NATIONAL UNIVERSITY OF SINGAPORE

Department of Mathematics

MA2108 Advanced Calculus II

2001/2002 Semester I Lecture Notes Part II

Chapter 3: Series of Functions

3.1. Sequence of Functions

Let I be an interval in $\mathbb{R}$, e.g. $(-1,1),[0,1]$, etc. For each $n \in \mathbb{N}$, let $F_n : I \to \mathbb{R}$ be a function. Then we say $\{F_n\}$ forms a *sequence of functions* on I.

Example 3.1.1.

1. $F_n(x) = x^n, \qquad 0 < x < 1.$
   Then $\{F_n\}$ forms a sequence of functions on (0,1).

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

2. $\{(1+\frac{x}{n})^n\}$ forms a sequence of functions on $(-\infty, \infty)$.

Write out some terms:

$F_1(x) =$

$F_2(x) =$

$F_3(x) =$

Know: if we fix the x, and let $n \to \infty$,

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} (1+\frac{x}{n})^n = e^x.$$

So for each x, we can define

$$F(x) = \lim_{n\to\infty} F_n(x).$$

Definition 3.1.1. A sequence $\{F_n\}$ is said to *converge pointwise* to a function F on I if

$$\lim_{n\to\infty} F_n(x) = F(x) \qquad \text{for each } x \in I,$$

i.e., for each $x \in I$ and given any $\epsilon > 0$, there exists an N (which depends on x and ϵ) such that

$$|F_n(x) - F(x)| < \epsilon \qquad \forall n > N.$$

Definition 3.1.2. $\{F_n\}$ is said to *converge pointwise on* I if $\{F_n\}$ converge pointwise to some function F.

Remark. The function F is called the *limiting function* of $\{F_n\}$ and is necessarily unique.

Example 3.1.2.

1. Does the sequence $\{x^n\}$ where $0 < x < 1$ have a limiting function?

Remarks on pointwise convergence

Suppose a sequence of functions $\{F_n\}$ converges pointwise to a function F on the interval $[a, b]$. We want to know whether the limiting function F inherits the properties of $\{F_n\}$. For example, we may ask the following questions:

Question 1: Suppose each F_n is a continuous function on $[a, b]$. Is it true that F is necessarily continuous on $[a, b]$?

Question 2: Is it true that

$$\lim_{n\to\infty} \int_a^b F_n(x)\, dx = \int_a^b F(x)\, dx?,$$

i.e. whether

$$\lim_{n\to\infty} \int_a^b F_n(x)\, dx = \int_a^b \big(\lim_{n\to\infty} F_n(x)\big)\, dx?$$

Answers: NO to both questions.

Counter-example to Question 1: Consider the functions

$$F_n(x) = x^n, \quad x \in [0, 1].$$

For each fixed $x \in [0, 1)$, we have

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} x^n = 0 \quad (\text{since } |x| < 1).$$

At $x = 1$, we have

$$\lim_{n\to\infty} F_n(1) = \lim_{n\to\infty} 1^n = 1.$$

Thus $\{F_n\}$ converges pointwise to the function F on the interval $[0, 1]$ given by

$$F(x) = \begin{cases} 0, & \text{for } x \in [0, 1), \\ 1, & x = 1. \end{cases}$$

Each F_n is continuous on the whole interval $[0, 1]$, but F is not continuous at $x = 1$.

Counter-example to Question 2: Consider the functions

$$F_n(x) = \begin{cases} n^2 x, & 0 < x < \dfrac{1}{n}, \\ 2n - n^2 x, & \dfrac{1}{n} \le x < \dfrac{2}{n}, \\ 0, & \dfrac{2}{n} \le x < 1. \end{cases}$$

For each fixed $x \in (0,1]$, one sees that $F_n(x) = 0$ whenever $n \geq \frac{2}{x}$, and hence

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} 0 = 0.$$

Also, at $x = 0$, we have

$$\lim_{n\to\infty} F_n(0) = \lim_{n\to\infty} n^2 \cdot 0 = 0.$$

Thus, $\{F_n\}$ converges pointwise to the zero function $F(x) \equiv 0$ on the interval $[0,1]$. For each $n \geq 1$, we have

$$\begin{aligned}
\int_0^1 F_n(x)\,dx &= \int_0^{1/n} n^2 x\,dx \\
&\quad + \int_{1/n}^{2/n} (2n - n^2 x)\,dx + \int_{2/n}^1 0\,dx \\
&= \frac{n^2x^2}{2}\Big|_0^{\frac{1}{n}} + \big(2nx - \frac{n^2x^2}{2})\Big|_{\frac{1}{n}}^{\frac{2}{n}} + 0 \\
&= \frac{1}{2} + (4-2) - (2 - \frac{1}{2}) + 0 = 1.
\end{aligned}$$

Thus we have

$$\begin{aligned}
&\lim_{n\to\infty} \int_0^1 F_n(x)\,dx = \lim_{n\to\infty} 1 = 1 \neq 0 = \int_0^1 F(x)\,dx, \quad \text{i.e.} \\
&\lim_{n\to\infty} \int_0^1 F_n(x)\,dx \neq \int_0^1 \Big(\lim_{n\to\infty} F_n(x)\Big)\,dx.
\end{aligned}$$

Reason: At different x, $F_n(x)$ converges to $F(x)$ at different pace (more specifically, in the definition of pointwise convergence, the choice of N depends on both ϵ and x).

3.2. Uniform Convergence

We define a slightly different concept of convergence.

Definition 3.2.1. $\{F_n\}$ is said to *converge uniformly* to a function F on an interval I if for every $\epsilon > 0$, there exists an N (which depends only on ϵ) such that

$$|F_n(x) - F(x)| < \epsilon$$

for ALL $x \in I$ whenever $n > N$.

Definition 3.2.2. $\{F_n\}$ is said to *converge uniformly on* I if $\{F_n\}$ converges uniformly to some F on I.

Theorem 3.2.3.
If $\{F_n\}$ *converges uniformly to* F *on* I*, then* $\{F_n\}$ *converges pointwise to* F *on* I*.*

Remark. The limiting function F is unique.

Two Criteria for Uniform Convergence of $\{F_n\}$.

The following theorem is useful (computationally) in determining whether a sequence of functions converges uniformly or not.

Theorem 3.2.4.
Suppose $\{F_n\}$ is a sequence of functions converging pointwise to a function F on an interval I, and let

$$T_n = \sup_{x \in I} |F_n(x) - F(x)|.$$

Then $\{F_n\}$ converges uniformly to F on I if and only if $\lim_{n \to 0} T_n = 0$.

Proof. First we prove the 'only if' part. Suppose that $\{F_n\}$ converges uniformly to F on I. Then for any given $\epsilon > 0$, there exists N such that

$$\begin{aligned} |F_n(x) - F(x)| &< \frac{\epsilon}{2} \quad \text{for all } n > N \text{ and } x \in I \\ \Rightarrow \; T_n &= \sup_{x \in I} |F_n(x) - F(x)| \\ &\le \frac{\epsilon}{2} < \epsilon \quad \text{for all } n > N \\ \Rightarrow \; |T_n - 0| &= T_n < \epsilon \quad \text{for all } n > N. \end{aligned}$$

Hence we have $\lim_{n \to 0} T_n = 0$.

Next we prove the 'only if' part. Suppose that $\lim_{n \to 0} T_n = 0$. Then for any given $\epsilon > 0$, there exists N such that

$$\begin{aligned} |T_n - 0| = T_n &< \epsilon \quad \text{for all } n > N \\ \Rightarrow \; \sup_{x \in I} |F_n(x) - F(x)| &< \epsilon \quad \text{for all } n > N \\ \Rightarrow \; |F_n(x) - F(x)| &< \epsilon \quad \text{for all } n > N \text{ and } x \in I. \end{aligned}$$

Hence $\{F_n\}$ converges uniformly to F on I. This finishes the proof of the theorem.

Example 3.2.1.

Theorem 3.2.5 [Cauchy's Criterion]. *A sequence of functions $\{F_n\}$ converges uniformly on an interval I if and only if given any $\epsilon > 0$, there exists a natural number N such that*

$$|F_n(x) - F_m(x)| < \epsilon \quad \textit{for all } x \in I \textit{ and all } m, n > N. \tag{1}$$

Remark: Here N does not depend on x.

Proof. First we prove the 'only if' part. Suppose that $\{F_n\}$ converges uniformly to the function F on I. Then given any $\epsilon > 0$, there exists N such that

$$|F_n(x) - F(x)| < \frac{\epsilon}{2} \quad \text{for all } x \in I \text{ and all } n > N.$$

Then for all $x \in I$ and $m, n > N$,

$$\begin{aligned} &|F_n(x) - F_m(x)| \\ &= |(F_n(x) - F(x)) - (F_m(x) - F(x))| \\ &\leq |F_n(x) - F(x)| + |F_m(x) - F(x)| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon. \end{aligned}$$

This finishes the proof of the 'only if' part.

Next we prove the 'if' part. Suppose that equation 1 holds. Then for each fixed point $x \in I$, $\{F_n(x)\}$ is a Cauchy sequence of real numbers, and thus by Theorem 1.10.2 (Cauchy's criterion for sequences), the sequence of real numbers $\{F_n(x)\}$ converges. For each $x \in I$, we denote the limit by $F(x) = \lim_{n\to\infty} F_n(x)$. Then $\{F(x)\}_{x\in I}$ forms a function on I, which we denote by F. Given any $\epsilon > 0$, by equation 1, there exists N such that

$$|F_n(x) - F_m(x)| < \frac{\epsilon}{2} \quad \text{for all } x \in I \text{ and all } m, n > N.$$

Then for each fixed $x \in I$ and $n > N$, we have

$$\begin{aligned}|F_n(x) - F(x)| &= |F_n(x) - \lim_{m\to\infty} F_m(x)| \\ &= \lim_{m\to\infty} |F_n(x) - F_m(x)| \\ &\le \lim_{m\to\infty} \frac{\epsilon}{2} \\ &= \frac{\epsilon}{2} < \epsilon.\end{aligned}$$

Thus $\{F_n\}$ converges uniformly to F, and this finishes the proof of the 'if' part.

Example 3.2.2.

1. Show that $F_n(x) = \frac{\sin^2 x}{n}$, $\qquad x \in (-\infty, +\infty)$, converges uniformly.

3.3. Importance of Uniform Convergence of $\{F_n\}$

We address the 2 questions raised in section 3.1 in the context of uniform convergence.

Suppose a sequence of continuous functions $\{F_n\}$ converges uniformly to a function F.

1. Will F be continuous?
2. Is $\lim_{n\to\infty} \int_a^b F_n(x)\,dx = \int_a^b F(x)\,dx$?

The answer to both questions is YES.

Question 1 is answered in Theorem 3.3.1, while Question 2 is answered in Theorem 3.4.2.

Theorem 3.3.1.

Let $\{F_n\}$ be a sequence of continuous functions on an interval I. Suppose that $\{F_n\}$ converges uniformly to a function F on I. Then F is continuous on I.

Proof. Fix any point $x_o \in I$. We are going to show that F is continuous at the point x_o. Given any $\epsilon > 0$, since $\{F_n\}$ converges uniformly to F on I, there exists N such that

$$|F_n(x) - F(x)| < \frac{\epsilon}{3} \quad \text{for all } x \in I \text{ and all } n > N. \tag{2}$$

Next we fix an $n > N$ (say, $n = [N] + 1$). Since F_n is continuous at x_o, there exists $\delta > 0$ (here δ depends on x_o and ϵ) such that for all x satisfying $|x - x_o| < \delta$, we have

$$|F_n(x) - F_n(x_o)| < \frac{\epsilon}{3}. \tag{3}$$

Then for all x satisfying $|x - x_o| < \delta$, we have

$$\begin{aligned}
&|F(x) - F(x_o)| \\
&= |F(x) - F_n(x) + F_n(x) - F_n(x_o) + F_n(x_o) - F(x_o)| \\
&\le |F(x) - F_n(x)| + |F_n(x) - F_n(x_o)| \\
&+ |F_n(x_o) - F(x_o)| \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&= \epsilon \quad \text{(by equations 2 and 3).}
\end{aligned}$$

Thus F is continuous at x_o. Since x_o is arbitrary, it follows that F is continuous on I. This finishes the proof of the theorem.

Example 3.3.1.

1. Find the pointwise limit F of the sequence

$$F_n = \frac{x^{2n}}{1 + x^{2n}}, \qquad x \in [0, 1].$$

Show using Theorem 3.3.1 that the convergence is not uniform.

3.4. Uniform Convergence and Riemann Integration

Before we go on, we first recall some facts about Riemann integrals (Reference: Chapter 6, p.150-160 of [1]).

Review of Riemann Integration

Let f be a bounded function on a finite interval $[a, b]$. A *partition* P of $[a, b]$ is a set of points $\{x_0, x_1, \cdots, x_n\}$ such that

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b.$$

For such a partition P and $i = 1, 2, \cdots n$, we denote

$$M_i(f) = \sup_{x \in [x_{i-1}, x_i]} f(x)$$

$$m_i(f) = \inf_{x \in [x_{i-1}, x_i]} f(x).$$

The *upper (Riemann) sum* of f with respect to the partition P is defined to be

$$U(P, f) := \sum_{i=1}^{n} M_i(f) \Delta x_i.$$

Here $\Delta x_i = x_i - x_{i-1}$. Similarly the *lower (Riemann) sum* of f with respect to P is defined to be

$$L(P, f) := \sum_{i=1}^{n} m_i(f) \Delta x_i.$$

Then f is said to be *Riemann integrable* on $[a, b]$ if and only if $\inf_P U(P, f) = \sup_P L(P, f)$, where the infimum and supremum are taken over all partitions P of $[a, b]$. When f is Riemann integrable, the common value of the above equality is called the *Riemann integral* of f over $[a, b]$, and it is denoted by $\displaystyle\int_a^b f(x)\, dx$. We recall

Theorem 3.4.1.

A function f is Riemann integrable on a finite interval $[a, b]$ if and only if given any $\epsilon > 0$, there exists a partition P of $[a, b]$ such that

$$U(P, f) - L(P, f) < \epsilon.$$

– End of Review –

Theorem 3.4.2.
Let $\{F_n\}$ be a sequence of Riemann integrable functions on a finite interval $[a,b]$. Suppose that $\{F_n\}$ converges uniformly to a function F on $[a,b]$. Then F is Riemann integrable on $[a,b]$, and

$$\lim_{n\to\infty}\int_a^b F_n(x)\,dx = \int_a^b F(x)\,dx \quad i.e.$$

$$\lim_{n\to\infty}\int_a^b F_n(x)\,dx = \int_a^b \Big(\lim_{n\to\infty} F_n(x)\Big)\,dx. \tag{4}$$

Proof. First we show that F is Riemann integrable on $[a,b]$ (using the criterion in Theorem 3.4.1. Since each F_n is a bounded function on $[a,b]$ and $\{F_n\}$ converges uniformly to F on $[a,b]$, it follows that F is also a bounded function on $[a,b]$. (Check this!) Given any $\epsilon > 0$, since $\{F_n\}$ converges uniformly to F on $[a,b]$, there exists N such that

$$|F_n(x) - F(x)| < \frac{\epsilon}{3(b-a)} \quad \text{for all } x \in [a,b] \text{ and all } n > N. \tag{5}$$

Fix an $n > N$ (say, $n = [N]+1$). Since F_n is Riemann integrable on $[a,b]$, it follows from Theorem 3.4.1 that there exists a partition P of $[a,b]$ such that

$$U(P,F_n) - L(P,F_n) < \epsilon.$$

Fix such a partition P, and write

$$M_i(F) = \sup_{x\in[x_{i-1},x_i]} F(x),$$
$$M_i(F_n) = \sup_{x\in[x_{i-1},x_i]} F_n(x),$$
$$m_i(F) = \inf_{x\in[x_{i-1},x_i]} F(x),$$
$$m_i(F_n) = \inf_{x\in[x_{i-1},x_i]} F_n(x).$$

(Check this!) Then it follows from equation 5 that for $i = 1, 2, \cdots, n$,

$$|M_i(F) - M_i(F_n)| \le \frac{\epsilon}{3(b-a)},$$

$$|m_i(F) - m_i(F_n)| \leq \frac{\epsilon}{3(b-a)}.$$

Then

$$\begin{aligned}
&U(P,F) - L(P,F) \\
&= \sum_{i=1}^{n} M_i(F)\Delta x_i - \sum_{i=1}^{n} m_i(F)\Delta x_i \\
&= \sum_{i=1}^{n} \big(M_i(F) - m_i(F)\big)\Delta x_i \\
&= \sum_{i=1}^{n} \big[M_i(F) - M_i(F_n) + M_i(F_n) \\
&\qquad\qquad - m_i(F_n) + m_i(F_n) - m_i(F)\big]\Delta x_i \\
&\leq \sum_{i=1}^{n} \big[|M_i(F) - M_i(F_n)| + M_i(F_n) \\
&\qquad\qquad - m_i(F_n) + |m_i(F_n) - m_i(F)|\big]\Delta x_i \\
&< \sum_{i=1}^{n} \big[\frac{\epsilon}{3(b-a)} + \frac{\epsilon}{3(b-a)} + \frac{\epsilon}{3(b-a)}\big]\Delta x_i \\
&= \frac{\epsilon}{b-a}\sum_{i=1}^{n} \Delta x_i \\
&= \frac{\epsilon}{b-a} \cdot (b-a) \\
&= \epsilon.
\end{aligned}$$

Using the criterion in Theorem 3.4.1, it follows that F is Riemann integrable on $[a,b]$.

Next we are going to prove equation 4. For any given $\epsilon > 0$, since $\{F_n\}$ converges uniformly to F on $[a,b]$, there exists N such that

$$|F_n(x) - F(x)| < \frac{\epsilon}{2(b-a)} \quad \text{for all } x \in [a,b] \text{ and all } n > N.$$

Then

$$
\begin{aligned}
&\Big| \int_a^b F_n(x)\,dx - \int_a^b F(x)\,dx \Big| \\
&= \Big| \int_a^b \big(F_n(x) - F(x)\big)\,dx \Big| \\
&\leq \int_a^b \big|F_n(x) - F(x)\big|\,dx \\
&\leq \int_a^b \frac{\epsilon}{2(b-a)}\,dx \\
&\leq \frac{\epsilon}{2(b-a)} \cdot (b-a) \\
&= \frac{\epsilon}{2} < \epsilon.
\end{aligned}
$$

Thus we have $\lim_{n\to\infty} \int_a^b F_n(x)\,dx = \int_a^b F(x)\,dx$, and this finishes the proof of the theorem.

Example 3.4.1.

1. Compute, justifying your answer,

$$\lim_{n\to\infty} \int_0^1 \frac{\sin nx}{n + x^2}\,dx.$$

3.5. Series of Functions

A *series of functions* on an interval I is of the form

$$\sum_{n=1}^{\infty} f_n(x) = f_1(x) + f_2(x) + \cdots ,$$

where each f_n is a function on I.

Example 3.5.1.

1. $\displaystyle\sum_{n=1}^{\infty} x^{n-1} = 1 + x + x^2 + x^3 + \cdots$
2. $\displaystyle\sum_{k=1}^{\infty} \frac{\sin kx}{k+x} = \frac{\sin x}{1+x} + \frac{\sin 2x}{2+x} + \frac{\sin 3x}{3+x} + \cdots , \qquad 0 \leq x \leq 1.$

As in chapter 2, we may form the partial sums

$$S_n(x) = \sum_{k=1}^{n} f_k(x) = f_1(x) + f_2(x) + \cdots + f_n(x).$$

Then $\{S_n\}$ forms a sequence of functions on I.

Definition 3.5.1. The series $\sum_{n=1}^{\infty} f_n$ is said to *converge pointwise* (to a function S) on I if $\{S_n\}$ converges pointwise (to S) on I, (i.e. $\lim_{n\to\infty} S_n(x) = S(x)$ for each $x \in I$.)

Example 3.5.2.

1. What is the pointwise limit of $\sum_{n=1}^{\infty} x^{n-1}$, where $x \in (-1, 1)$?

Definition 3.5.2. $\sum_{n=1}^{\infty} f_n$ is said to *converge uniformly (to S)* on I if $\{S_n\}$ converges uniformly (to S) on I.

Example 3.5.3.

1. Does $\sum_{n=1}^{\infty} x^{n-1}$ converge uniformly on $[0, \frac{1}{2}]$?

The following test is very useful in verifying that certain series of functions converge uniformly to some functions on an interval.

Theorem 3.5.3 [Weierstrass M-test].

Consider a series of functions $\sum_{k=1}^{\infty} f_k$ *on an interval* I. *Suppose that*

(i) $|f_k(x)| \leq M_k$ *for all* $x \in I$, $k = 1, 2, \cdots$, *and*

(ii) $\sum_{k=1}^{\infty} M_k$ *converges.*

Then $\sum_{k=1}^{\infty} f_k$ *converges uniformly (to some function) on* I.

Proof. For $n = 1, 2, \cdots$, let $S_n(x) = \sum_{k=1}^{n} f_k(x)$, $x \in I$. Since $\sum_{k=1}^{\infty} M_k$ converges (by (ii)), the sequence of partial sums $\{s_n = \sum_{k=1}^{n} M_k\}$ converges, and is thus a Cauchy sequence of real numbers (see Theorem 1.10.2). Thus, given any $\epsilon > 0$, there exists N such that

$$\begin{aligned} |s_n - s_m| < \epsilon \quad & \text{for all } n > m > N \\ \Rightarrow |\sum_{k=1}^{n} M_k - \sum_{k=1}^{m} M_k| < \epsilon \quad & \text{for all } n > m > N \\ \Rightarrow \quad \sum_{k=m+1}^{n} M_k < \epsilon \quad & \text{for all } n > m > N. \end{aligned} \tag{5}$$

Then for all $x \in I$, we have

$$\begin{aligned}
|S_n(x) - S_m(x)| &= \Big| \sum_{k=1}^{n} f_k(x) - \sum_{k=1}^{m} f_k(x) \Big| \\
&= \Big| \sum_{k=m+1}^{n} f_k(x) \Big| \\
&\le \sum_{k=m+1}^{n} |f_k(x)| \\
&\le \sum_{k=m+1}^{n} M_k \quad \text{(by (i))} \\
&< \epsilon \quad \text{(by equation 5).}
\end{aligned}$$

Thus by Theorem 3.2.2 (Cauchy criterion for functions), $\{S_n\}$ converges uniformly (to some function) on I, i.e. $\sum_{k=1}^{\infty} f_k$ converges uniformly (to that function) on I.

Example 3.5.3.

1. Show that $\displaystyle\sum_{n=1}^{\infty} \frac{\cos^n x}{n^2 + x}$ converges uniformly on $(0, \infty)$.

3.6. Importance of Uniform Convergence for Series of Functions

Corollary 3.6.1.
Suppose that $\sum_{k=1}^{\infty} f_k$ *converges uniformly to a function* S *on an interval* I. *Suppose that each* f_k *is continuous on* I. *Then* S *is also continuous on* I.

Proof. Consider the sequence of partial sums $\{S_n\}$ on I, where we have $S_n = \sum_{k=1}^{n} f_k$. Then $\{S_n\}$ converges uniformly to S on I. If each f_k is continuous on I, then each S_n is also continuous on I. Then by Theorem 3.3.1, S is also continuous on I.

Example 3.6.1.

1. Is $\displaystyle\sum_{n=1}^{\infty} \frac{x}{n^2 e^{nx}}, \quad x \in (0, \infty)$, a continuous function?

Corollary 3.6.2.
Suppose that $\sum_{k=1}^{\infty} f_k$ *converges uniformly to a function* S *on an interval* $[a,b]$*. Suppose that each* f_k *is Riemann integrable on* $[a,b]$*. Then* S *is also Riemann integrable on* $[a,b]$*, and*

$$\int_a^b S(x)\,dx = \sum_{k=1}^{\infty} \int_a^b f_k(x)\,dx, \quad i.e.$$
$$\int_a^b \sum_{k=1}^{\infty} f_k(x)\,dx = \sum_{k=1}^{\infty} \int_a^b f_k(x)\,dx.$$

Proof. Consider the sequence of partial sums $\{S_n\}$ on $[a,b]$, where $S_n = \sum_{k=1}^{n} f_k$. Then $\{S_n\}$ converges uniformly to S on $[a,b]$. If each f_k is Riemann integrable on $[a,b]$, then each S_n is also Riemann integrable on $[a,b]$. Then by Theorem 3.4.2, S is also Riemann integrable on $[a,b]$, and

$$\begin{aligned}\int_a^b S(x)\,dx &= \lim_{n\to\infty} \int_a^b S_n(x)\,dx \\ &= \lim_{n\to\infty} \int_a^b \sum_{k=1}^{n} f_k(x)\,dx \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \int_a^b f_k(x)\,dx \\ &= \sum_{k=1}^{\infty} \int_a^b f_k(x)\,dx.\end{aligned}$$

Example 3.6.2.

Theorem 3.6.3.
Let $\{F_n\}$ be a sequence of functions on $[a,b]$ such that
(i) each F_n' exists and is continuous on $[a,b]$,
(ii) $\{F_n\}$ converges pointwise to a function F on $[a,b]$, and
(iii) $\{F_n'\}$ converges uniformly on $[a,b]$.
Then F is differentiable on $[a,b]$, and for all $x \in [a,b]$,

$$F'(x) = \lim_{n\to\infty} F_n'(x), \ \ i.e.$$
$$\frac{d}{dx}\big(\lim_{n\to\infty} F_n(x)\big) = \lim_{n\to\infty}\big(\frac{d}{dx}F_n(x)\big).$$

Remark. Here the differentiability and continuity at the endpoints a and b refer to the one sided derivatives and limits respectively.

Proof. By (iii), there exists a function g such that $\{F_n'\}$ converges uniformly to g on $[a,b]$. In particular, $\lim_{n\to\infty} F_n'(x) = g(x)$ for all $x \in [a,b]$. By (i), since each F_n' is continuous on $[a,b]$, F_n' is also Riemann integrable on $[a,b]$, and by the fundamental theorem of calculus,

$$\int_a^x F_n'(t)\,dt = F_n(x) - F_n(a) \quad \text{for all } x \in [a,b].$$

Letting $n \to \infty$, we have, for all $x \in [a,b]$,

$$\begin{aligned}\lim_{n\to\infty}\int_a^x F_n'(t)\,dt &= \lim_{n\to\infty}\big(F_n(x) - F_n(a)\big)\\ &= F(x) - F(a). \end{aligned} \tag{6}$$

On the other hand, since $\{F_n'\}$ converges uniformly to g on $[a,b]$, it follows from Theorem 3.4.2 that

$$\begin{aligned}\lim_{n\to\infty}\int_a^x F_n'(t)\,dt &= \int_a^x \big(\lim_{n\to\infty} F_n'(t)\big)dt\\ &= \int_a^x g(t)\,dt.\end{aligned}$$

Together with equation 7, it follows that

$$F(x) - F(a) = \int_a^x g(t)\,dt \quad \text{for all } x \in [a,b]. \tag{7}$$

By (i) and Theorem 3.3.1, g is continuous on $[a,b]$. Then by the fundamental theorem of calculus, we have

$$\frac{d}{dx}\int_a^x g(t)\,dt = g(x).$$

Together with equation 7, it follows that F is also differentiable on $[a,b]$, and for all $x \in [a,b]$,

$$\begin{aligned}\frac{d}{dx}\big(F(x) - F(a)\big) &= \frac{d}{dx}\int_a^x g(t)\,dt = g(x), \quad \text{i.e.}\\ F'(x) &= g(x), \quad \text{i.e.}\\ \frac{d}{dx}\big(\lim_{n\to\infty} F_n(x)\big) &= \lim_{n\to\infty}\big(\frac{d}{dx}F_n(x)\big).\end{aligned}$$

Corollary 3.6.4.

Let $\sum_{k=1}^{\infty} f_k$ *be a series of functions on* $[a, b]$ *such that*

(i) *each* f'_k *exists and is continuous on* $[a, b]$*,*

(ii) $\sum_{k=1}^{\infty} f_k$ *converges pointwise to a function* S *on* $[a, b]$*, and*

(iii) $\sum_{k=1}^{\infty} f'_k$ *converges uniformly on* $[a, b]$*.*

Then S *is differentiable on* $[a, b]$*, and for all* $x \in [a, b]$*,*

$$S'(x) = \sum_{k=1}^{\infty} f'_k(x), \quad \textit{i.e.}$$

$$\frac{d}{dx}\Big(\sum_{k=1}^{\infty} f_k(x)\Big) = \sum_{k=1}^{\infty} \frac{d}{dx} f_k(x).$$

Proof. Consider the sequence of partial sums (of functions) $\{S_n\}$ on $[a, b]$, where $S_n = \sum_{k=1}^{n} f_k$ for each n. Then by (ii),

$$\lim_{n\to\infty} S_n(x) = S(x) \quad \text{for all } x \in [a, b].$$

The conditions (i), (ii), (iii) of the corollary imply that the sequence of functions $\{S_n\}$ satisfies the corresponding conditions (i), (ii), (iii) of Theorem 3.6.3. Thus by Theorem 3.6.3, the function S is differentiable on $[a, b]$, and for all $x \in [a, b]$,

$$S'(x) = \lim_{n\to\infty} S'_n(x), \quad \text{i.e.}$$

$$\begin{aligned}
\frac{d}{dx}\Big(\sum_{k=1}^{\infty} f_k(x)\Big) &= \lim_{n\to\infty} \Big(\sum_{k=1}^{n} f_k(x)\Big)' \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} f'_k(x) \\
&= \sum_{k=1}^{\infty} \frac{d}{dx} f_k(x).
\end{aligned}$$

Example 3.6.3.

3.7. Power Series

Definition 3.7.1. A *power series* in x is of the form

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots$$

Example 3.7.1.

1. $\displaystyle\sum_{n=0}^{\infty}(n+1)x^n = 1 + 2x + 3x^2 + 4x^3 + 5x^4 + \cdots$
2. $\displaystyle\sum_{n=0}^{\infty}\frac{x^n}{n!} =$

Definition 3.7.2 A *power series* in $x - x_0$ is of the form

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots$$

Example 3.7.2.

1. $\displaystyle\sum_{n=0}^{\infty}(x-1)^n = 1$
2. $\displaystyle\sum_{n=1}^{\infty} n^2(x+2)^n =$

Question: Given a power series $\displaystyle\sum_{n=0}^{\infty} a_n(x - x_0)^n$, when does it converge and when does it diverge?

Theorem 3.7.3.
Given any power series $\sum_{k=0}^{\infty} a_k(x-x_o)^k$, *there is an associated number* R, $0 \le R \le \infty$, *called the radius of convergence, with the following properties:*
(i) The series of real numbers $\sum_{k=0}^{\infty} a_k(x-x_o)^k$ *converges absolutely at each point* x *satisfying* $|x-x_o| < R$.
(ii) $\sum_{k=0}^{\infty} a_k(x-x_o)^k$ *diverges at each* x *satisfying* $|x-x_o| > R$.
(iii) The series of functions $\sum_{k=0}^{\infty} a_k(x-x_o)^k$ *converges uniformly on the interval* $|x-x_o| \le \rho$ *for any* ρ *satisfying* $0 < \rho < R$.
Moreover, R *is given by*

$$R = \frac{1}{\limsup |a_k|^{\frac{1}{k}}}. \tag{8}$$

($R = 0$ *if* $\limsup |a_k|^{\frac{1}{k}} = \infty$*; and* $R = \infty$ *if* $\limsup |a_k|^{\frac{1}{k}} = 0$*.) In addition, if* $\lim_{k\to\infty} \frac{|a_{k+1}|}{|a_k|}$ *exists, then* R *is also given by*

$$R = \frac{1}{\lim_{k\to\infty} \frac{|a_{k+1}|}{|a_k|}}. \tag{9}$$

Proof. Let $R = 1/\limsup |a_k|^{\frac{1}{k}}$ be as given in equation 8. First we are going to prove (i). For each point x satisfying $|x-x_o| < R$, we consider the series of (non-negative) real numbers $\sum_{k=0}^{\infty} |a_k(x-x_o)^k|$. We have

$$\begin{aligned}
\limsup \left(\left|a_k(x-x_o)^k\right|\right)^{\frac{1}{k}} &= \limsup \left(|a_k|^{\frac{1}{k}} |x-x_o|\right) \\
&= |x-x_o| \cdot \limsup |a_k|^{\frac{1}{k}} \\
&< R \cdot \frac{1}{R} \\
&= 1.
\end{aligned}$$

Thus by the root test, $\sum_{k=0}^{\infty} |a_k(x-x_o)^k|$ converges, i.e. $\sum_{k=0}^{\infty} a_k(x-x_o)^k$ converges absolutely.

Next we are going to prove (ii) by contradiction. Suppose that $\sum_{k=0}^{\infty} a_k(x-x_o)^k$ converges at a point x satisfying $|x-x_o| > R$. Then by Theorem 2.1.1, we have

$$\lim_{k\to\infty} a_k(x-x_o)^k = 0.$$

Letting $\epsilon = 1$. Then there exists N such that

$$\begin{aligned}
& |a_k(x-x_o)^k - 0| < 1 \quad \text{for all } k > N \\
\Rightarrow \quad & |a_k(x-x_o)^k|^{\frac{1}{k}} < 1 \quad \text{for all } k > N \\
\Rightarrow \quad & |a_k|^{\frac{1}{k}} < \frac{1}{|x-x_o|} \quad \text{for all } k > N \\
\Rightarrow \quad & \sup_{n\geq k} |a_k|^{\frac{1}{k}} \leq \frac{1}{|x-x_o|} \quad \text{for all } n > N \\
\Rightarrow \quad & \limsup |a_k|^{\frac{1}{k}} \leq \frac{1}{|x-x_o|} \\
\Rightarrow \quad & \frac{1}{R} \leq \frac{1}{|x-x_o|} < \frac{1}{R},
\end{aligned}$$

which is a contradiction. Hence we must have $\sum_{k=0}^{\infty} a_k(x-x_o)^k$ diverges at each x satisfying $|x-x_o| > R$.

Now we are going to prove (iii). Let ρ be a number such that $0 < \rho < R$. First, we have, for each k and all x satisfying $|x-x_o| \leq \rho$,

$$|a_k(x-x_o)^k| \leq |a_k| \cdot \rho^k. \tag{10}$$

Now we apply the root test to the series

$\sum_{k=0}^{\infty} |a_k| \cdot \rho^k$. We have

$$\begin{aligned}\limsup \left(|a_k| \cdot \rho^k\right)^{\frac{1}{k}} &= \limsup \left(\left(|a_k|\right)^{\frac{1}{k}} \cdot \rho\right) \\ &= \rho \cdot \limsup |a_k|^{\frac{1}{k}} \\ &< R \cdot \frac{1}{R} \\ &= 1.\end{aligned}$$

Thus by the root test, $\sum_{k=0}^{\infty} |a_k| \cdot \rho^k$ converges. Together with equation 10, it follows from the Weierstrass M-test that the series of functions $\sum_{k=0}^{\infty} a_k(x - x_o)^k$ converges uniformly on $[x_o - \rho, x_o + \rho]$.

Finally the proof of the Theorem under the condition (9) in place of (8) is similar (with the root test replaced by the ratio test in various places), and it will be left to the student as an exercise.

Example 3.7.3.

1. What is the radius of convergence for the series

$$1 + \frac{x}{3} + \frac{x^2}{4^2} + \frac{x^3}{3^3} + \frac{x^4}{4^4} + \frac{x^5}{3^5} + \frac{x^6}{4^6} + \cdots$$

Corollary 3.7.4

Given any $\sum_{k=0}^{\infty} a_k(x - x_o)^k$ *with radius of convergence* R*,*

(i) the series of real numbers $\sum_{k=0}^{\infty} a_k(x - x_o)^k$ *converges at each point* x *satisfying*
$|x - x_o| < R$*.*

(ii) $\sum_{k=0}^{\infty} a_k(x - x_o)^k$ *diverges at each* x *satisfying* $|x - x_o| > R$*.*

Proof. (i) follows from Theorem 3.7.3(i) and the fact that an absolutely convergent series is necessarily convergent. (ii) follows from Theorem 3.7.3(ii).

3.8. Interval of convergence.

In view of Corollary 3.7.4, for a power series $\sum_{k=0}^{\infty} a_k(x-x_o)^k$ with radius of convergence R, the set of points at which $\sum_{k=0}^{\infty} a_k(x-x_o)^k$ is convergent form an interval called the *interval of convergence*, which must be either

$$(x_o - R, x_o + R), \qquad (x_o - R, x_o + R],$$

$$[x_o - R, x_o + R) \quad \text{or} \quad [x_o - R, x_o + R].$$

Example 3.8.1.

1. Find the interval of convergence of $\displaystyle\sum_{n=1}^{\infty} \frac{(x-2)^n}{n^2}$.

Theorem 3.8.1.
Suppose that $\sum_{k=0}^{\infty} a_k(x-x_o)^k$ *has radius of convergence* $R > 0$ *with pointwise limiting function* $f(x)$ *on* $|x-x_o| < R$ *(i.e.* $f(x) = \sum_{k=0}^{\infty} a_k(x-x_o)^k$ *on* $|x-x_o| < R$*), then* $f(x)$ *has derivatives of all orders on* $|x-x_o| < R$*, and*

$$a_k = \frac{f^{(k)}(x_o)}{k!} \quad \textit{for all } k. \tag{11}$$

(i.e. we have $f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_o)}{k!}(x-x_o)^k$*.)*

Proof. First we are going to show that the series $\sum_{k=1}^{\infty} ka_k(x-x_o)^{k-1}$ also has radius of convergence R. To see this, we first observe easily that $|ka_k| \geq |a_k|$ for $k \geq 1$, and thus

$$\limsup |ka_k|^{\frac{1}{k}} \geq \limsup |a_k|^{\frac{1}{k}}. \tag{12}$$

For any $\epsilon > 0$, since $\lim_{k\to\infty} k^{\frac{1}{k}} = 1$, it follows that there exists N such that

$$\begin{aligned} |k^{\frac{1}{k}} - 1| &< 1+\epsilon \quad \text{for all } k > N \\ \Rightarrow \quad k^{\frac{1}{k}} &< 1+\epsilon \quad \text{for all } k > N \\ \Rightarrow \quad |ka_k|^{\frac{1}{k}} &< |a_k|^{\frac{1}{k}}(1+\epsilon) \quad \text{for all } k > N \\ \Rightarrow \quad \limsup |ka_k|^{\frac{1}{k}} &\leq \limsup |a_k|^{\frac{1}{k}}(1+\epsilon). \end{aligned}$$

Letting $\epsilon \to 0$, it follows that

$$\limsup |ka_k|^{\frac{1}{k}} \leq \limsup |a_k|^{\frac{1}{k}}.$$

Together with equation 12, it follows that

$$\limsup |ka_k|^{\frac{1}{k}} = \limsup |a_k|^{\frac{1}{k}} = 1/R.$$

Taking reciprocals, it follows that

$\sum_{k=1}^{\infty} ka_k(x-x_o)^{k-1}$ also has radius of convergence R.

Next we are going to apply Corollary 3.6.4 (with $f_k(x) = a_k(x-x_o)^k$) on any interval $|x-x_o| \leq \rho$ for any $\rho < R$. Obviously, each
$f_k'(x) = ka_k(x-x_o)^{k-1}$ exists and is continuous on $|x-x_o| \leq \rho$. Thus condition (i) of Corollary 3.6.4 is satisfied. Condition (ii) of Corollary 3.6.4 follows from Theorem 3.7.3 (i). Since the radius of $\sum_{k=1}^{\infty} ka_k(x-x_o)^{k-1}$ is R,
by Theorem 3.7.3 (iii) (applied to $\sum_{k=1}^{\infty} ka_k(x-x_o)^{k-1}$), it follows that
$\sum_{k=1}^{\infty} ka_k(x-x_o)^{k-1}$ converges uniformly on $|x-x_o| \leq \rho$, i.e. condition (iii) of Corollary 3.6.4 is also satisfied.

Thus by Corollary 3.6.4, $f(x)$ is differentiable on $|x-x_o| \leq \rho$, and for all x satisfying $|x-x_o| \leq \rho$, one has

$$f'(x) = \sum_{k=0}^{\infty} \frac{d}{dx} a_k(x-x_o)^k = \sum_{k=1}^{\infty} ka_k(x-x_o)^{k-1}. \tag{13}$$

In particular, by evaluating at $x = x_o$, one has

$$f'(x_o) = a_1.$$

For any number x satisfying $|x-x_o| < R$, we can choose a number $\rho < R$ such that
$|x-x_o| \leq \rho$, and thus $f(x)$ is differentiable at x. Hence $f(x)$ is differentiable and (i) holds everywhere on $|x-x_o| < R$.

Repeating the above argument (with $f(x) = \sum_{k=0}^{\infty} a_k(x-x_o)^k$ replaced by $f'(x) = \sum_{k=1}^{\infty} ka_k(x-x_o)^{k-1}$ on $|x-x_o| < R$), it follows that $f'(x)$ is differentiable on

$|x - x_o| < R$, and

$$f''(x_o) = \sum_{k=2}^{\infty} k(k-1)a_k(x - x_o)^{k-2}\Big|_{x=x_o} = 2 \cdot 1 \cdot a_2.$$

Repeating the above argument again and again, it follows that $f(x)$ is has derivatives of all order on $|x - x_o| < R$, and $f^k(x_o) = k! \cdot a_k$ for all k, which leads to equation 11. This finishes the proof of the Theorem.

3.9. Taylor Series

Definition 3.9.1. For any function $f(x)$ which has derivatives of all orders at a point $x = x_o$, we may construct the power series

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_o)}{k!}(x - x_o)^k,$$

which is called the *Taylor series* of f at $x = x_o$.

Example 3.9.1.

1. Find the Taylor series of e^x at $x_0 = 0$.

In view of Theorem 3.8.1, we may ask the following question:

Question: Does the equality

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_o)}{k!}(x - x_o)^k \tag{14}$$

hold for $|x - x_o| < R$?
(Here R is the radius of the convergence of the Taylor series.)

It turns out that in general, the answer is NO. (See Example 3.9.2 for an example of a function such that equation 14 does not hold.

However, the above equality does hold for some elementary functions such as $e^x, \sin x$, $\cos x$, $\ln(1 + x)$.

Definition 3.9.2. Functions for which equation 14 hold are called *analytic functions.*

Example 3.9.2.

1. Consider the function

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Then we will show that $f(x) \neq$ its Taylor series at $x_0 = 0$.

Some standard analytic functions and their Taylor series.

Some well known analytic functions and their Taylor series at $x = 0$ are given as follows:

$$\begin{aligned}
e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \\
&= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \qquad (|x| < \infty) \\
\sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \\
&= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots \qquad (|x| < \infty). \\
\cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \\
&= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots \qquad (|x| < \infty). \\
\ln(1+x) &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n} \\
&= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \qquad (|x| < 1). \\
\frac{1}{1-x} &= \sum_{n=0}^{\infty} x^n \\
&= 1 + x + x^2 + x^3 + \cdots \qquad (|x| < 1). \\
\frac{1}{1+x} &= \sum_{n=0}^{\infty} (-1)^n x^n \\
&= 1 - x + x^2 - x^3 + \cdots \qquad (|x| < 1).
\end{aligned}$$

Remark. We may use Theorem 3.8.1 and the standard Taylor series to find the Taylor series of certain analytic functions.

Example 3.9.3.

1. Compute the Taylor series for $\dfrac{1}{1+x^2}$ where $|x| < 1$ using the Taylor series for $\dfrac{1}{1+x}$ on the same domain.

Chapter 4: Ordinary Differential Equations (ODE)

4.1. Classification of Differential Equations

A differential equation (DE) is an equation involving an unknown function and its derivatives.

Example 4.1.1.

1. $4\frac{d^2y}{dt^2}+\frac{dy}{dt}+2y=0$ is a DE with y being the function (or dependent variable) and t being the independent variable.

Remark. The word 'ordinary' in the heading of this chapter means that the differential equation involves only 'ordinary' derivatives of the function (rather than partial derivatives of the function). If you had taken MA1104, then you should know that the derivatives of a function in two or more variables, say $f(x, y)$, are known as partial derivatives. In this course, we deal with functions of one variable only, thus the word 'ordinary' carries no special meaning to us.

Classifying DE by its order

The *order* of a differential equation (DE) is the order of its highest derivative.

Example 4.1.2.

1. $4\dfrac{d^2y}{dt^2} + \dfrac{dy}{dt} + 2y = 0$ is a   order DE.

2. $\dfrac{dy}{dt} - 2ty = t$ is a   order DE

3. $(\dfrac{dy}{dt})^3 - 2ty^3 = 0$ is a   order ODE.

Linear or non-linear DE

The *total power* of a term in a DE is the sum of all the powers of y and its derivatives in that term.

A *linear* DE is a DE where the total power of each term is at most 1.

Example 4.1.3.

1. $$4\frac{d^2y}{dt^2} + \frac{dy}{dt} + 2y = 0$$

2. $$\frac{dy}{dt} - 2ty = t:$$

3. $$y\frac{dy}{dt} - 2ty = t$$

Homogeneous or non-homogeneous DE

A *homogeneous* DE is a DE where the total power of each non-zero term is the same.

Example 4.1.4.

1. $$4\frac{d^2y}{dt^2} + \frac{dy}{dt} + 2y = 0.$$

2. $$\frac{dy}{dt} - 2ty = t.$$

3. $$y\frac{dy}{dt} - 2ty^2 = 0$$

We may combine the classifications:

Example 4.1.5.

1. $4\dfrac{d^2y}{dt^2} + \dfrac{dy}{dt} + 2y = 0$

   is a homogeneous/non-homogeneous, linear/non-linear, order ODE.

2. $\dfrac{dy}{dt} - 2ty = t$

3. $(\dfrac{dy}{dt})^3 - 2ty = t$

4.2 General solutions of homogeneous linear first order ODE

This is the easiest DE to solve:

$$\frac{dy}{dt} + g(t)y = 0. \tag{1}$$

Separation of variables. The method is to "separate the variables", y and t. Equation 1 can be rearranged as

$$\frac{dy}{y} = -g(t)dt.$$

Upon integrating both sides, we get

$$\begin{aligned} \int \frac{dy}{y} &= -\int g(t)dt \\ \ln|y| &= -\int g(t)dt + k, \\ |y| &= e^k \exp\Big(-\int g(t)dt\Big). \end{aligned}$$

The solution is then

$$y(t) = C \exp\Big(-\int g(t)dt\Big),$$

where C is an arbitrary constant. The above solution is said to be the *general solution* of the DE, equation 1, since every solution of equation 1 must be of this form.

Example 4.2.1.

1. Find the general solution of $\dfrac{dy}{dt} - 2ty = 0$.

4.3. General solutions of non-homogeneous linear first order ODE

The DE to consider is:

$$\frac{dy}{dt} + g(t)y = f(t). \tag{2}$$

There are essentially 2 methods to solve this:

Method 1: Multiply equation 2 by an Integrating Factor, $\mu(t)$, where

$$\mu(t) = exp(\int g(t)dt). \tag{3}$$

We get

$$\mu(t)\frac{dy}{dt} + \mu(t)g(t)y = \mu(t)f(t).$$

Using equation 3, it is easy to check that

$$\frac{d}{dt}\Big(\mu(t)y\Big) = \mu(t)\frac{dy}{dt} + \mu(t)g(t)y,$$

so we get

$$\begin{aligned} \frac{d}{dt}\Big(\mu(t)y\Big) &= \mu(t)f(t) \\ \Rightarrow \qquad \mu(t)y &= \int \mu(t)f(t)dt + C. \end{aligned}$$

Remark. This method is not generally applicable to higher order linear DE.

Example 4.3.1.

1. Find the general solution of $\dfrac{dy}{dt} - 2ty = t$.

Method 2: For a linear DE, say, equation 2, the general solution can be written as

$$y(t) = y_h(t) + y_p(t),$$

where y_h is the general solution of the corresponding homogeneous DE

$$\frac{dy}{dt} + g(t)y = 0,$$

while y_p is any one solution (called a *particular* solution) of equation 2.

Example 4.3.2.

1. Find the general solution of $\dfrac{dy}{dt} - 2ty = t$.

4.4. Some non-linear first order ODEs

We remark that the method of separation of variables can also be used to solve certain non-homogeneous non-linear first order ODE of the form

$$\frac{dy}{dt} + f(t)g(y) = 0, \quad \text{or}$$
$$\frac{dy}{dt} = f(\frac{y}{t}).$$

Example 4.4.1.

1. Find the general solution of $\dfrac{dy}{dt} = \dfrac{t}{y^2}$.

4.5. Homogeneous linear 2nd order ODE with constant coefficients

Homogeneous linear 2nd order ODE with constant coefficients are of the form

$$a\frac{d^2y}{dt^2} + b\frac{dy}{dt} + cy = 0, \tag{4}$$

where a, b, c are real constants.

Example 4.5.1.

1. $2\frac{d^2y}{dt^2} + 3\frac{dy}{dt} - 5y = 0.$

2. $y'' + 5y' + 4y = 0.$

Remark: The general solution of equation 4 is of the form

$$y(t) = Ay_1(t) + By_2(t),$$

where A, B are arbitrary constants, and $y_1(t)$, $y_2(t)$ are solutions of equation 4 such that $y_1(t)$ and $y_2(t)$ are not constant multiples of each other (see §2.1, p. 129-130 of [3]).

To solve equation 4, we look at the case of the homogeneous linear lst order ODE with constant coefficients, which is of the form

$$y' + ky = 0, \qquad k \in \mathbb{R}.$$

Its general solution is of the form $y(t) = Ae^{-kt}$.

Thus, to solve equation 5, we try

$$y(t) = e^{rt}, \qquad \text{where } r \in \mathbb{R}. \tag{5}$$

Then we have

$$y'(t) = re^{rt}, \quad y''(t) = r^2e^{rt}. \tag{6}$$

Substituting equations 5 and 6 into 4, equation 4 becomes

$$\begin{aligned} ay'' + by' + cy &= 0 \\ \Rightarrow ar^2e^{rt} + bre^{rt} + ce^{rt} &= 0 \\ \Rightarrow e^{rt}(ar^2 + br + c) &= 0 \\ \Rightarrow ar^2 + br + c &= 0. \end{aligned}$$

The quadratic equation

$$ar^2 + br + c = 0$$

is known as the *auxiliary* (or *characteristic*) equation of the homogeneous linear second order ODE, equation 4. We denote its two roots by r_1, r_2. The general solution $y(t)$ for equation 4 will depend on what kind of roots r_1 and r_2 are.

Case (a): r_1, r_2 are real and distinct ($r_1 \neq r_2$).

In this case, the general solution of 4 is

$$y(t) = Ae^{r_1 t} + Be^{r_2 t}.$$

By an earlier remark, it suffices to check that $y_1(t) = e^{r_1 t}$ and $y_2(t) = e^{r_2 t}$ are not constant multiples of each other and that $y_1(t)$ and $y_2(t)$ are indeed solutions of equation 4.

Check: $\dfrac{y_1(t)}{y_2(t)} = e^{(r_1 - r_2)t}$ is not a constant since $r_1 \neq r_2$. Hence $y_1(t)$ and $y_2(t)$ are not constant multiples of each other. Also, substituting $y_1(t)$ into equation 4, we get

$$\begin{aligned} LHS &= ay_1''(t) + by_1'(t) + cy_1(t) \\ &= ar_1^2 e^{r_1 t} + br_1 e^{r_1 t} + ce^{r_1 t} \\ &= e^{r_1 t}(ar_1^2 + br_1 + c) \\ &= 0 \\ &= RHS. \end{aligned}$$

Hence $y_1(t)$ is indeed a solution of equation 4. Similarly, $y_2(t)$ is also a solution of equation 4.

Case (b): r_1, r_2 are (non-real) complex numbers.

Write $r_1, r_2 = \alpha \pm i\beta$, where α, $\beta \in \mathbb{R}$ with $\beta \neq 0$.
In this case, the general solution of 4 is

$$y(t) = Ae^{\alpha t}\cos\beta t + Be^{\alpha t}\sin\beta t.$$

Again, it suffices to check that $y_1(t) = e^{\alpha t}\cos\beta t$ and $y_2(t) = e^{\alpha t}\sin\beta t$ are not constant multiples of each other, and $y_1(t)$ and $y_2(t)$ are solutions of equation 4.

Check: $\dfrac{y_2(t)}{y_1(t)} = \tan\beta t$ is not a constant since $\beta \neq 0$. Hence $y_1(t)$ and $y_2(t)$ are not constant multiples of each other. Since $\alpha \pm i\beta$ are roots of the characteristic equation, we have

$$\text{Sum of roots} = 2\alpha = -\frac{b}{a} \qquad (*)$$

$$\text{Product of roots} = \alpha^2 + \beta^2 = \frac{c}{a} \qquad (**)$$

Differentiating $y_1(t)$, we get

$$\begin{aligned}
y_1'(t) &= \alpha e^{\alpha t}\cos\beta t - \beta e^{\alpha t}\sin\beta t,\\
y_1''(t) &= \alpha^2 e^{\alpha t}\cos\beta t - 2\alpha\beta e^{\alpha t}\sin\beta t - \beta^2 e^{\alpha t}\cos\beta t.
\end{aligned}$$

Substituting $y_1(t)$ into equation 4, we get

$$\begin{aligned}
LHS &= ay_1''(t) + by_1'(t) + cy_1(t)\\
&= a[\alpha^2 e^{\alpha t}\cos\beta t - 2\alpha\beta e^{\alpha t}\sin\beta t - \beta^2 e^{\alpha t}\cos\beta t]\\
&\qquad + b[\alpha e^{\alpha t}\cos\beta t - \beta e^{\alpha t}\sin\beta t] + ce^{\alpha t}\cos\beta t\\
&= [a(\alpha^2 - \beta^2) + b\alpha + c]e^{\alpha t}\cos\beta t - (2a\alpha + b)\beta e^{\alpha t}\sin\beta t\\
&= [a(\alpha^2 - \beta^2) + (-2a\alpha)\alpha + c]e^{\alpha t}\cos\beta t - 0 \quad (\text{by } (*))\\
&= [-a(\alpha^2 + \beta^2) + c]e^{\alpha t}\cos\beta t\\
&= 0 \quad (\text{by } (**))\\
&= RHS.
\end{aligned}$$

Hence $y_1(t)$ is indeed a solution of equation 4. Similarly, $y_2(t)$ is also a solution of equation 4.

Case (c): $r_1 = r_2 (= r)$.

In this case, the general solution of equation 4 is

$$y(t) = (At + B)e^{rt}.$$

Again it suffices to check that $y_1(t) = e^{rt}$ and $y_2(t) = te^{rt}$ are not constant multiples of each other, and they are solutions of equation 4.

Check: $\dfrac{y_2(t)}{y_1(t)} = t$ is not a constant. Hence $y_1(t)$ and $y_2(t)$ are not constant multiples of each other. Also, one can show as in case (a) that $y_1(t)$ is a solution of equation 4. To check that $y_2(t)$ is a solution of equation 4, we first observe that

$$\text{Sum of roots} = 2r = -\frac{b}{a}. \qquad (*)$$

Differentiating $y_2(t)$, we get

$$\begin{aligned} y_2'(t) &= rte^{rt} + e^{rt}, \\ y_2''(t) &= r^2te^{rt} + 2re^{rt}. \end{aligned}$$

Substituting into equation 4, we get

$$\begin{aligned} LHS &= ay_2''(t) + by_2'(t) + cy_2(t) \\ &= a[r^2te^{rt} + 2re^{rt}] + b[rte^{rt} + e^{rt}] + cte^{rt} \\ &= (ar^2 + br + c)te^{rt} + (2ar + b)e^{rt} \\ &= 0 + 0 \quad \text{(by (*))} \\ &= RHS. \end{aligned}$$

Hence $y_2(t)$ is indeed a solution of equation 4.

4.6. Non-homogeneous linear 2nd order ODE with constant coefficients

Non-homogeneous linear 2nd order ODE with constant coefficients are of the general form

$$ay'' + by' + cy = g(t), \tag{7}$$

where $a, b, c \in \mathbb{R}$, and $g(t)$ is a function in t.

Example 4.6.1.

1. $y'' + y = t^2$.

Associated with the non-homogeneous equation 7 is a homogeneous equation given by

$$ay'' + by' + cy = 0. \tag{8}$$

Theorem 4.6.1. *The general solution of equation 7 is given by*

$$y(t) = y_h(t) + y_p(t),$$

where $y_h(t)$ *is the general solution of the associated homogeneous equation 8, and* $y_p(t)$ *is a particular solution of equation 7.*

Proof. Let $y(t)$ be any solution of equation 7, and let $y_p(t)$ be a particular solution of equation 8. Let $f(t) = y(t) - y_p(t)$. Then

$$\begin{aligned} & af'' + bf' + cf \\ &= a(y''(t) - y_p''(t)) + b(y'(t) - y_p'(t)) + c(y(t) - y_p(t)) \\ &= (ay''(t) + by'(t) + cy(t)) - (ay_p''(t) + by_p'(t) + cy_p(t)) \\ &= g(t) - g(t) \\ &= 0. \end{aligned}$$

Thus $f(t)$ satisfies equation 8.

Similarly, we can show that any function of the form $y_h(t) + y_p(t)$ satisfies equation 7.

Thus the general solution of equation 7 is of the form $y(t) = y_h(t) + y_p(t)$.

So the problem of solving the non-homogeneous 2nd order linear ODE with constant coefficients is reduced to finding $y_h(t)$ and $y_p(t)$.

$y_h(t)$ can be found by the method explained in section 4.5.

Finding $y_p(t)$ amounts to guesswork. But for certain kind of $g(t)$ (the non-homogeneous term of equation 7), the guesswork is quite systematic. This is explained below.

Judicious guessing/the method of undetermined coefficients

We want to find one particular solution $y_p(t)$ for the non-homogeneous equation

$$ay'' + by' + cy = g(t). \tag{9}$$

The homogeneous equation associated to equation 9 is

$$ay'' + by' + cy = 0, \tag{10}$$

with characteristic equation given by

$$ar^2 + br + c = 0. \tag{11}$$

Suppose r_1, r_2 are the roots of the characteristic equation 11. We look for particular solution $y_p(t)$ of equation 9 based on the form of $g(t)$.

Case (a): $g(t) = a_0 + a_1 t + \cdots + a_n t^n$.
Then $y_p(t)$ is of the form

$$y_p(t) = \begin{cases} A_0 + A_1 t + \cdots + A_n t^n, & c \neq 0 \\ t(A_0 + A_1 t + \cdots + A_n t^n), & c = 0,\ b \neq 0 \\ t^2(A_0 + A_1 t + \cdots + A_n t^n), & c = b = 0. \end{cases}$$

Case (b): $g(t) = (a_0 + a_1 t + \cdots + a_n t^n)e^{\alpha t}$.
The form of $y_p(t)$ will depend on whether α is a root of the characteristic equation.

(i) If $r_1, r_2 \neq \alpha$ (ie, α is not a root), try

$$y_p(t) = (A_0 + A_1 t + \cdots + A_n t^n)e^{\alpha t}.$$

(ii) If $r_1 = \alpha$, $r_2 \neq \alpha$ (ie α is a single root), try

$$y_p(t) = t(A_0 + A_1 t + \cdots + A_n t^n)e^{\alpha t}.$$

(iii) If $r_1, r_2 = \alpha$ (ie α is a double root), try

$$y_p(t) = t^2(A_0 + A_1 t + \cdots + A_n t^n)e^{\alpha t}.$$

Case (c):

$$g(t) = (a_0 + a_1 t + \cdots + a_n t^n)e^{\alpha t}\cos\beta t, \quad \text{or}$$
$$g(t) = (a_0 + a_1 t + \cdots + a_n t^n)e^{\alpha t}\sin\beta t.$$

The form of $y_p(t)$ will depend on whether $\alpha \pm i\beta$ are roots of the characteristic equation or not.

(i) If $r_1, r_2 \neq \alpha \pm i\beta$ (i.e. $\alpha \pm i\beta$ are not roots), try

$$\begin{aligned} y_p(t) = {} & (A_0 + A_1 t + \cdots + A_n t^n)e^{\alpha t}\cos\beta t \\ & + (B_0 + B_1 t + \cdots + B_n t^n)e^{\alpha t}\sin\beta t. \end{aligned}$$

(ii) If $r_1, r_2 = \alpha \pm i\beta$ (i.e. $\alpha \pm i\beta$ are roots), try

$$\begin{aligned} y_p(t) = {} & t(A_0 + A_1 t + \cdots + A_n t^n)e^{\alpha t}\cos\beta t \\ & + t(B_0 + B_1 t + \cdots + B_n t^n)e^{\alpha t}\sin\beta t. \end{aligned}$$

Remark.

1. Once you have determined the form of y_p, substitute into equation 9 to get equations in the unknown coefficients $A_0, \cdots, A_n$ and also $B_0, \cdots, B_n$ if necessary. It should be easy to solve for these coefficients.

2. Since the DE is linear, if $g(t)$ are sums of the above three types, we can still find y_p by adding the corresponding sums of the individual y_p.