Sequences, Series and Foundations

These notes by Mikhail Safonov serve as a supplementary material to the textbook by Weyne Richter "Sequences, Series and Foundations. Math 2283 and 3283W"

Chapter 1. Truth, Falsity and Mathematical Induction

1 Truth Tables

The mathematical statements P, Q, etc. can be treated as variables which can take one of two values: "T"="true" and "F"="false". If we further associate "T" with "1", and "F" with "0", then we have the following simple rules:

$$P \ \& \ Q = PQ, \qquad \neg P = 1 - P, \qquad P \vee Q = \neg((\neg P) \ \& \ (\neg Q)) = 1 - (1-P)(1-Q) = P + Q - PQ.$$

Exercise 1.1 *Simplify the statement $P \ \& \ Q = P \vee Q$.*

Solution. Since P and Q have values 1 and 0, we always have $P = P^2$ and $Q = Q^2$. Therefore, the given statement can be rewritten as $PQ = P^2 + Q^2 - PQ$, so that $(P - Q)^2 = 0$, and $P = Q$.

4 Mathematical Induction

Definition 4.1 *For $n \in \mathbb{N}$ and $k = 0, 1, \dots, n$, the* number of combinations *of k objects from a set of n objects is*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad \text{where} \quad 0! = 1! = 1, \quad \text{and} \quad n! = 1 \cdot 2 \cdot \ldots \cdot n \quad \text{for} \quad n \geq 2.$$

The following formula follows either by easy direct calculation, or by considering separately subsets of $\{1, 2, \dots, n, n+1)$ which (i) contain $n+1$, and (ii) do not contain $n+1$.

Lemma 4.2 *For $n \in \mathbb{N}$ and $k = 1, \dots, n$, we have*

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

In turn, using induction and this formula, one can get the following important *Newton's binomial formula.* For this reason $\binom{n}{k}$ are also called the *binomial coefficients.*

Theorem 4.3 (Newton's binomial formula) *For any $n \in \mathbb{N}$ and (real or complex) a and b, we have*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = a^n + \binom{n}{1} a^{n-1} b + \binom{n}{2} a^{n-2} b^2 + \cdots + \binom{n}{n-1} a b^{n-1} + b^n.$$

In particular, taking $a = b = 1$, we get

$$2^n = \sum_{k=0}^{n} \binom{n}{k}.$$

This equality has a simple interpretation: the number of all subsets of a set of n objects (including the empty set) is 2^n. Therefore, the notation 2^S ="all subsets of a given set S" makes sense even for infinite sets S.

Definition 4.4 *The sets S_1 and S_2 are* equivalent *if there is a one-to-one function $f : S_1 \to S_2$, i.e. (i) from $s', s'' \in S_1$ and $s' \neq s''$ it follows $f(s') \neq f(s'')$; (ii) for each $s_2 \in S_2$, there is $s_1 \in S_1$ such that $f(s_1) = s_2$.*

Theorem 4.5 (Cantor) *For an arbitrary set S, the sets S and 2^S are* **not** *equivalent.*

Proof. Suppose otherwise. Then there is a one-to-one function $f : S \to 2^S$. Denote

$$S_0 = \{s \in S : \ s \notin f(s)\} \in 2^S.$$

By our assumption, $S_0 = f(s_0)$ for some $s_0 \in S$. Consider two possible cases (i) $s_0 \in S_0$, and (ii) $s_0 \notin S_0$. In the case (i), $s_0 \in f(s_0)$, hence by definition of S_0, we have $s_0 \notin S_0 = f(s_0)$, i.e. case (ii). But in that case, we must have $s_0 \in S_0$. This contradiction proves the theorem. ■

A set S is *countable* is it is equivalent to the set of natural numbers $\mathbb{N}$, i.e. if S can be arranged as a sequence: $S = \{s_1, s_2, \dots, s_n, \dots\}$.

Corollary 4.6 *The set $[0, 1] = \{x \in \mathbb{R} : \ 0 \le x \le 1\}$ is not countable.*

Proof. Each $x \in [0, 1)$ can be represented in the *binary form*

$$x = \sum_{k=1}^{\infty} \frac{\varepsilon_k}{2^k}, \qquad \text{where} \qquad \varepsilon_k = 0 \text{ or } 1.$$

Therefore, each $x \in [0, 1]$ is associated with a subset $S_x = \{k \in \mathbb{N} : \ \varepsilon_k = 1\} \in 2^{\mathbb{N}}$ ($x = 0$ corresponds to the empty set $S_0 = \emptyset$). For each set $S \in 2^{\mathbb{N}}$, there is a corresponding $x \in [0, 1]$, but this is not a one-to-one correspondence. Indeed, for rational $x = m \cdot 2^{-n}$ with $m, n \in \mathbb{N}$, there are two representations, with $\varepsilon_k = 0$ for all large k, and $\varepsilon_k = 1$ for all large k (similarly to decimal representation $x = 0.5000\ldots = 0.4999\ldots$). If we assume that $[0, 1]$ is countable, since in the above representation each $x \in [0, 1]$ is counted not more that twice, then $2^{\mathbb{N}}$ must be countable. However, this contradicts to the previous Cantor's Theorem. ■

Example 4.7 *As another application of the binomial formula, we derive an explicit formula for $S_k = S_k(n) = 1^k + 2^k + \cdots + n^k$. We will do it for $k = 2$, based on the previous values*

$$S_0 = n, \quad S_1 = 1 + 2 + \cdots + n = \frac{n(n+1)}{2}.$$

Following this pattern, one can further evaluate S_3 based on the values for S_0, S_1, S_2, etc. By the binomial formula with $n = 3$, we have

$$(j+1)^3 - j^3 = 3j^2 + 3j + 1.$$

Using summation with respect to $j = 0, 1, \dots, n$, we get

$$\begin{aligned} (n+1)^3 &= 3S_2 + 3S_1 + n + 1, \\ 3(S_2 + S_1) &= (n+1)\left[(n+1)^2 - 1\right] = n(n+1)(n+2), \end{aligned}$$

and finally,

$$S_2 = \frac{n(n+1)(n+2)}{3} - \frac{n(n+1)}{2} = \frac{n(n+1)(2n+1)}{6}.$$

Chapter 3. Sequences

We start with some examples.

Example 0.1 *Let $a \neq 0, b, c$ be constants, and let s_1 and s_2 be given. For $n = 1, 2, \dots$, let s_{n+2} be defined from the relation*

$$as_{n+2} + bs_{n+1} + cs_n = 0, \qquad n = 1, 2, \dots .$$

In order to find an explicit expression for s_n, we first solve the characteristic equation $ar^2 + br + c = 0$. *This equation has roots*

$$r_{1,2} = \frac{-b \pm \sqrt{D}}{2a}, \qquad \textit{where} \qquad D = b^2 - 4ac.$$

Consider separately the cases **(i)** $r_1 \neq r_2$, *and* **(ii)** $r_1 = r_2$.

(i) $r_1 \neq r_2$. In this case, for $r = r_1$ and $r = r_2$, the sequence $s_n = r^n$ satisfies

$$as_{n+2} + bs_{n+1} + cs_n = ar^{n+2} + br^{n+1} + cr^n = r^n(ar^2 + br + c) = 0.$$

By linearity, any linear combination $s_n = c_1 r_1^n + c_2 r_2^n$ satisfies this equality for arbitrary constants c_1 and c_2. These two constants can be determined from the two "initial" conditions for s_1 and s_2.

The most famous example of this kind is the Fibonacci numbers

$$F_1 = 1, \quad F_2 = 1, \quad F_{n+2} = F_n + F_{n+1} \quad \textit{for} \quad n = 1, 2, \dots .$$

The beginning of this sequence is $1, 1, 2, 3, 5, 8, 13, 21$. *The corresponding characteristic equation $r^2 - r - 1 = 0$ has roots $r_1 = \varphi = \frac{1}{2}(1 + \sqrt{5})$ – the* golden ratio, *and $r_2 = \varphi = \frac{1}{2}(1 - \sqrt{5}) = 1 - \varphi$. Hence $F_n = c_1\varphi^n + c_2(1-\varphi)^n$ for $n = 1, 2, \dots$. From the conditions $F_1 = F_2 = 1$, we find $c_1 = -c_2 = 1/\sqrt{5}$, so that*

$$F_n = \frac{\varphi^n + (1-\varphi)^n}{\sqrt{5}} \qquad \textit{for} \quad n = 1, 2, \dots, \qquad \textit{where} \quad \varphi = \frac{1+\sqrt{5}}{2}.$$

(ii) $r_1 = r_2 = -b/2a$. *As in the previous argument,* $s_n = r_1^n$ *satisfies the given relation. We claim that the second "independent" solution is* $s_n = nr_1^n$. *We need to verify the equality*

$$a(n+2)r_1^{n+2} + b(n+1)r_1^{n+1} + cnr_1^n = 0.$$

One can write the left side in the form

$$A + B, \qquad \text{where} \quad A = (ar_1^2 + br_1 + c) \cdot nr_1^n, \qquad B = (2ar_1 + b)r_1^{n+1}.$$

Here $A = 0$ *because* r_1 *is a root of the characteristic equation, and* $B = 0$ *because we consider the case* $r_1 = r_2 = -b/2a$.

Example 0.2 *The sequence* $s_n = (1+\sqrt{2})^n + (1-\sqrt{2})^n$ *has "initial values"* $s_1 = 2$ *and* $s_2 = 6$. *We can write* $s_n = r_1^n + r_2^n$ *with* $r_{1,2} = 1 \pm \sqrt{2}$ *satisfying the quadratic equation*

$$(r - r_1)(r - r_2) = r^2 - 2r - 1 = 0.$$

This means $a = 1$. $b = -2$, $c = -1$, *so that* $s_{n+2} = 2s_{n+1} + s_n$ *for* $n = 1, 2, \ldots$.

3 Sequences and Continuous Functions

Theorem 3.1 *Suppose* $\lim a_n = L$, *and a function* f *is continuous at* L. *Then* $\lim f(a_n) = f(L)$. *Similarly, if* $g(t) \to L$ *as* $t \to t_0$, *and a function* f *is continuous at* L, *then* $f(g(t)) \to f(L)$ *as* $t \to t_0$.

Definition 3.2 *An* elementary *function is a function built from a finite number of exponentials, logarithms, trigonometric and inverse trigonometric functions, and constants, using the four elementary operations* $(+, -, \times, /\ )$ *and composition (taking function of function).*

Theorem 3.3 *Any elementary function is continuous in its domain, i.e. at each point where it is defined.*

Example 3.4 *(i) We define the irrational number*

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n \cong 2.718281828\ldots .$$

This existence of this limit follows from HW 1, Problem 6.

(ii) We claim that also

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

Indeed, for $x > 0$, *we can always write* $n \le x < n+1$ *with an integer* $n \ge 0$. *Then*

$$a_n = \left(1 + \frac{1}{n+1}\right)^n \le \left(1 + \frac{1}{x}\right)^x \le b_n = \left(1 + \frac{1}{n}\right)^{n+1} .$$

Here

$$\begin{aligned}\lim a_n &= \lim\left[\left(1+\frac{1}{n+1}\right)^{n+1}\left(1+\frac{1}{n+1}\right)^{-1}\right]\\ &= \lim\left(1+\frac{1}{n+1}\right)^{n+1}\cdot\lim\left(1+\frac{1}{n+1}\right)^{-1} = e\cdot 1 = e,\\ \lim b_n &= \lim\left[\left(1+\frac{1}{n}\right)^{n}\left(1+\frac{1}{n}\right)\right] = \lim\left(1+\frac{1}{n}\right)^{n}\cdot\lim\left(1+\frac{1}{n}\right) = e\cdot 1 = e.\end{aligned}$$

By a variant of Pinching Theorem, the desired property follows.

(iii) Moreover,

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

The notation $x\to\infty$ *means* $|x|\to+\infty$, *but* x *can be of arbitrary sign. The case* $x>0$ *was considered in the previous part (ii). In the case* $x<0$, *we have* $y=-x-1\to+\infty$, *so that*

$$\left(1+\frac{1}{x}\right)^{x} = \left(1-\frac{1}{y+1}\right)^{-y-1} = \left(1+\frac{1}{y}\right)^{y+1} = \left(1+\frac{1}{y}\right)^{y}\cdot\left(1+\frac{1}{y}\right)\to e\cdot 1 = e.$$

(iv) Substituting $h=1/x$ *in the previous limit, we also get*

$$e = \lim_{h\to 0}(1+h)^{1/h}.$$

(v) Now we apply Theorem 3.1 to the function $f(x)=\ln x=\log_e x$ *which is continuous at the point* $x=e$. *This implies*

$$\lim_{h\to 0}\frac{\ln(1+h)}{h} = \lim_{h\to 0}\ln\left[(1+h)^{1/h}\right] = \ln\left[\lim_{h\to 0}(1+h)^{1/h}\right] = \ln e = 1.$$

Theorem 3.5 (Bolzano-Weierstrass) *Any bounded convergent sequence* $\{s_n\}$ *in* $\mathbb{R}^m$ *contains a convergent subsequence.*

Proof. (i) One-dimensional case $m=1$. We define the upper and lower limits as follows

$$M = \limsup_{n\to\infty} s_n = \lim_{n\to\infty}\sup_{k\ge n} s_k, \qquad m = \liminf_{n\to\infty} s_n = \lim_{n\to\infty}\inf_{k\ge n} s_k.$$

Note that we always have $M\ge m$, and $M=m$ if and only if $\exists\lim s_n$. In our case, $M=\lim M_n$, where $M_n=\sup_{k\ge n} s_k$ is a bounded, non-increasing sequence. Therefore, we can write $M_n\searrow M$ as $n\to\infty$. For any integer $j\ge 1$, we can choose integers $n_j\ge 1$, and then $k_j\ge n_j$ such that

$$M+\frac{1}{j}\ge M_{n_j}\ge M, \qquad M_{n_j}\ge s_{k_j}\ge M_{n_j}-\frac{1}{j}.$$

This implies $|s_{k_j}-M|\le 1/j$. We can arrange this construction in such a way that $1\le k_1<k_2<\cdots$. Then $\{s_{k_j}\}$ is a subsequence convergent to M as $j\to\infty$.

(ii) Multi-dimensional case $m = 2$. We have a bounded sequence of vectors

$$s_n = \left(s_n^{(1)}, s_n^{(2)}, \ldots, s_n^{(m)}\right) \in \mathbb{R}^m.$$

This means $|s_n^{(j)}| \le C = const$ for all j, n. Relying on the one dimensional case, choose a subsequence $\{s_{n_{1,j}}\}$ of $\{s_n\}$ such that the first component $\{s_{n_{1,j}}^{(1)}\}$ converges. Further, $\{s_{n_{1,j}}\}$ contains another subsequence $\{s_{n_{2,j}}\}$ for which the second component $\{s_{n_{2,j}}^{(2)}\}$ converges. Continuing in a similar manner, we get a subsequence $\{s_{n_{m,j}}\}$ for which all m components converge. This is equivalent to the convergence of $\{s_{n_{m,j}}\}$ in $\mathbb{R}^m$. ■

Definition 3.6 **(i)** *U is an* open *set in $\mathbb{R}^m$ if $\forall x_0 \in U$, $\exists r > 0$ such that the ball $B_r(x_0) = \{x \in \mathbb{R}^m : \ |x - x_0| < r\}$ is contained in U.*

(ii) *F is a* closed *set in $\mathbb{R}^m$ if its complement $F^c = \mathbb{R}^m \setminus F$ is open.*

(iii) *s_0 is a* limit point *of a set $S \subseteq \mathbb{R}^m$ if there is a sequence $\{s_n\} \subseteq S$ which converges to s_0.*

Theorem 3.7 *A set $F \subseteq \mathbb{R}^m$ is closed if and only if it contains all its limit points, i.e. from $\{s_n\} \subseteq F$ and $s_n \to s_0$ as $n \to \infty$ it follows $s_0 \in F$.*

Proof. Suppose F is closed, $\{s_n\} \subseteq F$ and $s_n \to s_0$ as $n \to \infty$. If $s_0 \notin F$, then $s_0 \in F^c = \mathbb{R}^m \setminus F$ – an open set, so that $B_r(s_0) \subseteq F^c$ for some $r > 0$. Since $s_n \to s_0$, there exists $n_0 \ge 1$ such that $|s_n - s_0| < r$ for all $n \ge n_0$. For such n, we have $s_n \in B_r(s_0) \subseteq \mathbb{R}^m \setminus F$, so that $s_n \notin F$ in contradiction to the assumption $\{s_n\} \subseteq F$. Therefore, $s_0 \in F$, i.e. F contains its limit points.

Now suppose that F contains its limit points. If $F^c = \mathbb{R}^m \setminus F$ is not open, then $\exists x_0 \in F^c$ such that $\forall r > 0$, the ball $B_r(x_0)$ is not contained in F^c. Choose a sequence $r_n \searrow 0$. Since $B_{r_n}(x_0)$ is not contained in F^c, there is a point $x_n \in F \cap B_{r_n}(x_0)$. We have $|x_n - x_0| < r_n \to 0$, hence $x_0 = \lim x_n$ – a limit point of F. By our assumption, we must have $x_0 \in F$, in contradiction to the choice $x_0 \in F^c$. This contradiction shows that F^c must be open, i.e. F is closed. ■

Theorem 3.8 (Weierstrass) *If f is continuous on a bounded closed set (compact) F in $\mathbb{R}^m$, then it is bounded and attains its maximum and minimum values at some points.*

Proof. Suppose f is unbounded. Then one can choose a sequence $\{s_n\} \subseteq F$ such that $|f(s_n)| \ge n$ for all $n \in \mathbb{N}$. By Bolzano-Weierstrass theorem, there is a subsequence $\{s_{n_j}\} \subseteq F$ which is convergent to a point s_0. Since F is a closed set, we must have $s_0 \in F$. Moreover, by continuity of f at s_0, we also have $f(s_0) = \lim\limits_{j\to\infty} f(s_{n_j})$, which contradicts to $|f(s_{n_j})| \ge n_j$. Therefore, f is bounded on F.

Now we know that $M = \sup\limits_F f < \infty$. Once again, choose a sequence $\{s_n\} \subseteq F$ such that $f(s_n) \to M$ as $n \to \infty$, and then a subsequence $s_{n_j} \to s_0 \in F$ as $j \to \infty$. By continuity,

$$f(s_0) = \lim_{j\to\infty} f(s_{n_j}) = M,$$

i.e. the maximum of f is attained at the point $s_0 \in F$. Quite similarly (or replacing f by $-f$) one can show that the minimum of f is attained at some point in F. ■

4 L'Hôpital's Rule

Theorem 4.1 (Rolle's Theorem) *Let f be a continuous function on $[a, b]$, which is differentiable on (a, b), and $f(a) = f(b)$. Then there is a point $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. If $f = const$, then $f' \equiv 0$, and we can take an arbitrary point $c \in (a, b)$ with $f'(c) = 0$. If f is not constant, then it attains either its maximum or munimum at an interior point $c \in (a, b)$, and $f'(c) = 0$. ■

Theorem 4.2 (Cauchy's Mean Value Theorem) *Let f and g be continuous functions on $[a, b]$, which are differentiable on (a, b), and $g'(x) \neq 0$ for all x in (a, b). Then there is a point $c \in (a, b)$ such that*

$$K := \frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Proof. The function

$$F(x) = [f(x) - f(a)] - K \cdot [g(x) - g(a)]$$

satisfies all the assumptions of Rolle's theorem, with $F(a) = F(b) = 0$. Therefore, $\exists c \in (a, b)$ such that

$$F'(c) = f'(c) - K \cdot g'(c) = 0,$$

and the desired equality follows. ■

Theorem 4.3 (L'Hôpital's Rule) *Let $f(x)$ and $g(x)$ be continuous functions which are differentiable and $g'(x) \neq 0$ for $0 < |x - a| < h$. Suppose that either* **(i)** *$f(x),\ g(x) \to 0$ as $x \to a$, or* **(ii)** *$f(x),\ g(x) \to \infty$ as $x \to a$. (Indefinite forms $0/0$ or ∞/∞). Then*

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)},$$

if the limit in the right side exists.

Proof. **(i)** $0/0$. Define $f(a) = g(a) = 0$. Then f and g are continuous on $[a - h, a + h]$. By the previous theorem, for each $x \in [a - h, a + h],\ x \neq a$, we can write

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(c)}{g'(c)},$$

where c (depending on x) lies between a and x. Obviously, $c \to a$ as $x \to a$, and the desired property follows.

(ii) ∞/∞. Suppose that the limit in the right side is L. Then by definition of the limit, $\forall \varepsilon > 0$, $\exists \delta_1 \in (0, h]$ such that

$$\left| \frac{f'(x)}{g'(x)} - L \right| < \frac{\varepsilon}{2} \qquad \text{for all } x \ \text{ satisfying} \quad 0 < |x - a| \le \delta_1.$$

For certainty, consider the case $x > a$. Set $b = a + \delta_1$. As in the previous argument

$$A(x) = \frac{f(x) - f(b)}{g(x) - g(b)} = \frac{f'(c)}{g'(c)}, \qquad \text{where } a < x < c < b,$$

so that $|A(x) - L| \leq \varepsilon/2$. Further,

$$\frac{f(x)}{g(x)} = A(x) \cdot B(x), \qquad \text{where} \quad B(x) = \frac{1 - g(b)/g(x)}{1 - f(b)/f(x)}.$$

Since $f(x),\ g(x) \to \infty$ as $x \to a$, we have $B(x) \to 1$ as $x \to a$. Therefore, $\exists \delta \in (0, \delta_1]$ such that

$$\left| \frac{f(x)}{g(x)} - L \right| < \varepsilon \qquad \text{for all } x \in (a, a + \delta).$$

A similar estimate also holds for $x \in (a - \delta, a)$. This means that $f(x)/g(x) \to L$ as $x \to a$. ■

In the case $f(a) = f(b) = 0$, Rolle's theorem says that if a function f has at least two zeros, then its derivative f' has at least one zero. We will extend this result to functions with multiple zeros, some of which may coincide. Note that $x = a$ is a root (or zero) of a polynomial $P(x)$ of *multiplicity* $m \in \mathbb{N}$, if

$$P(x) = (x - a)^m Q(x), \qquad \text{where} \quad Q(x) \quad \text{is a polynomial with} \quad Q(a) \neq 0.$$

For general functions f this definition is equivalent to the following one.

Definition 4.4 *A function* $f(x)$ *has* zero of multiplicity $m \in \mathbb{N}$ *at the point* $x = a$*, if it has derivatives* $f^{(k)}(a)$ *for all* $k \leq m$*, and*

$$f(a) = f'(a) = f''(a) = \ldots = f^{(m-1)}(a) = 0 \neq f^{(m)}(a).$$

The following theorem can be considered as a generalizations of Rolle's theorem (which corresponds to the case $x_1 = a,\ x_2 = b,\ m_1 = m_2 = 1$).

Theorem 4.5 *Let* $x_1, x_2, \ldots, x_k$ *be distinct points in* $[a, b]$*, and let a function* $f(x)$ *have zeros of multiplicity* $m_j \geq 1$ *at* x_j *for each* $j = 1,\ 2,\ \ldots,\ k$*. Suppose that* $f(x)$ *is* $m - 1$ *times differentiable on* $[a, b]$*, where* $m = m_1 + m_2 + \cdots + m_k$*. Then there is a point* $c \in [a, b]$ *such that* $f^{(m-1)}(c) = 0$*.*

Proof. In the case $k = 1$, we have $c = x_1$ is a zero of multiplicity $m = m_1$, which implies $f^{(m-1)}(c) = 0$. Now suppose $k \geq 2$. We can always assume $a \leq x_1 < x_2 < \cdots < x_k \leq b$. Then by Rolle's theorem, in each of $k - 1$ intervals

$$(x_1, x_2), \quad (x_2, x_3), \quad \ldots, \quad (x_{k-1}, x_k),$$

the function $f'(x)$ has at least one zero, so there are at least $k - 1$ zeroes of $f'(x)$ in the union of these intervals (which do not include the points $x_1,\ x_2, \cdots, x_k$).

In addition, if $m_j \geq 2$ for some j, then $f'(x)$ has zero of multiplicity $m_j - 1$ at the point x_j. If we count each zero of $f'(x)$ as many times as its multiplicity, then $f'(x)$ has al least

$$(k - 1) + (m_1 - 1) + (m_2 - 1) + \cdots + (m_k - 1) = m - 1$$

zeros in $[a, b]$. Applying this argument to the function $f'(x)$, we conclude that $f''(x)$ has at least $m - 2$ zeros, etc., $f^{(m-2)}(x)$ has at least two zeros, and finally, $f^{(m-1)}(x)$ has at least one zero $x = c$. ■

Next, we consider approximation of a smooth enough function $f(x)$ by polynomials $P_n(x) = p_0 + p_1 x + p_2 x^2 + \cdots + p_n x^n$ of degree $\le n$. The most typical is approximation of $f(x)$ near $x = a$ by *Taylor polynomials*

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

which satisfy $P_n^{(k)}(a) = f^{(k)}(a)$ for all $k = 0, 1, 2, \ldots, n$. Another kind of approximation is given by *Lagrange interpolating polynomials*

$$L_n(x) = \sum_{j=1}^{n+1} f(x_j) \prod_{k \ne j} \frac{x - x_k}{x_j - x_k},$$

which satisfy $L_n(x_j) = f(x_j)$ for all $j = 1, 2, \ldots, n+1$, where $x_1, x_2, \ldots, x_n, x_{n+1}$ are fixed distinct numbers. It is possible to find a polynomial of degree $\le n$ in a more general situation, for which Taylor and Lagrange polynomials appear at the two extreme cases.

Lemma 4.6 *Let $k \in \mathbb{N}$, and for $j = 1, 2, \ldots, k$, let distinct numbers x_j and non-negative integers m_j be given. Then for any smooth enough function $f(x)$ there is a unique polynomial $P_n(x)$ of degree $\le n = m_1 + m_2 + \cdots + m_k + k - 1$, such that*

$$P_n^{(i)}(x_j) = f^{(i)}(x_j) \qquad \text{for all} \quad j = 1, 2, \ldots, k, \quad \text{and} \quad i = 0, 1, 2, \ldots, m_j.$$

Proof. Note that the number of conditions here is $\sum (m_j + 1) = n + 1$, which is same as the number of coefficients in the polynomial $P_n(x) = p_0 + p_1 x + p_2 x^2 + \cdots + p_n x^n$. Therefore, we have a system of $n+1$ linear equations with $n+1$ unknowns $p_0, p_1, \ldots, p_n$. It is known from Linear Algebra that such a system has a unique solution for arbitrary right sides if and only if the corresponding homogeneous system (with zero right sides) cannot have non-zero solution. In our case the homogeneous system is

$$P_n^{(i)}(x_j) = 0 \qquad \text{for all} \quad j = 1, 2, \ldots, k, \quad \text{and} \quad i = 0, 1, 2, \ldots, m_j.$$

This exactly means that each x_j is a root of P_n of multiplicity $m_j + 1$, so that the total number of roots, taking into account their multiplicities, is $\sum (m_j + 1) = n + 1$ – bigger than the degree of P_n. This is possible only if $P_n \equiv 0$, i.e. $p_0 = p_1 = \cdots = p_n = 0$. Therefore, the homogeneous system has only zero solution, and our statement follows. ■

Theorem 4.7 *Let $P_n(x)$ and $f(x)$ be as in the previous lemma. Suppose that $x_j \in [a, b]$ for all j, and the function $f(x)$ has all derivatives of order $\le n + 1$ on $[a, b]$ (the derivatives are one-sided at the points a and b). Then for each point $x \in [a, b]$, there is $c \in [a, b]$ (depending on x) such that*

$$f(x) - P_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \prod_{j=1}^{k} (x - x_j)^{m_j + 1}.$$

Proof. We will prove this representation for an arbitrary point $x = x_0 \in [a, b]$. If x_0 is one of the points $x_1, x_2, \ldots, x_k$, then both sides equal zero, and the equality holds with an arbitrary $c \in [a, b]$.

Now we consider a more interesting case, when x_0 is different from $x_1, x_2, \dots, x_k$. In this case, choose a constant K such that

$$f(x_0) - P_n(x_0) = \frac{K}{(n+1)!} \prod_{j=1}^{k} (x_0 - x_j)^{m_j+1},$$

and consider the function

$$F(x) = f(x) - P_n(x) - \frac{K}{(n+1)!} \prod_{j=1}^{k} (x - x_j)^{m_j+1}.$$

This function has zero of multiplicity $m_j + 1$ at x_j for each $j = 1, 2, \dots, k$, and one more zero at x_0. Therefore, the total number of zeros of $F(x)$ is at least $\sum(m_j + 1) + 1 = n + 2$. By Theorem 4.5 applied to the function $F(x)$ with $m = n + 1$, there is $c \in [a, b]$ such that $F^{(n+1)}(c) = 0$. Since the polynomial P_n has degree strictly less than $n + 1$, we have $P_n^{(n+1)} \equiv 0$. Finally, in the last product, the highest power of x is $\prod x^{m_j+1} = x^{n+1}$, its derivative of order $n + 1$ is $(n + 1)!$, and the lower powers of x disappear after $n + 1$ differentiations. This yields

$$0 = F^{(n+1)}(c) = f^{(n+1)}(c) - 0 - \frac{K}{(n+1)!} \cdot (n+1)! = f^{(n+1)}(c) = K,$$

i.e. $K = f^{(n+1)}(c)$. Theorem is proved. ■

Corollary 4.8 (Taylor Formula) *Let $f(x)$ be a function which has all derivatives of order $\le n+1$ in an interval $(a - R, a + R)$. Then for each $x \in (a - R, a + R)$, there is a point c between a and x (i.e. $c \in [a, x]$ is $x \ge a$, and $c \in [x, a]$ if $x \le a$) such that*

$$f(x) = P_n(x) + R_n(x) \qquad \text{where} \quad P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k, \quad \text{and} \quad R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

If an elementary function is defined in a neighborhood of a point a, then there is a maximal value $R \le \infty$, such that $R_n(x) \to 0$ as $n \to \infty$ for all $x \in (a - R, a + R)$. Then

$$f(x) = \lim_{n\to\infty} P_n(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \quad \text{for} \quad |x - a| < R$$ – the *Taylor series* for $f(x)$ at $x = a$.

The number R here is called the *radius of convergence* of the series. The radius of convergence of each of the following series will be discussed later.

Example 4.9 *(i) $f(x) = (1 + x)^m$, $a = 0$, and m is a real number. Then*

$$f(x) = (1+x)^m = \sum_{k=0}^{\infty} \binom{m}{k} x^k \qquad \text{for} \quad |x| < 1,$$

where

$$\binom{m}{0} = 1, \ \binom{m}{k} = \frac{m(m-1)\cdot\ldots\cdot(m-k+1)}{k!} \qquad \text{for} \quad k \ge 1.$$

(ii) *If $m \in \mathbb{N}$, then $\binom{m}{k} = 0$ starting from $k = m+1$. Then the above series is reduced to the binomial formula*

$$(1+x)^m = \sum_{k=0}^{m} \binom{m}{k} x^k,$$

which holds true for all x. By setting $x = b/a$, we can extend it to a more general form

$$(a+b)^m = a^m (1+x)^m = \sum_{k=0}^{m} \binom{m}{k} a^{m-k} b^k.$$

(iii) *Note that $\binom{-1}{k} = (-1)^k$. Then the formula in (i) with $m = -1$ yields the* geometric series

$$(1+x)^{-1} = \frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k = 1 - x + x^2 - \cdots \qquad \text{for} \quad |x| < 1.$$

Replacing x by $-x$, we get a more familiar expression

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \cdots \qquad \text{for} \quad |x| < 1.$$

(iv) $f(x) = \ln(1+x)$. *On can get the Taylor expansion of this function by a formal integrating of the series in (iii):*

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \qquad \text{for} \quad |x| < 1.$$

(v) $f(x) = \frac{1}{2}[\ln(1+x) - \ln(1-x)]$. *Using the previous series with $-x$ in place of x, we get*

$$f(x) = \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right) = x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots \qquad \text{for} \quad |x| < 1.$$

One can also derive this Taylor expansion by a formal integrating of the series

$$\frac{1}{1-x^2} = 1 + x^2 + x^4 + x^6 + \cdots \qquad \text{for} \quad |x| < 1.$$

7 Cauchy Sequences

Definition 7.1 (Cauchy sequence) *The sequence $\{s_n\}$ is a* Cauchy sequence *of $\forall \varepsilon > 0$, $\exists n_0$ such that $|s_m - s_n| < \varepsilon$ for all $m, n \geq n_0$.*

Theorem 7.2 *The sequence $\{s_n\}$ converges (to a finite number L) if and only if it is a Cauchy sequence.*

Proof. First suppose $s_n \to L$. Then $\forall \varepsilon > 0,\ \exists n_0$ such that $|s_n - L| < \frac{\varepsilon}{2}$ for all $n \geq n_0$. Therefore,

$$|s_m - s_n| = |(s_m - L) - (s_n - L)| \leq |s_m - L| + |s_n - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \text{for all} \quad m, n \geq n_0,$$

i.e. $\{s_n\}$ is a Cauchy sequence.

Now suppose that $\{s_n\}$ is a Cauchy sequence. Using Definition 7.1 with $\varepsilon = \varepsilon_k = 2^{-k},\ k = 1, 2, 3, \ldots$, we can find a sequence $n_1 < n_2 < n_3 < \cdots$, such that

$$|s_m - s_n| < \varepsilon_k = 2^{-k} \qquad \text{for all} \quad m, n \geq n_k, \qquad \text{where} \quad k = 1, 2, \ldots .$$

In particular, $|s_{n_{k+1}} - s_{n_k}| < \varepsilon_k$ for all $k = 1, 2, \ldots$. Introduce the sequences a_k and b_k as follows

$$\begin{aligned} a_k &= s_{n_k} - 2\varepsilon_k < s_{n_{k+1}} - \varepsilon_k = s_{n_{k+1}} - 2\varepsilon_{k+1} = a_{k+1}, \\ b_k &= s_{n_k} + 2\varepsilon_k > s_{n_{k+1}} + \varepsilon_k = s_{n_{k+1}} + 2\varepsilon_{k+1} = b_{k+1}. \end{aligned}$$

These sequences are monotone and bounded:

$$a_1 < a_2 < \cdots < a_k < \cdots < b_k < \cdots < b_2 < b_1.$$

Therefore, they converge, and their limit $L = \lim a_k = \lim b_k = \lim s_{n_k}$, because the differences between these quantities are of order $\varepsilon_k \to 0$.

In the previous expression, we can take $m = n_j$ with $j \geq k$. Then we have

$$|s_{n_j} - s_n| < \varepsilon_k \qquad \text{for all} \quad j \geq k \quad \text{and} \quad n \geq n_k, \qquad \text{where} \quad k = 1, 2, \ldots .$$

Since we already know that $s_{n_j} \to L$, we can take the limit as $j \to \infty$. The result is

$$|L - s_n| \leq \varepsilon_k \qquad \text{for all} \quad n \geq n_k, \qquad \text{where} \quad k = 1, 2, \ldots .$$

It is easy to see that the last property implies $\lim s_n = L$. ∎

8 Stirling's Formula

In this section we prove an interesting formula which is due to J. Stirling (Methodus differentialis, 1730). In a simplified form, it says that

$$n! \sim \sqrt{2\pi n} \cdot \left(\frac{n}{e}\right)^n .$$

Here the notation $a_n \sim b_n$ means $a_n / b_n \to 1$ as $n \to \infty$, but is does not specify the error in the approximation of a_n by b_n. Following an argument by H.E. Robbins in Amer. Math. Monthly (1955), 26–29, we will get a two-sided estimate for $n!$. We will use the following representation for π, which is due to John Wallis (1616-1703).

Lemma 8.1 *We have*

$$\frac{\pi}{2} = \lim_{k \to \infty} \frac{[(2k)!!]^2}{2k \cdot [(2k-1)!!]^2}.$$

Here $(2k)!! = 2 \cdot 4 \cdot 6 \cdot \cdots \cdot (2k), \qquad (2k+1)!! = 1 \cdot 3 \cdot 5 \cdot \cdots \cdot (2k+1).$

Proof. Consider the sequence of integrals

$$I_n = \int_0^{\pi/2} \sin^n t \; dt \qquad \text{for} \quad n = 0, 1, 2, \dots .$$

Obviously, $I_0 = \pi/2,\; I_1 = 1$. For $n \geq 2$, we use integration by parts:

$$\begin{aligned} I_n &= -\int_0^{\pi/2} \sin^{n-1} t \; d(\cos t) = (n-1)\int_0^{\pi/2} \sin^{n-2}\cos^2 t \; dt \\ &= (n-1)\int_0^{\pi/2} \sin^{n-2} t(1-\sin^2 t) \; dt = (n-1)(I_{n-2} - I_n). \end{aligned}$$

By iteration we obtain

$$I_n = \frac{n-1}{n} I_{n-2} = \frac{(n-1)(n-3)}{n(n-2)} I_{n-4} = \cdots .$$

Depending on the cases $n = 2k$ and $n = 2k+1$, we get different expressions:

$$\begin{aligned} I_{2k} &= \frac{(2k-1)(2k-3)\cdot\dots\cdot 3\cdot 1}{(2k)(2k-2)\cdots 4\cdot 2}\cdot I_0 = \frac{(2k-1)!!}{(2k)!!}\cdot\frac{\pi}{2}, \\ I_{2k+1} &= \frac{(2k)(2k-2)\cdot\dots\cdot 2}{(2k+1)(2k-1)\cdot\dots\cdot 3}\cdot I_1 = \frac{(2k)!!}{(2k+1)!!}. \end{aligned}$$

Further, since $\sin^{n+2} t \leq \sin^{n+1} t \leq \sin^n t$ we also have $I_{n+2} \leq I_{n+1} \leq I_n$, and

$$1 - \frac{1}{n+2} = \frac{I_{n+2}}{I_n} \leq \frac{I_{n+1}}{I_n} \leq 1.$$

Therefore, $I_{n+1}/I_n \to 1$ as $n \to \infty$, and

$$\frac{\pi}{2} = \lim_{k\to\infty} \frac{I_{2k+1}\cdot\pi/2}{I_{2k}} = \lim_{k\to\infty} \frac{[(2k)!!]^2}{[(2k+1)!!]\cdot[(2k-1)!!]} = \lim_{k\to\infty} \frac{[(2k)!!]^2}{2k\cdot[(2k-1)!!]^2}.$$

Lemma is proved. ■

Theorem 8.2 *For $n = 1, 2, 3, \dots$, we have*

$$\sqrt{2\pi n}\cdot\left(\frac{n}{e}\right)^n \cdot e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi n}\cdot\left(\frac{n}{e}\right)^n \cdot e^{\frac{1}{12n}}.$$

Proof. Denote

$$a_n = n!, \qquad b_n = n^{n+1/2}e^{-n}, \qquad c_n = \frac{a_n}{b_n}, \qquad d_n = \ln c_n.$$

We have

$$\begin{aligned} d_n &= \ln c_n - \ln b_n = \ln 1 + \ln 2 + \cdots + \ln n - \left(n + \frac{1}{2}\right)\ln n + n, \\ d_n - d_{n+1} &= -\ln(n+1) - \left(n+\frac{1}{2}\right)\ln n + \left(n + \frac{2}{2}\right)\ln(n+1) - 1 \\ &= \left(n+\frac{1}{2}\right)\ln\left(1+\frac{1}{n}\right) - 1. \end{aligned}$$

Set $x = (2n+1)^{-1}$. Using Example 4.9 (v), we can write

$$d_n - d_{n+1} = \frac{1}{2x} \ln\left(\frac{1+x}{1-x}\right) - 1 = \frac{x^2}{3} + \frac{x^4}{5} + \frac{x^6}{7} + \cdots .$$

Therefore,

$$\begin{aligned} d_n - d_{n+1} &< \frac{x^2}{3} \cdot (1 + x^2 + x^4 + \cdots) = \frac{x^2}{3(1-x^2)} = \frac{1}{3[(2n+1)^2 - 1]} \\ &= \frac{1}{12n(n+1)} = \frac{1}{12n} - \frac{1}{12(n+1)}, \\ d_n - \frac{1}{12n} &< d_{n+1} - \frac{1}{12(n+1)}, \end{aligned}$$

i.e. the sequence $\{d_n - (12n)^{-1}\}$ is non-decreasing. On the other hand, it is easy to verify that

$$d_n - d_{n+1} > \frac{x^2}{3} = \frac{1}{3(2n+1)^2} \geq \frac{1}{12n+1} - \frac{1}{12(n+1)+1}$$

for $n \geq 1$, so that

$$d_n - \frac{1}{12n+1} > d_{n+1} - \frac{1}{12(n+1)+1},$$

i.e. the sequence $\{d_n - (12n+1)^{-1}\}$ is non-increasing. As in the proof of Theorem 7.2, these sequences together with $\{d_n\}$ converge to the same limit C_0. From

$$d_n - \frac{1}{12n} \nearrow C_0 \qquad \text{and} \qquad d_n - \frac{1}{12n+1} \searrow C_0$$

it follows

$$C_0 + \frac{1}{12n+1} < d_n < C_0 + \frac{1}{12n}.$$

Since $n! = a_n = b_n e^{d_n}$, we get the two-sided estimate

$$C\sqrt{n} \cdot \left(\frac{n}{e}\right)^n \cdot e^{\frac{1}{12n+1}} < n! < C\sqrt{n} \cdot \left(\frac{n}{e}\right)^n \cdot e^{\frac{1}{12n}},$$

where $C = e^{C_0}$. Now in order to complete the proof of this theorem, we only need to show that $C = \sqrt{2\pi}$, using the fact that $n! \sim Cn^{n+1/2}e^{-n}$. We will use the previous lemma for this purpose.

Note that $(2k)! = (2k)!! \cdot (2k-1)!!$ and $(2k)!! = 2^k k!$. Then

$$\begin{aligned} \frac{\pi}{2} &= \lim_{k\to\infty} \frac{[(2k)!!]^2}{2k \cdot [(2k-1)!!]^2} = \lim_{k\to\infty} \frac{[(2k)!!]^4}{2k \cdot [(2k)!]^2} = \lim_{k\to\infty} \frac{2^{4k}(k!)^4}{2k \cdot [(2k)!]^2} \\ &= \lim_{k\to\infty} \frac{2^{4k}(Ck^{k+1/2}e^{-k})^4}{2k \cdot [C(2k)^{2k+1/2}e^{-2k}]^2} = \frac{1}{2}C^2, \end{aligned}$$

and $C = \sqrt{2\pi}$. Theorem is proved. ■

Chapter 4. Infinite Series.

2 Infinite Series

The following example is non-trivial. It essentially uses Stirling's formula.

Example 2.1 *We will evaluate the sum of the series*

$$S = \sum_{k=1}^{\infty} \left(k \ln \frac{2k-1}{2k+1} - 1 \right).$$

Note that the partial sum

$$\begin{aligned} S_n &= \sum_{k=1}^{n} \left(k \ln \frac{2k-1}{2k+1} - 1 \right) \\ &= 1 \cdot (\ln 3 - \ln 1) + 2 \cdot (\ln 5 - \ln 3) + \cdots + n \cdot [\ln(2n+1) - \ln(2n-1)] - n \\ &= -\ln 1 - \ln 3 - \ln 5 - \cdots - \ln(2n-1) + n \cdot \ln(2n+1) - n \\ &= -\ln[(2n-1)!!] + n \cdot \ln(2n+1) - n = -\ln \left[\frac{(2n)!}{2^n n!} \right] + n \cdot \ln(2n+1) - n. \end{aligned}$$

Using Stirling's formula $n! \sim \sqrt{2\pi n} \cdot (n/e)^n$ *(here* $a_n \sim b_n$ *means* $a_n/b_n \to 1$ *as* $n \to \infty$*), we obtain*

$$\begin{aligned} e^{S_n} &= \frac{2^n n!}{(2n)!} \cdot (2n+1)^n e^{-n} \sim \frac{2^n \sqrt{2\pi n} \cdot (n/e)^n \cdot (2n+1)^n e^{-n}}{\sqrt{4\pi n} \cdot (2n/e)^{2n}} \\ &= \frac{1}{\sqrt{2}} \cdot \frac{(2n+1)^n}{(2n)^n} = \frac{1}{\sqrt{2}} \cdot \left[\left(1 + \frac{1}{2n} \right)^{2n} \right]^{1/2} \to \sqrt{\frac{e}{2}} \qquad as \quad n \to \infty. \end{aligned}$$

Finally,

$$S = \lim_{n \to \infty} S_n = \ln \left(\sqrt{\frac{e}{2}} \right) = \frac{1 - \ln 2}{2}.$$

5 The Root Test and the Ratio Test

In the following two theorems, the assumption $a_n > 0$ is not needed.

Theorem 5.1 (The Ratio Test) *Suppose* $a_n \neq 0$ *for all* n.

i) If $\limsup_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = L < 1$, *then* $\sum_{n=1}^{\infty} a_n$ *is convergent.*

ii) If $\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = L > 1$, *then* $\sum_{n=1}^{\infty} a_n$ *is divergent.*

Theorem 5.2 (The Root Text)

i) If $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = L < 1$, *then* $\sum_{n=1}^{\infty} a_n$ *is convergent.*

ii) If $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = L > 1$, *then* $\sum_{n=1}^{\infty} a_n$ *is divergent.*

6 Absolute and Conditional Convergence

The following two theorems look similar.

Theorem 6.1 (Dirichlet) *Let $\{a_n\}$ be a monotone sequence convergent to 0, and let $\{b_n\}$ be another sequence, for which the partial sums $B_n = b_1+b_2+\cdots+b_n$ are bounded: $|B_n| \le C = const < \infty$ for all n. Then the series $\sum_{n=1}^{\infty} a_n b_n$ is convergent, and its sum*

$$S = \sum_{n=1}^{\infty} a_n b_n = \sum_{n=1}^{\infty} c_n, \qquad \textit{where} \quad c_n = (a_n - a_{n+1})B_n.$$

Proof. Without loss of generality, we may assume $a_1 \ge a_2 \ge \cdots \ge a_n \ge \cdots \ge 0$ (otherwise we could change a_n by $-a_n$). Then the N-th partial sum

$$\begin{aligned} S_N &= \sum_{n=1}^{N} a_n b_n = a_1 B_1 + a_2(B_2 - B_1) + \cdots + a_N(B_N - B_{N-1}) \\ &= (a_1 - a_2)B_1 + (a_2 - a_3)B_3 + \cdots + (a_{N-1} - a_N)B_{N-1} + a_N B_N. \end{aligned}$$

Here $|a_N B_N| \le C a_N \to 0$ as $N \to \infty$. Therefore, it suffices to show that the series $\sum c_n$ is convergent. In fact, it is convergent absolutely, because

$$\sum_{n=1}^{\infty} |c_n| = \sum_{n=1}^{\infty} (a_n - a_{n+1})|B_n| \le C \cdot \sum_{n=1}^{\infty} (a_n - a_{n+1}) = C a_1 < \infty.$$

■

Theorem 6.2 (Abel) *Let $\{a_n\}$ be a bounded monotone sequence, and let $\sum_{n=1}^{\infty} b_n$ be a convergent series. Then the series $\sum_{n=1}^{\infty} a_n b_n$ is convergent.*

Proof. Since $\{a_n\}$ be a bounded monotone sequence, it has a limit $L = \lim_{n\to\infty} a_n$. Then $\alpha_n = a_n - L \to 0$ as $n \to \infty$. We can write

$$\sum_{n=1}^{\infty} a_n b_n = \sum_{n=1}^{\infty} (L + \alpha_n) b_n = L \sum_{n=1}^{\infty} b_n + \sum_{n=1}^{\infty} \alpha_n b_n.$$

In the right side, the first series obviously converges because $\sum_{n=1}^{\infty} b_n$ converges to $B = \lim_{n\to\infty} B_n$. This also implies that the sequence $\{B_n\}$ is bounded, and by the previous theorem, the series $\sum_{n=1}^{\infty} \alpha_n b_n$ is also convergent, so that $\sum_{n=1}^{\infty} a_n b_n$ appears as the sun of two convergent series. ■

Example 6.3 *Consider the series*

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n}.$$

We are going to use Dirichlet's theorem with $a_n = 1/n \searrow 0$, *and* $b_n = b_n(x) = \sin nx$. *Using the formula*

$$2\sin\alpha\sin\beta = \cos(\alpha-\beta) - \cos(\alpha+\beta) \qquad \text{with} \quad \alpha = kx,\ \beta = \frac{x}{2},$$

we obtain

$$\begin{aligned} B_n(x) &= \sum_{k=1}^{n} \sin kx = \frac{1}{2\sin\left(\frac{x}{2}\right)} \sum_{k=1}^{n} \left[\cos\left(k - \frac{1}{2}\right)x - \cos\left(k + \frac{1}{2}\right)x\right] \\ &= \frac{\cos\left(\frac{x}{2}\right) - \cos\left(n + \frac{1}{2}\right)x}{2\sin\left(\frac{x}{2}\right)}. \end{aligned}$$

If x *is a multiple of* π, *i.e.* x/π *is an integer, we have* $B_n(x) = 0$ *for all* n. *If* x/π *is not an integer, then* $\sin\left(\frac{x}{2}\right) \neq 0$, *and we have* $|B_n(x)| \leq \left|\sin\left(\frac{x}{2}\right)\right|^{-1}$ *for all* n. *By Theorem 6.1, the given series converges.*

7 Power Series

For an arbitrary power series $\sum_{n=0}^{\infty} a_n x^n$, its radius of convergence $r = L^{-1}$, where $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$. The series converges (absolutely) for $|x| < r$, diverges for $|x| > r$. The bevavior of the series for $|x| = r$ may be different for different series.

8 Differentiation and Integration of Power Series

Definition 8.1 *A sequence of functions* S_n *converges* uniformly *on a set* D *to a function* S *if* $\sup_D |S_n - S| \to 0$ *as* $n \to \infty$. *A series* $\sum_{k=0}^{\infty} f_k$ *converges* uniformly *on a set* D *to a function* S *if the sequence of partial sums* $S_n = \sum_{k=0}^{n} f_k$ *converges uniformly.*

Theorem 8.2 *If a sequence of continuous functions* S_n *converges* uniformly *on an open set* D *to a function* S, *then* S *is continuous on* D.

Theorem 8.3 (Weierstrass) *If $|f_n| \le a_n = const$ on a set D, and $\sum_{n=0}^{\infty} a_n < \infty$, then the series $\sum_{n=0}^{\infty} f_n$ converges uniformly on a D.*

The proof of the following theorem is sketched in the textbook, p. 138–140.

Theorem 8.4 (Abel) *Assume $\sum_{n=0}^{\infty} a_n$ converges. Then the series*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \to \sum_{n=0}^{\infty} a_n \qquad as \quad x \to 1^-.$$

Example 8.5 *We will apply the previous Abel's theorem to evaluation of the series in Example 6.3, which corresponds to $a_n = \frac{1}{n}\sin nx$. For certainty, we fix $x \in (0, \pi)$. Since a_n depends on x, we will use the variable t in place of x. We have*

$$f(t) = \sum_{n=1}^{\infty} \frac{t^n \sin nx}{n} \to S(x) = \sum_{n=1}^{\infty} \frac{\sin nx}{n} \qquad as \quad t \to 1^-.$$

This means that $f(t)$ is a continuous function on $[0,1]$. Since this is a power series with respect to t, we can differentiate it for $|t| < 1$:

$$\begin{aligned} f'(t) &= \sum_{n=1}^{\infty} t^{n-1} \sin nx = \frac{1}{t} \operatorname{Im}\left(\sum_{n=0}^{\infty} t^n e^{inx} \right) = \operatorname{Im}\left(\frac{1}{t(1-te^{ix})} \right) \\ &= \operatorname{Im}\left(\frac{1+te^{ix}}{t(1-2t\cos x + t^2)} \right) = \frac{\sin x}{1-2t\cos x + t^2}. \end{aligned}$$

Then

$$\begin{aligned} S(x) &= f(1) = \int_0^1 f'(t)dt = \int_0^1 \frac{\sin x}{\sin^2 x + (t-\cos x)^2} dt = \arctan\left(\frac{t-\cos x}{\sin x} \right) \Bigg|_{t=0}^{t=1} \\ &= \arctan\left(\frac{1-\cos x}{\sin x} \right) - \arctan\left(\frac{-\cos x}{\sin x} \right) = \frac{x}{2} + \left(\frac{\pi}{2} - x \right) = \frac{\pi - x}{2}. \end{aligned}$$

Since $S(x)$ is an odd and 2π-periodic function, the above formula $S(x) = (\pi - x)/2$ is valid on the interval $(0, 2\pi)$. Aster being extended as an 2π-periodic function, we get a function which has value 0 and is discontinuous at the points $x = 0, \pm 1, \pm 2, \cdots$.

Example 8.6 *Consider the function*

$$f(x) = \sum_{k=0}^{\infty} f_k(x), \qquad where \quad f_k(x) = 2^{-k} \sin\left(4^k \pi x\right).$$

By the Weierstrass theorem (Theorem 8.3), this function is continuous. However, again due to Weierstrass, it is **nowhere differentiable**. *Indeed, suppose there is $f'(x_0)$ at some point x_0. Then there is a constant $K > 0$ such that*

$$|f(x) - f(x_0)| \le K \cdot |x - x_0| \qquad for\ all \quad x.$$

Fix a natural number n*, and select two points*

$$x_{1,2} = 4^{-n}\left(j \pm \frac{1}{2}\right), \qquad \textit{where} \quad j = 4j_0 + 2.$$

We can choose an integer j_0 *such that* $|x_{1,2} - x_0| < 5 \cdot 4^{-n}$*. Note that* $|x_1 - x_2| = 4^{-n}$*. Therefore,*

$$|f(x_1) - f(x_2)| \leq |f(x_1) - f(x_0)| + |f(x_2) - f(x_0)| \leq 10 \cdot 4^{-n}.$$

Further,

$$|f_n(x_1) - f_n(x_2)| = 2^{-n}\left|\sin\left(j + \frac{1}{2}\right)\pi - \sin\left(j - \frac{1}{2}\right)\pi\right| = 2^{1-n}.$$

Moreover, $f_k(x_1) = f_k(x_2) = 0$ *for* $k \geq n+1$*, and also*

$$\begin{aligned} |f_{n-1}(x_1) - f_{n-1}(x_2)| &= 2^{1-n}\left|\sin\left(\frac{1}{4}\left(j + \frac{1}{2}\right)\pi\right) - \sin\left(\frac{1}{4}\left(j - \frac{1}{2}\right)\pi\right)\right| \\ &= 2^{1-n}\left|\sin\left(j_0 + \frac{1}{2} + \frac{1}{8}\right)\pi - \sin\left(j_0 + \frac{1}{2} - \frac{1}{8}\right)\pi\right| = 0, \end{aligned}$$

because x_1 *and* x_2 *are symmetric with respect to the point* $\left(j_0 + \frac{1}{2}\right)\pi$ *at which* $f_{n-1} = 1$ *or* -1*. By the triangle inequality,*

$$|f(x_1) - f(x_2)| \geq |f_n(x_1) - f_n(x_2)| - \sum_{k=0}^{n-2} |f_k(x_1) - f_k(x_2)|.$$

The last expression can be evaluated by the mean value theorem:

$$|f_k(x_1) - f_k(x_2)| \leq \max|f_k'| \cdot |x_1 - x_2| = 2^k\pi \cdot 4^{-n}.$$

Then

$$\begin{aligned} \sum_{k=0}^{n-2} |f_k(x_1) - f_k(x_2)| &\leq 4^{-n}\pi \sum_{k=0}^{n-2} 2^k < 2^{-n-1}\pi, \\ |f(x_1) - f(x_2)| &\geq 2^{1-n} - 2^{-n-1}\pi = (4-\pi)2^{-n-1}. \end{aligned}$$

For large n*, this estimate contradicts to the above estimate* $|f(x_1) - f(x_2) \leq 10 \cdot 4^{-n}$*. This contradiction shows that* $f'(x_0)$ *cannot exist.*

Remark 8.7 *One can make the above argument a bit shorter with* $f_k(x) = 2^{-k}\sin(6^k\pi x)$*. Then the case* $k = n-1$ *is treated together with* $k \leq n-2$*.*

9 Taylor Polynomials

Example 9.1 *One can also use power series to derive (and solve) certain differential equations. Consider the* Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2} \left(e^{-x^2}\right)^{(n)} \qquad \text{for} \quad n = 0, 1, 2, \dots .$$

It is easy to see that each H_n is a polynomial of degree n. For $k < n$, integrating by parts, we obtain

$$\int_{-\infty}^{\infty} e^{-x^2} H_k H_n dx = \int_{-\infty}^{\infty} H_k (-1)^n \left(e^{-x^2}\right)^{(n)} dx = \int_{-\infty}^{\infty} H_k^{(n)} e^{-x^2} dx = 0,$$

which means that the polynomials $\{H_n\}$ are orthogonal on $\mathbb{R}^1$ with weight e^{-x^2}.

Further, introduce the generating function for $\{H_n\}$:

$$F(t,x) = \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x) = e^{x^2} \sum_{n=0}^{\infty} \frac{\left(e^{-x^2}\right)^{(n)}}{n!} (-t)^n.$$

One can show that the Taylor expansion

$$f(x+h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} h^n \qquad \text{is valid for} \quad f(x) = e^{-x^2}, \; h = -t.$$

Hence

$$F(t,x) = e^{x^2} \cdot e^{-(x-t)^2} = e^{2tx - t^2}.$$

We have the following equalities between the partial derivatives of F :

$$F''_{xx} - 2xF'_x = 4t(t-x)F = -2tF'_t,$$

which implies

$$\sum_{n=0}^{\infty} \frac{t^n}{n!} \left(H''_n - 2xH'_n + 2nH_n\right) = 0.$$

Comparing the coefficients of t^n, we see that the function $y = H_n$ satisfies the equation

$$y'' - 2xy' + 2ny = 0,$$

or equivalently,

$$e^{x^2} \left(e^{-x^2} y'\right)' + 2ny = 0.$$

Example 9.2 *The following formula is an extension of Newton's binomial formula for $(1+x)^n$ to the case or an arbitrary real a in place of n.*

$$f(x) = (1+x)^a = \sum_{k=0}^{\infty} \binom{a}{k} x^k \qquad \text{for} \quad |x| < 1, \qquad \text{where} \quad \binom{a}{k} = \frac{a(a-1)(a-2) \cdot \ldots \cdot (a-k+1)}{k!}.$$

In order to prove this expansion, note that by Taylor's formula on p. 141,

$$(1+x)^a = \sum_{k=0}^{n} \binom{a}{k} x^k + R_n(x), \qquad \textit{where} \quad R_n(x) = \frac{1}{n!}\int_0^x f^{(n+1)}(t)\ (x-t)^n dt.$$

We need to show that $R_n(x) \to 0$ as $n \to \infty$, if $|x| < 1$. Substitution $t = \theta x$, and then using the mean value theorem, we represent R_n in the Cauchy form:

$$R_n(x) = \frac{1}{n!}\int_0^1 f^{(n+1)}(\theta x)\ (1-\theta)^n dt \cdot x^{n+1} = \frac{f^{(n+1)}(\theta_1 x)\cdot(1-\theta_1)^n}{n!} \cdot x^{n+1}, \qquad \textit{where} \quad 0 < \theta_1 < 1.$$

In our case,

$$R_n(x) = c_n (1-\theta_1)^n (1+\theta_1 x)^{a-n-1} x^{n+1}, \qquad \textit{where} \quad c_n = \frac{\alpha(a-1)(a-2)\cdot\ldots\cdot(a-n)}{n!}.$$

Here

$$\left|\frac{c_{n+1}}{c_n}\right| = \left|\frac{a-n-1}{n+1}\right| \to 1 \qquad \textit{as} \quad n \to \infty.$$

Therefore, the series $\sum c_n x^{n+1}$ is absolutely convergent for $|x| < 1$. Furthermore, since $1 - \theta_1 \leq 1 + \theta_1 x$, the additional factor in $R_n(x)$ does not exceed $(1+\theta_1 x)^a$. This means

$$|R_n(x)| \leq |c_n x^{n+1}| \cdot (1+\theta_1 x)^a \to 0 \qquad \textit{as} \quad n \to \infty, \qquad \textit{provided} \quad |x| < 1.$$

This gives us the desired expansion of $(1+x)^a$.

Remark 9.3 *Alternatively, one can prove that equality*

$$y_1(x) := (1+x)^a = y_2(x) := \sum_{k=0}^{\infty} \binom{a}{k} x^k \qquad \textit{for} \quad |x| < 1,$$

by showing that both functions y_1 and y_2 satisfy the same differential equation with initial condition:

$$(1+x)y' = ay \qquad \textit{for} \quad |x| < 1, \qquad y(0) = 1,$$

and relying on the uniqueness of solution to this problem.