APPENDIX A. Rudiments of Analysis

Warning. This is not a substitute of a standard textbook of elementary calculus, but covers most topics every undergraduate analysis course must cover. This is only a summary or a check list of the reader's knowledge. Scan the titles of the numbered entries, and if she finds a somewhat unfamiliar concept, read the entry. Try to form vivid mental image of defined concepts. Try to be able to explain why the statements are plausible intuitively. If you feel a theorem to be obvious, you need not prove it. The following material heavily relies on K. Kodaira, *Introductory Calculus* I-IV (Iwanami 1986), and *Encyclopedic Dictionary of Mathematics* (Iwanami 1985, 3rd edition). J. D. DePree and C. W. Swartz, *Introduction to Real Analysis* (Wiley, 1988) may be recommended as an introductory textbook.

Table of Standard Symbols
I. Point sets and limits
II. Functions
III. Differentiation
IV. Integration
V. Infinite Series
VI. Functions of two variables
VII. Fourier series and Fourier transform
VIII. Ordinary differential equations
IX. Vector analysis

Table of Standard Symbols

$\forall$	all, any, arbitrary.
$\exists$	there exist(s)
$\Rightarrow$	$A \Rightarrow B$ means A implies B.
$\Longleftrightarrow$	if and only if (iff)
$\equiv$	$A \equiv B$ means "A is defined by B."
$\in$	$a \in A$ implies that a is an element of A.
$\boldsymbol{C}$	the set of all complex numbers.
$\boldsymbol{N}$	the set of all nonnegative integers
$\boldsymbol{Q}$	the set of all rational numbers
$\boldsymbol{R}$	the set of all real numbers.
$\boldsymbol{Z}$	the set of all the integers
C^r	the set of all the r-times continuously differentiable functions.
C^0	the set of all the continuous functions
C^∞	the set of all the infinite times differentiable functions.
C^ω	the set of all (real) analytic functions
$L_1(A, \rho)$	Lebesgue integrable functions on A with weight ρ.
$L_2(A, \rho)$	Square Lebesgue integrable functions on A with weight ρ.
inf	infimum
sup	supremum
supp	support
L(R)HS	Left (right) hand side

A1 Point Set and Limit

The properties of reals (=real numbers) such as their continuity are assumed to be known.

A1.1 Sequence. Let a_1, a_2, $\cdots$ be reals. $a_1, a_2, a_3, \cdots$ is called a *sequence* and is denoted as $\{a_n\}$. Each real in the sequence $\{a_n\}$ is called a *term*.

A1.2 Convergence, limit. A sequence is said to *converge* to α if for any positive ϵ, there is a positive integer $N(\epsilon)$ such that

$$n > N(\epsilon) \Rightarrow |a_n - \alpha| < \epsilon. \tag{A1.1}$$

α is called the *limit* of the sequence $\{a_n\}$, and is often written as $a_n \to \alpha$.

A1.3 Theorem [Cauchy]. A necessary and sufficient condition for a (real) sequence $\{a_n\}$ to converge is that for any positive number ϵ there is a positive integer $N(\epsilon)$ such that

$$n > N(\epsilon), \;\; m > N(\epsilon) \Rightarrow |a_n - a_m| < \epsilon. \tag{A1.2}$$

□
Such a sequence is called a *Cauchy sequence.* [In an infinite dimensional space, a Cauchy sequence may not converge.]

A1.4 Symbol 'O' and 'o'.
(1) $f = O[g]$ means that the quantity f is of order g in the appropriate limit in the context. That is $\lim f/g$ is not divergent. For example, $1 - \cos x = O[x^2]$ in the $x \to 0$ limit. That is, $\lim_{x \to 0}(1 - \cos x)/x^2 < +\infty$, which is, of course, correct.
(2) $f = o[g]$ means that the quantity f is 'much smaller' than g in the appropriate limit in the context. For example, $\sin(x^2) = o[x]$ in the $x \to 0$ limit.

A1.5 Limit and arithmetic operations commute. Let $a_n \to \alpha$ and $b_n \to \beta$. Then,
(i) If $a_n \geq b_n$ for infinitely many n, then $\alpha \geq \beta$.
(ii) $a_n \pm b_n \to \alpha \pm \beta$.
(iii) $a_n b_n \to \alpha\beta$.
(iv) If $a_n \neq 0$ and $\alpha \neq 0$, then $b_n/a_n \to \beta/\alpha$.

A1.6 Lower and upper bound, supremum and infimum. Let $S \subset \boldsymbol{R}$. If any element in S does not exceed a real μ (i.e., $s \leq \mu$ for any $s \in S$) [resp., is not exceeded by a real number μ (i.e., $s \geq \mu$ for any $s \in S$)], we say S is *bounded to the above* [resp., *bounded to the below*] and μ is called an *upper bound* [resp., *lower bound*] of S. The smallest upper bound [resp., the largest lower bound] of S is called the *supreme* [resp., *infimum*] of S, and is written as $\sup_{s \in S} s$ [resp., $\inf_{s \in S} s$]. If S is bounded to the above and to the below, S is said to be *bounded.*

A1.7 Monotone sequences. If $a_1 < a_2 < \cdots < a_n < \cdots$ [resp., $a_1 > a_2 > \cdots > a_n > \cdots$], $\{a_n\}$ is called a *monotone increasing* [resp., *monotone decreasing*] sequence. If $a_1 \geq a_2 \geq \cdots \geq a_n \geq \cdots$ [resp.,

$a_1 \le a_2 \le \cdots \le a_n \le \cdots$], $\{a_n\}$ is called a *monotone non-decreasing* [resp., *monotone non-increasing*] sequence.

A1.8 Theorem [Bounded monotone sequences converge].
A monotone non-decreasing [resp., non-increasing] sequence bounded to the above [resp., to the below] converges to its supremum [resp., its infimum]. □

A1.9 Divergence to ± infinity. If a monotone non-decreasing sequence [resp., non-increasing sequence] is not bounded to the above [resp., to the below], we say it diverges to positive infinity [resp., negative infinity] and write $\lim_{n\to\infty} a_n = +\infty$ [resp., $\lim_{n\to\infty} a_n = -\infty$].

A1.10 Limsup and liminf. Suppose $\{a_n\}$ is a bounded sequence. Let $\sup_n a_{n+m} = \alpha_m$ for $m = 1, 2, 3, \cdots$. Then $\{\alpha_n\}$ is a bounded monotone non-increasing sequence. Hence, Theorem **A1.8** tells us that $\lim_{n\to\infty} \alpha_n$ exists. This is called the *superior limit* of the sequence $\{a_n\}$, and is written as $\limsup_{n\to\infty} a_n$. Analogously, the limit $\lim_{m\to\infty} \inf_n a_{n+m}$ exists, which is called the *inferior limit* of the sequence $\{a_n\}$, and is written as $\liminf_{n\to\infty} a_n$.
(i) For any positive ϵ there are only finitely many a_n larger than $\limsup_{n\to\infty} a_n + \epsilon$, but there are infinitely many a_n larger than $\limsup_{n\to\infty} a_n - \epsilon$.
(ii) For any positive ϵ there are only finitely many a_n smaller than $\liminf_{n\to\infty} a_n - \epsilon$, but there are infinitely many a_n smaller than $\liminf_{n\to\infty} a_n + \epsilon$.
(iii) A necessary and sufficient condition for $\{a_n\}$ to converge is $\limsup a_n = \liminf a_n$.

A1.11 Infinite series. For a sequence $\{a_n\}$, $a_1 + a_2 + a_3 + \cdots + a_n + \cdots$ is called an *infinite series*, and is often written as $\sum_{n=1}^{\infty} a_n$. The convergence of the series is defined by the convergence of the sequence $\{s_n\}$ consisting of its *partial sums*: $s_n \equiv a_1 + \cdots + a_n$. $\lim_{n\to\infty} s_n$, if it converges, is called the *sum* of the infinite series $\sum_{n=1}^{\infty} a_n$. If $\{s_n\}$ does not converge, the series is said to be divergent.
If $\sum_{n=1}^{\infty} a_n$ converges, then a_n converges to zero.

A1.12 Absolute convergence. If $\sum_{n=1}^{\infty} |a_n|$ converges, $\sum_{n=1}^{\infty} a_n$ is said to be *absolutely convergent.*
(i) If $\{a_n\}$ converges absolutely, $\{a_n\}$ converges.
(ii) Suppose $\sum_{n=1}^{\infty} r_n$ is convergent and $r_n \ge 0$. If $|a_n| \le r_n$ for all n larger than some integer m, then $\sum_{n=1}^{\infty} a_n$ converges absolutely.

A1.13 Power series. A series of the form $\sum_{n=0}^{\infty} a_n(x-b)^n$ is called a *power series*, where b is a constant.

A1.14 Conditional convergence, alternating series. If a convergent series is not absolutely convergent, it is said to *converge conditionally.* If positive and negative terms appear alternatingly, the series is called an *alternating series.*
If $\{a_n\}$ $(a_n > 0)$ is a monotone decreasing sequence converging to zero, then the alternating series $a_1 - a_2 + a_3 - a_4 + \cdots$ converges (→[AV7]).

A1.15 Theorem [Nested sequence of intervals shrinking to a point share the point]. If a sequence of closed intervals $\{I_n\}$ such that $I_n = [a_n, b_n]$ satisfies (i) $I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots$ and (ii) $\lim_{n\to\infty}(b_n - a_n) = 0$, then there is a unique real c which is in all I_n.□
For this theorem it is crucial that I_n are closed intervals.

A1.16 Denumerability. An infnite set for which we can make a one-to-one correspondence with nonnegative integers $\boldsymbol{N}$ is called a *countable set* or *denumerable set.* An infinite set which is not countable is called an *uncountable set* or *nondenumerable set.*
The set of rational numbers $\boldsymbol{Q}$ is countable.

A1.17 Cantor's Theorem [Continuum is not denumerable]. A closed interval $I = [a, b]$ is nondenumerable. □

A1.18 n-space, distance, ϵ-neighborhood. The totality of the n-tuples $(x_1, x_2, \cdots, x_n)$ is a direct product set $\boldsymbol{R} \times \cdots \times \boldsymbol{R} \equiv \boldsymbol{R}^n$ and is called the *n-space.* The (Euclidean) distance between two points $(x_1, \cdots, x_n)$ and $(y_1, \cdots, y_n)$ is defined by $[(x_1-y_1)^2+\cdots+(x_n-y_n)^2]^{1/2}$. The (Euclidean) distance between point P and Q is denoted by $|PQ|$. The totality of the points which are within the distance ϵ of point P is called the *ϵ-neighborhood* (ϵ-nbh) of P (and is denoted by $U_\epsilon(P)$ in this Appendix).

A1.19 Inner point, boundary, accumulating point, closure, open kernel. Let S be a subset of $\boldsymbol{R}^n$.
Inner point: P is an *inner point* of S if there is $\epsilon > 0$ such that $U_\epsilon(P) \subset S$.
Boundary point: If for any $\epsilon > 0$ $U_\epsilon(P) \subset_{\neq} S$ and $U_\epsilon(P) \cap S \neq \emptyset$, P is called a *boundary point* of S.
Boundary: The totality of the boundary points of S is called the *boundary* of S and is denoted by ∂S.
Closure: $S \cup \partial S$ is called the *closure* of S and is denoted by $[S]$. If $T \subset S$, then $[T] \subset [S]$.
Open kernel: $S \setminus \partial S$ is called the *open kernel* of S and is denoted by

S°.
Dense: Let T be a subset of S. If $[T] \supset S$, T is said to be *dense* in S.
Accumulating point: If $U_\epsilon(P) \cap S$ contains infinitely many points of S for any positive ϵ, we say P is an *accumulating point* of S.
Isolated point: If a point in S is not an accumulating point of S, the point is called an *isolated point*.
(i) A necessary and sufficient condition for a point Q to be in $[S]$ is that for any positive ϵ $U_\epsilon(Q) \cap S \neq \emptyset$.
(ii) The totality of rational numbers $\boldsymbol{Q}$ has no inner point and $[\boldsymbol{Q}] = \boldsymbol{R}$.
(iii) All the inner points of S are accumulating points of S. An accumulating point of S is its inner point or its boundary point. If a boundary point of S is not in S, it is an accumulating point of S.
(iv) A necessary and sufficient condition for a point P to be an isolated point of S is that there is a positive ϵ such that $U_\epsilon(P) \cap S = \emptyset$.

A1.20 Open set, closed set. If S contains only its inner points, that is, if $S = S^\circ$, then S is called an *open set*. If all the boundary points are included in S, that is, if $S = [S]$, S is called a *closed set*. The empty set $\emptyset$ is simultaneously open and closed, so is $\boldsymbol{R}$.
(i) The intersection of finite or infinite closed sets is a closed set.
(ii) The union of finite or infinite open sets is an open set.
(iii) The intersection of finitely many open sets is an open set.
(iv) The union of finitely many closed sets is a closed set.

A1.21 Limit of point sequence. A sequence of points $\{P_n\}$ ($P_n \in \boldsymbol{R}^n$) is called a *point sequence*. If there is a point A such that $\lim_{n\to\infty} |P_n A| = 0$, we say the point sequence $\{P_n\}$ converges to A. and write $\lim_{n\to\infty} P_n = A$.

A1.22 Bounded set, diameter. If the distance between any point $P \in S$ and the origin O is bounded to the above ($\to$**A1.6**), then the set S is called a *bounded set*. When S is a bounded set we can define its diameter $\delta(S)$ as $\delta(S) \equiv \sup_{P,Q\in S} |PQ|$. There is a theorem analogous to **A1.15**:

A1.23 Theorem [Shrinking nested sequence of bounded closed sets]. If a sequence of nonempty bounded closed sets $\{S_n\}$ satisfies the following two conditions (i) and (ii), then there is a unique point P shared by all of the closed sets S_n: (i) $S_1 \supset S_2 \supset \cdots \supset S_n \supset \cdots$, (ii) $\lim_{n\to\infty} \delta(S_n) = 0$.

A1.24 Covering. Let $\mathcal{U}$ be a set of sets. The joint set of all the members of $\mathcal{U}$ is written as $\cup_{U\in\mathcal{U}} U$. If a set S satisfies $S \subset \cup_{U\in\mathcal{U}} U$, then $\mathcal{U}$ is called a *covering* of S. If all the elements of $\mathcal{U}$ is open, it is called an *open covering* of S. If a covering $\mathcal{U}$ contains only a finite number of elements, $\mathcal{U}$ is called a *finite covering*. If a subset $\mathcal{V}$ of $\mathcal{U}$ is also a covering of S, $\mathcal{V}$ is called a *subcovering* of $\mathcal{U}$.

A1.25 Compact set. If any open covering of S has a finite subcovering, S is called a *compact set.*

A1.26 Theorem [Compactness is equivalent to bounded closedness]. S is compact if and only if S is a bounded closed set.$\square$
The only-if part is called the *Heine-Borel covering theorem.* This is true only if the space is finite dimensional.

[27] Theorem [Bolzano and Weierstrass]. A bounded infinite set must have an accumulating point($\rightarrow$**A1.19**).$\square$
Theorem. A bounded point sequence has a converging subsequence.$\square$[477]

A2 Function

A2.1 Function, domain, range, independent and dependent variables. Let $D \subset \boldsymbol{R}$. A rule f corresponding a single real η to each $\xi \in D$ is called a *function* f.[478] $\eta = f(\xi)$ is called the *value* of f at ξ. D is called its *domain* and $f(D) \equiv \{f(\xi)|\xi \in D\}$ is called the *range* of f. Usually, f is described as $f(x)$, and x is called the *variable*, and $f(x)$ is called a function of x. When we write $y = f(x)$, x is called the *independent variable* and y the *dependent variable.*

A2.2 Limit of function. Let $f(x)$ be a function whose domain is D. We say $f(x)$ converges to α in the limit $x \rightarrow a$, if for any positive ϵ, there is a positive number $\delta(\epsilon)$ such that

$$|x - a| < \delta(\epsilon),\ x \in D \Rightarrow |f(x) - \alpha| < \epsilon.$$

and we write $\lim_{x\rightarrow a} f(x) = \alpha$. $\lim_{x\rightarrow a}$ and arithmetic operations are commutative as **A1.5**. We have a theorem analogous to **A1.3**:

[477] These theorems assume that we can always choose one point from each member of a family of infinitely many sets. From the constructive point of view, this is not always possible. That is, we may not be able to write a computer program to do so. In the usual mathematics, we postulate this possibility as an axiom called the Axiom of Choice.
[478] This is often called a *map* as well.

A2.3 Cauchy's criterion. Let f be a function whose domain is D. A necessary and sufficient condition for f to be convergent in the $x \to a$ limit is: For any positive ϵ there is a positive $\delta(\epsilon)$ such that for $x, y \in D$

$$|x - a| < \delta(\epsilon),\ |y - a| < \delta(\epsilon) \Rightarrow |f(x) - f(y)| < \epsilon.$$

□

A2.4 Graph of a function. The graph G_f of a function f is a set $G_f = \{(x, f(x)) | x \in D\}$.

A2.5 Continuity. A function f is *continuous* at a, if $\lim_{x \to a} f(x) = f(a)$.
If the definition of the limit is spelled out completely as in **A2.2**, we say: f is continuous at a, if for $x \in D$ and for any positive ϵ there is a positive $\delta(\epsilon)$ such that

$$|x - a| < \delta(\epsilon) \Rightarrow |f(x) - f(a)| < \epsilon.$$

Theorem. If the domain of a continuous function f is a closed interval, then its range is again a closed interval. □

A2.6 Left and right continuity. When taking the $x \to a$ limit, if x is always smaller (resp., larger) than a, we write this limiting procedure as $\lim_{x \to a-0}$ (resp., $\lim_{x \to a+0}$) and is called the *left limit* (resp., right limit). If $\lim_{x \to a-0} f(x) = f(a)$ (resp., $\lim_{x \to a-0} f(x) = f(a)$), we say f is *left (resp., right) continuous* at a.

A2.7 Theorem of middle value. Let a function f be continuous in a closed interval $[a, b]$, and $f(a) \neq f(b)$. There is a real c such that $a < c < b$ and $f(c) = \mu$ for any μ between $f(a)$ and $f(b)$. □
The image of a finite interval by a continuous map is again a finite interval.

A2.8 Uniform continuity. A function f is *uniformly continuous* in D if for any positive ϵ, there is a positive constant $\delta(\epsilon)$ such that

$$|x - y| < \delta(\epsilon), x \in D, y \in D \Rightarrow |f(x) - f(y)| < \epsilon.$$

Theorem. A continuous function defined on a closed interval is uniformly continuous on the interval. □

A2.9 Maximum and minimum. Let f be a function whose domain is D. If $f(D)$ is bounded, we say f is bounded. If there is a maximum (resp., minimum) value in $f(D)$, then it is called the *maximum* (resp.,

minimum) of f.
Theorem [Maximum value theorem]. A continuous function defined on a closed interval has a maximum and minimum values. □

A2.10 Composite function. Let f be a function whose domain is D, and g is a function whose domain is in the range of f, $f(D)$. Then $h(x) = g(f(x))$ is called the *composite function* of f and g, and is denoted by $g \circ f$.

A2.11 Monotone function. Let f be a function whose domain is D. If for any $x, y \in D$ $x < y$ implies $f(x) < f(y)$ (resp., $f(x) > f(y)$), f is called a *monotone increasing function* (resp., *monotone decreasing function*). If for any $x, y \in D$ $x < y$ implies $f(x) \leq f(y)$ (resp., $f(x) \geq f(y)$), f is called a *monotone non-decreasing function* (resp., *monotone non-increasing function*).

A2.12 Inverse function. Let f be a function whose domain is D. If there is only one x such that $f(x) = y$ for each $y \in f(D)$, the correspondence $y \to x$ defines a function. This function, denoted by f^{-1}, is called the *inverse function* of f.
The symbol f^{-1} is used generally to denote the *preimage* of a point. Thus $f^{-1}(x) = \{y | f(y) = x, y \in D\}$, where D is the domain of f. f^{-1} becomes the inverse function, if $f^{-1}(x)$ is a single point for all x in the range of f.
Theorem. If f is a monotone increasing (resp., decreasing) function defined on an interval, then f has the inverse function which is monotone increasing (resp., decreasing). □

A2.13 Even and odd functions. If a function f has a domain invariant under $x \to -x$, and
(i) $f(x) = f(-x)$, we say f is an *even function*,
(ii) $f(x) = -f(-x)$, we say f is an *odd function.*

A3 Differentiation

A3.1 Differentiability, derivative. Let f be a function defined on an interval I, and $a \in I$. If the following limit, denoted by $f'(a)$, exists, we say f is *differentiable* at a:

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

$f'(a)$ is called the *differential coefficient* of f at a. If f is differentiable for any $x \in I$, we say that f is differentiable in I, and $f'(x)$ becomes a function on I. f' is called the *derivative* of f. To obtain f' from f is said to *differentiate* f. Recognize that the existence of the limit implies that the limit does not depend on how the point a is reached.

A3.2 Theorem [Differentiability implies continuity]. If f is differentiable at a, then f is continuous there. If f is differentiable in an interval I, it is continuous in the interval. □
Warning. However, continuity does not guarantee differentiability. See **A3.12**.

A3.3 Increment, differential quotient]. Let f be as in **A3.1** and write $y = f(x)$, and $\Delta y \equiv f(x + \Delta x) - f(x)$. Δx and Δy are called *increments*. Then

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x},$$

so that the derivative is also called the *differential quotient* and is denoted by dy/dx. If f is differentiable, then we may write

$$\Delta y = \frac{dy}{dx}\Delta x + o[\Delta x],$$

For o see **A1.4**.

A3.4 Right or left differentiable. If the right limit ($\to$**A2.6**) $\lim_{x\to a+0}(f(x) - f(a))/(x-a)$ exists, then we say f is *right differentiable* at a, and the limit, called *right differential coefficient* at a, is denoted by $D^+f(a)$. Analogously the left differential coefficient $D^-f(a)$ can be defined.

A3.5 Differentiation and arithmetic operations commute. Let f, g be differentiable in some interval, and c_1, c_2 be constants. Then 'arithmetic operations do not destroy differentiability':
(i) $\frac{d}{dx}(c_1 f(x) + c_2 g(x)) = c_1 f'(x) + c_2 g'(x)$.
(ii) $\frac{d}{dx}(f(x)g(x)) = f'(x)g(x) + f(x)g'(x)$.
(iii) If g is not zero, then $\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$.

A3.6 Derivative of composite function. Let f be a differentiable function on an interval I, and g be a differentiable function on an interval J containing $f(I)$. Then, $g \circ f$ ($\to$**A2.10**) is differentiable and

$$\frac{d}{dx}g(f(x)) = g'(f(x))f'(x).$$

A3.7 Derivative of inverse function. Let f be a differentiable monotone function on an interval I. Then its inverse function (→**A2.12**) is differentiable and

$$\frac{d}{dx}f^{-1}(x) = 1/f'(f^{-1}(x)).$$

A3.8 Theorem [Mean-value theorem]. Let f be a continuous function on the closed interval $[a, b]$. If f is differentiable in (a, b), then there is $\xi \in (a, b)$ such that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

□

A special case of this theorem is:

A3.9 Theorem [Rolle's theorem]. Let f be continuous in $[a, b]$. If f is differentiable in (a, b) and $f(a) = f(b)$, then there is $\xi \in (a, b)$ such that $f'(\xi) = 0$. □

A3.10 Theorem [Generalization of mean-value theorem]. Let f and g be continuous functions on a closed interval $[a, b]$, and are differentiable on (a, b). If f' and g' do not simultaneously vanish in (a, b) and $g(a) \neq g(b)$, then there is $\xi \in (a, b)$ such that

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

□

A3.11 Theorem [Condition for monotonicity]. A necessary and sufficient condition for a differentiable function defined on an interval I is monotone increasing (→**A2.11**) is that $f'(x) \geq 0$ on I and $f'(x) > 0$ on a dense (→**A1.19**) subset of I. □

A3.12 Counterexamples.
(i) $f(x) = x \sin(1/x)$ is continuous at $x = 0$ but not differentiable there.
(ii) $f(x) = \sum_{n=1}^{\infty} 2^{-n} |\sin(\pi n! x)|$ is continuous on $\boldsymbol{R}$, but not differentiable on $\boldsymbol{Q}$.
(iii) $f(x) = \sum_{n=1}^{\infty} 2^{-n} \cos(k^n \pi x)$ (k is an odd integer larger than 13) is continuous on $\boldsymbol{R}$, but is nowhere differentiable.

A3.13 Higher order derivatives. Suppose f is a differentiable function on an interval I. If f' is again differentiable on I, then we can define the second derivative df'/dx. If the function f is sufficiently smooth, then we can define higher-order derivatives like the *n-th derivative*, which is denoted by $f^{(n)}(x)$, $d^n f/dx^n$, $D^n f(x)$ or $(d/dx)^n f(x)$. Arithmetic operations do not destroy higher order differentiability as **A3.5**. The composite function of n-times differentiable functions is n-times differentiable as [6].

A3.14 Leibniz' formula.

$$\begin{aligned}\frac{d^n}{dx^n}(f(x)g(x)) &= f^{(n)}(x)g(x) + \binom{n}{1} f^{(n-1)}(x)g'(x) + \cdots \\ &\quad + \binom{n}{k} f^{(n-k)}(x)g^{(k)}(x) + \cdots + f(x)g^{(n)}(x).\end{aligned}$$

A3.15 Taylor's formula, remainder. Let f be a n-times differentiable function on an interval I, and $a \in I$. Then for any $x \in I$ there is a point ξ between a and x such that

$$f(x) = f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n)}(\xi)}{n!}(x-a)^n. \tag{A3.1}$$

The last term is called the *remainder*, and is written as R_n. □
For $n = 1$ this is the mean-value theorem (→**A3.8**), and this theorem is regarded as an extension of the mean-value theorem.
The remainder can be written as follows: Let $\xi = a + \theta(x-a)$ $(0 < \theta < 1)$.
(i) Schlömilch's remainder: Choosing an integer q $(0 \le q \le n-1)$,

$$R_n = \frac{f^{(n)}(\xi)(1-\theta)^q}{(n-1)!(n-q)}(x-a)^n.$$

(ii) Cauchy's remainder. This is a special case of (i) with $q = n-1$:

$$R_n = \frac{f^{(n)}(\xi)}{(n-1)!}(1-\theta)^{n-1}(x-a)^n.$$

(iii) The remainder in (A3.1) is another special case of (i) with $q = 0$, and is called Lagrange's remainder.

A3.16 Taylor's series. If f is infinite times differentiable, and $\{R_n\}$ in **A3.15** converges to zero, then f can be expanded in a *Taylor series* about a:

$$f(x) = f(a) + \sum_{k=1}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

A3.17 Convex and concave function. Let f be a function whose domain is I. Let $x_1, x_2 \in I$ and λ and μ be positive reals satisfying $\lambda + \mu = 1$. If

$$f(\lambda x_1 + \mu x_2) \leq \lambda f(x_1) + \mu f(x_2), \tag{A3.2}$$

we say f is *convex* on I, and f is called a *convex function*. If there is no equality in (A3.2), then we say f is *strictly convex*, and f is called a *strictly convex function*. If $-f$ is (strictly) convex, we say f is (strictly) *concave*, and f is called a (strictly) concave function.

A3.18 Theorem [Convexity and second derivative]. Let f be a twice differentiable function on an interval I.
(i) A necessary and sufficient condition that f is convex on I is that $f''(x) \geq 0$ for all the inner points of I.
(ii) If $f''(x) > 0$ for all the inner points of I, then f is strictly convex. □
An analogous theorem for concave functions should be self-evident. Simply switch f to $-f$.
Remark. (i) assumes that f is twice differentiable. Convex functions must be continuous, but need not even be differentiable once.

A3.19 Local maximum, minimum. Let f be a continuous function on an interval I, and a be an inner point of I. $f(a)$ is a *local maximum* (resp., *local minimum*), if for some positive number ϵ $0 < |x - a| < \epsilon$ implies $f(x) < f(a)$ (resp., $f(x) > f(a)$). These are collectively called *local extrema*.
Theorem. If f is a differentiable function on an interval I, and has a local extremum at $a \in I^\circ$,[479] then $f'(a) = 0$.□
Theorem. Let f be a function which is n-times ($n \geq 2$) differentiable and $f'(a) = f''(a) = \cdots = f^{(n-1)}(a) = 0$ at some inner point of I.
(i) If n is odd, then $f(a)$ is not an extremum of f.
(ii) If n is even, and $f^{(n)}(a) > 0$, then $f(a)$ is a local minimum of $f(x)$.
(iii) If n is even, and $f^{(n)}(a) < 0$, then $f(a)$ is a local maximum of $f(x)$.□

A3.20 Stationary value. Suppose f is n-times differentiable in an interval I, and for some inner point a of I $f'(a) = f''(a) = \cdots = f^{(n-1)}(a) = 0$ but $f^{(n)}(a) \neq 0$. If $n \geq 3$ and odd, $f(a)$ is called a *stationary value* of f, and a is called a *stationary point* of f.
Theorem. For the f in this item,
(i) If $f^{(n)}(a) > 0$, then there is a positive number ϵ such that f is strictly convex in $[a, a + \epsilon]$ and strictly concave in $[a - \epsilon, a]$.

---

[479] For $^\circ$ see **A1.19**.

(ii) If $f^{(n)}(a) < 0$, then there is a positive number ϵ such that f is strictly concave in $[a, a+\epsilon]$ and strictly convex in $[a-\epsilon, a]$.□

A3.21 Class C^n. Let f be a function defined on an interval I. If f is n-times differentiable and $f^{(n)}$ is continuous on I, then f is called a *function of class C^n* (or a *C^n-function*). If f is infinite-times differentiable, it is called a *C^∞-function.*
Theorem. Let f and g be C^n-functions on an interval I.
(i) Arithmetic operations do not destroy C^n-functions.
(ii) $g \circ f$ is again a C^n-function.
(iii) If for all $x \in I$ $f'(x) \neq 0$, and f is monotone, then its inverse function f^{-1} is a monotone C^n-function.□
These statements hold for C^∞ functions as well.

A3.22 Class C^ω. Let f be a C^∞-function in an open interval I. If f can be Taylor-expanded (→**A3.16**) in the neighborhood of each $a \in I$, then f is said to be *real analytic* in I and is called a *real analytic function* or a C^ω-function.
Warning. A C^∞-function need not be a real analytic function. A typical example is

$$\begin{aligned}\psi(x) &= 0 \text{ for } x \le 0,\\ &= e^{-1/x} \text{ for } x > 0.\end{aligned}$$

Its derivatives at $x = 0$ all vanish, so that Taylor series formally constructed becomes identically zero, but this contradicts the fact that $\psi(x) > 0$ for positive x. Hence, this function is not real analytic.
This is an important function to be used to 'mollify' functions through convolution.

A3.23 Theorem [Existence of mollifier]. Let a and b be two arbitrary points ($a < b$) in $\boldsymbol{R}$. There is a C^∞-function $\rho(x)$ on $\boldsymbol{R}$ such that $\rho(x) = 0$ for $x \le a$, $\rho(x) = 1$ for $x \ge b$ and $0 \le \rho(x) \le 1$.□
Corollary. Let f and g be C^∞-functions on $\boldsymbol{R}$, and a and b are the same as in the theorem. There is a C^∞-function h such that $h(x) = f(x)$ for $x \le a$, $h(x) = g(x)$ for $x \ge b$ (and h interpolates f and g between a and b). □
Thus C^∞-functions can be deformed freely. In contradistinction, C^ω-functions cannot be deformed freely as shown in the following

A3.24 Theorem [Identity theorem]. Let f and g be C^ω-functions defined on an open set I. If f and g coincide in some neighborhood of a point $a \in I$, then f and g are identical on I.□

A3.25 Complex analysis. Real analytic functions are best understood as special complex-valued functions defined on the complex plain. "The shortest path between two truths in the real domain passes through the complex domain." (J. Hadamard). H. A. Priestley, *Introduction to Complex Analysis* (Oxford UP, 1990, revised edition) is a convenient introduction to the topic. See also my notes for Physics 413, which is much more complete than the book with a sizable chapter on conformal mapping and its application to boundary value problems.

A4 Integration

A4.1 Definite integral (Riemann integral). Let f be a continuous function defined on a closed interval $I = [a, b]$. Let $a = x_0 < x_1 < x_2 < \cdots < x_k < \cdots < x_{m-1} < x_m = b$, and partition $[a, b]$ into m intervals $[x_{k-1}, x_k]$ $(k = 1, 2, \cdots, m)$. The partion determined by the set $\Delta \equiv \{x_0, x_1, \cdots, x_m\}$ is called the *partition* Δ. Let the maximum of $|x_k - x_{k-1}|$ $(k = 1, 2, \cdots, m)$ be $\delta(\Delta)$. The following limit exists (remember that f is assumed to be continuous) and called the *definite integral* of f on $[a, b]$:

$$\int_a^b f(x)dx \equiv \lim_{\delta(\Delta)\to 0} \sum_k f(\xi_k)(x_k - x_{k-1}),$$

where $\xi_k \in [x_{k-1}, x_k]$. The limit does not depend on the choice of ξ_k. f is called the *integrand*, x the *integration variable*, and b (resp., a) the *upper limit* (resp., *lower limit*) of integration. The integration variable x is a dummy variable in the sense that we may freely replace it with any letter.

We define for $b > a$ $\int_b^a f(x)dx \equiv -\int_a^b f(x)dx$, and $\int_a^a f(x)dx = 0$.

Sometimes, the definite integral is written as

$$\int_a^b dx\, f(x).$$

This notation clearly shows that integration is an operation applied to f.

A4.2 Riemann-integrability. Integration can be defined even if f is not continuous. Let f be a bounded function on $[a, b]$. For the partition Δ in **A4.1**, define

$$S_\Delta \equiv \sum_{k=1}^m M_k(x_k - x_{k-1}), \quad s_\Delta \equiv \sum_{k=1}^m m_k(x_k - x_{k-1}),$$

where M_k (resp., m_k) is the maximum (resp., minimum) value of f in $[x_{k-1}, x_k]$. Let $S \equiv \sup_\Delta S_\Delta$ and $s \equiv \inf_\Delta s_\Delta$ (Here the supremum (infimum) is looked for over all the possible finite partitions of $[a, b]$. If $S = s$, we say f is *Riemann integrable* on $[a, b]$. In this case, $S = s$ is the definition of $\int_a^b f(x)dx$. Even if f has finitely many discontinuous points in I, f is Riemann-integrable.

A4.3 Basic properties of definite integral. Let f and g be Riemann-integrable on the closed interval $[a, b]$.
(i) For $c \in (a, b)$

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

(ii) For arbitrary constant c_1 and c_2,

$$\int_a^b [c_1 f(x) + c_2 g(x)]dx = c_1 \int_a^b f(x)dx + c_2 \int_a^b g(x)dx.$$

(iii) If $f \geq 0$ on $[a, b]$, then $\int_a^b f(x)dx \geq 0$. If, furthermore, f is continuous and is not identically zero, then the integral is strictly positive.
(iv)

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx.$$

A4.4 Theorem [Mean value theorem].
(1) If f is a continuous function defined on a closed interval $[a, b]$, there exists a point $\xi \in (a, b)$ such that

$$\frac{1}{b-a} \int_a^b f(x)dx = f(\xi).$$

(2) If f and g are continuous on the closed interval [a,b], and if $g > 0$ on the open interval (a, b), then there exists $\xi \in (a, b)$ such that

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx.$$

□

A4.5 Fundamental theorem of calculus, primitive function, indefinite integral. If f is integrable on a closed interval $I = [a, b]$, then for $x \in [a, b]$, we can define the definite integral of f on $[a, x]$:

$$F(x) = \int_a^x f(t)dt$$

which is a function of x on I.
Theorem [Fundamental theorem of calculus].

$$\int_a^b f(x) = F(b) - F(a)$$

or $F'(x) = f(x)$.□
Any function such that $F'(x) = f(x)$ on I is called a *primitive function* of f. A primitive function for f, if any, is not unique; it is unique up to an additive constant. Thus any primitive function of f, if any, can be written as

$$F(x) = \int_a^x f(t)dx + C,$$

where C is called the *integration constant.*
The *indefinite integral* of f is defined as a primitive function of f, and is denoted by

$$\int f(x)dx.$$

A4.6 Integration by parts. Let f and g be C^1-functions ($\rightarrow$**A3.21**) on an interval I. Then

$$\int_a^b f(x)g'(x)dx = f(x)g(x)|_a^b - \int_a^b f'(x)g(x)dx,$$

where $h(x)|_a^b \equiv h(b) - h(a)$.

A4.7 Improper integral. When

$$\lim_{c \to b} \int_a^c f(x)dx$$

exists, we write this $\int_a^b f(x)dx$ even if f is not integrable on $[a, b)$ in the sense of **A4.2**, and call it an *improper integral.* b may be a discontinuous point of f or $\pm\infty$. It is easy to construct the Cauchy convergence criterion ($\rightarrow$**A1.3**) for improper integrals.
Improper integrals satisfy **A4.3** (i) and (ii), and if the improper integral of $|f|$ is definable (we say f is *absolutely integrable*; absolutely integrable functions are integrable.), (iii) holds as well.
Also the fundamental theorem of calculus ($\rightarrow$**A4.5**), and the mean value theorem ($\rightarrow$**A4.4**) are valid.

A4.8 Change of integration variables. Let f be a continuous function on an interval $I = [a, b]$, $\varphi(t)$ be a continuous function defined on an interval J whose range is in I. $\alpha, \beta \in J$ $(\alpha \neq \beta)$, $a = \varphi(\alpha)$ and $b = \varphi(\beta)$. Then

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt.$$

If f is an even function (→[AII13]), then

$$\int_a^b f(x)dx = \int_{-b}^{-a} f(x)dx.$$

If f is an odd function (→AII13]), then

$$\int_a^b f(x)dx = -\int_{-b}^{-a} f(x)dx.$$

A5 Infinite Series

A5.1 Changing the order of summation in infinite series. Absolutely convergent series and conditionally convergent series (→**A1.12**, **A1.14**) have diametrically different properties with respect to the rearrangement of the terms in the summation:
Theorem.
(i) The sum of an absolutely convergent series does not depend on the order of summation of the terms in the series.
(ii) If a series $\sum_{n=1}^\infty a_n$ is conditionally convergent, then for any given real ξ there is a reordering of the series $\{a_{\gamma(n)}\}$ such that

$$\sum_{n=1}^\infty a_{\gamma(n)} = \xi.$$

There is also a reordering to make the series divergent to $\pm\infty$. $\square$

A5.2 Product of two series. The product of two absolutely convergent series (→**A1.12**) can be computed via distributive law: Let $s = \sum a_n$ and $t = \sum b_n$, and both are absolutely convergent. Then

$$st = a_1b_1 + a_2b_1 + a_1b_2 + a_3b_1 + a_2b_2 + a_1b_3 + \cdots.$$

This is not necessarily true for conditionally convergent series, e.g., consider $a_n = b_n = (-1)^n/\sqrt{n}$.

A5.3 Theorem [Comparison theorem I. comparison with improper integral]. Let $r(x) > 0$ be a continuous monotone decreasing function (→**A2.11**) on $[k, +\infty)$ with k being a positive integer such that $\lim_{x\to\infty} r(x) = 0$. Let $r_n \equiv r(n)$. $\sum_{n=k}^{\infty} r_n$ converges (resp., diverges), if $\int_k^{\infty} r(x)dx$ converges (resp., diverges). □
Examples:
(i) $\sum_{n=1}^{\infty} n^{-s}$ $(s > 0)$ converges for $s > 1$ and diverges for $s \geq 1$.
(ii) $\sum_{n=2}^{\infty}\{1/[n(\log n)^s]\}$ $(s > 0)$ converges for $s > 1$ and diverges for $s \geq 1$.
(iii) $\sum_{n=3}^{\infty}\{1/[n \log n(\log\log n)^s]\}$ $(s > 0)$ converges for $s > 1$ and diverges for $s \geq 1$.

A5.4 Theorem [Comparison theorem II. comparison of series]. Let $\sum_{n=1}^{\infty} u_n$ and $\sum_{n=1}^{\infty} v_n$ be positive term series, and there is a positive integer n_0 such that for $n > n_0$

$$u_n/u_{n+1} \geq v_n/v_{n+1}.$$

Then
(i) If $\sum v_n$ converges, then so is $\sum u_n$.
(ii) If $\sum u_n$ diverges, then so is $\sum v_n$. □
From this theorem, we get useful convergence criteria:

A5.5 Cauchy's convergence criterion. Let $\sum a_n$ be a positive term series. Suppose the limit $\rho = \lim_{n\to\infty}(a_n/a_{n+1})$ exists. If $\rho < 1$, then the series converges, and if $\rho \geq 1$, the series diverges.

A5.6 Gauss' convergence criterion. For a positive term series $\sum a_n$ with

$$\frac{a_n}{a_{n+1}} = 1 + \frac{\sigma}{n} + O\left[\frac{1}{n^{1+\delta}}\right],$$

where δ is positive.[480] Then the series converges if $\sigma > 1$, and diverges if $\sigma \leq 1$.

A5.7 Abel's formula. Let the partial sums $s_m = \sum_{n=1}^{m} a_n$ and $t_m = \sum_{n=1}^{m} b_n$. Then

$$\sum_{n=k}^{m} a_n t_n = [s_m t_m - s_{k-1} t_k] - \sum_{n=k}^{m} s_n b_{n+1}.$$

This is a discrete analogue of integration by parts (→**A4.6**).
This transformation implies the following criteria:
(i) If $\sum_{n=1}^{\infty} a_n$ converges and $\sum_{n=2}^{\infty}(t_n - t_{n-1})$ converges absolutely, then

[480] For the symbol O see **A1.4**.

$\sum_{n=1}^{\infty} a_n t_n$ converges.
(ii) The sequence $\{s_n\}$ is bounded and $\{t_n\}$ is a monotone decreasing positive sequence converging to zero, then $\sum a_n t_n$ converges.
For example, (ii) implies that $\sum t_n \cos(nx)$ and $\sum t_n \sin(nx)$ converge, if $\{t_n\}$ is a monotone decreasing positive sequence converging to zero. This is an extension of [AI14] on alternating series.

A5.8 Function sequence, convergence. A sequence of functions $\{f_n(x)\}$ is called a *function sequence* defined on I, if the domains of all the functions in the sequence are identically I. For a fixed $x = \xi \in I$, if the sequence $\{f_n(\xi)\}$ converges, we say the function sequence converges at $x = \xi$. If the function sequence converges at every point of I, we say that the sequence converges on I. The limit for each x may be written as $f(x)$, which is regarded as the limit function of the function sequence, and we say the function sequence $\{f_n(x)\}$ converges to $f(x)$. More formally, we say that the function sequence $\{f_n(x)\}$ converges to $f(x)$ if for each $x \in I$ and for any positive number ϵ, there is a positive integer $n_0(\epsilon, x)$ such that

$$n > n_0(\epsilon, x) \Rightarrow |f_n(x) - f(x)| < \epsilon. \tag{A5.1}$$

A5.9 Uniform convergence. Let $\{f_n(x)\}$ be a function sequence defined on an interval I. If in (A5.1) $n_0(\epsilon, x)$ is independent of $x \in I$, we say the function sequence $\{f_n\}$ is *uniformly convergent* to f on I. That $\{f_n\}$ is uniformly convergent to f on I is equivalent to

$$\lim_{n\to\infty} \sup_{x\in I} |f_n(x) - f(x)| = 0.$$

A5.10 Theorem [Cauchy's criterion for uniform convergence]. Let $\{f_n(x)\}$ be a function sequence defined on an interval I. A necessary and sufficient condition for the sequence to be uniformly convergent is that there is a positive integer $n_0(\epsilon)$ such that for any $x \in I$

$$n > n_0(\epsilon), \quad m > n_0(\epsilon) \Rightarrow |f_n(x) - f_m(x)| < \epsilon.$$

□

A5.11 Function series, convergence, uniform convergence, maximal convergence. $\sum_{n=1}^{\infty} f_n(x)$ is called a function series. Let its partial sum be $s_m(x) \equiv \sum_{n=1}^{m} f_n(x)$. If the function sequence $\{s_n(x)\}$ (uniformly) converges to $s(x)$, we say the series $\sum_{n=1}^{\infty} f_n(x)$ (uniformly) converges to $s(x)$, which is called the *sum* of the series. If $\sum_{n=1}^{\infty} f_n(x)$ is uniformly and absolutely convergent, we say the series is *maximally convergent.*

A5.12 Theorem [Uniform convergence preserves continuity]. Let $\{f_n(x)\}$ be a function sequence of continuous functions defined on an interval I.
(i) If the sequence uniformly converges to f on I, then f is continuous in I.
(ii) If the series $\sum_{n=1}^{\infty} f_n(x)$ converges uniformly, then its sum is a continuous function on I.□

A5.13 Theorem [Dini's theorem]. Let $\{f_n(x)\}$ be a sequence of continuous functions defined on the closed interval $[a, b]$. Suppose the sequence is monotonically decreasing: for any $x \in [a, b]$ $f_1(x) \geq f_2(x) \geq \cdots \geq f_n(x) \geq \cdots$. If the sequence $\{f_n(x)\}$ converges on $[a, b]$ to a continuous function $f(x)$, then the sequence uniformly converges to f on $[a, b]$. □.

A5.14 Theorem [Comparison theorem]. Let $\sum_{n=1}^{\infty} a_n$ be a convergent positive term series. For a sequence of $\{f_n(x)\}$, suppose $|f_n(x)| \leq a_n$ for all n on an interval I. Then the infinite sum $\sum_{n=1}^{\infty} f_n(x)$ is maximally convergent.□

A5.15 Theorem [Exchange of limit and integration]. Let $\{f_n(x)\}$ be a sequence of continuous functions defined on $[a, b]$, uniformly convergent to $f(x)$ there. Then

$$\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx.$$

□
A more general theorem (Arzelá's theorem) will be given in **A5.17**. The theorem implies that a uniformly convergent series of continuous functions is termwisely integrable:

$$\int_c^x \sum_{n=1}^{\infty} f_n(x)dx = \sum_{n=1}^{\infty} \int_c^x f_n(x)dx.$$

A5.16 Theorem [Exchange of limit and differentiation]. Let $f_n(x)$ be a C^1-function ($\to$**A3.21**). If $\sum_{n=1}^{\infty} f_n(x)$ converges on I, and $\sum_{n=1}^{\infty} f_n'(x)$ converges uniformly on I, then the sum of the series is differentiable and

$$\frac{d}{dx} \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} f_n'(x)$$

□

A5.17 Theorem [Arzelá's theorem]. Let $f_n(x)$ be a continuous function defined on a closed set [a,b] (actually this need not be a closed interval) and uniformly bounded, i.e., there is a positive number M independent of n such that $|f_n(x)| < M$ on the interval. If the function sequence $\{f_n(x)\}$ converges to a continuous function $f(x)$ on [a,b], then

$$\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx.$$

□

A5.18 Majorant. For a function sequence $\{f_n(x)\}$ defined on an interval I, a function $\sigma(x)$ such that $|f_n(x)| < \sigma(x)$ is called a *majorant* of the sequence.
Theorem. If a majorant $\sigma(x)$ is integrable on the interval, then the order of integration and $\lim_{n\to\infty}$ can be exchanged.□

A5.19 Convergence radius of power series. For a power series $\sum_{n=0}^{\infty} a_n x^n$,

$$r \equiv \frac{1}{\limsup_{n\to\infty} |a_n|^{1/n}}$$

is called the *convergence radius* of the power series (the reason for the name is seen from the following theorem **A5.20**. The formula is called the *Cauchy-Hadamard formula*). Here, if the lim sup diverges to $+\infty$, then we define $r = 0$, and if lim sup converges to zero, then we define $r = +\infty$.

A5.20 Theorem [Power series is termwisely differentiable]. The power series $\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent for $|x| < r$, and is divergent for $|x| > r$, where r is the convergence radius (→**A5.19**). For any $0 < \rho < r$, the series is uniformly convergent (→**A5.11**) in $[-\rho, \rho]$ to a continuous function (*cf.* **A5.12**), so that the series is termwisely differentiable there. □

A5.21 Theorem [Power series defines a real analytic function]. The power series $\sum_{n=0}^{\infty} a_n x^n$ whose convergence radius is r uniquely determines a C^ω-function (→**A3.22**) f in the open interval $(-r, r)$. Actually, the power series is the Taylor series (→**A3.16**) for f. □

A5.22 Theorem [Continuity at $x = r$ or $-r$]. Let r be the convergence radius of the power series $\sum_{n=0}^{\infty} a_n x^n = f(x)$. If $\sum_{n=0}^{\infty} a_n r^n$ is convergent, then $f(x)$ is continuous in $(-r, r]$. If $\sum_{n=0}^{\infty} a_n (-r)^n$ is convergent, then $f(x)$ is continuous in $[-r, r)$. □

A5.23 Infinite product. For a sequence $\{a_n\}$ ($a_n \neq 0$) $a_1 a_2 \cdots a_n \cdots$ is called an *infinite product*, and is denoted by $\prod_{n=1}^{\infty} a_n$. $p_n = a_1 a_2 \cdots a_n$ is called the *partial product*.

A5.24 Convergence of infinite product. Let $\prod_{n=1}^{\infty} a_n$ be an infinite product and its partial product sequence be $\{p_n\}$. If this sequence converges, and $p = \lim_{n\to\infty} p_n$ is not zero, we say the infinite product *converges* to p: $p = \prod_{n=1}^{\infty} a_n$. Else, we say the infinite product is *divergent.*

A5.25 Theorem [Convergence condition for infinite product].
(i) A necessary and sufficient condition for the infinite product $\prod_{n=1}^{\infty}(1+u_n)$ $(u_n > -1)$ to be convergent is that the infinite series $\sum_{n=1}^{\infty} \log(1+u_n)$ converges.
(ii) If $\sum_{n=1}^{\infty} u_n$ $(u_n > -1)$ converge absolutely, then $\prod_{n=1}^{\infty}(1+u_n)$ converges. (In this case we say the infinite product converges absolutely, and the product does not depend on the order of its terms.)
(iii) If $\sum u_n$ $(u_n > -1)$ and $\sum u_n^2$ both converges, then the infinite product $\prod_{n=1}^{\infty}(1+u_n)$ converges. □

A5.26 Conditional convergence of infinite product. If an infinite product $\prod_{n=1}^{\infty}(1+u_n)$ converges but does not converge absolutely (→**A5.25**(ii)), we can reorder the product to converge to any positive number. □
This is quite parallel to a similar theorem for conditionally convergent series (→A5.1(ii)).

A6 Function of Two Variables

Since real valued-functions of two variables illustrate complications due to the existence of many independent variables, in this rudimentary part, we discuss only a function defined on a point set D in $\boldsymbol{R}^2$.

A6.1 Rudiments of topology.
(i) If an open set $U \in \boldsymbol{R}^2$ is not a join of two open sets, U is said to be *connected.*
(ii) **Theorem.** A necessary and sufficient condition for an open set U to be connected is that any points $P, Q \in U$ can be connected by a piecewise straight curve in U. □
(iii) A connected open set is called a *region*, and its closure (→**A1.19**) is called a *closed region.*

(iv) Distance $\rho(x, y)$ of two points x and y in $\boldsymbol{R}^2$ is defined as

$$\rho(x, y) = |x - y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2},$$

where x (or y) is identified with its coordinate expression, say, (x_1, x_2).

A6.2 Function, domain, range. Let D be a point set in $\boldsymbol{R}^2$. A rule which defines a correspondence of each $x \in D$ to some real is called a function from $D \subset \boldsymbol{R}^2$ to $\boldsymbol{R}$. D is called its *domain* and $f(D) \subset \boldsymbol{R}$ is called its *range*.[481] We write $f : D \to \boldsymbol{R}$.

A6.3 Limit. Let D be a point set in $\boldsymbol{R}^2$. For $f : D \to \boldsymbol{R}$, we say $\lim_{P \to A} f(P) = \alpha \in \boldsymbol{R}$, if for any positive number ϵ there is a positive number $\delta(\epsilon)$ such that

$$\rho(P, A) < \delta(\epsilon) \Rightarrow |f(P) - \alpha| < \epsilon.$$

Notice that the limit should not depend on how P approaches A.
It is easy to write down Cauchy's criterion for the convergence (*cf.* **A1.3**).

A6.4 Continuity. Let D be a point set in $\boldsymbol{R}^2$. A function $f : D \to \boldsymbol{R}$ is continuous at an accumulation point ($\to$**A1.19**) $P \in D$, if $\lim_{Q \to P} f(Q) = f(P)$. (To discuss the continuity on the point of D which are not accumulating points of D is uninteresting.)
Uniform continuity can also be defined quite analogously as in the one-variable function case (*cf.* **A2.8**).

A6.5 Theorem [Maximum value theorem]. A real-valued continuous function defined on a bounded closed set $D \subset \boldsymbol{R}^2$ has a maximum and minimum values on D. The range of f is a closed interval. □
(*cf.* [AII9]).

A6.6 Partial differentiation. Let $f(x, y)$ be a real-valued function defined in a region $D \subset \boldsymbol{R}^2$, and $(a, b) \in D$. If $f(x, b)$ is differentiable at a with respect to x, we say that $f(x, y)$ is *partially differentiable* with respect to x at (a, b), and the derivative is denoted by $f_x(a, b)$. More generally, if f is partial differentiable in D with respect to x, we may define $f_x(x, y)$:

$$f_x(x, y) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}.$$

If we write $z = f(x, y)$, $f_x(x, y)$ is written as $\partial z / \partial x$. $f_x(x, y)$ is called the *partial derivative* of f with respect to x. We say that we *partial-differentiate* f with respect to x to obtain $f_x(x, y)$. Similarly, we can

[481] The set $\{f(x) : x \in D\}$ is often written as $f(D)$.

define the partial derivative with respect to y of f. Analogously, we can define higher-order (mixed) partial derivatives like f_{xxy}.
Warning. Even if f_x and f_y exists at a point, f need not be continuous at the point. □
This implies that the 'differentiability' of f must be defined separately from its partial differentiability.

A6.7 Differentiability, total differential. Let $f(x, y)$ be a real-valued function defined in a region $D \subset \boldsymbol{R}^2$, and $(a, b) \in D$. We say f is *differentiable* at (a, b) if there is constants A and B such that

$$f(x, y) = f(a, b) + A(x - a) + B(y - b) + o[\sqrt{(x - a)^2 + (y - b)^2}].$$

Theorem. If f above is differentiable at (a, b), then f is continuous there, and is partially differentiable with respect to x and y with $A = f_x(a, b)$, $B = f_y(a, b)$. □.
$dz = f_x dx + f_y dy$ is called the *total differential* of f.
We say that f is differentiable in D, if f is differentiable at every point in D.
Intuitively, if a local linear approximation is reliable, we say the function is differentiable.
A6.8 Theorem [Partial differentiability and differentiability]. Let f be a function defined in a region $D \subset \boldsymbol{R}^2$. If f_x and f_y exist and are continuous in D, then f is differentiable in D. □

A6.9 Theorem [Order of partial differentiation]. Let f be a function defined in a region $D \subset \boldsymbol{R}^2$. If partial derivatives f_x, f_y, f_{xy} and f_{yx} exist and if f_{xy} and f_{yx} are continuous, then $f_{xy} = f_{yx}$. □

A6.10 Theorem [Young's theorem]. Let f be a function defined in a region $D \subset \boldsymbol{R}^2$. If f_x and f_y exist and f is differentiable, then $f_{xy} = f_{yx}$. □

A6.11 Theorem [Schwarz' theorem]. Let f be a function defined in a region $D \subset \boldsymbol{R}^2$. If partial derivatives f_x, f_y and f_{xy} exist and if f_{xy} is continuous, then f_{yx} exists and $f_{xy} = f_{yx}$. □

A6.12 $\boldsymbol{f_{xy} = f_{yx}}$ is not always correct. Let $f(x, y) = xy(x^2 - y^2)/(x^2 + y^2)$ except for $(0, 0)$, where f is defined to be 0. Then $f_{xy}(0, 0) \neq f_{yx}(0, 0)$; the left-hand-side is -1, and the right-hand-side is 1.

A6.13 $\boldsymbol{C^n}$-class function. If f has all the partial derivatives of order n, which are all continuous, we say that f is a *C^n-function.* If derivatives of any order exists, we say that the function is a C^∞-function.

A6.14 Composite function. Let $\varphi(t)$ and $\psi(t)$ be continuous functions defined on an interval I such that $(\varphi(t), \psi(t)) \in D$ for all $t \in I$. If $f(x, y)$ is a continuous function defined on D, then $f(\varphi(t), \psi(t))$ is continuous.
If $\varphi(t)$ and $\psi(t)$ are differentiable with respect to t, and $f(x, y)$ is differentiable in D, then $f(\varphi(t), \psi(t))$ is differentiable with respect to t, and

$$\frac{d}{dt} f(\varphi(t), \psi(t)) = f_x(\varphi(t), \psi(t))\varphi'(t) + f_y(\varphi(t), \psi(t))\psi'(t).$$

If $\varphi(t)$ and $\psi(t)$ are C^n-functions of t, and $f(x, y)$ is C^n in D, then $f(\varphi(t), \psi(t))$ is again C^n.
These propositions hold even if we replace the function of t with functions of s and t. For example, If $\varphi(s, t)$ and $\psi(s, t)$ are C^n-functions of s and t in a domain D_1, $(\varphi(s, t), \psi(t, s)) \in D$, and $f(x, y)$ is C^n in D, then $f(\varphi(s, t), \psi(s, t))$ is again C^n in D_1.

A6.15 Taylor's formula. Let $f(x, y)$ be a C^n-class function defined on a region D, $(a, b) \in D$, and the line segment AP with $P = (a+h, b+k)$ be in D. Then

$$f(a + f, b + k) = f(a, b) + \sum_{m=1}^{n-1} \frac{1}{m!} \left(f\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} \right)^m f(a, b) + R_n$$

with

$$R_n \equiv \frac{1}{n!} \left(f\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} \right)^n f(a + \theta h, b + \theta k)$$

for some $\theta \in (0, 1)$. R_n is called the residue.
If f is a C^∞-function ($\rightarrow$[13]) in a region D and if in some subregion D_A of D $\lim_{n\to\infty} R_n = 0$, then we say f is *Taylor-expandable* in D_A:

$$f(x, y) = f(a, b) = \sum_{n=1}^{\infty} \sum_{p+q=n} \frac{\partial^{p+q} f(a, b)}{\partial x^p \partial y^q}(x - a)^p (y - b)^q.$$

If f is Taylor-expandable, we say f is a real analytic function (C^ω-function) of two variables.
This is a double series, so we need some general theory of double series and double sequences.

A6.16 Limit of double sequence. Let $\{a_{nm}\}$ be a double sequence. If for any positive ϵ, there is a positive integer $N(\epsilon)$ such that

$$m > N(\epsilon),\ n > N(\epsilon) \Rightarrow |a_{mn} - \alpha| < \epsilon,$$

then we say the double sequence converges to α, and write $\lim_{m,n\to\infty} a_{mn} = \alpha$. It is easy to state Cauchy's convergence criterion ($\to$**A1.3**) for a double series.

A6.17 Warning. $\lim_{m,n\to\infty}$ and $\lim_{m\to\infty}\lim_{n\to\infty}$ or $\lim_{m\to\infty}\lim_{n\to\infty}$ are distinct. For example, if $a_{mn} = 2mn/(m^2+n^2)$, $\lim_{m\to\infty}\lim_{n\to\infty} a_{mn} = 0$ and $\lim_{n\to\infty}\lim_{m\to\infty} a_{mn} = 0$, but $\lim_{m,n\to\infty} a_{mn}$ does not exist. If $a_{mn} = (-1)^n/m + (-1)^m/n$, then $\lim_{m,n\to\infty} a_{mn} = 0$ but the other limits do not exist.

A6.18 Theorem [Exchange of limits]. Suppose $\lim_{m,n\to\infty} a_{mn} = \alpha$ exists. If for each n $\lim_{m\to\infty} a_{mn}$ exists, then $\lim_{n\to\infty}\lim_{m\to\infty} a_{mn} = \alpha$. If for each m $\lim_{n\to\infty} a_{mn}$ exists, then $\lim_{m\to\infty}\lim_{n\to\infty} a_{mn} = \alpha$. □

A6.19 Double series, convergence. For a double sequence $\{a_{mn}\}$, $\sum_{m,n=1}^{\infty} a_{mn}$ is called a *double series.* We say that the double series *converges* if the double sequence $\{s_{mn}\}$ of its partial sums $s_{mn} = \sum_{p=1}^{m}\sum_{q=1}^{n} a_{pq}$ converges. Its absolute convergence can also be defined analogously as in the ordinary series case ($\to$**A1.12**).
Theorem. If a double sequence $\sum_{m,n=1}^{\infty} a_{mn}$ converges absolutely, then $\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} a_{mn} = \sum_{m,n=1}^{\infty} a_{mn}$. □

A6.20 Power series of two variables. The set G such that $\sum_{m,n=0}^{\infty} a_{mn}x^m y^n$ for $\forall(x,y) \in G$ is absolutely convergent is called the convergence domain of the double power series.
Theorem. If for $(\xi,\eta) \neq (0,0)$ the double series $\sum a_{mn}\xi^m\eta^n$ is bounded, then for $|x| < |\xi|$ and $|y| < |\eta|$ the double power series $\sum_{m,n=0}^{\infty} a_{mn}x_m y_n$ is absolutely convergent.□

A6.21 Exchange of order of limits, uniform convergence. If for any positive ϵ there is $N(\epsilon)$ independent of m such that

$$n > N(\epsilon) \Rightarrow |a_{mn} - \alpha_m| < \epsilon,$$

we say $\{a_{mn}\}$ converges to α_m *uniformly* with respect to m.
Theorem. If $\{a_{mn}\}$ converges to α_m uniformly with respect to m in the $n \to \infty$ limit, and if α_m converges to α in the $m \to \infty$ limit, then $\lim_{m,n\to\infty} a_{mn} = \alpha$. □
Theorem. If $\lim_{n\to\infty}\lim_{m\to\infty} a_{mn}$ exists, $\{a_{mn}\}$ converges to α_m uniformly with respect to m in the $n \to \infty$ limit, and if α_m converges to α in the $m \to \infty$ limit, then $\lim_{m\to\infty}\lim_{n\to\infty} a_{mn} = \lim_{n\to\infty}\lim_{m\to\infty} a_{mn} = \alpha$. □
In contrast to **A6.18** here the existence of $\lim_{m,n\to\infty} a_{mn}$ is not assumed.

A6.22 Counterexample.

(i) For $a_{mn} = (-1)^n m/(m+n)$, $\lim_{m\to\infty} \lim_{n\to\infty} a_{mn} = 0$, $\lim_{n\to\infty} \lim_{m\to\infty} a_{mn}$ does not exist.
(ii) For $a_{mn} = m/(m+n)$ both limits exist but not identical.

A6.23 Theorem [Differentiation and integration within integration]. Let $f(x, y)$ be a bounded function defined on a rectangle $K = \{(x, y) | x \in [a, b], y \in [c, d]\}$. Assume that f is continuous as a function of x (resp., y) for each y (resp., x). Then
(i) $\int_a^b dx f(x, y)$ is a continuous function of y in $[c, d]$.
(ii) If $f(x, y)$ is partially differentiable with respect to y, and if $f_y(x, y)$ is bounded on K, and continuous as a function of x for each y, then

$$\frac{d}{dy} \int_a^b dx f(x, y) = \int_a^b dx \frac{\partial}{\partial y} f(x, y).$$

(iii)

$$\frac{d}{du} \int_a^u dx f(x, y) = f(u, y).$$

(iv)

$$\int_c^d dy \int_a^b dx f(x, y) = \int_a^b dx \int_c^d dy f(x, y).$$

□

A6.24 Theorem [Differentiation and integration within improper integration]. Let $f(x, y)$ be a bounded function defined on a rectangle $K = \{(x, y) | x > a, y \in [c, d]\}$. Assume that f is continuous as a function of x (resp., y) for each y (resp., x). Assume, furthermore, that there is a nonnegative continuous function $\sigma(x)$ such that $|f(x, y)| \le \sigma(x)$ and $\int_a^{+\infty} dx \sigma(x) < +\infty$. Then
(i) $\int_a^{+\infty} dx f(x, y)$ is a continuous function of y in $[c, d]$.
(ii) If $f(x, y)$ is partially differentiable with respect to y and if there is a nonnegative continuous function $\sigma(x)$ such that $|f_y(x, y)| \le \sigma_1(x)$ and $\int_a^{+\infty} dx \sigma_1(x) < +\infty$, then

$$\frac{d}{dy} \int_a^{+\infty} dx f(x, y) = \int_a^{+\infty} dx \frac{\partial}{\partial y} f(x, y).$$

(iii)

$$\int_c^d dy \int_a^{+\infty} dx f(x, y) = \int_a^{+\infty} dx \int_c^d dy f(x, y).$$

□

A7 Fourier Series and Fourier Transform

In this section all the integrals are Riemann integrals [AIV1]. Thus integrable or absolutely integrable means Riemann-integrable and absolutely Riemann integrable.

A7.1 Fourier series.{Fourier series Let f be a function on $\boldsymbol{R}$ with period 2π.[482] Assume that the following integrals exist[483]

$$a_n = \frac{1}{\pi}\int_0^{2\pi} dx f(x)\cos nx \ \text{ for } n = 0, 1, 2, \cdots,$$

$$b_n = \frac{1}{\pi}\int_0^{2\pi} dx f(x)\sin nx \ \text{ for } n = 1, 2, 3, \cdots.$$

Then

$$S[f] = \frac{1}{2}a_0 + \sum_{n+1}^{\infty}(a_n \cos nx + b_n \sin nx)$$

is called the *Fourier series* of f. To construct $S[f]$ is said to *Fourier-expand* f.
Notice that the Fourier series converges uniformly ($\rightarrow$[AV11]) if $\sum_{n=0}^{\infty}|a_n|$ and $\sum_{n=0}^{\infty}|b_n|$ both converge.

A7.2 Theorem. Let f be a 2π periodic function which has at most finitely many discontinuities, and is absolutely integrable on $[0, 2\pi]$. If $S[f]$ converges uniformly, then $S[f](x_0)$ converges to $f(x_0)$ if f is continuous at x_0. Specifically, if f is 2π-periodic continuous function, then $S[f] = f$.$\square$
This theorem uses the property of the Fourier series (its uniform convergence), so it is not very satisfactory. **A7.8** below tells us that we cannot remove of this extra condition from this theorem.

A7.3 Complex Fourier series. Let f be a function on $\boldsymbol{R}$ with period 2π. Assume that the following integrals exist

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} dx f(x)e^{-ikx} \ \text{ for } k = \cdots, -2, -1, 0, 1, 2, \cdots.$$

Then

$$S[f] \equiv \sum_{k=-\infty}^{\infty} c_n e^{ikx}$$

[482]That is, $f(x + 2\pi) = f(x)$ for any $x \in \boldsymbol{R}$.
[483]The integration range can be $[-\pi, \pi]$.

is called the *complex Fourier series* of f.
Needless to say, a theorem corresponding to **A7.2** holds.

A7.4 Theorem [Bessel's inequality]. If f is 2π-periodic and square integrable on $[0, 2\pi]$, then

$$2\pi \sum_{k=-\infty}^{\infty} |c_k|^2 \leq \int_{-\pi}^{\pi} dx |f(x)|^2.$$

□

A7.5 Theorem [Parseval's equality]. If f is a 2π-periodic continuous function, and f' is square integrable (especially f is a 2π-periodic C^1-function (→**A3.21**)), then $S[f]$ uniformly converges to f. In this case the following equality holds

$$2\pi \sum_{k=-\infty}^{\infty} |c_k|^2 = \int_{-\pi}^{\pi} dx |f(x)|^2,$$

which is called *Parseval's equality.*□
Warning. The continuity of f is not sufficient even for pointwise convergence of $S[f]$ to f. See **A7.8**.

A7.6 L^2-convergence. A function sequence f_n defined on $(-\pi, \pi)$ is said to *L^2-converge* to f (or to converge in the square mean), if

$$\int_{-\pi}^{\pi} dx |f_n(x) - f(x)|^2 \to 0$$

as $n \to \infty$.

A7.7 Theorem. If f is a 2π-periodic continuous function, then $S[f]$ L^2-converges to f, and Parceval's equality (→**A7.5**) holds. □

A7.8 Theorem [duBois-Reymond]. For a 2π-periodic function f, its continuity does not guarantee the pointwise convergence of $S[f]$ to f. [Counterexamples exist.] □
However,

A7.9 Theorem [Fejér]. Let S_n be the partial sum of the Fourier series up to the n-th term. Define

$$\sigma_n \equiv \frac{1}{n+1} \sum_{k=0}^{n} S_k.$$

If f is a 2π-periodic continuous function, then σ_n uniformly converges to f. □

A7.10 Piecewise C^1-function. A function f is said to be piecewise C^1, if there are finitely many points $\gamma_1 < \gamma_2 < \cdots < \gamma_m$ such that on each open interval (γ_l, γ_{l+1}) f and f' are continuous and bounded. Notice that at each γ_l right and left limits ($\to$**A2.6**) of f (denoted by $f(\gamma_l + 0)$ and $f(\gamma_l - 0)$) exist.

A7.11 Theorem. If f is piecewisely C^1, then $S[f]$ converges to $[f(x+0) + f(x-0)]/2$ for all x. The same holds if f' is piecewise continuous and square-integrable (i.e., its boundedness need not be assumed). The convergence is uniform except in the arbitrarily small neighborhood of the discontinuities of f. □

A7.12 Theorem. If f is a 2π-periodic function, integrable on $(-\pi, \pi)$ and is of bounded variation,[484] then the conclusion of **A7.11** holds. □

A7.13 Theorem [Locality of convergence]. Let f_1 and f_2 be piecewise 2π-periodic functions integrable on $(-\pi, \pi)$. If there is a neighborhood of x_0 such that $f_1 \equiv f_2$ on it, then $S[f_1]$ converges (resp., diverges) at x_0 if and only if $S[f_2]$ converges (resp., diverges) at x_0. When they converge, the limits are identical. □

A7.14 Fourier transform. Let f be an integrable function on $\boldsymbol{R}$. If the following integral exists

$$\hat{f} = \int_{-\infty}^{\infty} dx f(x) e^{-ikx},$$

it is called the *Fourier transform* of f. Mathematicians often multiply $1/\sqrt{2\pi}$ to this definition to symmetrize the formulas. However, this makes the convolution formula **A7.20**(iv) awkward. For physicists and practitioners, the definition here is the most convenient.
If a function $f : \boldsymbol{R} \to \boldsymbol{C}$ is continuous except for finitely many points, and absolutely integrable, then its Fourier transform $\hat{f} : \boldsymbol{R} \to \boldsymbol{C}$ is a bounded continuous function such that $\lim_{k\to\infty} f(\pm k) = 0$.
Also we have an important relation

$$\hat{f'} = ik\hat{f}.$$

A7.15 Rapidly decreasing function. A function $f : \boldsymbol{R} \to \boldsymbol{C}$ is called a *rapidly decreasing function*, if the following two conditions hold:
(i) f is a C^∞-function ($\to$**A3.21**).

[484]That is, f can be written as a difference of two monotone increasing functions ($\to$**A2.11**).

(ii) For any $k, l \in \boldsymbol{N}$, $x^l f^{(k)} \to 0$ in the $|x| \to \infty$ limit.
The function is also called a *Schwartz*-class function (or $\mathcal{S}$-function).

A7.16 Inverse Fourier transform. If f is a rapidly decreasing function, then the following *inversion formula* holds:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk \hat{f}(k) e^{ikx}.$$

□

A7.17 Theorem. If $f : \boldsymbol{R} \to \boldsymbol{C}$ is continuous (and bounded), and both f and $\hat{f}$ are absolutely integrable, then the inversion formula holds. □

A7.18 Parseval's equality. If the inversion formula holds and if f is square integrable, we have

$$\int_{-\infty}^{\infty} dx |f(x)|^2 = 2\pi \int_{-\infty}^{\infty} dk |\hat{f}(k)|^2.$$

□

A7.19 Convolution. Let f and g be integrable function defined on $\boldsymbol{R}$. The following $h(x)$ is called the *convolution* of f and g and is denoted by $f * g$:

$$h(x) = (f * g)(x) \equiv \int_{-\infty}^{\infty} dy f(x-y) g(y).$$

A7.20 Properties of convolution.
(i) The definition is symmetric with respect to f and g, that is, $f * g = g * f$.
(ii) If f and g are rapidly decreasing, then so is h.
(iii) $h^{(k)} = f^{(k)} * g$
(iv) $\widehat{f * g} = \hat{f}\hat{g}$.

A7.21 Theorem [Inversion formula for piecewise C^1-function]. Let f be piecewise C^1-function ($\to$**A7.10**) on $\boldsymbol{R}$. Then

$$\frac{1}{2}[f(x_0 - 0) + f(x_0 + 0)] = \frac{1}{2\pi} \text{p.v.} \int_{-\infty}^{\infty} dk e^{ikx_0} \hat{f}(k).$$

Here p.v. implies Cauchy's principal value of the integral. □
We can write the formula as

$$\frac{1}{2}[f(x_0 - 0) + f(x_0 + 0)] = \lim_{\lambda \to \infty} \int_{-\infty}^{\infty} d\xi \frac{\sin(\lambda(x_0 - \xi))}{x_0 - \xi} f(\xi).$$

□

A7.22 Multidimensional case. It is easy to generalize the rapidly decreasing property to multidimensional cases. If a function is rapidly decreasing, then formal generalization of the above results to multidimensional cases are legitimate.

A8 Ordinary Differential Equation

Practical advice. See Schaum's outline series *Differential Equations* by R. Bronson for elementary methods and practice. To learn the theoretical side, V. I. Arnold, *Ordinary differential equations* (MIT Press 1973; there is a new version from Springer) is highly recommended.

A8.1 Ordinary differential equation. Let y be a n-times differentiable function of $x \in \boldsymbol{R}$. A funcitonal relation $f(x, y, y', \cdots, y^{(n)}) = 0$ among x,y, y', $\cdots$, $y^{(n)}$ is called an *ordinary differential equation* (ODE) for $y(x)$, and n is called its *order*, where the domain of f is assumed to be appropriate. Such $y(x)$ that satisfies $f = 0$ is called a *solution* to the ODE.

A8.2 General solution, singular solution. The solution $y = \varphi(x, c_1, c_2, \cdots, c_n)$ to $f = 0$ in **A8.1** which contains n arbitrary constants c_1, $\cdots$, c_n (which are called *integral constants*) is called the *general solution* of $f = 0$. A solution which can be obtained from this by specifying the arbitrary constants is called a *particular solution.* A solution which cannot be obtained as a particular solution is called a *singular solution.*

A8.3 Normal form. If the highest order derivative of y is explicitly solved as $y^{(n)}(x) = F(x, y', \cdots, y^{(n-1)})$, we say the ODE is in the *normal form.* Notice that not normal ODE's may have many pathological phenomena.

A8.4 Initial value problem of first order ODE. Consider the following first order ODE

$$\frac{dy}{dx} = f(x, y), \tag{A8.1}$$

where f is defined in a region $D \subset \boldsymbol{R}^2$. To solve this under the condition that $y(x_0) = y_0$ ($(x_0, y_0) \in D$) is called a *initial value problem.*

A8.5 Theorem [Cauchy-Peano]. If for (A8.1) f is continuous on a region $D \subset \mathbf{R}^2$, then for any $(x_0, y_0) \in D$ there is a solution $y(x)$ of (A8.1) passing through this point whose domain is an open interval (α, ω) $(-\infty \leq \alpha < \omega \leq \infty)$, and in the limits $x \to \alpha$ and $x \to \omega$ $y(x)$ approaches the boundary of D or the solution becomes unbounded. □

A8.6 Lipschitz condition. Let $f(x, y)$ be a continuous function whose domain is a region $D \subset \mathbf{R}^2$. For any compact set (→[AI25]) $K \subset D$, if for any (x, y), $(x', y') \in K$ there is a positive constant L_K (which is usually dependent on K) such that

$$|f(x,y) - f(x', y')| \leq L_K |y - y'|,$$

then f is said to satisfy a *Lipschitz condition* on D for y.
If f and f_y are both continuous in D, then f satisfies a Lipschitz condition on D.

A8.7 Theorem [Cauchy-Lipschitz uniqueness theorem]. For (A8.1), if f satisfies a Lipschitz condition on D for y, then if there is a solution passing through $(x_0, y_0) \in D$, it is unique. □

A8.8 Theorem. Let $f : \boldsymbol{R} \to \boldsymbol{R}$ be a continuous and monotone decreasing function. Then the initial value problem $dy/dx = f(y)$ (for $x > x_0$) with $y(x_0) = y_0$ has a unique solution for $x \geq x_0$. □

A8.9 Method of quadrature. To solve an ODE by a finite number of indefinite integrals is called the *method of quadrature.* Representative examples are given in **A8.10-A8.13**.

A8.10 Separation of variables. The first order equation of the following form

$$\frac{dy}{dx} = p(x)q(y),$$

where p and q are continuous functions, is solvable by the separation of variables: Let $Q(y)$ be a primitive function (→[AIV5]) of $1/q(y)$ and P that of p. Then $Q(y) = P(x) + C$ is the general solution, where C is the integration constant.

A8.11 Perfect differential equation, integrating factor. The first order ODE of the following form

$$\frac{dy}{dx} = -\frac{P(x,y)}{Q(x,y)},$$

where $Q \neq 0$. If there is a function Φ such that $\Phi_x = P$ and $\Phi_y = Q$, then $\Phi(x, y) = C$, C being the integral constant, is the general solution. Even if P and Q may not have such a 'potential' Φ, P and Q times some function I called *integrating factor* may have a 'potential.' However, it is generally not easy to find such a factor except for some special cases.

A8.12 Linear first order equation, variation of parameter. The first order equation

$$\frac{dy}{dx} = p(x)y + q(x)$$

is called a *linear equation.* The equation can be solved by the method of variation of parameters. Let $y(x) = C(x)e^{\int^x p(s)ds}$. Then the equation for C can be integrated easily. As we will see in **A8.14**, the method of variation of parameters *always* works for linear equations.

A8.13 Bernoulli equation. The first order equation of the following form is called a *Bernoulli equation*:

$$\frac{dy}{dx} = p(x)y + Q(x)y^n,$$

where n is a real number. Introducing the new variable $z(x) = y(x)^{1-n}$, we can reduce this equation to the case [12] for $z(x)$.

A8.14 Linear ODE with constant coefficients, characteristic equation. Consider

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = 0, \tag{A8.2}$$

where a and b are constants.

$$P(\lambda) = \lambda^2 + a\lambda + b$$

is called its *characteristic equation*, and its roots are called *characteristic roots.*

A8.15 Theorem [General solution to (A8.2)]. If the characteristic roots of (A8.2) are α and β ($\neq \alpha$), then its general solution is the linear combination of $\varphi_1(x) = e^{\alpha x}$ and $\varphi_2(x) = e^{\beta x}$. If $\alpha = \beta$, then the general solution is the linear combination of $\varphi_1(x) = e^{\alpha x}$ and $\varphi_2(x) = xe^{\alpha x}$ (the characteristic roots need not be real.) □
$\varphi_1(x)$ and $\varphi_2(x)$ are called *fundamental solutions* and $\{\varphi_1(x), \varphi_2(x)\}$ is called the *system of fundamental solutions* for (A8.2).

A8.16 Inhomogeneous equation, Lagrange's method of variation of constants. An ODE

$$\frac{d^2y}{dx^2} + a\frac{dy}{dx} + by = f(x) \tag{A8.3}$$

with nonzero f is called an *inhomogeneous* ODE (the one without nonzero f is called a homogeneous equation). The general solution is given by the sum of the general solution to the corresponding homogeneous equation and one particular solution to the inhomogeneous problem. A method to find one solution to (A8.3) is the *Lagrange's method of variation of constants.* Let $\varphi_i(x)$ be the fundamental solutions and determine the functions $C_i(x)$ to satisfy (A8.3):

$$u(x) = C_1(x)\varphi_1(x) + C_2(x)\varphi_2(x).$$

One solution can be obtained from

$$\frac{dC_1}{dx} = -\frac{f(x)\varphi_2(x)}{W(x)}, \quad \frac{dC_2}{dx} = \frac{f(x)\varphi_1(x)}{W(x)},$$

where $W(x) = \varphi_1(x)\varphi_2'(x) - \varphi_2(x)\varphi_1'(x)$, the Wronskian of the fundamental system. □

If the two characteristic roots α and β are distinct, then such a u is given by

$$u(x) = \frac{1}{\alpha - \beta}\left(\int_0^t ds f(s)e^{\alpha(t-s)} - \int_0^t ds f(s)e^{\beta(t-s)}\right).$$

A9 Vector Analysis

A9.1 Gradient. Suppose we have a sufficiently smooth function $f : D \to \boldsymbol{R}$, where $D \subset \boldsymbol{R}^2$ is a region. We may imagine that $f(P)$ for $P \in D$ is the altitude of the point P on the island D. Since we assume the landscape to be sufficiently smooth, at each point on D there is a well defined direction $\boldsymbol{n}$ of the steepest ascent and the slope (magnitude) $s(\geq 0)$. That is, at each point on D, we may define the *gradient vector* $s\boldsymbol{n}$, which will be denoted by a vector field $grad\, f$.

A9.2 Coordinate expression of *grad f*. Although *grad f* is meaningful without any specific coordinate system, in actual calculations, introduction of a coordinate system is often useful. Choose a Cartesian coordinate system O-xy. Then the vector has the following representation:

$$grad\, f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right),$$

or

$$grad\, f = \boldsymbol{i}\frac{\partial f}{\partial x} + \boldsymbol{j}\frac{\partial f}{\partial y}. \tag{A9.1}$$

A9.3 Remark. Note that to represent *grad f* in terms of numbers, we need two devices: one is the coordinate system to specify the point in D with two numbers, which allow us to describe f as a function of two independent variables, and two vectors to span the two dimensional vector '*grad f*' at each point on D. In principle any choice is fine, but practically, it is wise to choose these base vectors to be parallel to the coordinate directions at each point. In the choice **A9.2**, the coordinate system has globally the same coordinate direction at every point on D, and the basis vectors are chosen to be parallel to these directions, so again globally uniformly chosen. Nonuniformity in space of representation schemes may cause complications. Especially when we formally use operators as explained below, we must be very careful ($\to$**A9.7**,**A9.9** for a warning).

A9.4 Nabla or del. (A9.1) suggests that *grad* is a map which maps f to the gradient vector at each point in its domain (if f is once partially differentiable). We often write this linear operator as ∇, which is called *nabla*,[485] but is often read 'del' in the US. We write $grad\, f = \nabla f$. ∇ has the following expression if we use the Cartesian coordinates (read [3])

$$\nabla \equiv \sum_{k=1}^{n} \boldsymbol{i}_k \frac{\partial}{\partial x_k}, \tag{A9.2}$$

where x_k is the k-th coordinate and $\boldsymbol{i}_k$ is the unit directional vector in the k-th coordinate direction.

A9.5 Divergence. Suppose we have a flow field (velocity field) $\boldsymbol{u}$ on a domain $D \in \boldsymbol{R}^3$. Let us consider a convex domain[486] $V \subset \boldsymbol{R}^3$ which may be imagined to be covered by area elements $d\boldsymbol{S}$ whose area is $|d\boldsymbol{S}|$, and whose outward normal unit vector is $d\boldsymbol{S}/|d\boldsymbol{S}|$. Then $\boldsymbol{u} \cdot d\boldsymbol{S}$

[485] 'Nabla' is a kind of harp (Assyrian harp).

[486] A set is said to be *convex* if the segment connecting any two points in the set is entirely included in the same set.

is the rate of the volume of fluid going out through the area element in the unit time. Hence the area integral

$$\int_{\partial V} d\boldsymbol{S} \cdot \boldsymbol{u}$$

is the total amount of the volume of the fluid lost from the domain V. The following limit, if exists, is called the *divergence* of the vector field $\boldsymbol{u}$ at point P and is written as $div\, \boldsymbol{u}$:

$$div\, \boldsymbol{u} \equiv \lim_{|V| \to 0} \frac{\int_{\partial V} \boldsymbol{u} \cdot d\boldsymbol{S}}{|V|}, \tag{A9.3}$$

where the limit is taken over a nested sequence of convex volumes converging to a unique point P. Thus its meaning is clear: $div\, \boldsymbol{u}$ is the rate of loss of the quantity carried by the flow field $\boldsymbol{u}$ per unit volume.

A9.6 Cartesian expression of *div.* From (A9.3) assuming the existence of the limit, we may easily derive the Cartesian expression for *div.* Choose as V a tiny cube whose surfaces are perpendicular to the Cartesian coordinates of O-xyz. We immediately get

$$div\, \boldsymbol{u} = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} + \frac{\partial u_z}{\partial z}. \tag{A9.4}$$

A9.7 Operator *div.* (A9.4) again suggests that *div* is an operator which maps a vector field to a scalar field. Comparing (A9.2) and (A9.4) allows us to write

$$div\, \boldsymbol{u} = \nabla \cdot \boldsymbol{u}.$$

This 'abuse' of the symbol is allowed only in the Cartesian coordinates. Generalization to n-space is straightforward.

A9.8 Curl. Let $\boldsymbol{u}$ be a vector field as in **A9.5**. Take a singly connected compact surface S in $\boldsymbol{R}^3$ whose boundary is smooth. The boundary closed curve with the orientation according to the right-hand rule is denoted by ∂S (see Fig.). Consider the following line integral along ∂S:

$$\int_{\partial S} \boldsymbol{u} \cdot d\boldsymbol{l},$$

where $d\boldsymbol{l}$ is the line element along the boundary curve. Ket us imagine a straight vortex line and take S to be a disc perpendicular to the line and its center is on the line. Immediately we see that the integral is the strength of the vortex whose center (singular point) goes through

S. Thus the following limit, if exists, describes the 'area' density of the $\boldsymbol{n}$-component of the vortex (as in the case of angular velocity, the direction of vortex is the direction of the axis of rotation with the right-hand rule):

$$\boldsymbol{n} \cdot curl\, \boldsymbol{u} = \lim_{|S| \to 0} \frac{\int_{\partial S} \boldsymbol{u} \cdot d\boldsymbol{l}}{|S|}, \tag{A9.5}$$

where the limit is over the sequence of smooth surfaces which converges to point P with its orientation in the $\boldsymbol{n}$-direction. If the limit exists, then obviously there is a vector $curl\, \boldsymbol{u}$ called *curl* of the vector filed $\boldsymbol{u}$.

A9.9 Cartesian expression of curl. If we assume the existence of the limit (A9.5), we can easily derive the Cartesian expression for $curl\, \boldsymbol{u}$. We have

$$curl\, \boldsymbol{u} = \left(\frac{\partial u_z}{\partial y} - \frac{\partial u_y}{\partial z}, \frac{\partial u_x}{\partial z} - \frac{\partial u_z}{\partial x}, \frac{\partial u_x}{\partial y} - \frac{\partial u_y}{\partial x} \right), \tag{A9.6}$$

or

$$curl\, \boldsymbol{u} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \partial_x & \partial_y & \partial_z \\ u_x & u_y & u_z \end{vmatrix} = \nabla \times \boldsymbol{u}. \tag{A9.7}$$

This 'abuse' of the nabla symbol is admissible only with the Cartesian coordinates.

A9.10 Potential field, potential, solenoidal field, irrotational field. If a vector field $\boldsymbol{u}$ allows an expression $\boldsymbol{u} = grad\, \phi$, then the field is called a *potential field* and ϕ is called its *potential.* A field without divergence $div\, \boldsymbol{u} = 0$ is called a divergenceless or *solenoidal field.* The field without curl $curl\, \boldsymbol{u} = 0$ is called an *irrotational field.*

A9.11.
(i) $curl\, grad\, \phi = 0$ (Potential fields are irrotational).
(ii) $div\, curl\, \boldsymbol{u} = 0$.
(iii) If a vector field is irrotational on a singly connected domain,[487] then the field is a potential field.
(iv) If a vector field $\boldsymbol{u}$ is solenoidal in a singly connected domain, then there is a vector field $\boldsymbol{A}$ on the domain such that $\boldsymbol{u} = curl\, \boldsymbol{A}$. $\boldsymbol{A}$ is called a *vector potential.*

[487] A domain is *singly connected,* if, for any given pair of points in the domain, any two curves connecting them are homotopic. That is, they can be smoothly deformed into each other without going out of the domain.

A9.12 Theorem [Gauss-Stokes-Green's theorem]. From our definitions of divergence and curl, we have
(i) Gauss' theorem.

$$\int_{\partial V} \boldsymbol{u} \cdot d\boldsymbol{S} = \int_V div\, \boldsymbol{u} d\tau, \tag{A9.8}$$

where V is a domain in the 3-space and $d\tau$ is the volume element.
(ii) Stokes' theorem.

$$\int_{\partial S} \boldsymbol{u} \cdot d\boldsymbol{l} = \int_S curl\, \boldsymbol{u} \cdot d\boldsymbol{S}, \tag{A9.9}$$

where S is a compact surface in 3-space.

A9.13 Laplacian. The operator Δ defined by $\Delta f \equiv div\, grad\, f$ is called the *Laplacian*, and is often written as ∇^2. Δ is defined for a scalar function.

A9.14 Laplacian for vector fields. If we formally calculate *curl curl* $\boldsymbol{u}$ in the Cartesian coordinates, then we have

$$curl\, curl\, \boldsymbol{u} = grad\, div\, \boldsymbol{u} - \nabla^2 \boldsymbol{u}.$$

Since the formal calculation treating ∇ as a vector is legitimate only in the Cartesian coordinate system, this calculation is meaningful only in the Cartesian system. Thus, in particular $\nabla^2 \boldsymbol{u} = (\Delta u_x, \Delta u_y, \Delta u_z)$ is meaningful only in this coordinate system. However, the other two terms are coordinate-free expressions. Hence, we *define* $\Delta \boldsymbol{u}$ as

$$\Delta \boldsymbol{u} \equiv grad\, div\, \boldsymbol{u} - curl\, curl\, \boldsymbol{u}. \tag{A9.10}$$