcal{E}$ is also equicontinuous.

Suppose there exists a subsequence $\{f_{n_j}\}$ of $\{f_n\}$ converges to the same $f \in C(\overline{G})$ in d_∞. In other words, $f_{n_j} \to f$ uniformly on $\overline{G}$ implies pointwise convergence $f_{n_j}(x) \to f(x)$ for all $x \in \overline{G}$. However, for fixed x, $f_n(x) = 0$ for all $n \geq x$, so it is expected to have

$$\lim_{j \to +\infty} f_{n_j}(x) \to 0$$

which shows that $f(x) = 0$ for all $x \in \overline{G}$. This is a contraction since

$$0 < \|\varphi\|_{\infty,[0,1]} = \left\|f_{n_j}\right\|_{\infty,\overline{G}} = \left\|f_{n_j} - f\right\|_{\infty,\overline{G}} \to 0$$

Therefore $\mathcal{E}$ is bounded and equicontinuous, but Ascoli's Theorem doesn't hold.

4.1.4 Arzela's Theorem

Below is the **Arzela's Theorem**, which is the converse of Ascoli's Theorem:

Theorem 4.16. Suppose G is a bounded nonempty open set in $\mathbb{R}^m$, then every precompact set in $C(\overline{G})$ must be bounded and equicontinuous.

4.2 Applications to Ordinary Differential Equations

4.2.1 Improvement to Picard-Lindelof Theorem

Consider the initial value problem

$$\begin{cases} \mathrm{d}x/\,\mathrm{d}t = f(t,x) \\ x(t_0) = x_0 \end{cases}$$

with f being continuous (but not necessary Lipschitz) on $R = [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$. Of course this is not expected to give a unique result, but existence can be proved. The idea of proof is as follows:

(1) By Weierstrass Approximation Theorem (on $\mathbb{R}^2$), there exists a sequence $\{p_n\}$ of polynomials such that $d_\infty(p_n, f) \to 0$ (in $C(R)$).

(2) By Picard-Lindelof Theorem, since every p_n satisfies Lipschitz condition (uniform in t), there exists $a'_n > 0$ with

$$a'_n = \min\left\{a, \frac{b}{M_n}, \frac{1}{L_n}\right\}$$

where $M_n = \|p_n\|_{\infty,R}$ and L_n be Lipschitz constant of p_n on R, such that there exists a unique solution $x_n \in C^1[t_0 - a'_n, t_0 + a'_n]$ to the approximated initial value problem

$$\begin{cases} \mathrm{d}x_n/\,\mathrm{d}t = p_n(t, x_n) \\ x_n(t_0) = x_0 \end{cases}$$

for all $t \in [t_0 - a'_n, t_0 + a'_n]$.

(3) By Ascoli's Theorem, there exists a convergent subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \to x$ for some function $x(t)$. It is hoped that such x is the required solution.

However, since f is not assumed to satisfy the Lipschitz condition, one cannot expect $\{L_n\}$ is bounded. In fact, $\{L_n\}$ is unbounded, otherwise f satisfies Lipschitz condition. Here

$$a'_n = \min\left\{a, \frac{b}{M_n}, \frac{1}{L_n}\right\} \to 0$$

then there is no proper interval for existence of the solution. On the other hand, as $p_n \to f$ in $(C(R), d_\infty)$, $M_n \leq M$ for some $M > 0$. Therefore, in order to implement the plan above, it is required to improve Picard-Lindelof Theorem:

Proposition 4.17. Under the setting of Picard-Lindelof Theorem, there exists a unique solution $x(t)$ on the interval $[t_0 - a', t_0 + a']$ with $x(t) \in [x_0 - b, x_0 + b]$, where a' is any number satisfying

$$0 < a' < a^* = \min\left\{a, \frac{b}{M}\right\}$$

4.2.2 Cauchy-Peano Theorem

Below is the **Cauchy-Peano Theorem**:

Theorem 4.18. Consider the initial value problem

$$\begin{cases} \mathrm{d}x/\,\mathrm{d}t = f(t,x) \\ x_{t_0} = x_0 \end{cases}$$

where f is continuous on $R = [t_0 - a, t_0 + a] \times [x_0 - b, x_0 + b]$, then there exists $a' \in (0,a)$ and a C^1 function

$$x : [t_0 - a, t_0 + a] \to [x_0 - b, x_0 + b]$$

solving the initial value problem.

4.3 Baire Category Theorem

4.3.1 Denseness

Definition 4.19. Let (X,d) be a metric space, then a set $E \subset X$ is said to be **dense** if for all $x \in X$ and $\epsilon > 0$, $B_\epsilon(x) \cap E \neq \emptyset$.

Note that X is naturally dense in (X,d), and if E is dense in X, its closure $\overline{E} = X$.

Definition 4.20. Let (X,d) be a metric space, then a set $E \subset X$ is said to be **nowhere dense** if its closure does not contain any ball. In other words, $\overline{E}$ has empty interior.

Example 4.21. Set of integers $\mathbb{Z}$ is nowhere dense in $\mathbb{R}$. However, although set of rationals $\mathbb{Q}$ has empty interior, its closure $\overline{\mathbb{Q}} = \mathbb{R}$ has nonempty interior, so $\mathbb{Q}$ is not nowhere dense.

Proposition 4.22. Let (X,d) be a metric space and $E \subset X$ be a set, then E is nowhere dense if and only if $X \setminus \overline{E}$ is dense in X.

Proof. If E is nowhere dense, for all $x \in X$ and any $r > 0$, $B_r(x) \not\subset \overline{E}$, which implies $B_r(x) \cap (X \setminus \overline{E}) \neq \emptyset$, so $X \setminus \overline{E}$ is dense. The converse follows the reverse order.

Definition 4.23. Let (X, d) be a metric space, then a point $x \in X$ is called an **isolated point** if $\{x\}$ is open in X.

Note that $\{x\}$ is always closed in a metric space. Therefore, $\{x\}$ is both open and closed if and only if x is an isolated point.

Proposition 4.24. Let (X, d) be a metric space, then the following applies:

(a) If E is nowhere dense in X, $\overline{E}$ is nowhere dense in X. Also, E' is nowhere dense in X if $E' \subset E$.

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

(c) If (X, d) has no isolated point, then every finite set is nowhere dense.

Now consider the following example in infinite dimensional normed spaces:

Example 4.25. Let $M[a, b]$ be a space of bounded functions on $[a, b]$, then

$$\|f\|_\infty = \sup_{[a,b]} |f(x)|$$

is well-defined and is a norm on $M[a, b]$. It is clear that $(C[a, b], d_\infty)$ is a metric (and vector) subspace of $(M[a, b], d_\infty)$. Show that $C[a, b]$ is nowhere dense in $M[a, b]$.

Answer. Note that $C[a, b]$ is closed because uniform limit of continuous functions is continuous. It is left to show that for all $B_\epsilon^\infty(f) \subset M[a, b]$,

$$B_\epsilon^\infty(f) \cap (M[a, b] \setminus C[a, b]) \neq \emptyset$$

If $f \in M[a, b] \setminus C[a, b]$, the result is already achieved. For any $f \in C[a, b]$, let

$$g(x) = \begin{cases} f(x) + \epsilon/2 & \text{if } x \in [a, b] \cap \mathbb{Q} \\ f(x) - \epsilon/2 & \text{if } x \in [a, b] \setminus \mathbb{Q} \end{cases}$$

such that $\|g - f\|_\infty = \epsilon/2$ implies $g \in B_\epsilon^\infty(f)$. Since both $[a, b] \cap \mathbb{Q}$ and $[a, b] \setminus \mathbb{Q}$ are dense in $[a, b]$,

$$\limsup_{x \to a} g(x) = f(a) + \frac{\epsilon}{2}$$

and

$$\liminf_{x \to a} g(x) = f(a) - \frac{\epsilon}{2}$$

shows that $g \in M[a,b] \setminus C[a,b]$. Therefore $B_\epsilon^\infty(f) \cap (M[a,b] \setminus C[a,b]) \neq \emptyset$, and $C[a,b]$ is nowhere dense in $M[a,b]$.

4.3.2 First Category and Second Category

Definition 4.26. Let (X, d) be a metric space, then a set $E \subset X$ is called **first category** (or **meager**) if it can be expressed as a countable union of nowhere dense sets. If E is not of first category, then E is called **second category**.

E is said to be **residual** if its complement is of first category.

Proposition 4.27. Let (X, d) be a metric space, then the following applies:

(a) Every subset of a set of first category is of first category.

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

(c) If (X, d) has no isolated point, every countable subset of X is of first category.

With the proposition above, a similar proposition for residual sets can be made by taking complements:

Proposition 4.28. Let (X, d) be a metric space, then the following applies:

(a) Every subset containing a residual set is residual.

(b) The intersection of countable many residual sets is residual.

(c) If (X, d) has no isolated point, complement of any countable set is residual.

Example 4.29. Let $(\mathbb{R}, d_1)$ be a metric space. Since $\mathbb{R}$ has no isolated points, $\{q\}$ is nowhere dense for any $q \in \mathbb{Q}$, so $\mathbb{Q}$ is of first category since it is a countable union of $\{q\}$. On the other hand, the set of irrational numbers $\mathbb{I} = \mathbb{R} \setminus \mathbb{Q}$ is residual in $\mathbb{R}$.

4.3.3 General Theorem

Below is the **Baire Category Theorem**:

Theorem 4.30. Any set of first category in a complete metric space has empty interior. In other words, any countable intersection of open dense sets in a complete metric space is dense.

With Baire Category Theorem, there are some corollaries to follow:

Corollary. Let (X, d) be a complete metric space. Suppose that $X = \bigcup_{n=1}^{\infty} E_n$ with E_n are closed subsets. Then at least one of there E_n has nonempty interior.

Corollary. A set of first category in a complete metric space cannot be a residual set, and vice versa.

4.3.4 Applications of Baire Category Theorem

Proposition 4.31. Let $f \in C[a, b]$ be differentiable at x, then it is Lipschitz continuous at x.

Proof. By assumption, for any $\epsilon = 1 > 0$, there exists $\delta_0 > 0$ such that for all $y \in (x - \delta_0, x + \delta_0) \setminus \{x\}$ and $y \in [a, b]$,

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

implies

$$|f(y) - f(x)| \leq (1 + |f'(x)|) |y - x|$$

for all $y \in (x - \delta_0, x + \delta_0) \cap [a, b]$. If $[a, b] \setminus (x - \delta_0, x + \delta_0) = \emptyset$, it is already done. Consider for $y \in [a, b] \setminus (x - \delta_0, x + \delta_0)$ if such set is nonempty, $|y - x| \geq \delta_0$, hence

$$\begin{aligned} |f(y) - f(x)| &\leq |f(y)| + |f(x)| \\ &\leq 2 \|f\|_\infty \\ &\leq \frac{2 \|f_\infty\|}{\delta_0} |y - x| = L' |y - x| \end{aligned}$$

Finally, let $L = \max \{1 + |f'(x)|, L'\}$, then $|f(y) - f(x)| \leq L |y - x|$ for all $y \in [a, b]$.

With the proposition above, the following theorem can be introduced:

Theorem 4.32. The set of all continuous, nowhere differentiable functions forms a residual set in $C[a, b]$ and hence dense in $C[a, b]$.

References

The following are the references of the context of this document:

(a) Professor(s) associated to *MATH3060: Mathematical Analysis III*

(b) Elias M. Stein, Rami Shakarchi, *Fourier Analysis: An Introduction (Princeton Lectures in Analysis)*, Princeton, 2003

(c) Walter Rudin, *Principles of Mathematical Analysis*, McGraw-Hill (3rd Edition), 1976