410 (VI.21) Surfaces

A. The Notion of a Surface

The notion of a surface may be roughly expressed by saying that by moving a curve we get a surface or that the boundary of a solid body is a surface. But these propositions cannot be considered mathematical definitions of a surface. We also make a distinction between surfaces and planes in ordinary language, where we mean by surfaces only those that are not planes. In mathematical language, however, planes are usually included among the surfaces.

A surface can be defined as a 2-dimensional †continuum, in accordance with the definition of a curve as a 1-dimensional continuum. However, while we have a theory of curves based on this definition, we do not have a similar theory of surfaces thus defined (→ 93 Curves).

What is called a surface or a curved surface is usually a 2-dimensional †topological manifold, that is, a topological space that satisfies the †second countability axiom and of which every point has a neighborhood †homeomorphic to the interior of a circular disk in a 2-dimensional Euclidean space. In the following sections, we mean by a surface such a 2-dimensional topological manifold.

B. Examples and Classification

The simplest examples of surfaces are the 2-dimensional †simplex and the 2-dimensional †sphere. Surfaces are generally †simplicially decomposable (or triangulable) and hence homeomorphic to 2-dimensional polyhedra (T. Radó, *Acta Sci. Math. Szeged.* (1925)). A †compact surface is called a **closed surface**, and a noncompact surface is called an **open surface**. A closed surface is decomposable into a finite number of 2-simplexes and so can be interpreted as a †combinatorial manifold. A 2-dimensional topological manifold having a boundary is called a **surface with boundary**. A 2-simplex is an example of a surface with boundary, and a sphere is an example of a closed surface without boundary.

Surfaces are classified as †orientable and †nonorientable. In the special case when a surface is †embedded in a 3-dimensional Euclidean space E^3, whether the surface is orientable or not depends on its having two sides (the "surface" and "back") or only one side. Therefore, in this special case, an orientable surface is called **two-sided**, and a nonorientable surface, **one-sided**. A nonorientable closed surface without boundary cannot be embedded in the Euclidean space E^3 (→ 56 Characteristic Classes, 114 Differential Topology).

The first example of a nonorientable surface (with boundary) is the so-called **Möbius strip** or **Möbius band**, constructed as an †identification space from a rectangle by twisting through 180° and identifying the opposite edges with one another (Fig. 1).

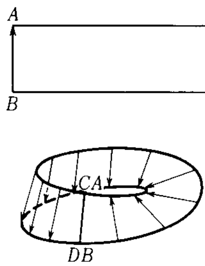


Fig. 1

As illustrated in Fig. 2, from a rectangle $ABCD$ we can obtain a closed surface homeomorphic to the product space $S^1 \times S^1$ by identifying the opposite edges AB with DC and BC with AD. This surface is the so-called 2-dimensional **torus** (or **anchor ring**). In this case, the four vertices A, B, C, D of the rectangle correspond to one point p on the surface, and the pairs of edges AB, DC and BC, AD correspond to closed curves a' and b' on the surface. We use the notation $aba^{-1}b^{-1}$ to represent a torus. This refers to the fact that the torus is obtained from an oriented four-sided polygon by identifying the first side and the third (with reversed orientation), the second side and the fourth (with reversed orientation). Similarly, aa^{-1} represents a sphere (Fig. 3), and $a_1b_1a_1^{-1}b_1^{-1}a_2b_2a_2^{-1}b_2^{-1}$ represents the closed surface shown in Fig. 4.

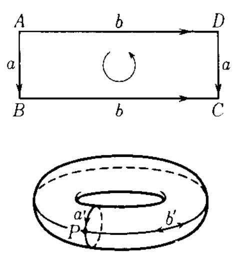


Fig. 2

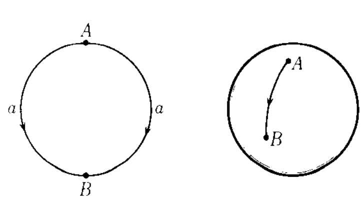


Fig. 3

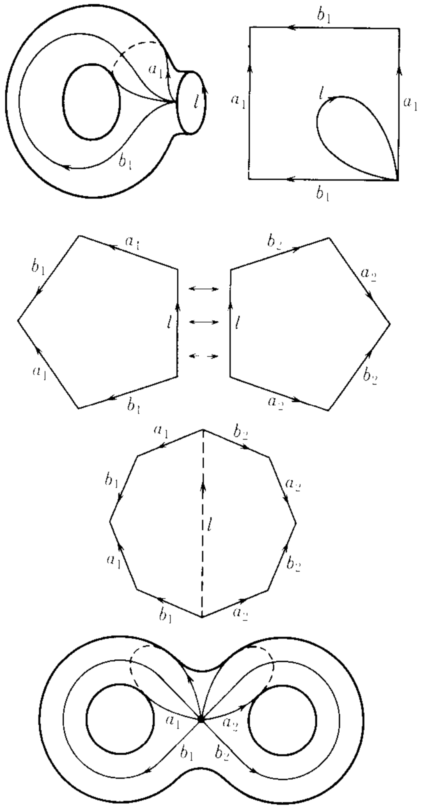


Fig. 4

All closed surfaces without boundary are constructed by identifying suitable pairs of sides of a $2n$-sided polygon in a Euclidean plane E^2. Furthermore, a closed orientable surface without boundary is homeomorphic to the surface represented by aa^{-1} or

$$a_1 b_1 a_1^{-1} b_1^{-1} \dots a_p b_p a_p^{-1} b_p^{-1}. \qquad (1)$$

The 1-dimensional †Betti number of this surface is $2p$, the 0-dimensional and 2-dimensional †Betti numbers are 1, the †torsion coefficients are all 0, and p is called the **genus** of the surface. Also, a closed orientable surface of genus p with boundaries $c_1, \dots, c_k$ is represented by

$$w_1 c_1 w_1^{-1} \dots w_k c_k w_k^{-1} a_1 b_1 a_1^{-1} b_1^{-1} \dots a_p b_p a_p^{-1} b_p^{-1} \qquad (2)$$

(Fig. 5). A closed nonorientable surface without boundary is represented by

$$a_1 a_1 a_2 a_2 \dots a_q a_q. \qquad (3)$$

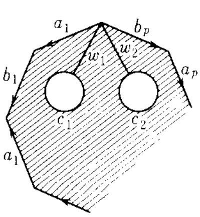


Fig. 5

The 1-dimensional Betti number of this surface is $q-1$, the 0-dimensional and 2-dimensional Betti numbers are 1 and 0, respectively, the 1-dimensional torsion coefficient is 2, the 0-dimensional and 2-dimensional torsion coefficients are 0, and q is called the **genus** of the surface. A closed nonorientable surface of genus q with boundaries $c_1, \dots, c_k$ is represented by

$$w_1 c_1 w_1^{-1} \dots w_k c_k w_k^{-1} a_1 a_1 \dots a_q a_q. \qquad (4)$$

Each of forms (1)–(4) is called the **normal form** of the respective surface, and the curves a_i, b_j, w_k are called the **normal sections** of the surface. To explain the notation in (3), we first take the simplest case, aa. In this case, the surface is obtained from a disk by identifying each pair of points on the circumference that are endpoints of a diameter (Fig. 6). The surface aa is then homeomorphic to a †projective plane of which a decomposition into a complex of triangles is illustrated in Fig. 7. On the other hand, $aabb$ represents a surface like that shown in Fig. 8, called the **Klein bottle**. Fig. 9 shows a **handle**, and Fig. 10 shows a **cross cap**.

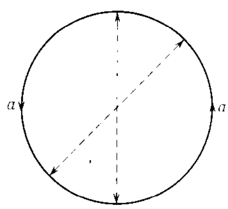


Fig. 6

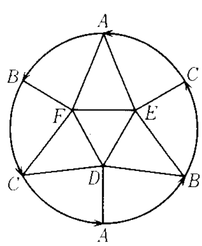


Fig. 7

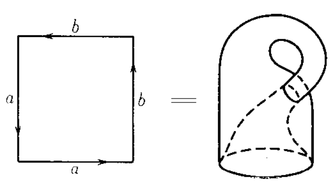


Fig. 8

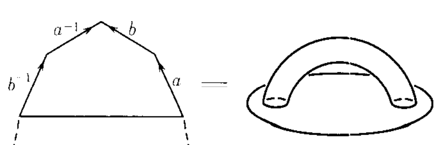


Fig. 9

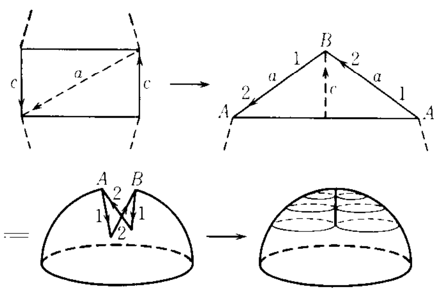


Fig. 10

The last two surfaces have boundaries; a handle is orientable, while a cross cap is nonorientable and homeomorphic to the Möbius strip. If we delete p disks from a sphere and replace them with an equal number of handles, then we obtain a surface homeomorphic to the surface represented in (1), while if we replace the disks by cross caps instead of by handles, then the surface thus obtained is homeomorphic to that represented in (3). Now we decompose the surfaces (1) and (3) into triangles and denote the number of i-dimensional simplexes by α_i $(i=0,1,2)$. Then in view of the †Euler-Poincaré formula, the surfaces (1) and (3) satisfy the respective formulas

$$\alpha_0-\alpha_1+\alpha_2=2(1-p),$$

$$\alpha_0-\alpha_1+\alpha_2=2-q.$$

The †Riemann surfaces of †algebraic functions of one complex variable are always surfaces of type (1), and their genera p coincide with those of algebraic functions.

All closed surfaces are homeomorphic to surfaces of types (1), (2), (3), or (4). A necessary and sufficient condition for two surfaces to be homeomorphic to each other is coincidence of the numbers of their boundaries, their orientability or nonorientability, and their genera (or †Euler characteristic $\alpha^0-\alpha^1+\alpha^2$). This proposition is called the **fundamental theorem of the topology of surfaces**. The †homeomorphism problem of closed surfaces is completely solved by this theorem. The same problem for n $(n\geqslant 3)$ manifolds, even if they are compact, remains open. (For surface area → 246 Length and Area. For the differential geometry of surfaces → 111 Differential Geometry of Curves and Surfaces.)

References

[1] B. Kerékjártó, Vorlesungen über Topologie, Springer, 1923.
[2] H. Seifert and W. Threlfall, Lehrbuch der Topologie, Teubner, 1934 (Chelsea, 1945).
[3] S. Lefschetz, Introduction to topology, Princeton Univ. Press, 1949.
[4] D. Hilbert and S. Cohn-Vossen, Anschauliche Geometrie, Springer, 1932; English translation, Geometry and the imagination, Chelsea, 1952.
[5] W. S. Massey, Algebraic topology: An introduction, Springer, 1967.
[6] E. E. Moise, Geometric topology in dimensions 2 and 3, Springer, 1977.

411 (I.4) Symbolic Logic

A. General Remarks

Symbolic logic (or **mathematical logic**) is a field of logic in which logical inferences commonly used in mathematics are investigated by use of mathematical symbols.

The **algebra of logic** originally set forth by G. Boole [1] and A. de Morgan [2] is actually an algebra of sets or relations; it did not reach the same level as the symbolic logic of today. G. Frege, who dealt not only with the logic of propositions but also with the first-order predicate logic using quantifiers (→ Sections C and K), should be regarded as the real originator of symbolic logic. Frege's work, however, was not recognized for some time. Logical studies by C. S. Peirce, E. Schröder, and G. Peano appeared soon after Frege, but they were limited mostly to propositions and did not develop Frege's work. An essential development of Frege's method was brought about by B. Russell, who, with the collaboration of A. N. Whitehead, summarized his results in *Principia mathematica* [4], which seemed to have completed the theory of symbolic logic at the time of its appearance.

B. Logical Symbols

If A and B are propositions, the propositions (A and B), (A or B), (A implies B), and (not A) are denoted by

$$A\wedge B, \quad A\vee B, \quad A\rightarrow B, \quad \neg A,$$

respectively. We call $\neg A$ the **negation** of A, $A\wedge B$ the **conjunction** (or **logical product**), $A\vee B$ the **disjunction** (or **logical sum**), and $A\rightarrow B$ the **implication** (or B by A). The proposition $(A\rightarrow B)\wedge(B\rightarrow A)$ is denoted by $A\leftrightarrow B$ and is read "A and B are **equivalent**." $A\vee B$ means that at least one of A and B holds. The propositions (For all x, the proposition $F(x)$ holds) and (There exists an x such that $F(x)$ holds) are denoted by $\forall xF(x)$ and $\exists xF(x)$, respectively. A proposition of the form $\forall xF(x)$

is called a **universal proposition**, and one of the form $\exists x F(x)$, an **existential proposition**. The symbols $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\neg$, $\forall$, $\exists$ are called **logical symbols**.

There are various other ways to denote logical symbols, including:

$A \wedge B$: $A \& B$, $A \cdot B$,

$A \vee B$: $A + B$,

$A \rightarrow B$: $A \supset B$, $A \Rightarrow B$,

$A \leftrightarrow B$: $A \rightleftarrows B$, $A \equiv B$, $A \sim B$, $A \supset\subset B$, $A \Leftrightarrow B$,

$\neg A$: $\sim A$, $\bar{A}$,

$\forall x F(x)$: $(x)F(x)$, $\prod x F(x)$, $\bigwedge x F(x)$,

$\exists x F(x)$: $(Ex)F(x)$, $\sum x F(x)$, $\bigvee x F(x)$.

C. Free and Bound Variables

Any function whose values are propositions is called a **propositional function**. $\forall x$ and $\exists x$ can be regarded as operators that transform any propositional function $F(x)$ into the propositions $\forall x F(x)$ and $\exists x F(x)$, respectively. $\forall x$ and $\exists x$ are called **quantifiers**; the former is called the **universal quantifier** and the latter the **existential quantifier**. $F(x)$ is transformed into $\forall x F(x)$ or $\exists x F(x)$ just as a function $f(x)$ is transformed into the definite integral $\int_0^1 f(x)dx$; the resultant propositions $\forall x F(x)$ and $\exists x F(x)$ are no longer functions of x. The variable x in $\forall x F(x)$ and in $\exists x F(x)$ is called a **bound variable**, and the variable x in $F(x)$, when it is not bound by $\forall x$ or $\exists x$, is called a **free variable**. Some people employ different kinds of symbols for free variables and bound variables to avoid confusion.

D. Formal Expressions of Propositions

A formal expression of a proposition in terms of logical symbols is called a **formula**. More precisely, formulas are constructed by the following **formation rules**: (1) If $\mathfrak{A}$ is a formula, $\neg\mathfrak{A}$ is also a formula. If $\mathfrak{A}$ and $\mathfrak{B}$ are formulas, $\mathfrak{A} \wedge \mathfrak{B}$, $\mathfrak{A} \vee \mathfrak{B}$, $\mathfrak{A} \rightarrow \mathfrak{B}$ are all formulas. (2) If $\mathfrak{F}(a)$ is a formula and a is a free variable, then $\forall x \mathfrak{F}(x)$ and $\exists x \mathfrak{F}(x)$ are formulas, where x is an arbitrary bound variable not contained in $\mathfrak{F}(a)$ and $\mathfrak{F}(x)$ is the result of substituting x for a throughout $\mathfrak{F}(a)$.

We use formulas of various scope according to different purposes. To indicate the scope of formulas, we fix a set of formulas, each element of which is called a **prime formula** (or **atomic formula**). The scope of formulas is the set of formulas obtained from the prime formulas by formation rules (1) and (2).

E. Propositional Logic

Propositional logic is the field in symbolic logic in which we study relations between propositions exclusively in connection with the four logical symbols $\wedge$, $\vee$, $\rightarrow$, and $\neg$, called **propositional connectives**.

In propositional logic, we deal only with operations of **logical operators** denoted by propositional connectives, regarding the variables for denoting propositions, called **proposition variables**, only as prime formulas. We examine problems such as: What kinds of formulas are identically true when their proposition variables are replaced by any propositions, and what kinds of formulas can sometimes be true?

Consider the two symbols $\curlyvee$ and $\curlywedge$, read **true** and **false**, respectively, and let $\mathbf{A} = \{\curlyvee, \curlywedge\}$. A univalent function from $\mathbf{A}$, or more generally from a Cartesian product $\mathbf{A} \times \ldots \times \mathbf{A}$, into $\mathbf{A}$ is called a **truth function**. We can regard $\wedge$, $\vee$, $\rightarrow$, $\neg$ as the following truth functions: (1) $A \wedge B = \curlyvee$ for $A = B = \curlyvee$, and $A \wedge B = \curlywedge$ otherwise; (2) $A \vee B = \curlywedge$ for $A = B = \curlywedge$, and $A \vee B = \curlyvee$ otherwise; (3) $A \rightarrow B = \curlywedge$ for $A = \curlyvee$ and $B = \curlywedge$, and $A \rightarrow B = \curlyvee$ otherwise; (4) $\neg A = \curlywedge$ for $A = \curlyvee$, and $\neg A = \curlyvee$ for $A = \curlywedge$.

If we regard proposition variables as variables whose domain is $\mathbf{A}$, then each formula represents a truth function. Conversely, any truth function (of a finite number of independent variables) can be expressed by an appropriate formula, although such a formula is not uniquely determined. If a formula is regarded as a truth function, the value of the function determined by a combination of values of the independent variables involved in the formula is called the **truth value** of the formula.

A formula corresponding to a truth function that takes only $\curlyvee$ as its value is called a **tautology**. For example, $\mathfrak{A} \vee \neg\mathfrak{A}$ and $((\mathfrak{A} \rightarrow \mathfrak{B}) \rightarrow \mathfrak{A}) \rightarrow \mathfrak{A}$ are tautologies. Since a truth function with n independent variables takes values corresponding to 2^n combinations of truth values of its variables, we can determine in a finite number of steps whether a given formula is a tautology. If $\mathfrak{A} \leftrightarrow \mathfrak{B}$ is a tautology (that is, $\mathfrak{A}$ and $\mathfrak{B}$ correspond to the same truth function), then the formulas $\mathfrak{A}$ and $\mathfrak{B}$ are said to be **equivalent**.

F. Propositional Calculus

It is possible to choose some specific tautologies, designate them as axioms, and derive all tautologies from them by appropriately given rules of inference. Such a system is called a **propositional calculus**. There are many ways to stipulate axioms and rules of inference for a propositional calculus.

The abovementioned propositional calculus corresponds to the so-called classical propositional logic (→ Section L). By choosing appropriate axioms and rules of inference we can also formally construct intuitionistic or other propositional logics. In intuitionistic logic the law of the †excluded middle is not accepted, and hence it is impossible to formalize intuitionistic propositional logic by the notion of tautology. We therefore usually adopt the method of propositional calculus, instead of using the notion of tautology, to formalize intuitionistic propositional logic. For example, V. I. Glivenko's theorem [5], that if a formula $\mathfrak{A}$ can be proved in classical logic, then $\neg\neg\mathfrak{A}$ can be proved in intuitionistic logic, was obtained by such formalistic considerations. A method of extending the classical concepts of truth value and tautology to intuitionistic and other logics has been obtained by S. A. Kripke. There are also studies of logics intermediate between intuitionistic and classical logic (T. Umezawa).

G. Predicate Logic

Predicate logic is the area of symbolic logic in which we take quantifiers in account. Mainly propositional functions are discussed in predicate logic. In the strict sense only single-variable propositional functions are called **predicates**, but the phrase **predicate of** n **arguments** (or n**-ary predicate**) denoting an n-variable propositional function is also employed. Single-variable (or unary) predicates are also called **properties**. We say that a has the property F if the proposition $F(a)$ formed by the property F is true. Predicates of two arguments are called **binary relations**. The proposition $R(a,b)$ formed by the binary relation R is occasionally expressed in the form aRb. Generally, predicates of n arguments are called n**-ary relations**. The domain of definition of a unary predicate is called the **object domain**, elements of the object domain are called **objects**, and any variable running over the object domain is called an **object variable**. We assume here that the object domain is not empty. When we deal with a number of predicates simultaneously (with different numbers of variables), it is usual to arrange things so that all the independent variables have the same object domain by suitably extending their object domains.

Predicate logic in its purest sense deals exclusively with the general properties of quantifiers in connection with propositional connectives. The only objects dealt with in this field are **predicate variables** defined over a certain common domain and object variables running over the domain. Propositional variables are regarded as predicates of no variables. Each expression $F(a_1, \ldots, a_n)$ for any predicate variable F of n variables $a_1, \ldots, a_n$ (object variables designated as free) is regarded as a prime formula ($n = 0, 1, 2, \ldots$), and we deal exclusively with formulas generated by these prime formulas, where bound variables are also restricted to object variables that have a common domain. We give no specification for the range of objects except that it be the common domain of the object variables.

By designating an object domain and substituting a predicate defined over the domain for each predicate variable in a formula, we obtain a proposition. By substituting further an object (object constant) belonging to the object domain for each object variable in a proposition, we obtain a proposition having a definite truth value. When we designate an object domain and further associate with each predicate variable as well as with each object variable a predicate or an object to be substituted for it, we call the pair consisting of the object domain and the association a **model**. Any formula that is true for every model is called an **identically true formula** or **valid formula**. The study of identically true formulas is one of the most important problems in predicate logic.

H. Formal Representations of Mathematical Propositions

To obtain a formal representation of a mathematical theory by predicate logic, we must first specify its object domain, which is a nonempty set whose elements are called **individuals**; accordingly the object domain is called the **individual domain**, and object variables are called **individual variables**. Secondly we must specify **individual symbols**, **function symbols**, and **predicate symbols**, signifying specific individuals, functions, and †predicates, respectively. Here a function of n arguments is a univalent mapping from the Cartesian product $D \times \ldots \times D$ of n copies of the given set to D. Then we define the notion of **term** as in the next paragraph to represent each individual formally. Finally we express propositions formally by formulas.

Definition of terms (formation rule for terms): (1) Each individual symbol is a term. (2) Each free variable is a term. (3) $f(t_1, \ldots, t_n)$ is a term if $t_1, \ldots, t_n$ are terms and f is a function symbol of n arguments. (4) The only terms are those given by (1)–(3).

As a prime formula in this case we use any formula of the form $F(t_1, \dots, t_n)$, where F is a predicate symbol of n arguments and $t_1, \dots, t_n$ are arbitrary terms. To define the notions of term and formula, we need logical symbols, free and bound individual variables, and also a list of individual symbols, function symbols, and predicate symbols.

In pure predicate logic, the individual domain is not concrete, and we study only general forms of propositions. Hence, in this case, predicate or function symbols are not representations of concrete predicates or functions but are **predicate variables** and **function variables**. We also use free individual variables instead of individual symbols. In fact, it is now most common that function variables are dispensed with, and only free individual variables are used as terms.

I. Formulation of Mathematical Theories

To formalize a theory we need **axioms** and **rules of inference**. Axioms constitute a certain specific set of formulas, and a rule of inference is a rule for deducing a formula from other formulas. A formula is said to be **provable** if it can be deduced from the axioms by repeated application of rules of inference. Axioms are divided into two types: **logical axioms**, which are common to all theories, and **mathematical axioms**, which are peculiar to each individual theory. The set of mathematical axioms is called the **axiom system** of the theory.

(I) Logical axioms: (1) A formula that is the result of substituting arbitrary formulas for the proposition variables in a tautology is an axiom. (2) Any formula of the form

$$\forall x\mathfrak{F}(x)\rightarrow\mathfrak{F}(t) \quad \text{or} \quad \mathfrak{F}(t)\rightarrow\exists x\mathfrak{F}(x)$$

is an axiom, where $\mathfrak{F}(t)$ is the result of substituting an arbitrary term t for x in $\mathfrak{F}(x)$.

(II) Rules of inference: (1) We can deduce a formula $\mathfrak{B}$ from two formulas $\mathfrak{A}$ and $\mathfrak{A}\rightarrow\mathfrak{B}$ (**modus ponens**). (2) We can deduce $\mathfrak{A}\rightarrow\forall x\mathfrak{F}(x)$ from a formula $\mathfrak{A}\rightarrow\mathfrak{F}(a)$ and $\exists x\mathfrak{F}(x)\rightarrow\mathfrak{A}$ from $\mathfrak{F}(a)\rightarrow\mathfrak{A}$, where a is a free individual variable contained in neither $\mathfrak{A}$ nor $\mathfrak{F}(x)$ and $\mathfrak{F}(a)$ is the result of substituting a for x in $\mathfrak{F}(x)$.

If an axiom system is added to these logical axioms and rules of inference, we say that a **formal system** is given.

A formal system S or its axiom system is said to be **contradictory** or to contain a **contradiction** if a formula $\mathfrak{A}$ and its negation $\neg\mathfrak{A}$ are provable; otherwise it is said to be **consistent**. Since

$$(\mathfrak{A}\wedge\neg\mathfrak{A})\rightarrow\mathfrak{B}$$

is a tautology, we can show that any formula is provable in a formal system containing a contradiction. The validity of a proof by **reductio ad absurdum** lies in the fact that

$$(\mathfrak{A}\rightarrow(\mathfrak{B}\wedge\neg\mathfrak{B}))\rightarrow\neg\mathfrak{A}$$

is a tautology. An affirmative proposition (formula) may be obtained by reductio ad absurdum since the formula (of propositional logic) representing the **discharge of double negation**

$$\neg\neg\mathfrak{A}\rightarrow\mathfrak{A}$$

is a tautology.

J. Predicate Calculus

If a formula has no free individual variable, we call it a **closed formula**. Now we consider a formal system S whose mathematical axioms are closed. A formula $\mathfrak{A}$ is provable in S if and only if there exist suitable mathematical axioms $\mathfrak{C}_1, \dots, \mathfrak{C}_n$ such that the formula

$$(\mathfrak{C}_1\wedge\dots\wedge\mathfrak{C}_n)\rightarrow\mathfrak{A}$$

is provable without the use of mathematical axioms. Since any axiom system can be replaced by an equivalent axiom system containing only closed formulas, the study of a formal system can be reduced to the study of pure logic.

In the following we take no individual symbols or function symbols into consideration and we use predicate variables as predicate symbols in accordance with the commonly accepted method of stating properties of the pure predicate logic; but only in the case of **predicate logic with equality** will we use predicate variables and the equality predicate $=$ as a predicate symbol. However, we can safely state that we use function variables as function symbols.

The formal system with no mathematical axioms is called the **predicate calculus**. The formal system whose mathematical axioms are the equality axioms

$$a=a, \quad a=b \rightarrow (\mathfrak{F}(a)\rightarrow\mathfrak{F}(b))$$

is called the **predicate calculus with equality**.

In the following, by being provable we mean being provable in the predicate calculus.

(1) Every provable formula is valid.

(2) Conversely, any valid formula is provable (K. Gödel [6]). This fact is called the **completeness** of the predicate calculus. In fact, by Gödel's proof, a formula $\mathfrak{A}$ is provable if $\mathfrak{A}$ is always true in every interpretation whose individual domain is of †countable cardinality. In another formulation, if $\neg\mathfrak{A}$ is not provable, the formula $\mathfrak{A}$ is a true proposition in some interpretation (and the individual domain in this case is of countable cardinality). We can

extend this result as follows: If an axiom system generated by countably many closed formulas is consistent, then its mathematical axioms can be considered true propositions by a common interpretation. In this sense, **Gödel's completeness theorem** gives another proof of the †Skolem-Löwenheim theorem.

(3) The predicate calculus is consistent. Although this result is obtained from (1) in this section, it is not difficult to show it directly (D. Hilbert and W. Ackermann [7]).

(4) There are many different ways of giving logical axioms and rules of inference for the predicate calculus. G. Gentzen gave two types of systems in [8]; one is a natural deduction system in which it is easy to reproduce formal proofs directly from practical ones in mathematics, and the other has a logically simpler structure. Concerning the latter, Gentzen proved **Gentzen's fundamental theorem**, which shows that a formal proof of a formula may be translated into a "direct" proof. The theorem itself and its idea were powerful tools for obtaining consistency proofs.

(5) If the proposition $\exists x\mathfrak{A}(x)$ is true, we choose one of the individuals x satisfying the condition $\mathfrak{A}(x)$, and denote it by $\varepsilon x\mathfrak{A}(x)$. When $\exists x\mathfrak{A}(x)$ is false, we let $\varepsilon x\mathfrak{A}(x)$ represent an arbitrary individual. Then

$$\exists x\mathfrak{A}(x)\rightarrow\mathfrak{A}(\varepsilon x\mathfrak{A}(x)) \qquad (1)$$

is true. We consider εx to be an operator associating an individual $\varepsilon x\mathfrak{A}(x)$ with a proposition $\mathfrak{A}(x)$ containing the variable x. Hilbert called it the **transfinite logical choice function**; today we call it **Hilbert's ε-operator** (or **ε-quantifier**), and the logical symbol ε used in this sense **Hilbert's ε-symbol**. Using the ε-symbol, $\exists x\mathfrak{A}(x)$ and $\forall x\mathfrak{A}(x)$ are represented by

$$\mathfrak{A}(\varepsilon x\mathfrak{A}(x)), \quad \mathfrak{A}(\varepsilon x \neg \mathfrak{A}(x)),$$

respectively, for any $\mathfrak{A}(x)$. The system of predicate calculus adding formulas of the form (1) as axioms is essentially equivalent to the usual predicate calculus. This result, called the **ε-theorem**, reads as follows: When a formula $\mathfrak{C}$ is provable under the assumption that every formula of the form (1) is an axiom, we can prove $\mathfrak{C}$ using no axioms of the form (1) if $\mathfrak{C}$ contains no logical symbol ε (D. Hilbert and P. Bernays [9]). Moreover, a similar theorem holds when axioms of the form

$$\forall x(\mathfrak{A}(x)\leftrightarrow\mathfrak{B}(x))\rightarrow\varepsilon x\mathfrak{A}(x)=\varepsilon x\mathfrak{B}(x) \qquad (2)$$

are added (S. Maehara [10]).

(6) For a given formula $\mathfrak{A}$, call $\mathfrak{A}'$ a normal form of $\mathfrak{A}$ when the formula

$$\mathfrak{A}\leftrightarrow\mathfrak{A}'$$

is provable and $\mathfrak{A}'$ satisfies a particular condition. For example, for any formula $\mathfrak{A}$ there is a normal form $\mathfrak{A}'$ satisfying the condition: $\mathfrak{A}'$ has the form

$$Q_1x_1 \ldots Q_nx_n\mathfrak{B}(x_1, \ldots, x_n),$$

where Qx means a quantifier $\forall x$ or $\exists x$, and $\mathfrak{B}(x_1, \ldots, x_n)$ contains no quantifier and has no predicate variables or free individual variables not contained in $\mathfrak{A}$. A normal form of this kind is called a **prenex normal form**.

(7) We have dealt with the classical first-order predicate logic until now. For other predicate logics (→ Sections K and L) also, we can consider a predicate calculus or a formal system by first defining suitable axioms or rules of inference. Gentzen's fundamental theorem applies to the intuitionistic predicate calculus formulated by V. I. Glivenko, A. Heyting, and others. Since Gentzen's fundamental theorem holds not only in classical logic and intuitionistic logic but also in several systems of first-order predicate logic or propositional logic, it is useful for getting results in modal and other logics (M. Ohnishi, K. Matsumoto). Moreover, Glivenko's theorem in propositional logic [5] is also extended to predicate calculus by using a rather weak representation (S. Kuroda [12]). G. Takeuti expected that a theorem similar to Gentzen's fundamental theorem would hold in higher-order predicate logic also, and showed that the consistency of analysis would follow if that conjecture could be verified [13]. Moreover, in many important cases, he showed constructively that the conjecture holds partially. The conjecture was finally proved by M. Takahashi [14] by a nonconstructive method. Concerning this, there are also contributions by S. Maehara, T. Simauti, M. Yasuhara, and W. Tait.

K. Predicate Logics of Higher Order

In ordinary predicate logic, the bound variables are restricted to individual variables. In this sense, ordinary predicate logic is called **first-order predicate logic**, while predicate logic dealing with quantifiers $\forall P$ or $\exists P$ for a predicate variable P is called **second-order predicate logic**.

Generalizing further, we can introduce the so-called **third-order predicate logic**. First we fix the individual domain D_0. Then, by introducing the whole class D_1^n of predicates of n variables, each running over the object domain D_0, we can introduce predicates that have D_1^n as their object domain. This kind of predicate is called a **second-order predicate** with respect to the individual domain D_0. Even when we restrict second-order predicates to one-variable predicates, they are divided into various types, and the domains of independent variables do not coincide in the case of more than two variables. In contrast, predicates having D_0 as their object domain are called **first-order predicates**. The logic having quantifiers that admit first-order predicate variables is second-order predicate logic, and the logic having quantifiers that admit up to second-order predicate variables is third-order predicate logic. Similarly, we can define further **higher-order predicate logics**.

Higher-order predicate logic is occasionally called **type theory**, because variables arise that are classified into various types. Type theory is divided into **simple type theory** and **ramified type theory**.

We confine ourselves to variables for single-variable predicates, and denote by P such a bound predicate variable. Then for any formula $\mathfrak{F}(a)$ (with a a free individual variable), the formula

$$\exists P \forall x (P(x) \leftrightarrow \mathfrak{F}(x))$$

is considered identically true. This is the point of view in simple type theory.

Russell asserted first that this formula cannot be used reasonably if quantifiers with respect to predicate variables occur in $\mathfrak{F}(x)$. This assertion is based on the point of view that the formula in the previous paragraph asserts that $\mathfrak{F}(x)$ is a first-order predicate, whereas any quantifier with respect to first-order predicate variables, whose definition assumes the totality of the first-order predicates, should not be used to introduce the first-order predicate $\mathfrak{F}(x)$. For this purpose, Russell further classified the class of first-order predicates by their **rank** and adopted the axiom

$$\exists P^k \forall x (P^k(x) \leftrightarrow \mathfrak{F}(x))$$

for the predicate variable P^k of rank k, where the rank i of any free predicate variable occurring in $\mathfrak{F}(x)$ is $\leqslant k$, and the rank j of any bound predicate variable occurring in $\mathfrak{F}(x)$ is $< k$. This is the point of view in ramified type theory, and we still must subdivide the types if we deal with higher-order propositions or propositions of many variables. Even Russell, having started from his ramified type theory, had to introduce the **axiom of reducibility** afterwards and reduce his theory to simple type theory.

L. Systems of Logic

Logic in the ordinary sense, which is based on the **law of the excluded middle** asserting that every proposition is in principle either true or false, is called **classical logic**. Usually, propositional logic, predicate logic, and type theory are developed from the standpoint of classical logic. Occasionally the reasoning of intuitionistic mathematics is investigated using symbolic logic, in which the law of the excluded middle is not admitted (→ 156 Foundations of Mathematics). Such logic is called **intuitionistic logic**. Logic is also subdivided into propositional logic, predicate logic, etc., according to the extent of the propositions (formulas) dealt with.

To express **modal propositions** stating **possibility, necessity**, etc., in symbolic logic, J. Łukaszewicz proposed a propositional logic called **three-valued logic**, having a third truth value, neither true nor false. More generally, **many-valued logics** with any number of truth values have been introduced; classical logic is one of its special cases, **two-valued logic** with two truth values, true and false. Actually, however, many-valued logics with more than three truth values have not been studied much, while various studies in **modal logic** based on classical logic have been successfully carried out. For example, studies of **strict implication** belong to this field.

References

[1] G. Boole, An investigation of the laws of thought, Walton and Maberly, 1854.
[2] A. de Morgan, Formal logic, or the calculus of inference, Taylor and Walton, 1847.
[3] G. Frege, Begriffsschrift, eine der arithmetischen nachgebildete Formalsprache des reinen Denkens, Halle, 1879.
[4] A. N. Whitehead and B. Russell, Principia mathematica I, II, III, Cambridge Univ. Press, 1910–1913; second edition, 1925–1927.
[5] V. Glivenko, Sur quelques points de la logique de M. Brouwer, Acad. Roy. de Belgique, Bulletin de la classe des sciences, (5) 15 (1929), 183–188.
[6] K. Gödel, Die Vollständigkeit der Axiome des logischen Funktionenkalküls, Monatsh. Math. Phys., 37 (1930), 349–360.
[7] D. Hilbert and W. Ackermann, Grundzüge der theoretischen Logik, Springer, 1928, sixth edition, 1972; English translation, Principles of mathematical logic, Chelsea, 1950.
[8] G. Gentzen, Untersuchungen über das logische Schliessen, Math. Z., 39 (1935), 176–210, 405–431.
[9] D. Hilbert and P. Bernays, Grundlagen der Mathematik II, Springer, 1939; second edition, 1970.
[10] S. Maehara, Equality axiom on Hilbert's ε-symbol, J. Fac. Sci. Univ. Tokyo, (I), 7 (1957), 419–435.

[11] A. Heyting, Die formalen Regeln der intuitionistischen Logik I, S.-B. Preuss. Akad. Wiss., 1930, 42–56.
[12] S. Kuroda, Intuitionistische Untersuchungen der formalistischen Logik, Nagoya Math. J., 2 (1951), 35–47.
[13] G. Takeuti, On a generalized logic calculus, Japan. J. Math., 23 (1953), 39–96.
[14] M. Takahashi, A proof of the cut-elimination theorem in simple type-theory, J. Math. Soc. Japan, 19 (1967), 399–410.
[15] S. C. Kleene, Mathematical logic, Wiley, 1967.
[16] J. R. Shoenfield, Mathematical logic, Addison-Wesley, 1967.
[17] R. M. Smullyan, First-order logic, Springer, 1968.

412 (IV.13) Symmetric Riemannian Spaces and Real Forms

A. Symmetric Riemannian Spaces

Let M be a †Riemannian space. For each point p of M we can define a mapping σ_p of a suitable neighborhood U_p of p onto U_p itself so that $\sigma_p(x_t)=x_{-t}$, where x_t ($|t|<\varepsilon, x_0=p$) is any †geodesic passing through the point p. We call M a **locally symmetric Riemannian space** if for any point p of M we can choose a neighborhood U_p so that σ_p is an †isometry of U_p. In order that a Riemannian space M be locally symmetric it is necessary and sufficient that the †covariant differential (with respect to the †Riemannian connection) of the †curvature tensor of M be 0. A locally symmetric Riemannian space is a †real analytic manifold. We say that a Riemannian space M is a **globally symmetric Riemannian space** (or simply **symmetric Riemannian space**) if M is connected and if for each point p of M there exists an isometry σ_p of M onto M itself that has p as an isolated fixed point (i.e., has no fixed point except p in a certain neighborhood of p) and such that σ_p^2 is the identity transformation on M. In this case σ_p is called the **symmetry** at p. A (globally) symmetric Riemannian space is locally symmetric and is a †complete Riemannian space. Conversely, a †simply connected complete locally symmetric Riemannian space is a (globally) symmetric Riemannian space.

B. Symmetric Riemannian Homogeneous Spaces

A †homogeneous space G/K of a connected †Lie group G is a **symmetric homogeneous space** (with respect to θ) if there exists an **involutive automorphism** (i.e., automorphism of order 2) θ of G satisfying the condition $K_\theta^0\subset K\subset K_\theta$, where K_θ is the closed subgroup consisting of all elements of G left fixed by θ and K_θ^0 is the connected component of the identity element of K_θ. In this case, the mapping $aK\rightarrow\theta(a)K$ ($a\in G$) is a transformation of G/K having the point K as an isolated fixed point; more generally, the mapping $\theta_{a_0}: aK\rightarrow a_0\theta(a_0)^{-1}\theta(a)K$ is a transformation of G/K that has an arbitrary given point a_0K of G/K as an isolated fixed point. If there exists a G-invariant Riemannian metric on G/K, then G/K is a symmetric Riemannian space with symmetries $\{\theta_{a_0}\,|\,a_0\in G\}$ and is called a **symmetric Riemannian homogeneous space**. A sufficient condition for a symmetric homogeneous space G/K to be a symmetric Riemannian homogeneous space is that K be a compact subgroup. Conversely, given a symmetric Riemannian space M, let G be the connected component of the identity element of the Lie group formed by all the isometries of M; then M is represented as the symmetric Riemannian homogeneous space $M=G/K$ and K is a compact group. In particular, a symmetric Riemannian space can be regarded as a Riemannian space that is realizable as a symmetric Riemannian homogeneous space.

The Riemannian connection of a symmetric Riemannian homogeneous space G/K is uniquely determined (independent of the choice of G-invariant Riemannian metric), and a geodesic $x_t(|t|<\infty, x_0=a_0K)$ passing through a point a_0K of G/K is of the form $x_t=(\exp tX)a_0K$. Here X is any element of the Lie algebra $\mathfrak{g}$ of G such that $\theta(X)=-X$, where θ also denotes the automorphism of $\mathfrak{g}$ induced by the automorphism θ of G and $\exp tX$ is the †one-parameter subgroup of G defined by the element X. The covariant differential of any G-invariant tensor field on G/K is 0, and any G-invariant †differential form on G/K is a closed differential form.

C. Classification of Symmetric Riemannian Spaces

The †simply connected †covering Riemannian space of a symmetric Riemannian space is also a symmetric Riemannian space. Therefore the problem of classifying symmetric Riemannian spaces is reduced to classifying simply connected symmetric Riemannian spaces M and determining †discontinuous groups of isometries of M. When we take the †de Rham decomposition of such a space M and represent M as the product of a real Euclidean space and a number of simply connected irreducible Riemannian spaces, all the factors are symmetric Riemannian spaces. We say that M is an **irreducible symmetric Riemannian space** if it is a symmetric Riemannian space and is irreducible as a Riemannian space.

A simply connected irreducible symmetric Riemannian space is isomorphic to one of the following four types of symmetric Riemannian homogeneous spaces (here Lie groups are always assumed to be connected):

(1) The symmetric Riemannian homogeneous space $(G \times G)/\{(a,a) \mid a \in G\}$ of the direct product $G \times G$, where G is a simply connected compact †simple Lie group and the involutive automorphism of $G \times G$ is given by $(a,b) \rightarrow (b,a)$ $((a,b) \in G \times G)$. This space is isomorphic, as a Riemannian space, to the space G obtained by introducing a two-sided invariant Riemannian metric on the group G; the isomorphism is induced from the mapping $G \times G \ni (a,b) \rightarrow ab^{-1} \in G$.

(2) A symmetric homogeneous space G/K_θ of a simply connected compact simple Lie group G with respect to an involutive automorphism θ of G. In this case, the closed subgroup $K_\theta = \{a \in G \mid \theta(a) = a\}$ of G is connected. We assume here that θ is a member of the given complete system of representatives of the †conjugate classes formed by the elements of order 2 in the automorphism group of the group G.

(3) The homogeneous space $G^{\mathbf{C}}/G$, where $G^{\mathbf{C}}$ is a complex simple Lie group whose †center reduces to the identity element and G is an arbitrary but fixed maximal compact subgroup of $G^{\mathbf{C}}$.

(4) The homogeneous space G_0/K, where G_0 is a noncompact simple Lie group whose center reduces to the identity element and which has no complex Lie group structure, and K is a maximal compact subgroup of G. In Section D we shall see that (3) and (4) are actually symmetric homogeneous spaces. All four types of symmetric Riemannian spaces are actually irreducible symmetric Riemannian spaces, and G-invariant Riemannian metrics on each of them are uniquely determined up to multiplication by a positive number. On the other hand, (1) and (2) are compact, while (3) and (4) are homeomorphic to Euclidean spaces and not compact. For spaces of types (1) and (3) the problem of classifying simply connected irreducible symmetric Riemannian spaces is reduced to classifying †compact real simple Lie algebras and †complex simple Lie algebras, respectively, while for types (2) and (4) it is reduced to the classification of noncompact real simple Lie algebras (→ Section D) (for the result of classification of these types → Appendix A, Table 5.II). On the other hand, any (not necessarily simply connected) irreducible symmetric Riemannian space defines one of (1)–(4) as its †universal covering manifold; if the covering manifold is of type (3) or (4), the original symmetric Riemannian space is necessarily simply connected.

D. Symmetric Riemannian Homogeneous Spaces of Semisimple Lie Groups

In Section C we saw that any irreducible symmetric Riemannian space is representable as a symmetric Riemannian homogeneous space G/K on which a connected semisimple Lie group G acts †almost effectively (→ 249 Lie Groups). Among symmetric Riemannian spaces, such a space $M = G/K$ is characterized as one admitting no nonzero vector field that is †parallel with respect to the Riemannian connection. Furthermore, if G acts effectively on M, G coincides with the connected component $I(M)^0$ of the identity element of the Lie group formed by all the isometries of M.

We let $M = G/K$ be a symmetric Riemannian homogeneous space on which a connected semisimple Lie group G acts almost effectively. Then G is a Lie group that is †locally isomorphic to the group $I(M)^0$, and therefore the Lie algebra of G is determined by M. Let $\mathfrak{g}$ be the Lie algebra of G, $\mathfrak{k}$ be the subalgebra of $\mathfrak{g}$ corresponding to K, and θ be the involutive automorphism of G defining the symmetric homogeneous space G/K. The automorphism of $\mathfrak{g}$ defined by θ is also denoted by θ. Then $\mathfrak{k} = \{X \in \mathfrak{g} \mid \theta(X) = X\}$. Putting $\mathfrak{m} = \{X \in \mathfrak{g} \mid \theta(X) = -X\}$, we have $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$ (direct sum of linear spaces), and $\mathfrak{m}$ can be identified in a natural way with the tangent space at the point K of G/K. The †adjoint representation of G gives rise to a representation of K in $\mathfrak{g}$, which induces a linear representation $\mathrm{Ad}_{\mathfrak{m}}(k)$ of K in $\mathfrak{m}$. Then $\{\mathrm{Ad}_{\mathfrak{m}}(k) \mid k \in K\}$ coincides with the †restricted homogeneous holonomy group at the point K of the Riemannian space G/K.

Now let φ be the †Killing form of $\mathfrak{g}$. Then $\mathfrak{k}$ and $\mathfrak{m}$ are mutually orthogonal with respect to φ, and denoting by $\varphi_{\mathfrak{k}}$ and $\varphi_{\mathfrak{m}}$ the restrictions of φ to $\mathfrak{k}$ and $\mathfrak{m}$, respectively, $\varphi_{\mathfrak{k}}$ is a negative definite quadratic form on $\mathfrak{k}$. If $\varphi_{\mathfrak{m}}$ is also a negative definite quadratic form on $\mathfrak{m}$, $\mathfrak{g}$ is a compact real semisimple Lie algebra and G/K is a compact symmetric Riemannian space; in this case we say that G/K is of **compact type**. In the opposite case, where $\varphi_{\mathfrak{m}}$ is a †positive definite quadratic form, G/K is said to be of **noncompact type**. In this latter case, G/K is homeomorphic to a Euclidean space, and if the center of G is finite, K is a maximal compact subgroup of G. Furthermore, the group of isometries $I(G/K)$ of G/K is canonically isomorphic to the automorphism group of the Lie algebra $\mathfrak{g}$. When G/K is of compact type (noncompact type), there exists one and only one G-invariant Riemannian metric on G/K, which induces in the tangent space $\mathfrak{m}$ at the point K the positive definite inner product $-\varphi_{\mathfrak{m}}(\varphi_{\mathfrak{m}})$.

A symmetric Riemannian homogeneous space G/K_θ of compact type defined by a simply connected compact semisimple Lie group G with respect to an involutive automorphism θ is simply connected. Let $\mathfrak{g}=\mathfrak{m}+\mathfrak{k}_\theta$ be the decomposition of the Lie algebra $\mathfrak{g}$ of G with respect to the automorphism θ of $\mathfrak{g}$, and let $\mathfrak{g}^{\mathbf{C}}$ be the †complex form of $\mathfrak{g}$. Then the real subspace $\mathfrak{g}_\theta=\sqrt{-1}\,\mathfrak{m}+\mathfrak{k}_\theta$ in $\mathfrak{g}^{\mathbf{C}}$ is a real semisimple Lie algebra and a †real form of $\mathfrak{g}^{\mathbf{C}}$. Let G_θ be the Lie group corresponding to the Lie algebra $\mathfrak{g}_\theta$ with center reduced to the identity element, and let K be the subgroup of G_θ corresponding to $\mathfrak{k}_\theta$. Then we get a (simply connected) symmetric Riemannian homogeneous space of noncompact type G_θ/K.

When we start from a symmetric Riemannian space of noncompact type G/K instead of the symmetric Riemannian space of compact type G/K_θ and apply the same process as in the previous paragraphs, taking a simply connected G_θ as the Lie group corresponding to $\mathfrak{g}_\theta$, we obtain a simply connected symmetric Riemannian homogeneous space of compact type. Indeed, each of these two processes is the reverse of the other, and in this way we get a one-to-one correspondence between simply connected symmetric Riemannian homogeneous spaces of compact type and those of noncompact type. This relationship is called **duality** for symmetric Riemannian spaces; when two symmetric Riemannian spaces are related by duality, each is said to be the **dual** of the other.

If one of the two symmetric Riemannian spaces related by duality is irreducible, the other is also irreducible. The duality holds between spaces of types (1) and (3) and between those of types (2) and (4) described in Section C. This fact is based on the following theorem in the theory of Lie algebras, where we identify isomorphic Lie algebras. (i) Complex simple Lie algebras $\mathfrak{g}^{\mathbf{C}}$ and compact real simple Lie algebras $\mathfrak{g}$ are in one-to-one correspondence by the relation that $\mathfrak{g}^{\mathbf{C}}$ is the complex form of $\mathfrak{g}$. (ii) Form the Lie algebra $\mathfrak{g}_\theta$ in the above way from a compact real simple Lie algebra $\mathfrak{g}$ and an involutive automorphism θ of $\mathfrak{g}$. We assume that θ is a member of the given complete system of representatives of conjugate classes of involutive automorphisms in the automorphism group of $\mathfrak{g}$. Then we get from the pair $(\mathfrak{g},\theta)$ a noncompact real simple Lie algebra $\mathfrak{g}_\theta$, and any noncompact real simple Lie algebra is obtained by this process in one and only one way.

Consider a Riemannian space given as a symmetric Riemannian homogeneous space $M=G/K$ with a semisimple Lie group G, and let K be the †sectional curvature of M. Then if M is of compact type the value of K is $\geqslant 0$, and if M is of noncompact type it is $\leqslant 0$. On the other hand, the **rank** of M is the (unique) dimension of a commutative subalgebra of $\mathfrak{g}$ that is contained in and maximal in $\mathfrak{m}$. (For results concerning the group of isometries of M, distribution of geodesics on M, etc. → [3].)

E. Symmetric Hermitian Spaces

A connected †complex manifold M with a †Hermitian metric is called a **symmetric Hermitian space** if for each point p of M there exists an isometric and †biholomorphic transformation of M onto M that is of order 2 and has p as an isolated fixed point. As a real analytic manifold, such a space M is a symmetric Riemannian space of even dimension, and the Hermitian metric of M is a †Kähler metric. Let $I(M)$ be the (not necessarily connected) Lie group formed by all isometries of M, and let $A(M)$ be the subgroup consisting of all holomorphic transformations in $I(M)$. Then $A(M)$ is a closed Lie subgroup of $I(M)$. Let G be the connected component $A(M)^0$ of the identity element of $A(M)$. Then G acts transitively on M, and M is expressed as a symmetric Riemannian homogeneous space G/K.

Under the de Rham decomposition of a simply connected symmetric Hermitian space (regarded as a Riemannian space), all the factors are symmetric Hermitian spaces. The factor that is isomorphic to a real Euclidean spaces as a Riemannian space is a symmetric Hermitian space that is isomorphic to the complex Euclidean space $\mathbf{C}^n$. A symmetric Hermitian space defining an irreducible symmetric Riemannian space is called an **irreducible symmetric Hermitian space**. The problem of classifying symmetric Hermitian spaces is thus reduced to classifying irreducible symmetric Hermitian spaces.

In general, if the symmetric Riemannian space defined by a symmetric Hermitian space M is represented as a symmetric Riemannian homogeneous space G/K by a connected semisimple Lie group G acting effectively on M, then M is simply connected, G coincides with the group $A(M)^0$ introduced in the previous paragraph, and the center of K is not a †discrete set. In particular, an irreducible symmetric Hermitian space is simply connected. Moreover, in order for an irreducible symmetric Riemannian homogeneous space G/K to be defined by an irreducible symmetric Hermitian space M, it is necessary and sufficient that the center of K not be a discrete set. If G acts effectively on M, then G is a simple Lie group whose center is reduced to the identity element, and the center of K is of dimension 1. For a space G/K satisfying these conditions, there are two kinds of structures of symmetric Hermitian spaces defining the Riemannian structure of G/K.

As follows from the classification of irreducible symmetric Riemannian spaces, an irreducible Hermitian space defines one of the following symmetric Riemannian homogeneous spaces, and conversely, each of these homogeneous spaces is defined by one of the two kinds of symmetric Hermitian spaces.

(I) The symmetric homogeneous space G/K of a compact simple Lie group G with respect to an involutive automorphism θ such that the center of G reduces to the identity element and the center of K is not a discrete set. Here θ may be assumed to be a representative of a conjugate class of involutive automorphisms in the automorphism group of G.

(II) The homogeneous space G_0/K of a noncompact simple Lie group G_0 by a maximal compact subgroup K such that the center of G_0 reduces to the identity element and the center of K is not a discrete set.

An irreducible symmetric Hermitian space of type (I) is compact and is isomorphic to a †rational algebraic variety. An irreducible symmetric Hermitian space of type (II) is homeomorphic to a Euclidean space and is isomorphic (as a complex manifold) to a bounded domain in $\mathbf{C}^n$ (Section F).

By the same principle as for irreducible symmetric Riemannian spaces, a duality holds for irreducible symmetric Hermitian spaces which establishes a one-to-one correspondence between the spaces of types (I) and (II). Furthermore, an irreducible symmetric Hermitian space M_b of type (II) that is dual to a given irreducible symmetric Hermitian space $M_a = G/K$ of type (I) can be realized as an open complex submanifold of M_a in the following way. Let $G^{\mathbf{C}}$ be the connected component of the identity element in the Lie group formed by all the holomorphic transformations of M_a. Then $G^{\mathbf{C}}$ is a complex simple Lie group containing G as a maximal compact subgroup, and the complex Lie algebra $\mathfrak{g}^{\mathbf{C}}$ of $G^{\mathbf{C}}$ contains the Lie algebra $\mathfrak{g}$ of G as a real form. Let θ be the involutive automorphism of G defining the symmetric homogeneous space G/K, and let $\mathfrak{g} = \mathfrak{m} + \mathfrak{k}$ be the decomposition of $\mathfrak{g}$ determined by θ. We denote by G_0 the real subgroup of $G^{\mathbf{C}}$ corresponding to the real form $\mathfrak{g}_0 = \sqrt{-1}\,\mathfrak{m} + \mathfrak{k}$ of $\mathfrak{g}^{\mathbf{C}}$. Then G_0 (i) is a closed subgroup of $G^{\mathbf{C}}$ whose center reduces to the identity element and (ii) contains K as a maximal compact subgroup. By definition the space M_b is then given by G_0/K. Now the group G_0 acts on M_a as a subgroup of $G^{\mathbf{C}}$, and the orbit of G_0 containing the point K of M_a is an open complex submanifold that is isomorphic to M_b (as a complex manifold). M_a regarded as a complex manifold can be represented as the homogeneous space $G^{\mathbf{C}}/U$ of the complex simple Lie group $G^{\mathbf{C}}$.

F. Symmetric Bounded Domains

We denote by D a bounded domain in the complex Euclidean space $\mathbf{C}^n$ of dimension n. We call D a **symmetric bounded domain** if for each point of D there exists a holomorphic transformation of order 2 of D onto D having the point as an isolated fixed point. On the other hand, the group of all holomorphic transformations of D is a Lie group, and D is called a **homogeneous bounded domain** if this group acts transitively on D. A symmetric bounded domain is a homogeneous bounded domain. The following theorem gives more precise results: On a bounded domain D, †Bergman's kernel function defines a Kähler metric that is invariant under all holomorphic transformations of D. If D is a symmetric bounded domain, D is a symmetric Hermitian space with respect to this metric, and its defining Riemannian space is a symmetric Riemannian homogeneous space of noncompact type G/K with semisimple Lie group G. Conversely, any symmetric Hermitian space of noncompact type is isomorphic (as a complex manifold) to a symmetric bounded domain. When D is isomorphic to an irreducible symmetric Hermitian space, we call D an **irreducible symmetric bounded domain**. A symmetric bounded domain is simply connected and can be decomposed into the direct product of irreducible symmetric bounded domains.

The connected component of the identity element of the group of all holomorphic transformations of a symmetric bounded domain D is a semisimple Lie group that acts transitively on D. Conversely, D is a symmetric bounded domain if a connected semisimple Lie group, or more generally, a connected Lie group admitting a two-sided invariant †Haar measure, acts transitively on D. Homogeneous bounded domains in $\mathbf{C}^n$ are symmetric bounded domains if $n \leqslant 3$ but not necessarily when $n \geqslant 4$.

G. Examples of Irreducible Symmetric Riemannian Spaces

Here we list irreducible symmetric Riemannian spaces of types (2) and (4) (→ Section C) that can be represented as homogeneous spaces of classical groups, using the notation introduced by E. Cartan. We denote by $M_u = G/K$ a simply connected irreducible symmetric Riemannian space of type (2), where G is a group that acts almost effectively on M_u and K is the subgroup given by $K = K_\theta^0$ for an involutive automorphism θ of G. For such an M_u, the space of type (4) that is dual to M_u is denoted by $M_\theta = G_\theta/K$. Clearly $\dim M_u = \dim M_\theta$. (For the dimension and rank of M_u and for those M_u that are represented as homogeneous spaces of simply connected †exceptional compact simple Lie groups → Appendix A, Table 5.III.) In this section (and also in Appendix A, Table 5.III), $O(n)$, $U(n)$, $Sp(n)$, $SL(n, \mathbf{R})$, and $SL(n, \mathbf{C})$ are the †orthogonal group of degree n, the †unitary group of degree n, the †symplectic group of degree $2n$, and the real and complex †special linear groups of degree n, respectively. Let $SO(n) = SL(n, \mathbf{R}) \cap O(n)$ and $SU(n) = SL(n, \mathbf{C}) \cap U(n)$. We put

$$I_{p,q} = \begin{pmatrix} -I_p & 0 \\ 0 & I_q \end{pmatrix}, \quad J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

$$K_{p,q} = \begin{pmatrix} -I_p & 0 & 0 & 0 \\ 0 & I_q & 0 & 0 \\ 0 & 0 & -I_p & 0 \\ 0 & 0 & 0 & I_q \end{pmatrix},$$

where I_p is the $p \times p$ unit matrix.

Type AI. $M_u = SU(n)/SO(n)$ $(n > 1)$, where $\theta(s) = \bar{s}$ (with $\bar{s}$ the complex conjugate matrix of s). $M_\theta = SL(n, \mathbf{R})/SO(n)$.

Type AII. $M_u = SU(2n)/Sp(n)$ $(n > 1)$, where $\theta(s) = J_n s J_n^{-1}$. $M_\theta = SU^*(2n)/Sp(n)$. Here $SU^*(2n)$ is the subgroup of $SL(2n, \mathbf{C})$ formed by the matrices that commute with the transformation $(z_1, \dots, z_n, z_{n+1}, \dots, z_{2n}) \to (\bar{z}_{n+1}, \dots, \bar{z}_{2n}, -\bar{z}_1, \dots, -\bar{z}_n)$ in $\mathbf{C}^n$; $SU^*(2n)$ is called the **quaternion unimodular group** and is isomorphic to the commutator group of the group of all regular transformations in an n-dimensional vector space over the quaternion field $\mathbf{H}$.

Type AIII. $M_u = SU(p+q)/S(U_p \times U_q)$ $(p \geqslant q \geqslant 1)$, where $S(U_p \times U_q) = SU(p+q) \cap (U(p) \times U(q))$, with $U(p) \times U(q)$ being canonically identified with a subgroup of $U(p+q)$, and $\theta(s) = I_{p,q} s I_{p,q}$. This space M_u is a †complex Grassmann manifold. $M_\theta = SU(p,q)/S(U_p \times U_q)$, where $SU(p,q)$ is the subgroup of $SL(p+q, \mathbf{C})$ consisting of matrices that leave invariant the Hermitian form $z_1\bar{z}_1 + \dots + z_p\bar{z}_p - z_{p+1}\bar{z}_{p+1} - \dots - z_{p+q}\bar{z}_{p+q}$.

Type AIV. This is the case $q = 1$ of type AIII. M_u is the $(n-1)$-dimensional complex projective space, and M_θ is called a **Hermitian hyperbolic space**.

Type BDI. $M_u = SO(p+q)/SO(p) \times SO(q)$ $(p \geqslant q \geqslant 1, p > 1, p+q \neq 4)$, where $\theta(s) = I_{p,q} s I_{p,q}$. M_u is the †real Grassmann manifold formed by the oriented p-dimensional subspaces in $\mathbf{R}^{p+q}$. $M_\theta = SO_0(p,q)/SO(p) \times SO(q)$, where $SO(p,q)$ is the subgroup of $SL(n, \mathbf{R})$ consisting of matrices that leave invariant the quadratic form $x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_{p+q}^2$, and $SO_0(p,q)$ is the connected component of the identity element.

Type BDII. This is the case $q = 1$ of type BDI. M_u is the $(n-1)$-dimensional sphere, and M_θ is called a **real hyperbolic space**.

Type DIII. $M_u = SO(2n)/U(n)$ $(n > 2)$, where $U(n)$ is regarded as a subgroup of $SO(2n)$ by identifying $s \in U(n)$ with

$$\begin{pmatrix} \operatorname{Re} s & \operatorname{Im} s \\ -\operatorname{Im} s & \operatorname{Re} s \end{pmatrix} \in SO(2n),$$

and $\theta(s) = J_n s J_n^{-1}$. $M_\theta = SO^*(2n)/U(n)$. Here $SO^*(2n)$ denotes the group of all complex orthogonal matrices of determinant 1 leaving invariant the skew-Hermitian form $z_1\bar{z}_{n+1} - z_{n+1}\bar{z}_1 + z_2\bar{z}_{n+2} - z_{n+2}\bar{z}_2 + \dots + z_n\bar{z}_{2n} - z_{2n}\bar{z}_n$; this group is isomorphic to the group of all linear transformations leaving invariant a nondegenerate skew-Hermitian form in an n-dimensional vector space over the quaternion field $\mathbf{H}$.

Type CI. $M_u = Sp(n)/U(n)$ $(n \geqslant 1)$, where $U(n)$ is considered as a subgroup of $Sp(n)$ by the identification $U(n) \subset SO(2n)$ explained in type DIII and $\theta(s) = \bar{s} (= J_n s J_n^{-1})$. $M_\theta = Sp(n, \mathbf{R})/U(n)$, where $Sp(n, \mathbf{R})$ is the real symplectic group of degree $2n$.

Type CII. $M_u = Sp(p+q)/Sp(p) \times Sp(q)$ $(p \geqslant q \geqslant 1)$, where $Sp(p) \times Sp(q)$ is identified with a subgroup of $Sp(p+q)$ by the mapping

$$\left(\begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix}, \begin{pmatrix} A_2 & B_2 \\ C_2 & D_2 \end{pmatrix} \right) \to \begin{pmatrix} A_1 & 0 & B_1 & 0 \\ 0 & A_2 & 0 & B_2 \\ C_1 & 0 & D_1 & 0 \\ 0 & C_2 & 0 & D_2 \end{pmatrix}$$

and $\theta(s) = K_{p,q} s K_{p,q}$. $M_\theta = Sp(p,q)/Sp(p) \times Sp(q)$. Here $Sp(p,q)$ is the group of complex symplectic matrices of degree $2(p+q)$ leaving invariant the Hermitian form $(z_1, \dots, z_{p+q}) K_{p,q} {}^t(\bar{z}_1, \dots, \bar{z}_{p+q})$; this group is interpreted as the group of all linear transformations leaving invariant a nondegenerate Hermitian form of index p in a $(p+q)$-dimensional vector space over the quaternion field $\mathbf{H}$. For $q = 1$, M_u is the quaternion projective space, and M_θ is called the **quaternion hyperbolic space**.

Among the spaces introduced here, there are some with lower p, q, n that coincide (as Riemannian spaces) (→ Appendix A, Table 5.III).

H. Space Forms

A Riemannian manifold of †constant curvature is called a **space form**; it is said to be **spherical**,

Euclidean, or **hyperbolic** according as the constant curvature K is positive, zero, or negative. A space form is a locally symmetric Riemannian space; a simply connected complete space form is a sphere if $K>0$, a real Euclidean space if $K=0$, and a real hyperbolic space if $K<0$. More generally, a complete spherical space form of even dimension is a sphere or a projective space, and one of odd dimension is an orientable manifold. A complete 2-dimensional Euclidean space form is one of the following spaces: Euclidean plane, cylinder, torus, †Möbius strip, †Klein bottle. Except for these five spaces and the 2-dimensional sphere, any †closed surface is a 2-dimensional hyperbolic space form (for details about space forms → [6]).

I. Examples of Irreducible Symmetric Bounded Domains

Among the irreducible symmetric Riemannian spaces described in Section H, those defined by irreducible symmetric Hermitian spaces are of types AIII, DIII, BDI ($q=2$), and CI. We list the irreducible symmetric bounded domains that are isomorphic to the irreducible Hermitian spaces defining these spaces. Positive definiteness of a matrix will be written $\gg 0$.

Type $\mathbf{I}_{m,m'}$ ($m'\geqslant m\geqslant 1$). The set of all $m\times m'$ complex matrices Z satisfying the condition $I_{m'}-{}^t\bar{Z}Z\gg 0$ is a symmetric bounded domain in $\mathbf{C}^{mm'}$, which is isomorphic (as a complex manifold) to the irreducible symmetric Hermitian space defined by M_θ of type AIII ($p=m$, $q=m'$).

Type $\mathbf{II}_m$ ($m\geqslant 2$). The set of all $m\times m$ complex †skew-symmetric matrices Z satisfying the condition $I_m-{}^t\bar{Z}Z\gg 0$ is a symmetric bounded domain in $\mathbf{C}^{m(m-1)/2}$ corresponding to the type DIII ($n=m$).

Type $\mathbf{III}_m$ ($m\geqslant 1$). The set of all $m\times m$ complex symmetric matrices satisfying the condition $I_m-{}^t\bar{Z}Z\gg 0$ is a symmetric bounded domain in $\mathbf{C}^{m(m+1)/2}$ corresponding to the type CI ($n=m$). This bounded domain is holomorphically isomorphic to the †Siegel upper half-space of degree m.

Type $\mathbf{IV}_m$ ($m\geqslant 1, m\neq 2$). This bounded domain in $\mathbf{C}^m$ is formed by the elements $(z_1,\ldots,z_m)$ satisfying the condition $|z_1|^2+\ldots+|z_m|^2<(1+|z_1^2+\ldots+z_m^2|)/2<1$, and corresponds to the type BDI ($p=m, q=2$).

Among these four types of bounded domains, the following complex analytic isomorphisms hold: $\mathrm{I}_{1,1}\cong\mathrm{II}_2\cong\mathrm{III}_1\cong\mathrm{IV}_1$, $\mathrm{II}_3\cong\mathrm{I}_{1,3}$, $\mathrm{IV}_3\cong\mathrm{III}_2$, $\mathrm{IV}_4\cong\mathrm{I}_{2,2}$, $\mathrm{IV}_6\cong\mathrm{II}_4$. (For details about these symmetric bounded domains → [2].) There are two more kinds of irreducible symmetric bounded domains, which are represented as homogeneous spaces of exceptional Lie groups.

J. Weakly Symmetric Riemannian Spaces

A generalization of symmetric Riemannian space is the notion of weakly symmetric Riemannian space introduced by Selberg. Let M be a Riemannian space. M is called a **weakly symmetric Riemannian space** if a Lie subgroup G of the group of isometries $I(M)$ acts transitively on M and there exists an element $\mu\in I(M)$ satisfying the relations (i) $\mu G\mu^{-1}=G$; (ii) $\mu^2\in G$; and (iii) for any two points x, y of M, there exists an element m of G such that $\mu x=my$, $\mu y=mx$. A symmetric Riemannian space M becomes a weakly symmetric Riemannian space if we put $G=I(M)$ and $\mu=$ the identity transformation; as the element m in condition (iii) we can take the symmetry σ_p at the midpoint p on the geodesic arc joining x and y. There are, however, weakly symmetric Riemannian spaces that do not have the structure of a symmetric Riemannian space. An example of such a space is given by $M=G=SL(2,\mathbf{R})$ with a suitable Riemannian metric, where μ is the inner automorphism defined by

$$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

(Selberg [4]). On a weakly symmetric Riemannian space, the ring of all G-invariant differential-integral operators is commutative; this fact is useful in the theory of spherical functions (→ 437 Unitary Representations).

References

[1] S. Helgason, Differential geometry, Lie groups, and symmetric spaces, Academic Press, 1978.
[2] C. L. Siegel, Analytic functions of several complex variables, Princeton Univ. Press, 1950.
[3] J.-L. Koszul, Exposés sur les espaces homogènes symétriques, São Paulo, 1959.
[4] A. Selberg, Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with application to Dirichlet series, J. Indian Math. Soc., 20 (1956), 47–87.
[5] E. Cartan, Sur certaines formes riemanniennes remarquables des géométries à groupe fondamental simple, Ann. Sci. Ecole Norm. Sup., 44 (1927), 345–467 (Oeuvres complètes, Gauthier-Villars, 1952, pt. I, vol. 2, 867–989).
[6] E. Cartan, Sur les domaines bornés homogènes de l'espace de n-variables complexes, Abh. Math. Sem. Univ. Hamburg, 11 (1936), 116–162 (Oeuvres complètes, Gauthier-Villars, 1952, pt. I, vol. 2, 1259–1305).

[7] J. A. Wolf, Spaces of constant curvature, McGraw-Hill, 1967.
[8] S. Kobayashi and K. Nomizu, Foundations of differential geometry II, Interscience, 1969.
[9] O. Loos, Symmetric spaces. I, General theory; II, Compact spaces and classification, Benjamin, 1969.

413 (VII.7) Symmetric Spaces

A †Riemannian manifold M is called a **symmetric Riemannian space** if M is connected and if for each $p \in M$ there exists an involutive †isometry σ_p of M that has p as an isolated fixed point. For the classification and the group-theoretic properties of symmetric Riemannian spaces → 412 Symmetric Riemannian Spaces and Real Forms. We state here the geometrical properties of a symmetric Riemannian space M. Let M be represented by G/K, a †symmetric Riemannian homogeneous space. The †Lie algebras of G and K are denoted by $\mathfrak{g}$ and $\mathfrak{k}$ respectively. Let us denote by τ_a the †left translation of M defined by $a \in G$, and by X^* the vector field on M generated by $X \in \mathfrak{g}$. We denote by θ the differential of the involutive automorphism of G defining G/K and identify the subspace $\mathfrak{m} = \{X \in \mathfrak{g} \mid \theta(X) = -X\}$ of $\mathfrak{g}$ with the tangent space $T_o(M)$ of M at the origin $o = K$ of M. The †representation of $\mathfrak{k}$ on $\mathfrak{m}$ induced from the †adjoint representation of $\mathfrak{g}$ is denoted by $\mathrm{ad}_{\mathfrak{m}}$.

A. Riemannian Connections

M is a complete real analytic †homogeneous Riemannian manifold. If M is a †symmetric Hermitian space, it is a †homogeneous Kählerian manifold. The †Riemannian connection ∇ of M is the †canonical connection of the homogeneous space G/K and satisfies $\nabla_Y X^* = [X, Y]$ $(Y \in \mathfrak{m})$ for each $X \in \mathfrak{k}$ and $\nabla_Y X^* = 0$ $(Y \in \mathfrak{m})$ for each $X \in \mathfrak{m}$. For each $X \in \mathfrak{m}$, the curve γ_X of M defined by $\gamma_X(t) = (\exp tX)o$ $(t \in \mathbf{R})$ is a †geodesic of M such that $\gamma_X(0) = o$ and $\dot{\gamma}_X(0) = X$. In particular, the †exponential mapping Exp_o at o is given by $\mathrm{Exp}_o X = (\exp X)o$ $(X \in \mathfrak{m})$. For each $X \in \mathfrak{m}$, the †parallel translation along the geodesic arc $\gamma_X(t)$ $(0 \leqslant t \leqslant t_0)$ coincides with the differential of $\tau_{\exp t_0 X}$. If M is compact, for each $p \in M$ there exists a smooth simply closed geodesic passing through p. Any G-invariant tensor field on M is †parallel with respect to ∇. Any G-invariant †differential form on M is closed. The Lie algebra $\mathfrak{h}$ of the †restricted homogeneous holonomy group of M at o coincides with $\mathrm{ad}_{\mathfrak{m}}[\mathfrak{m}, \mathfrak{m}]$. If the group $I(M)$ of all isometries of M is †semisimple, one has $\mathfrak{h} = \{A \in \mathfrak{gl}(\mathfrak{m}) \mid A \cdot g_o = 0, A \cdot R_o = 0\} = \mathrm{ad}_{\mathfrak{m}}\mathfrak{k}$. Here, g_o and R_o denote the values at o of the Riemannian metric g and the †Riemannian curvature R of M, respectively, and $A \cdot$ is the natural action of A on the tensors over $\mathfrak{m}$. If, moreover, M is a symmetric Hermitian space, the value J_0 at o of the †almost complex structure J of M belongs to the center of $\mathfrak{h}$. In general, $\mathfrak{h} = \{0\}$ if and only if M is †flat, and $\mathfrak{h}$ has no nonzero invariant on $\mathfrak{m}$ if and only if $I(M)$ is semisimple.

B. Riemannian Curvature Tensors

The Riemannian curvature tensor R of M is parallel and satisfies $R_0(X, Y) = -\mathrm{ad}_{\mathfrak{m}}[X, Y]$ $(X, Y \in \mathfrak{m})$. Assume that $\dim M \geqslant 2$ in the following. Let P be a 2-dimensional subspace of $\mathfrak{m}$, and $\{X, Y\}$ an orthonormal basis of P with respect to g_o. Then the †sectional curvature $K(P)$ of P is given by $K(P) = g_o([[X, Y], X], Y)$. $K = 0$ everywhere if and only if M is flat. If M is of †compact type (resp. of †noncompact type), then $K \geqslant 0$ (resp. $K \leqslant 0$) everywhere. $K > 0$ (resp. $K < 0$) everywhere if and only if the †rank of M is 1 and M is of compact type (resp. of noncompact type). For any four points p, q, p', q' of a manifold M of any of these types satisfying $d(p, q) = d(p', q')$, d being the †Riemannian distance of M, there exists a $\phi \in I(M)$ such that $\phi(p) = p'$ and $\phi(q) = q'$. Other than the aforementioned M's, the only Riemannian manifolds having this property are circles and Euclidean spaces. If $K > 0$ everywhere, any geodesic of M is a smooth simply closed curve and all geodesics are of the same length. For a symmetric Hermitian space M, the †holomorphic sectional curvature H satisfies $H = 0$ (resp. $H > 0$, $H < 0$) everywhere if and only if M is flat (resp. of compact type, of noncompact type).

C. Ricci Tensors

The †Ricci tensor S of M is parallel. If $\varphi_{\mathfrak{m}}$ denotes the restriction to $\mathfrak{m} \times \mathfrak{m}$ of the †Killing form φ of $\mathfrak{g}$, the value S_o of S at o satisfies $S_o = -\frac{1}{2}\varphi_{\mathfrak{m}}$. If M is †irreducible, it is an †Einstein space. $S = 0$ (resp. positive definite, negative definite, nondegenerate) everywhere if and only if M is flat (resp. M is of compact type, M is of noncompact type, $I(M)$ is semisimple). If M is a †symmetric bounded domain and g is the †Bergman metric of M, one has $S = -g$.

D. Symmetric Riemannian Spaces of Noncompact Type

Let M be of noncompact type. For each $p \in M$, p is the only fixed point of the †symmetry σ_p, and the exponential mapping at p is a diffeomorphism from $T_p(M)$ to M. In particular, M is diffeomorphic to a Euclidean space. For each pair $p, q \in M$, a geodesic arc joining p and q is unique up to parametrization. For each $p \in M$ there exists neither a †conjugate point nor a †cut point of p. If M is a symmetric Hermitian space, that is, if it is a symmetric bounded domain, then it is a †Stein manifold and holomorphically homeomorphic to a †Siegel domain.

E. Groups of Isometries

The isotropy subgroup at o in $I(M)$ is denoted by $I_o(M)$. Then the smooth mapping $I_o(M) \times \mathfrak{m} \rightarrow I(M)$ defined by the correspondence $\phi \times X \mapsto \phi\tau_{\exp X}$ is surjective, and it is a diffeomorphism if M is of noncompact type. If M is of noncompact type, $I(M)$ is isomorphic to the group $A(\mathfrak{g})$ of all automorphisms of $\mathfrak{g}$ in a natural way, and $I_o(M)$ is isomorphic to the subgroup $A(\mathfrak{g}, \mathfrak{k}) = \{\phi \in A(\mathfrak{g}) \mid \phi(\mathfrak{k}) = \mathfrak{k}\}$ of $A(\mathfrak{g})$, provided that G acts almost effectively on M. Moreover, in this case the center of the identity component $I(M)^0$ of $I(M)$ reduces to the identity, and the isotropy subgroup at a point in $I(M)^0$ is a maximal compact subgroup of $I(M)^0$. If $I(M)$ is semisimple, any element of $I(M)^0$ may be represented as a product of an even number of symmetries of M. In the following, let M be a symmetric Hermitian space, and denote by $A(M)$ (resp. $H(M)$) the group of all holomorphic isometries (resp. all holomorphic homeomorphisms) of M, and by $A(M)^0$ and $H(M)^0$ their identity components. All these groups act transitively on M. If M is compact or if $I(M)$ is semisimple, one has $A(M)^0 = I(M)^0$. If $I(M)$ is semisimple, M is simply connected and the center of $I(M)^0$ reduces to the identity. If M is of compact type, M is a †rational †projective algebraic manifold, and $H(M)^0$ is a complex semisimple Lie group whose center reduces to the identity, and it is the †complexification of $I(M)^0$. In this case, the isotropy subgroup at a point in $H(M)^0$ is a †parabolic subgroup of $H(M)^0$. If M is of noncompact type, one has $H(M)^0 = I(M)^0$. In the following we assume that G is compact.

F. Cartan Subalgebras

A maximal Abelian †Lie subalgebra in $\mathfrak{m}$ is called a **Cartan subalgebra** for M. Cartan subalgebras are conjugate to each other under the †adjoint action of K. Fix a Cartan subalgebra $\mathfrak{a}$ and introduce an inner product $(\ ,\ )$ on $\mathfrak{a}$ by the restriction to $\mathfrak{a} \times \mathfrak{a}$ of g_o. For an element α of the dual space $\mathfrak{a}^*$ of $\mathfrak{a}$, we put $\mathfrak{m}_\alpha = \{X \in \mathfrak{m} \mid [H, [H, X]] = -\alpha(H)^2 X$ for any $H \in \mathfrak{a}\}$. The subset $\Sigma = \{\alpha \in \mathfrak{a}^* - \{0\} \mid \mathfrak{m}_\alpha \neq \{0\}\}$ of $\mathfrak{a}^*$ is called the **root system** of M (relative to $\mathfrak{a}$). We write $m_\alpha = \dim \mathfrak{m}_\alpha$ for $\alpha \in \Sigma$. The subset $D = \{H \in \mathfrak{a} \mid \alpha(H) \in \pi\mathbf{Z}$ for some $\alpha \in \Sigma\}$ of $\mathfrak{a}$ is called the **diagram** of M. A connected component of $\mathfrak{a} - D$ is called a **fundamental cell** of M. The quotient group W of the normalizer of $\mathfrak{a}$ in K modulo the centralizer of $\mathfrak{a}$ in K is called the **Weyl group** of M. W is identified with a finite group of orthogonal transformations of $\mathfrak{a}$.

G. Conjugate Points

For a geodesic arc γ with the initial point o, any †Jacobi field along γ that vanishes at o and the end point of γ is obtained as the restriction to γ of the vector field X^* generated by an element $X \in \mathfrak{k}$. For $H \in \mathfrak{a} - \{0\}$, $\mathrm{Exp}_o H$ is a conjugate point to o along the geodesic γ_H if and only if $\alpha(H) \in \pi\mathbf{Z} - \{0\}$ for some $\alpha \in \Sigma$. In this case, the multiplicity of the conjugate point $\mathrm{Exp}_o H$ is equal to $\frac{1}{2}\sum_{\alpha \in \Sigma, \alpha(H) \in \pi\mathbf{Z} - \{0\}} M_\alpha$. From this fact and Morse theory (→ 279 Morse Theory), we get a †cellular decomposition of the †loop space of M. The set of all points conjugate to o coincides with $K\,\mathrm{Exp}_o D$ and is stratified to a disjoint union of a finite number of connected regular submanifolds with dimension $\leqslant \dim M - 2$.

H. Cut Points

We define a †lattice group Γ of $\mathfrak{a}$ by $\Gamma = \{A \in \mathfrak{a} \mid \mathrm{Exp}_o A = o\}$, and put $C_\mathfrak{a} = \{H \in \mathfrak{a} \mid \mathrm{Max}_{A \in \Gamma - \{0\}} 2(H, A)/(A, A) = 1\}$. Then, for $H \in \mathfrak{a} - \{0\}$, $\mathrm{Exp}_o H$ is a cut point of o along the geodesic γ_H if and only if $H \in C_\mathfrak{a}$. The set C_o of all cut points of o coincides with $K\,\mathrm{Exp}_o C_\mathfrak{a}$ and is stratified to a disjoint union of a finite number of connected regular submanifolds with dimension $\leqslant \dim M - 1$. The set of all points †first conjugate to o coincides with C_o if and only if M is simply connected.

I. Fundamental Groups

Let Γ_0 denote the subgroup of $\mathfrak{a}$ generated by $\{(2\pi/(\alpha, \alpha))\alpha \mid \alpha \in \Sigma\}$, identifying $\mathfrak{a}^*$ with $\mathfrak{a}$ by means of the inner product $(\ ,\ )$ of $\mathfrak{a}$. This is a subgroup of Γ. We regard Γ as a subgroup of the group $I(\mathfrak{a})$ of all motions of $\mathfrak{a}$ by parallel translations. The subgroup $\tilde{W} = W\Gamma$ of $I(\mathfrak{a})$ generated by Γ and the Weyl group W is called the **affine Weyl group** of M. $\tilde{W}$ leaves the diagram D invariant and acts transitively on the set of all fundamental cells of M. Take a fundamental cell σ such that its closure $\bar{\sigma}$ contains 0, and put $\tilde{W}_\sigma = \{w \in \tilde{W} \mid w(\sigma) = \sigma\}$. Then the fundamental group $\pi_1(M)$ of M is an †Abelian group isomorphic to the groups $\tilde{W}_\sigma$ and Γ/Γ_0. $\pi_1(M)$ is a finite group if and only if M is of compact type. In this case, the order of $\pi_1(M)$ is equal to the cardinality of the set $\Gamma \cap \bar{\sigma}$ as well as to the index $[\Gamma : \Gamma_0]$. Moreover, if we denote by $\tilde{W}_\sigma^*$ the group $\tilde{W}_\sigma$ for the symmetric Riemannian space $M^* = G^*/K^*$ defined by the †adjoint group G^* of G and $K^* = \{a \in G^* \mid a\theta = \theta a\}$, then $\tilde{W}_\sigma$ is isomorphic to a subgroup of $\tilde{W}_\sigma^*$. If M is irreducible, $\tilde{W}_\sigma^*$ is isomorphic to a subgroup of the group of all automorphisms of the †extended Dynkin diagram of the root system Σ.

J. Cohomology Rings

Let $P(\mathfrak{g})$ (resp. $P(\mathfrak{k})$) be the †graded linear space of all †primitive elements in the †cohomology algebra $H(\mathfrak{g})$ of $\mathfrak{g}$ (resp. $H(\mathfrak{k})$ of $\mathfrak{k}$), and $P(\mathfrak{g}, \mathfrak{k})$ the intersection of $P(\mathfrak{g})$ with the image of the natural homomorphism $H(\mathfrak{g}, \mathfrak{k}) \to H(\mathfrak{g})$, where $H(\mathfrak{g}, \mathfrak{k})$ denotes the relative cohomology algebra for the pair $(\mathfrak{g}, \mathfrak{k})$. Then one has $\dim P(\mathfrak{g}, \mathfrak{k}) + \dim P(\mathfrak{k}) = \dim P(\mathfrak{g})$. Denote by $\Lambda P(\mathfrak{g}, \mathfrak{k})$ the exterior algebra over $P(\mathfrak{g}, \mathfrak{k})$. The †graded algebra of all G-invariant polynomials on $\mathfrak{g}$ (resp. all K-invariant polynomials on $\mathfrak{k}$) is denoted by $I(G)$ (resp. $I(K)$), where the degree of a homogeneous polynomial with degree p is defined to be $2p$. We denote by $I^+(G)$ the ideal of $I(G)$ consisting of all $f \in I(G)$ such that $f(0) = 0$, and regard $I(K)$ as an $I^+(G)$-module through the restriction homomorphism. Then the †real cohomology ring $H(M)$ of M is isomorphic to the tensor product $\Lambda P(\mathfrak{g}, \mathfrak{k}) \otimes (I(K)/I^+(G)I(K))$. If K is connected and the †Poincaré polynomials of $P(\mathfrak{g})$, $P(\mathfrak{k})$, and $P(\mathfrak{g}, \mathfrak{k})$ are $\sum_{i=1}^r t^{2m_i-1}$, $\sum_{i=1}^s t^{2n_i-1}$, and $\sum_{i=s+1}^r t^{2m_i-1}$, respectively, then the Poincaré polynomial of $H(M)$ is given by $\prod_{i=s+1}^r (1 + t^{2m_i-1}) \prod_{i=1}^s (1 - t^{2m_i}) \prod_{i=1}^s (1 - t^{2n_i})^{-1}$.

References

[1] E. Cartan, Sur certaines formes riemannienes remarquables des géométries à groupe fondamental simple, Ann. Sci. Ecole Norm. Sup., 44 (1927), 345–467.
[2] S. Helgason, Differential geometry, Lie groups, and symmetric spaces, Academic Press, 1978.
[3] S. Kobayashi and K. Nomizu, Foundations of differential geometry II, Interscience, 1969.
[4] H. C. Wang, Two point homogeneous spaces, Ann. Math., (2) 55 (1952), 177–191.
[5] A. Korányi and J. A. Wolf, Realization of Hermitian symmetric spaces as generalized half-planes, Ann. Math., (2) 81 (1965), 265–288.
[6] R. Bott and H. Samelson, Applications of the theory of Morse to symmetric spaces, Amer. J. Math., 80 (1958), 964–1029.
[7] T. Sakai, On cut loci of compact symmetric spaces, Hokkaido Math. J., 6 (1977), 136–161.
[8] M. Takeuchi, On conjugate loci and cut loci of compact symmetric spaces I, Tsukuba J. Math., 2 (1978), 35–68.
[9] R. Crittenden, Minimum and conjugate points in symmetric spaces, Canad. J. Math., 14 (1962), 320–328.
[10] J. L. Koszul, Sur un type d'algèbre différentielles avec la transgression, Colloque de Topologie (Espaces fibrès), Brussels, 1950, 73–81.

414 (XX.1) Systems of Units

A. International System of Units

Units representing various physical quantities can be derived from a certain number of **fundamental (base) units**. By a **system of units** we mean a system of fundamental units. Various systems of units have been used in the course of the development of physics. Today, the standard is set by the **international system of units** (système international d'unités; abbreviated SI) [1], which has been developed in the spirit of the meter-kilogram system. This system consists of the seven fundamental units listed in Table 1, units induced from them, and unit designations with prefixes representing the powers of 10 where necessary. It also contains two **auxiliary units** for plane and solid angles, and a large number of derived units [1].

B. Systems of Units in Mechanics

Units in mechanics are usually derived from length, mass, and time, and SI uses the meter, kilogram, and second as base units. Neither the CGS system, derived from centimeter, gram, and second, nor the **system of gravitational units**, derived from length, force, and time, are recommended for general use by

Table 1

Quantity	SI unit	Symbol	Description
Length	meter	m	The meter is the length equal to 1,650,763.73 wavelengths in vacuum of the radiation corresponding to the transmission between the levels $2p^{10}$ and $5d^5$ of the krypton-86 atom.
Mass	kilogram	kg	The kilogram is equal to the mass of the international prototype of the kilogram.
Time	second	s	The second is the duration of 9,192,631,770 periods of the radiation corresponding to the transmission between the two hyperfine levels of the ground state of the cesium-133 atom.
Intensity of electric current	ampere	A	The ampere is the intensity of the constant current maintained in two parallel, rectilinear conductors of infinite length and of negligible circular section, placed 1 m apart in vacuum, and producing a force between them equal to 2×10^{-7} newton ($m \cdot kg \cdot s^{-2}$) per meter of length.
Temperature	kelvin	K	The kelvin, the unit of thermodynamical temperature, is 1/273.16 of the thermodynamical temperature of the triple point of water.
Amount of substance	mole	mol	The mole is the amount of substance of a system containing as many elementary entities as there are atoms in 0.012 kg of carbon-12.
Luminous intensity	candela	cd	The candela is the luminous intensity in a given direction of a source emitting monochromatic radiation of frequency 540×10^{12} hertz ($= s^{-1}$), the radiant intensity of which in that direction is 1/683 watt per steradian. (This revised definition of candela was adopted in 1980.)

Table 2

Quantity	SI unit	Symbol	Unit in terms of SI base or derived units
Frequency	hertz	Hz	$1\ Hz = 1\ s^{-1}$
Force	newton	N	$1\ N = 1\ kg \cdot m/s^2$
Pressure and stress	pascal	Pa	$1\ Pa = 1\ N/m^2$
Work, energy, quantity of heat	joule	J	$1\ J = 1\ N \cdot m$
Power	watt	W	$1\ W = 1\ J/s$
Quantity of electricity	coulomb	C	$1\ C = 1\ A \cdot s$
Electromotive force, potential difference	volt	V	$1\ V = 1\ W/A$
Electric capacitance	farad	F	$1\ F = 1\ C/V$
Electric resistance	ohm	Ω	$1\ \Omega = 1\ V/A$
Electric conductance	siemens	S	$1\ S = 1\ \Omega^{-1}$
Flux of magnetic induction magnetic flux	weber	Wb	$1\ Wb = 1\ V \cdot s$
Magnetic induction, magnetic flux density	tesla	T	$1\ T = 1\ Wb/m^2$
Inductance	henry	H	$1\ H = 1\ Wb/A$
Luminous flux	lumen	lm	$1\ lm = 1\ cd \cdot sr$
Illuminance	lux	lx	$1\ lx = 1\ lm/m^2$
Activity	becquerel	Bq	$1\ Bq = 1\ s^{-1}$
Adsorbed dose	gray	Gy	$1\ Gy = 1\ J/kg$
Radiation dose	sievert	Sv	$1\ Sv = 1\ J/kg$

the SI Committee. Besides the base units, minute, hour, and day, degree, minute, and second (angle), liter, and ton have been approved by the SI Committee. Units such as the electron volt, atomic mass unit, astronomical unit, and parsec (not SI) are empirically defined and have been approved. Several other units, such as nautical mile, knot, are (area), and bar, have been provisionally approved.

C. System of Units in Thermodynamics

The base unit for temperature is the degree Kelvin (°K; formerly called the absolute temperature). Degree Celsius (°C), defined by $t = T - 273.15$, where T is in °K, is also used. The unit of heat is the joule J, the same as the unit for other forms of energy. Formerly, one calorie was defined as the quantity of heat that must be supplied to one gram of water to raise its temperature from 14.5°C to 15.5°C; now one calorie is defined by 1 cal = 4.1855 J.

D. Systems of Units in Electricity and Magnetism

Three distinct systems of units have been developed in the field of electricity and magnetism: the electrostatic system, which originates from Coulomb's law for the force between two electric charges and defines magnetic quantities by means of the Biot-Savart law; the electromagnetic system, which originates from Coulomb's law for magnetism; and the Gaussian system, in which the dielectric constant and permeability are taken to be nondimensional. At present, however, the rationalized MKSA system of units is adopted as the international standard. It uses the **derived units** listed in Table 2 (taken from [2]), where the derived units with proper names in other fields are also listed.

E. Other Units

In the field of photometry, the following definition was adopted in 1948: One candela (cd) ($\fallingdotseq 0.98$ old candle) is defined as $1/(6 \times 10^5)$ of the luminous intensity in the direction normal to a plane surface of 1 m^2 area of a black body at the temperature of the solidifying point of platinum. The total luminous flux emanating uniformly in all directions from a source of luminous intensity 1 cd is defined as 4π lumen (lm). One lux (lx) is defined as the illuminance on a surface area of 1 m^2 produced by a luminous flux of 1 cd uniformly incident on the surface. In 1980, the definition was revised as shown in Table 1.

For theoretical purposes, a system of units called the absolute system of units is often used, in which units of mass, length, and time are chosen so that the values of universal constants, such as the universal gravitational constant, speed of light, Planck's constant, and Boltzmann's constant, are equal to 1.

References

[1] Bureau International des Points et Mesures, Le système international d'unités, 1970, fourth revised edition, 1981.
[2] R. G. Lerner and G. L. Trigg (eds.), Encyclopedia of physics, Addison-Wesley, 1981.

T

415 (XXI.41) Takagi, Teiji

Teiji Takagi (April 21, 1875–February 28, 1960) was born in Gifu Prefecture, Japan. After graduation from the Imperial University of Tokyo in 1897, he continued his studies in Germany, first with Frobenius in Berlin and then with Hilbert in Göttingen. He returned to Japan in 1901 and taught at the Imperial University of Tokyo until 1936, when he retured. He died in Tokyo of cerebral apoplexy.

Since his student years he had been interested in Kronecker's conjecture on †Abelian extensions of imaginary quadratic number fields. He solved it affirmatively for the case of $\mathbf{Q}(\sqrt{-1})$ while still in Göttingen and presented this result as his doctoral thesis. During World War I, he pursued his research in the theory of numbers in isolation from Western countries. It developed into †class field theory, a beautiful general theory of Abelian extensions of algebraic number fields. This was published in 1920, and was complemented by his 1922 paper on the †reciprocity law of power residues and then by †Artin's general law of reciprocity published in 1927. Besides these arithmetical works, he also published papers on algebraic and analytic subjects and on the foundations of the theories of natural numbers and of real numbers. His book (in Japanese) on the history of mathematics in the 19th century and his *General course of analysis* (also in Japanese) as well as his teaching and research activities at the University exercised great influence on the development of mathematics in Japan.

Reference

[1] S. Kuroda (ed.), The collected papers of Teiji Takagi, Iwanami, 1973.

416 (XI.16) Teichmüller Spaces

Consider the set $\mathbf{M}_g$ consisting of the conformal equivalence classes of closed Riemann surfaces of genus g. In 1859 Riemann stated, without rigorous proof, that $\mathbf{M}_g$ is parametrized by $m(g)$ ($=0$ if $g=0$, $=1$ if $g=1$, $=3g-3$ if $g\geqslant 2$) complex parameters (→ 11 Algebraic Functions). Later, the introduction of a topology and $m(g)$-dimensional complex structure on $\mathbf{M}_g$ were discussed rigorously in various ways. The following explanation of these methods is due to O. Teichmüller [1, 2], L. V. Ahlfors [3, 4], and L. Bers [5–7]. For the algebraic-geometric approach → 9 Algebraic Curves.

The trivial case $g=0$ is excluded, since $\mathbf{M}_0$ consists of a single point. Take a closed Riemann surface $\mathfrak{R}_0$ of genus $g\geqslant 1$, and consider the pairs $(\mathfrak{R}, H)$ consisting of closed Riemann surfaces $\mathfrak{R}$ of the same genus g and the †homotopy classes H of orientation-preserving homeomorphisms of $\mathfrak{R}_0$ into $\mathfrak{R}$. Two pairs $(\mathfrak{R}, H)$ and $(\mathfrak{R}', H')$ are defined to be conformally equivalent if the homotopy class $H'H^{-1}$ contains a conformal mapping. The set $\mathbf{T}_g$ consisting of the conformal equivalence classes $\langle\mathfrak{R}, H\rangle$ is called the **Teichmüller space** (with center at $\mathfrak{R}_0$). Let $\mathfrak{H}_g$ be the group of homotopy classes of orientation-preserving homeomorphisms of $\mathfrak{R}_0$ onto itself. $\mathfrak{H}_g$ is a transformation group acting on $\mathbf{T}_g$ in the sense that each $\eta\in\mathfrak{H}_g$ induces the transformation $\langle\mathfrak{R}, H\rangle\rightarrow\langle\mathfrak{R}, H\eta\rangle$. It satisfies $\mathbf{T}_g/\mathfrak{H}_g=\mathbf{M}_g$. The set $\mathfrak{J}_g$ of elements of $\mathfrak{H}_g$ fixing every point of $\mathbf{T}_g$ consists only of the unity element if $g\geqslant 3$ and is a normal subgroup of order 2 if $g=1, 2$. For the remainder of this article we assume that $g\geqslant 2$. The case $g=1$ can be discussed similarly, and the result coincides with the classical one: $\mathbf{T}_1$ can be identified with the upper half-plane and $\mathfrak{H}_1/\mathfrak{J}_1$ is the †modular group.

Denote by $B(\mathfrak{R}_0)$ the set of measurable invariant forms $\mu\, d\bar{z}\, dz^{-1}$ with $\|\mu\|_\infty<1$. For every $\mu\in B(\mathfrak{R}_0)$ there exists a pair $(\mathfrak{R}, H)$ for which some $h\in H$ satisfies $h_{\bar{z}}=\mu h_z$ (→ 352 Quasiconformal Mappings). This correspondence determines a surjection $\mu\in B(\mathfrak{R}_0)\mapsto\langle\mathfrak{R}, H\rangle\in\mathbf{T}_g$. Next, if $Q(\mathfrak{R}_0)$ denotes the space of holomorphic quadratic differentials $\varphi\, dz^2$ on $\mathfrak{R}_0$, a mapping $\mu\in B(\mathfrak{R}_0)\mapsto\varphi\in Q(\mathfrak{R}_0)$ is obtained as follows: Consider μ on the universal covering space U ($=$upper half-plane) of $\mathfrak{R}_0$. Extend it to U^* ($=$lower half-plane) by setting $\mu=0$, and let f be a quasiconformal mapping f of the plane onto itself satisfying $f_{\bar{z}}=\mu f_z$. Take the †Schwarzian derivative $\psi=\{f, z\}$ of the holomorphic function f in U^*. The desired φ is given by $\varphi(z)=\overline{\psi(\bar{z})}$ on U. It has been verified that two μ induce the same φ if and only if the same $\langle\mathfrak{R}, H\rangle$ corresponds to μ. Consequently, an injection $\langle\mathfrak{R}, H\rangle\in\mathbf{T}_g\mapsto\varphi\in Q(\mathfrak{R}_0)$ is obtained. Since $Q(\mathfrak{R}_0)=\mathbf{C}^{m(g)}$ by the Riemann-Roch theorem, this injection yields an embedding $\mathbf{T}_g\subset\mathbf{C}^{m(g)}$, where $\mathbf{T}_g$ is shown to be a domain.

As a subdomain of $\mathbf{C}^{m(g)}$, the Teichmüller space is an $m(g)$-dimensional complex analytic manifold. It is topologically equivalent to the unit ball in real $2m(g)$-dimensional space and is a bounded †domain of holomorphy in $\mathbf{C}^{m(g)}$.

Let $\{\alpha_1, \ldots, \alpha_{2g}\}$ be a 1-dimensional homology basis with integral coefficients in $\mathfrak{R}_0$ such that the intersection numbers are $(\alpha_i, \alpha_j)=(\alpha_{g+i}, \alpha_{g+j})=0$, $(\alpha_i, \alpha_{g+j})=\delta_{ij}$, $i, j=1, \ldots, g$.

Given an arbitrary $\langle\mathfrak{R}, H\rangle \in \mathbf{T}_g$, consider the †period matrix Ω of $\mathfrak{R}$ with respect to the homology basis $H\alpha_1, \ldots, H\alpha_{2g}$ and the basis $\omega_1, \ldots, \omega_g$ of †Abelian differentials of the first kind with the property that $\int_{H\alpha_i} \omega_j = \delta_{ij}$. Then Ω is a holomorphic function on $\mathbf{T}_g$. Furthermore, the analytic structure of the Teichmüller space introduced previously is the unique one (with respect to the topology defined above) for which the period matrix is holomorphic.

$\mathfrak{H}_g$ is a properly discontinuous group of analytic transformations, and therefore $\mathbf{M}_g$ is an $m(g)$-dimensional normal †analytic space. $\mathfrak{H}_g$ is known to be the whole group of the holomorphic automorphisms of $\mathbf{T}_g$ (Royden [8]); thus $\mathbf{T}_g$ is not a †symmetric space.

To every point τ of the Teichmüller space, there corresponds a Jordan domain $D(\tau)$ in the complex plane in such a way that the fiber space $\mathbf{F}_g = \{(\tau, z) \mid z \in D(\tau), \tau \in \mathbf{T}_g \subset \mathbf{C}^{m(g)}\}$ has the following properties: $\mathbf{F}_g$ is a bounded domain of holomorphy of $\mathbf{C}^{m(g)+1}$. It carries a properly discontinuous group $\mathfrak{G}_g$ of holomorphic automorphisms, which preserves every fiber $D(\tau)$ and is such that $D(\tau)/\mathfrak{G}_g$ is conformally equivalent to the Riemann surface corresponding to τ. $\mathbf{F}_g$ carries holomorphic functions $F_j(\tau, z)$, $j = 1, \ldots, 5g-5$ such that for every τ the functions $F_j/F_1, j = 2, \ldots, 5g-5$ restricted to $D(\tau)$ generate the meromorphic function field of the Riemann surface $D(\tau)/\mathfrak{G}_g$.

By means of the †extremal quasiconformal mappings, it can be verified that $\mathbf{T}_g$ is a complete metric space. The metric is called the **Teichmüller metric**, and is known to be a Kobayashi metric.

The Teichmüller space also carries a naturally defined Kähler metric, which for $g=1$ coincides with the †Poincaré metric if $\mathbf{T}_1$ is identified with the upper half-plane. The †Ricci curvature, †holomorphic sectional cruvature, and †scalar curvature are all negative (Ahlfors [9]).

By means of the quasiconformal mapping f, which we considered previously in order to construct the correspondence $\mu \mapsto \varphi$, it is possible to regard the Teichmüller space as a space of quasi-Fuchsian groups (→ 234 Kleinian Groups). To the boundary of $\mathbf{T}_g$, it being a bounded domain in $\mathbf{C}^{m(g)}$, there correspond various interesting Kleinian groups, which are called †boundary groups (Bers [10], Maskit [11]).

The definition of Teichmüller spaces can be extended to open Riemann surfaces $\mathfrak{R}_0$ and, further, to those with signatures. A number of propositions stated above are valid to these cases as well. In particular, the Teichmüller space for the case where $\mathfrak{R}_0$ is the unit disk is called the **universal Teichmüller space**. It is a bounded domain of holomorphy in an infinite-dimensional Banach space and is a symmetric space. Every Teichmüller space is a subspace of the universal Teichmüller space.

References

[1] O. Teichmüller, Extremale quasikonforme Abbildungen und quadratische Differentiale, Abh. Preuss. Akad. Wiss., 1939.
[2] O. Teichmüller, Bestimmung der extremalen quasikonformen Abbildung bei geschlossenen orientierten Riemannschen Flächen, Abh. Preuss. Akad. Wiss., 1943.
[3] L. V. Ahlfors, The complex analytic structure of the space of closed Riemann surfaces, Analytic functions, Princeton Univ. Press, 1960, 45–66.
[4] L. V. Ahlfors, Lectures on quasiconformal mappings, Van Nostrand, 1966.
[5] L. Bers, Spaces of Riemann surfaces. Proc. Intern. Congr. Math., Edinburgh, 1958, 349–361.
[6] L. Bers, On moduli of Riemann surfaces, Lectures at Forschungsinstitut für Mathematik, Eidgenössische Technische Hochschule, Zürich, 1964.
[7] L. Bers, Uniformization, moduli, and Kleinian groups, Bull. London Math. Soc., 4 (1972), 257–300.
[8] H. L. Royden, Automorphisms and isometries of Teichmüller spaces, Advances in the Theory of Riemann Surfaces, Princeton Univ. Press, 1971, 369–383.
[9] L. V. Ahlfors, Curvature properties of Teichmüller's space, J. Analyse Math., 9 (1961), 161–176.
[10] L. Bers, On boundaries of Teichmüller spaces and on Kleinian groups I, Ann. Math., (2) 91 (1970), 570–600.
[11] B. Maskit, On boundaries of Teichmüller spaces and on Kleinian groups II, Ann. Math., (2) 91 (1970), 608–638.

417 (VII.5) Tensor Calculus

A. General Remarks

In a †differentiable manifold with an †affine connection (in particular, in a †Riemannian manifold), we can define an important operator on tensor fields, the operator of covariant differentiation. The **tensor calculus** is a differential calculus on a differentiable manifold that deals with various geometric objects and differential operators in terms of covariant differentiation, and it provides an important tool for studying geometry and analysis on a differentiable manifold.

B. Covariant Differential

Let M be an n-dimensional smooth manifold. We denote by $\mathfrak{F}(M)$ the set of all smooth functions on M and by $\mathfrak{X}^r_s(M)$ the set of all smooth tensor fields of type (r,s) on M. $\mathfrak{X}^1_0(M)$ is the set of all smooth vector fields on M, and we denote it simply by $\mathfrak{X}(M)$.

In the following we assume that an affine connection ∇ is given on M. Then we can define the **covariant differential** of tensor fields on M with respect to the connection (→ 80 Connections). We denote the **covariant derivative** of a tensor field K in the direction of a vector field X by $\nabla_X K$ and the covariant differential of K by ∇K. The operator ∇_X maps $\mathfrak{X}^r_s(M)$ into itself and has the following properties:
(1) $\nabla_{X+Y}=\nabla_X+\nabla_Y$, $\nabla_{fX}=f\nabla_X$,
(2) $\nabla_X(K+K')=\nabla_X K+\nabla_X K'$,
(3) $\nabla_X(K\otimes K')=(\nabla_X K)\otimes K'+K\otimes(\nabla_X K')$,
(4) $\nabla_X f=Xf$,
(5) ∇_X commutes with contraction of tensor fields, where K and K' are tensor fields on M, X, $Y\in\mathfrak{X}(M)$ and $f\in\mathfrak{F}(M)$.

The **torsion tensor** T and the **curvature tensor** R of the affine connection ∇ are defined by

$$T(X,Y)=\nabla_X Y-\nabla_Y X-[X,Y],$$

$$R(X,Y)Z=\nabla_X(\nabla_Y Z)-\nabla_Y(\nabla_X Z)-\nabla_{[X,Y]}Z$$

for vector fields X, Y, and Z. The torsion tensor is of type $(1,2)$, and the curvature tensor is of type $(1,3)$. Some authors define $-R$ as the curvature tensor. We here follow the convention used in [1–6], while in [7, 8] the sign of the curvature tensor is opposite. The torsion tensor and the curvature tensor satisfy the identities

$$T(X,Y)=-T(Y,X),\qquad R(X,Y)=-R(Y,X),$$

$$R(X,Y)Z+R(Y,Z)X+R(Z,X)Y$$
$$=(\nabla_X T)(Y,Z)+(\nabla_Y T)(Z,X)+(\nabla_Z T)(X,Y)$$
$$+T(T(X,Y),Z)+T(T(Y,Z),X)$$
$$+T(T(Z,X),Y),$$

$$(\nabla_X R)(Y,Z)+(\nabla_Y R)(Z,X)+(\nabla_Z R)(X,Y)$$
$$=R(X,T(Y,Z))+R(Y,T(Z,X))$$
$$+R(Z,T(X,Y)).$$

The last two identities are called the **Bianchi identities**.

The operators ∇_X and ∇_Y for two vector fields X and Y are not commutative in general, and they satisfy the following formula, the **Ricci formula**, for a tensor field K:

$$\nabla_X(\nabla_Y K)-\nabla_Y(\nabla_X K)-\nabla_{[X,Y]}K=R(X,Y)\cdot K,$$

where in the right-hand side $R(X,Y)$ is regarded as a derivation of the tensor algebra $\sum_{r,s}\mathfrak{X}^r_s(M)$.

A **moving frame** of M on a neighborhood U is, by definition, an ordered set $(e_1,\ldots,e_n)$ of n vector fields on U such that $e_1(p),\ldots,e_n(p)$ are linearly independent at each point $p\in U$. For a moving frame $(e_1,\ldots,e_n)$ of M on a neighborhood U we define n differential 1-forms $\theta^1,\ldots,\theta^n$ by $\theta^i(e_j)=\delta^i_j$, and we call them the **dual frame** of $(e_1,\ldots,e_n)$. For a tensor field K of type (r,s) on M, we define n^{r+s} functions $K^{i_1\ldots i_r}_{j_1\ldots j_s}$ on U by

$$K^{i_1\ldots i_r}_{j_1\ldots j_s}=K(e_{j_1},\ldots,e_{j_s},\theta^{i_1},\ldots,\theta^{i_r})$$

and call these functions the components of K with respect to the moving frame $(e_1,\ldots,e_n)$.

Since the covariant differentials ∇e_j are tensor fields of type $(1,1)$, n^2 differential 1-forms ω^i_j are defined by

$$\nabla e_j=\omega^i_j\otimes e_i,$$

where in the right-hand side (and throughout the following) we adopt **Einstein's summation convention**: If an index appears twice in a term, once as a superscript and once as a subscript, summation has to be taken on the range of the index. (Some authors write the above equation as $de_j=\omega^i_j e_i$ or $De_j=\omega^i_j e_i$.) We call these 1-forms ω^i_j the **connection forms** of the affine connection with respect to the moving frame $(e_1,\ldots,e_n)$. The torsion forms Θ^i and the curvature forms Ω^i_j are defined by

$$\Theta^i=d\theta^i+\omega^i_j\wedge\theta^j,\qquad \Omega^i_j=d\omega^i_j+\omega^i_k\wedge\omega^k_j.$$

These equations are called the **structure equation** of the affine connection ∇. If we denote the components of the torsion tensor and the curvature tensor with respect to $(e_1,\ldots,e_n)$ by T^i_{jk} and R^i_{jkl} $(=\theta^i(R(e_k,e_l)e_j))$, respectively, then they satisfy the relations

$$\Theta^i=\frac{1}{2}T^i_{jk}\theta^j\wedge\theta^k,\qquad \Omega^i_j=\frac{1}{2}R^i_{jkl}\theta^k\wedge\theta^l.$$

Using these forms, the Bianchi identities are written as

$$d\Theta^i+\omega^i_j\wedge\Theta^j=\Omega^i_j\wedge\theta^j,$$

$$d\Omega^i_j+\omega^i_k\wedge\Omega^k_j-\omega^k_j\wedge\Omega^i_k=0.$$

Let K be a tensor field of type (r,s) on M and $K^{i_1\ldots i_r}_{j_1\ldots j_s}$ be the components of K with respect to $(e_1,\ldots,e_n)$. We define the covariant differential $DK^{i_1\ldots i_r}_{j_1\ldots j_s}$ and the covariant derivative $K^{i_1\ldots i_r}_{j_1\ldots j_s,k}$ by

$$DK^{i_1\ldots i_r}_{j_1\ldots j_s}=K^{i_1\ldots i_r}_{j_1\ldots j_s,k}\theta^k=dK^{i_1\ldots i_r}_{j_1\ldots j_s}+\sum_{v=1}^{r}K^{i_1\ldots a\ldots i_r}_{j_1\ldots j_s}\omega^{i_v}_a$$
$$-\sum_{v=1}^{s}K^{i_1\ldots i_r}_{j_1\ldots a\ldots j_s}\omega^a_{j_v},$$

Then $K^{i_1\dots i_r}_{j_1\dots j_s,k}$ are the components of ∇K with respect to the moving frame $(e_1,\dots,e_n)$. Some authors write $\nabla_k K^{i_1\dots i_r}_{j_1\dots j_s}$ instead of $K^{i_1\dots i_r}_{j_1\dots j_s}$ [5, 6].

Using components, the Bianchi identities are written as

$$R^h_{ijk}+R^h_{jki}+R^h_{kij}=T^h_{ij,k}+T^h_{jk,i}+T^h_{ki,j}$$
$$+T^h_{ai}T^a_{jk}+T^h_{aj}T^a_{ki}+T^h_{ak}T^a_{ij},$$

$$R^h_{ijk,l}+R^h_{ikl,j}+R^h_{ilj,k}=R^h_{iak}T^a_{jl}+R^h_{iaj}T^a_{lk}+R^h_{ial}T^a_{kj}.$$

The Ricci formula is written as

$$K^{i_1\dots i_r}_{j_1\dots j_s,kl}-K^{i_1\dots i_r}_{j_1\dots j_s,lk}=\sum_{v=1}^{r}R^{i_v}_{alk}K^{i_1\dots a\dots i_r}_{j_1\dots\dots\dots j_s}$$
$$-\sum_{v=1}^{s}R^a_{j_v lk}K^{i_1\dots\dots\dots i_r}_{j_1\dots a\dots j_s}$$
$$+T^a_{kl}K^{i_1\dots i_r}_{j_1\dots j_s,a}.$$

Let $(x^1,\dots,x^n)$ be a local coordinate system defined on a neighborhood U of M. Then $(\partial/\partial x^1,\dots,\partial/\partial x^n)$ is a moving frame of M on U, and we call it the **natural moving frame** associated with the coordinate system $(x^1,\dots,x^n)$. Components of a tensor field with respect to the natural moving frame $(\partial/\partial x^1,\dots,\partial/\partial x^n)$ are often called components with respect to the coordinate system $(x^1,\dots,x^n)$. We define an n^3 function Γ^i_{kj} on U by $\omega^i_j=\Gamma^i_{kj}dx^k$, where ω^i_j are the connection forms for the natural moving frame. Γ^i_{kj} are called the coefficients of the affine connection ∇. The components of the torsion tensor and the curvature tensor with respect to $(x^1,\dots,x^n)$ are given by

$$T^i_{jk}=\Gamma^i_{jk}-\Gamma^i_{kj},$$

$$R^h_{ijk}=\partial_j\Gamma^h_{ki}-\partial_k\Gamma^h_{ji}+\Gamma^a_{ki}\Gamma^h_{ja}-\Gamma^a_{ji}\Gamma^h_{ka},$$

where $\partial_i=\partial/\partial x^i$.

With respect to the foregoing coordinate system, the components $K^{i_1\dots i_r}_{j_1\dots j_s,k}$ of the covariant differential ∇K of a tensor field K of type (r,s) are given by

$$K^{i_1\dots i_r}_{j_1\dots j_s,j}=\partial_j K^{i_1\dots i_r}_{j_1\dots j_s}+\sum_{v=1}^{r}\Gamma^{i_v}_{ja}K^{i_1\dots a\dots i_r}_{j_1\dots j_s}$$
$$-\sum_{v=1}^{s}\Gamma^a_{jj_v}K^{i_1\dots i_r}_{j_1\dots a\dots j_s}.$$

C. Covariant Differential of Tensorial Forms

A tensorial p-form of type (r,s) on a manifold M is an alternating $\mathfrak{F}(M)$-multilinear mapping of $\overbrace{\mathfrak{X}(M)\times\dots\times\mathfrak{X}(M)}^{p}$ to $\mathfrak{X}^r_s(M)$. A tensorial p-form of type $(0,0)$ is a differential p-form in the usual sense. A tensorial p-form of type $(1,0)$ is often called a **vectorial p-form**.

If an affine connection ∇ is provided on M, we define the covariant differential of tensorial forms. Let α be a tensorial p-form of type (r,s). The covariant differential $D\alpha$ of α is a tensorial $(p+1)$-form of type (r,s) and is defined by

$$(p+1)D\alpha(X_1,\dots,X_{p+1})$$
$$=\sum_{i=1}^{p+1}(-1)^{i-1}\nabla_{X_i}(\alpha(X_1,\dots,\hat{X}_i,\dots,X_{p+1}))$$
$$+\sum_{i<j}(-1)^{i+j}\alpha([X_i,X_j],$$
$$X_1,\dots,\hat{X}_i,\dots,\hat{X}_j,\dots,X_{p+1}),$$

where $\hat{X}_i$ means that X_i is deleted. If α is of type $(0,0)$, $D\alpha$ coincides with the usual exterior differential $d\alpha$.

The simplest example of a tensorial form is the identity mapping of $\mathfrak{X}(M)$, which will be denoted by θ. Some authors write this vectorial form as dp or dx, where p or x expresses an arbitrary point of a manifold. We call θ the **canonical vectorial form** of M. The torsion tensor T can be regarded as a vectorial 2-form, and we have $2D\theta=T$. The curvature tensor R can be regarded as a tensorial 2-form of type $(1,1)$, i.e., $(X,Y)\rightarrow R(X,Y)\in\mathfrak{X}^1_1(M)$, and the Bianchi identities are written as $DT=R\wedge\theta$, $DR=0$, where the exterior product $R\wedge\alpha$ of R and a tensorial p-form α is defined by

$$(p+1)(p+2)(R\wedge\alpha)(X_1,\dots,X_{p+2})$$
$$=2\sum_{i<j}(-1)^{i+j-1}R(X_i,X_j)\alpha(X_1,\dots,\hat{X}_i,\dots,\hat{X}_j,$$
$$\dots,X_{p+2}).$$

In general, $2D^2\alpha=R\wedge\alpha$ holds for an arbitrary tensorial form α.

Let $(e_1,\dots,e_n)$ be a moving frame of M on a neighborhood U and $\theta^1,\dots,\theta^n$ be its dual frames. A tensorial p-form α of type (r,s) is written as

$$\alpha=\alpha^{i_1\dots i_r}_{j_1\dots j_s}\otimes e_{i_1}\otimes\dots\otimes e_{i_r}\otimes\theta^{j_1}\otimes\dots\otimes\theta^{j_s},$$

on U, where the $\alpha^{i_1\dots i_r}_{j_1\dots j_s}$ are the usual differential p-forms on U. We call them the components of α with respect to $(e_1,\dots,e_n)$. Then the components of $D\alpha$, which we denote by $D\alpha^{i_1\dots i_r}_{j_1\dots j_s}$, are given by

$$D\alpha^{i_1\dots i_r}_{j_1\dots j_s}=d\alpha^{i_1\dots i_r}_{j_1\dots j_s}+\sum_{v=1}^{r}\omega^{i_v}_a\wedge\alpha^{i_1\dots a\dots i_r}_{j_1\dots j_s}$$
$$-\sum_{v=1}^{s}\omega^a_{j_v}\wedge\alpha^{i_1\dots i_r}_{j_1\dots a\dots j_s}.$$

Then we have

$$D^2\alpha^{i_1\dots i_r}_{j_1\dots j_s}=\sum_{v=1}^{r}\Omega^{i_v}_a\wedge\alpha^{i_1\dots a\dots i_r}_{j_1\dots j_s}-\sum_{v=1}^{s}\Omega^a_{j_v}\wedge\alpha^{i_1\dots i_r}_{j_1\dots a\dots j_s}.$$

This is an expression of $2D^2\alpha=R\wedge\alpha$ in terms of components. The components of the canonical vectorial form θ are the dual forms $\theta^1,\dots,\theta^n$ of $(e_1,\dots,e_n)$, and we have $D\theta^i=\Theta^i$, which means that the components of $D\theta$ are the torsion forms Θ^i.

D. Tensor Fields on a Riemannian Manifold

Let (M, g) be an n-dimensional Riemannian manifold (→ 364 Riemannian Manifolds). The fundamental tensor g defines a one-to-one correspondence between vector fields and differential 1-forms. A differential 1-form α which corresponds to a vector field X is defined by $\alpha(Y)=g(X, Y)$ for any vector field Y. This correspondence is naturally extended to a one-to-one correspondence between $\mathfrak{X}_s^r(M)$ and $\mathfrak{X}_{s'}^{r'}(M)$, where $r+s=r'+s'$. Let $(e_1, \ldots, e_n)$ be a moving frame of M on a neighborhood U and g_{ij} be the components of g with respect to the moving frame. Let (g^{ij}) be the inverse matrix of the matrix (g_{ij}). The g^{ij} are the components of a symmetric contravariant tensor field of order 2. Let X^i be the components of a vector field X and α_i be the components of the differential 1-form α corresponding to X. Then X^i and α_i satisfy the relations $\alpha_i=g_{ij}X^j$ and $X^i=g^{ij}\alpha_j$. If K_{ij}^h are the components of a tensor field K of type (1, 2) (here taken for simplicity), then

$$K_{hij}=K_{ij}^a g_{ah}, \quad K_j^{hi}=K_{aj}^h g^{ai},$$

$$K^{hij}=K_{ab}^h g^{ai} g^{bj}, \ldots,$$

are the components of a tensor field of type (0, 3), (2, 1), (3, 0), ..., respectively, all of which correspond to K. We call this process of obtaining the components of the corresponding tensor fields from the components of a given tensor field **raising the subscripts** and **lowering the superscripts** by means of the fundamental tensor g.

On a Riemannian manifold, we use the †Riemannian connection, unless otherwise stated. The covariant derivative with respect to the Riemannian connection is given by

$$\begin{aligned}2g(\nabla_X Y, Z)=&Xg(Y,X)+Yg(X,Z)-Zg(X,Y)\\&+g([X,Y],Z)-g([X,Z],Y)\\&-g(X,[Y,Z])\end{aligned}$$

for vector fields X, Y, and Z. The coefficients of the Riemannian connection with respect to a local coordinate system $(x^1, \ldots, x^n)$ are usually written as $\left\{{i \atop kj}\right\}$, called the **Christoffel symbols**, which are given by $\left\{{i \atop kj}\right\}=g^{ia}(\partial_k g_{ja}+\partial_j g_{ka}-\partial_a g_{kj})/2$. The curvature tensor R of the Riemannian connection satisfies the identities

$$R(X,Y)Z+R(Y,Z)X+R(Z,X)Y=0,$$

$$(\nabla_X R)(Y,Z)+(\nabla_Y R)(Z,X)+(\nabla_Z R)(X,Y)=0,$$

$$R(X,Y)=-R(Y,X),$$

$$\begin{aligned}g(R(X,Y)Z,W)&=g(R(Z,W)X,Y)\\&=-g(Z,R(X,Y)W),\end{aligned}$$

$$\begin{aligned}g(R(X,Y)Z,W)&+g(R(X,Z)W,Y)\\&+g(R(X,W)Y,Z)=0.\end{aligned}$$

In terms of the components, these identities are

$$R_{ijk}^h+R_{jki}^h+R_{kij}^h=0,$$

$$R_{ijk,l}^h+R_{ikl,j}^h+R_{ilj,k}^h=0,$$

$$R_{ijk}^h=-R_{ikj}^h, \; R_{hijk}=R_{jkhi}=-R_{ihjk},$$

$$R_{hijk}+R_{hjki}+R_{hkij}=0,$$

where $R_{hijk}=R_{ijk}^a g_{ah}$.

The †Ricci tensor S of the Riemannian manifold is a tensor field of type (0, 2) defined by

$$S(X,Y)=\text{trace of the mapping } Z\to R(Z,X)Y$$

for vector fields X and Y. The components S_{ji} of the Ricci tensor are given by $S_{ji}=R_{jai}^a$. The †scalar curvature k of the Riemannian manifold M is a scalar on M defined by $k=g^{ji}S_{ji}$. The Ricci tensor and the scalar curvature satisfy the identities

$$S(X,Y)=S(Y,X) \quad \text{or} \quad S_{ji}=S_{ij},$$

$$S_{ij,k}-S_{ik,j}=R_{ikj,a}^a, \qquad 2g^{jk}S_{ij,k}=\partial_i k.$$

For a moving frame of a Riemannian manifold, it is convenient to use an **orthonormal moving frame**. A moving frame $(e_1, \ldots, e_n)$ is orthonormal if $e_1, \ldots, e_n$ satisfy $g(e_i, e_j)=\delta_{ij}$. Since the components of the fundamental tensor with respect to an orthonormal moving frame are δ_{ij}, raising or lowering the indices does not change the values of the components. Some authors write all the indices as subscripts. Also they write the dual 1-forms, the connection forms, and the curvature forms as θ_i, ω_{ji}, and Ω_{ji}, respectively, instead of θ^i, ω_j^i, and Ω_j^i. With respect to an orthonormal moving frame, the connection forms ω_j^i and the curvature forms Ω_j^i satisfy

$$\omega_j^i+\omega_i^j=0 \quad \text{and} \quad \Omega_j^i+\Omega_i^j=0.$$

On a Riemannian manifold, the divergence of a vector field and the operators d, δ, and Δ on differential forms (→ 194 Harmonic Integrals) can be expressed by using the covariant derivatives with respect to the Riemannian connection.

If X^i are the components of a vector field X with respect to a local coordinate system $(x^1, \ldots, x^n)$, the divergence $\operatorname{div} X$ of X is given by $\operatorname{div} X=X^i{}_{,i}$.

Let α be a differential p-form on M. α is written locally in the form $\alpha=(1/p!)\alpha_{i_1 \ldots i_p} dx^{i_1} \wedge \ldots \wedge dx^{i_p}$, where the coefficients $\alpha_{i_1 \ldots i_p}$ are skew-symmetric in all the indices. We call $\alpha_{i_1 \ldots i_p}$ the components of α with respect to the coordinate system. Since α is regarded as an alternating tensor field of type $(0, p)$, we can define the covariant differential $\nabla\alpha$ of α. Then the components of $d\alpha$, $\delta\alpha$, and $\Delta\alpha$ are given by

$$(d\alpha)_{i_1\dots i_{p+1}}=\sum_{v=1}^{p+1}(-1)^{v-1}\alpha_{i_1\dots\hat{i}_v\dots i_{p+1},i_v},$$

$$(\delta\alpha)_{i_1\dots i_{p-1}}=-g^{ab}\alpha_{ai_1\dots i_{p-1},b},$$

$$(\Delta\alpha)_{i_1\dots i_p}=-g^{ab}\left[\alpha_{i_1\dots i_p,ab}-\sum_{v=1}^{p}S_{i_va}\alpha_{i_1\dots b\dots i_p}\right.$$
$$\left.-\sum_{v<w}R^c_{ai_vi_w}\alpha_{i_1\dots b\dots c\dots i_p}\right].$$

For a smooth function f and a differential 1-form β we have

$$\Delta f=-\frac{1}{\sqrt{g}}\partial_i(g^{ij}\sqrt{g}\,\partial_j f),$$

$$(\Delta\beta)_i=-g^{ab}[\beta_{i,ab}-S_{ia}\beta_b],$$

where $g=\det(g_{ij})$.

E. Van der Waerden–Bortolotti Covariant Differential

Let E be a finite dimensional smooth †vector bundle over a smooth manifold M and $\Gamma(E)$ be an $\mathfrak{F}(M)$-module of all smooth sections of E. A connection ∇' in E is a mapping of $\mathfrak{X}(M)\times\Gamma(E)$ to $\Gamma(E)$ such that

(1) $\nabla'_X(\xi+\eta)=\nabla'_X\xi+\nabla'_X\eta$,

(2) $\nabla'_X(f\xi)=Xf\cdot\xi+f\nabla'_X\xi$,

(3) $\nabla'_{X+Y}\xi=\nabla'_X\xi+\nabla'_Y\xi$,

(4) $\nabla'_{fX}\zeta=f\nabla'_X\xi$,

for X, $Y\in\mathfrak{X}(M)$, $\xi,\eta\in\Gamma(E)$, and $f\in\mathfrak{F}(M)$. $\nabla'_X\xi$ is called the covariant derivative of ξ in the direction X.

An element K of $\mathfrak{X}^r_s(M)\otimes\Gamma(E)$ is called a **tensor field of type** (r,s) **with values in** E (or simply an E-valued tensor field of type (r,s)). K can be regarded as an $\mathfrak{F}(M)$-linear mapping of $\mathfrak{X}^s_r(M)$ to $\Gamma(E)$ or an $\mathfrak{F}(M)$-multilinear mapping of $\overbrace{\mathfrak{X}(M)\times\dots\times\mathfrak{X}(M)}^{s}$ to $\mathfrak{X}^r_0(M)\otimes\Gamma(E)$. For a given $\xi\in\Gamma(E)$, a mapping $X\to\nabla'_X\xi$ defines a tensor field of type (0, 1) with values in E which we call the covariant differential of ξ, denoted by $\nabla'\xi$.

The curvature tensor R' of ∇' is a tensor field of type (0, 2) with values in $E^*\otimes E$ (E^* is the dual vector bundle of E), and is defined by

$$R'(X,Y)\xi=\nabla'_X(\nabla'_Y\xi)-\nabla'_Y(\nabla'_X\xi)-\nabla'_{[X,Y]}\xi$$

for any vector fields X and Y and any $\xi\in\Gamma(E)$.

If an affine connection ∇ is given on M, we can define the **van der Waerden–Bortolotti covariant derivative** $\bar{\nabla}_XK$ for ∇ and ∇' of a tensor field K of type (r,s) with values in E. It is defined by

$$(\bar{\nabla}_XK)(S)=\nabla'_X(K(S))-K(\nabla_XS)$$

for any $S\in\mathfrak{X}^s_r(M)$. If we regard $\xi\in\Gamma(E)$ as an E-valued tensor field of type (0, 0), we have $\bar{\nabla}_X\xi=\nabla'_X\xi$. The covariant derivative $\bar{\nabla}_XR'$ of the curvature tensor R' of ∇' is a tensor field of type (0, 2) with values in $E^*\otimes E$ is defined by

$$(\bar{\nabla}_XR')(Y,Z)\xi=\nabla'_X(R'(Y,Z)\xi)-R'(\nabla_XY,Z)\xi$$
$$-R'(Y,\nabla_XZ)\xi-R'(Y,Z)\nabla'_X\xi.$$

The Bianchi identity is written as

$$(\bar{\nabla}_XR')(Y,Z)+(\bar{\nabla}_YR')(Z,X)+(\bar{\nabla}_ZR')(X,Y)$$
$$=R'(X,T(Y,Z))+R'(Y,T(Z,X))$$
$$+R'(Z,T(X,Y)),$$

where T is the torsion tensor of ∇. The Ricci formula is given by

$$(\bar{\nabla}_X(\bar{\nabla}_YK))(S)-(\bar{\nabla}_Y(\bar{\nabla}_XK))(S)-(\bar{\nabla}_{[X,Y]}K)(S)$$
$$=R'(X,Y)\cdot K(S)-K(R(X,Y)\cdot S),$$

where R is the curvature tensor of ∇, $K\in\mathfrak{X}^r_s(M)\otimes\Gamma(E)$ and $S\in\mathfrak{X}^s_r(M)$.

In the following we assume that the fiber of E is of finite dimension m. A moving frame of E on a neighborhood U of M is an ordered set $(\xi_1,\dots,\xi_m)$ of local sections $\xi_1,\dots,\xi_m$ on U such that $\xi_1(p),\dots,\xi_m(p)$ are linearly independent at each point p of U. Let $(e_1,\dots,e_n)$ be a moving frame of M on U. Then an E-valued tensor field K of type (r,s) is locally written as

$$K^{i_1\dots i_r\alpha}_{j_1\dots j_s}e_{i_1}\otimes\dots\otimes e_{i_r}\otimes\theta^{j_1}\otimes\dots\otimes\theta^{j_s}\otimes\xi_\alpha,$$

where $\theta^1,\dots,\theta^n$ are the dual 1-forms of $(e_1,\dots,e_n)$. The $n^{r+s}m$ functions $K^{i_1\dots i_r\alpha}_{j_1\dots j_s}$ on U are called the components of K with respect to $(e_1,\dots,e_n)$ and $(\xi_1,\dots,\xi_m)$. We define the connection forms ω'^α_β of the connection ∇' by $\nabla'\xi_\beta=\omega'^\alpha_\beta\otimes\xi_\alpha$. Then the curvature forms Ω'^α_β are defined by

$$\Omega'^\alpha_\beta=d\omega'^\alpha_\beta+\omega'^\alpha_\lambda\wedge\omega'^\lambda_\beta=\tfrac{1}{2}R^\alpha_{\beta ji}\theta^j\wedge\theta^i,$$

where $R^\alpha_{\beta ji}$ are the components of the curvature tensor R', i.e., $R'(e_j,e_i)\xi_\beta=R^\alpha_{\beta ji}\xi_\alpha$.

For a given tensor field K of type (r,s) with values in E, the mapping $X\to\bar{\nabla}_XK$ defines a tensor field $\bar{\nabla}K$ of $(r,s+1)$ with values in E which we call the van der Waerden–Bortollotti covariant differential of K. Then if $K^{i_1\dots i_r\alpha}_{j_1\dots j_s}$ are the components of K with respect to $(e_1,\dots,e_n)$ and $(\xi_1,\dots,\xi_m)$, the components $K^{i_1\dots i_r\alpha}_{j_1\dots j_s,k}$ of $\bar{\nabla}K$ are given by

$$K^{i_1\dots i_r\alpha}_{j_1\dots j_s,k}\theta^k=dK^{i_1\dots i_r\alpha}_{j_1\dots j_s}+\sum_{v=1}^{r}K^{i_1\dots a\dots i_r\alpha}_{j_1\dots j_s}\omega^{i_v}_a$$
$$-\sum_{v=1}^{s}K^{i_1\dots i_r\alpha}_{j_1\dots a\dots j_s}\omega^a_{j_v}+K^{i_1\dots i_r\beta}_{j_1\dots j_s}\omega'^\alpha_\beta.$$

Let f be a smooth mapping of M into a smooth manifold M'. The differential f_*(or df) can be regarded as a tensor field of type (0, 1) with values in $f^*T(M')$. Assume that M (resp. M') has a Riemannian metric g (resp. g'). We denote the Riemannian connection of M by ∇.

From the Riemannian connection of M' a connection ∇' in $f^*T(M')$ can be defined. Let $(y^1, \ldots, y^m)$ be a local coordinate system of M' on a neighborhood V and $(x^1, \ldots, x^n)$ be a local coordinate system on a neighborhood U of M such that $f(U) \subset V$. Put $\xi_\alpha(p) = (\partial/\partial y^\alpha)(f(p))$ for a point $p \in U$. Then $(\xi_1, \ldots, \xi_m)$ is a moving frame of $f^*T(M')$. The components of f_* with respect to $(\partial/\partial x^1, \ldots, \partial/\partial x^n)$ and $(\xi_1, \ldots, \xi_m)$ are given by $f^\alpha(p) = (\partial y^\alpha/\partial x^i)(p)$. The Laplacian Δf of the mapping f is a tensor field of type $(0,0)$ with values in $f^*T(M')$ and is defined by $(\Delta f)^\alpha = g^{ij} f^\alpha_{i,j}$. If $\Delta f = 0$, the mapping f is called a harmonic mapping (→ 195 Harmonic Mappings).

F. Tensor Fields on a Submanifold

Consider an n-dimensional smooth manifold M immersed in an $(n+m)$-dimensional Riemannian manifold $(\overline{M}, \overline{g})$. If we denote the immersion $M \to \overline{M}$ by f, then $g = f^*\overline{g}$ is a Riemannian metric on M, and we denote its Riemannian connection by ∇. The induced bundle $f^*T(\overline{M})$ splits into the sum of the tangent bundle $T(M)$ of M and the normal bundle $T^\perp(M)$. The Riemannian connection on $\overline{M}$ induces connections in $f^*T(\overline{M})$ and in $T^\perp(M)$ which are denoted by $\overline{\nabla}$ and $\nabla^\perp$, respectively. The van der Waerden–Bortolotti covariant derivative for ∇ and $\nabla^\perp$ is denoted by $\tilde{\nabla}$.

For vector fields X and Y on M, the tangential part of $\overline{\nabla}_X Y$ (here we regard Y as a section of $f^*T(\overline{M})$) is $\nabla_X Y$, and we denote the normal part of $\overline{\nabla}_X Y$ by $h(X, Y)$. Then h is a symmetric tensor field of type $(0,2)$ with values in $T^\perp(M)$, and we call h the **second fundamental tensor** of the immersion f. For $\xi \in \Gamma(T^\perp(M))$, the tangential part of $\overline{\nabla}_X \xi$ (here ξ is also regarded as a section of $f^*T(\overline{M})$) is denoted by $-A_\xi X$ and the normal part of $\overline{\nabla}_X \xi$ is $\nabla^\perp_X \xi$. Thus we have

$$\overline{\nabla}_X Y = \nabla_X Y + h(X, Y), \quad \overline{\nabla}_X \xi = -A_\xi X + \nabla_X \xi.$$

h and A are related by

$$\overline{g}(h(X, Y), \xi) = g(A_\xi X, Y).$$

We have the following formulas, called the equations of Gauss, Codazzi, and Ricci:

$$\overline{g}(\overline{R}(X, Y)Z, W) = g(R(X, Y)Z, W) + \overline{g}(h(X, Z), h(Y, W)) - \overline{g}(h(X, W), h(Y, X)),$$

$$\overline{g}(\overline{R}(X, Y)Z, \xi) = \overline{g}((\tilde{\nabla}_X h)(Y, Z), \xi) - \overline{g}((\tilde{\nabla}_Y h)(X, Z), \xi),$$

$$\overline{g}(\overline{R}(X, Y)\xi, \eta) = \overline{g}(R^\perp(X, Y)\xi, \eta) + g([A_\xi, A_\eta]X, Y),$$

for $X, Y, Z, W \in X(M)$ and $\xi, \eta \in \Gamma(T^\perp(M))$, where R, $\overline{R}$, and $R^\perp$ are the curvature tensors of ∇, $\overline{\nabla}$, and $\nabla^\perp$, respectively.

For the manifold M immersed in $\overline{M}$, we use a moving frame $(e_1, \ldots, e_n, \xi_1, \ldots, \xi_m)$ such that $(e_1, \ldots, e_n)$ is an orthonormal moving frame of M on a neighborhood U and $(\xi_1, \ldots, \xi_m)$ is a moving frame of $T^\perp(M)$ on U with $\overline{g}(\xi_\alpha, \xi_\beta) = \delta_{\alpha\beta}$. Then we can define the connection forms ω^i_j for ∇ and ω^α_β for $\nabla^\perp$. If we extend $(e_1, \ldots, e_n, \xi_1, \ldots, \xi_m)$ to an orthonormal moving frame $(\overline{e}_1, \ldots, \overline{e}_{n+m})$ of $\overline{M}$ such that $\overline{e}_i(p) = e_i(p)$ $(i = 1, \ldots, n)$ and $\overline{e}_{n+\alpha}(p) = \xi_\alpha(p)$ $(\alpha = 1, \ldots, m)$ for $p \in U$, then the restriction $f^*\overline{\theta}^A$ and $f^*\overline{\omega}^A_B$ of the dual 1-forms and the connection forms of $\overline{M}$ with respect to $(\overline{e}_1, \ldots, \overline{e}_{n+m})$ satisfy the relations

$$f^*\overline{\theta}^i = \theta^i, \quad f^*\overline{\theta}^{n+\alpha} = 0, \quad f^*\overline{\omega}^i_j = \omega^i_j,$$

$$f^*\overline{\omega}^{n+\alpha}_{n+\beta} = \omega^\alpha_\beta, \quad f^*\overline{\omega}^{n+\alpha}_i = \sum_j h^\alpha_{ij}\theta^j,$$

where h^α_{ij} are the components of the second fundamental tensor h with respect to $(e_1, \ldots, e_n, \xi_1, \ldots, \xi_m)$.

The components $h^\alpha_{ij,k}$ of the covariant differential $\tilde{\nabla}h$ of h are defined by

$$h^\alpha_{ij,k}\theta^k = dh^\alpha_{ij} - h^\alpha_{aj}\omega^a_i - h^\alpha_{ia}\omega^a_j + h^\beta_{ij}\omega^\alpha_\beta.$$

In terms of the components, the equations of Gauss, Codazzi, and Ricci are given by

$$\overline{R}_{hijk} = R_{hijk} + \sum_\alpha (h^\alpha_{ij}h^\alpha_{hk} - h^\alpha_{ik}h^\alpha_{hj}),$$

$$\overline{R}^\alpha_{ijk} = h^\alpha_{ik,j} - h^\alpha_{ij,k},$$

$$\overline{R}^\alpha_{\beta jk} = R^\perp{}^\alpha{}_{\beta jk} - \sum_a (h^\alpha_{ja}h^\beta_{ak} - h^\beta_{ja}h^\alpha_{ak}).$$

Let $(x^1, \ldots, x^n)$ be a local coordinate system on a neighborhood U of M and $(y^1, \ldots, y^{n+m})$ be a local coordinate system on a neighborhood V of $\overline{M}$ such that $f(U) \subset V$. Regarding the differential f_* of the immersion f as a tensor field of type $(0,1)$ with values in $f^*T(\overline{M})$, we denote the components of f_* with respect to $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^{n+m})$ by B^A_i $(i = 1, \ldots, n;\ A = 1, \ldots, n+m)$. Then we have $B^A_i = \partial y^A/\partial x^i$. We denote by ∇' the van der Waerden–Bortolotti covariant derivative for ∇ and $\overline{\nabla}$. Then the components $B^A_{i,j}$ of $\nabla' f_*$ are given by

$$B^A_{i,j} = \partial_j B^A_i - \left\{ {a \atop ji} \right\} B^A_a + B^C_j B^B_i \overline{\left\{ {A \atop CB} \right\}},$$

where $\partial_j = \partial/\partial x^j$, $\left\{ {h \atop ji} \right\}$, and $\overline{\left\{ {A \atop CB} \right\}}$ are the Christoffel symbols of the Riemannian metrics g and $\overline{g}$, respectively.

Let $(\xi_1, \ldots, \xi_m)$ be an orthonormal moving frame of $T^\perp(M)$ on U and ξ^A_α be the components of ξ_α with respect to $(y^1, \ldots, y^{n+m})$. Then we have

$$B^A_{i,j} = h^\alpha_{ij}\xi^A_\alpha,$$

where h_{ij} are the components of the second fundamental tensor with respect to $(\partial/\partial x^1, \ldots, \partial/\partial x^n)$ and $(\xi_1, \ldots, \xi_m)$.

A tensor field K with values in $T^{\perp}(M)$ can be regarded as a tensor field with values in $f^*T(M)$, and $\tilde{\nabla}K$ is the normal component of $\nabla' K$. For example, if we regard the second fundamental tensor h as a tensor field with values in $f^*T(\overline{M})$, the components of h with respect to the coordinates $(x^1, \ldots, x^n)$ and $(y^1, \ldots, y^{n+m})$ are equal to $B^A_{i,j}$, and we have

$h^{\alpha}_{ij,k} = B^A_{i,jk} \xi^B_{\alpha} \bar{g}_{AB}$.

References

[1] E. Cartan, Leçon sur la géometrie des espaces de Riemann, Gauthier-Villars, second edition, 1946; English translation, Geometry of Riemannian Spaces, Math-Sci Press, 1983.
[2] B.-Y. Chen, Geometry of submanifolds, Dekker, 1973.
[3] S. Helgason, Differential geometry, Lie groups, and symmetric spaces, Academic Press, 1978.
[4] S. Kobayashi and K. Nomizu, Foundations of differential geometry, Interscience, I, 1963; II, 1969.
[5] J. A. Schouten, Ricci-calculus, Springer, second edition, 1954.
[6] K. Yano, Integral formulas in Riemannian geometry, Dekker, 1970.
[7] G. de Rham, Variétés différentiables, Actualités Sci. Ind., Hermann, second edition, 1960.
[8] L. P. Eisenhart, Riemannian geometry, Princeton Univ. Press, second edition, 1949.
[9] R. L. Bishop and S. I. Goldberg, Tensor analysis on manifolds, Macmillan, 1968.

418 (IX.20) Theory of Singularities

A. Introduction

Let $f_1, f_2, \ldots, f_r$ be †holomorphic functions defined in an open set U of the complex space $\mathbf{C}^n$. Let X be the analytic set $f_1^{-1}(0) \cap \ldots \cap f_r^{-1}(0)$. Let $z_0 \in X$, and let $g_1, \ldots, g_s$ be a system of generators of the ideal $\mathscr{I}(X)_{z_0}$ of the germs of the holomorphic functions which vanish identically on a neighborhood of z_0 in X. z_0 is called a **simple point** of X if the matrix $(\partial g_i/\partial z_j)$ attains its maximal rank, say k, at $z = z_0$. In this case, X is a †complex manifold of dimension $n-k$ near z_0. Otherwise, z_0 is called a **singular point** of X.

B. Resolution of Singularities

Let X be a complex analytic space, and let Y be its singular locus. A **resolution of the singularity** of X is a pair of a complex manifold $\tilde{X}$ and a proper surjective holomorphic mapping $\pi: \tilde{X} \to X$ such that the restriction $\pi|_{\tilde{X} - \pi^{-1}(Y)}$ is biholomorphic and $\tilde{X} - \pi^{-1}(Y)$ is dense in $\tilde{X}$. H. Hironaka proved that there exists a resolution for any X such that $\pi^{-1}(Y)$ is a divisor in $\tilde{X}$ with only †normal crossings [16, 17].

Suppose that a compact connected analytic subset $\tilde{Y}$ of a complex manifold $\tilde{X}$ has a †strongly pseudoconvex neighborhood in $\tilde{X}$. Then the contraction $\tilde{X}/\tilde{Y}$ naturally has a structure of a †normal complex analytic variety such that the projection $\tilde{X} \to \tilde{X}/\tilde{Y}$ is a resolution of $\tilde{X}/\tilde{Y}$ (H. Grauert [14]).

C. Two-Dimensional Singularities

Let X be a normal 2-dimensional analytic space. Then the singular points of X are discrete.

Among the resolutions of X, there exists a unique resolution $\pi: \tilde{X} \to X$ with the following universal property: For any resolution $\pi': \tilde{X}' \to X$, there exists a unique mapping $\rho: \tilde{X}' \to \tilde{X}$ with $\pi' = \pi \circ \rho$. This resolution is called the **minimal resolution**.

Let $\pi: \tilde{X} \to X$ be a resolution of a singular point x of X, and let A_i $(i = 1, \ldots, m)$ be the irreducible components of $\pi^{-1}(x)$. The matrix $(A_i \cdot A_j)$ of the †intersection numbers is known to be negative definite (P. Du Val [12]).

The resolution $\pi: \tilde{X} \to X$ is called **good** if (i) each A_i is nonsingular, (ii) $A_i \cap A_j$ $(i \neq j)$ is at most one point and the intersection is transverse and (iii) no three A_i's meet at a point. For a given good resolution $\pi: \tilde{X} \to X$, we associate a diagram in which the vertices v_i $(i = 1, \ldots, m)$ correspond to A_i $(i = 1, \ldots, m)$ and v_i and v_j are joined by a segment if and only if $A_i \cap A_j \neq \varnothing$.

The **geometric genus** $p_g(X, x)$ of a singular point $x \in X$ is the dimension of the †stalk at x of the first direct image sheaf $R^1\pi_*\mathscr{O}_{\tilde{X}}$, where $\pi: \tilde{X} \to X$ is a resolution of $x \in X$ and $\mathscr{O}_{\tilde{X}}$ is the †structure sheaf of $\tilde{X}$. The definition is independent of the choice of the resolution, and $p_g(X, x)$ is a finite integer.

Among the positive cycles of the form $Z = \sum_{i=1}^{n} n_i A_i$ (i.e., $n_i \geqslant 0$) such that $Z \cdot A_i < 0$ for each $i = 1, \ldots, m$, there exists a smallest one Z_0, which is called the **fundamental cycle** [3].

(1) **Rational singularities**. A singular point x of X is called **rational** if $p_g(X, x) = 0$. (The singularity (X, x) is also called rational even when $\dim X \geqslant 3$ if the direct image sheaf $R^i\pi_*\mathscr{O}_{\tilde{X}} = 0$ for $i > 0$.)

For a rational singularity $x \in X$, the †multiplicity of X at x equals $-Z_0^2$ and the local embedding dimension of X at x is $-Z_0^2+1$. Hence a rational singularity with multiplicity 2, which is called a **rational double point**, is a hypersurface singularity. The following weighted homogeneous polynomials (→ Section D) give the complete list of the defining equations up to analytic isomorphism:

$A_n: x^{n+1}+y^2+z^2$,

weights $(1/(n+1),\ 1/2,\ 1/2)$, $n \geqslant 1$;

$D_n: x^{n-1}+xy^2+z^2$,

weights $(1/(n-1),\ (n-2)/2(n-1),\ 1/2)$, $n \geqslant 4$;

$E_6: x^4+y^3+z^2$,

weights $(1/4,\ 1/3,\ 1/2)$;

$E_7: x^3y+y^3+z^2$,

weights $(2/9,\ 1/3,\ 1/2)$;

$E_8: x^5+y^3+z^2$,

weights $(1/5,\ 1/3,\ 1/2)$,

where the labels appearing at the left are given according to the coincidence of the diagram of the respective minimal resolutions and the †Dynkin diagrams. Rational double points have many different characterizations [11].

The generic part of the singular locus of the unipotent variety of a †complex simple Lie group G (= the orbit of the subregular †unipotent elements in G) is locally expressed as the product of a rational double point and a polydisk. The †universal deformation of a rational double point and its †simultaneous resolution are constructed by restricting the following diagram on a transverse slice to the subregular unipotent orbit (Brieskorn [7]; [34]):

$$\begin{array}{ccc} Y & \longrightarrow & G \\ \downarrow & & \downarrow \\ T & \longrightarrow & T/W \end{array}$$

where T is a †Cartan subgroup of G with the action of the Weyl group W, $G \to T/W$ is the quotient mapping by the †adjoint action of G and $Y = \{(x, B) \mid x \in G$ and B is a †Borel subgroup of G with $x \in B\}$, and other morphisms are defined naturally so that the diagram commutes. Here, $Y \to T$ is the simultaneous resolution of the morphism $G \to T/W$.

(2) **Quotient singularities**. A singular point $x \in X$ is called a **quotient singularity** if there exists a neighborhood of x which is analytically isomorphic to an orbit space U/G, where U is a neighborhood of 0 in $\mathbf{C}^2$ and G is a finite group of analytic automorphisms of U with the unique fixed point 0. The quotient singularities are rational, and their resolutions have been well studied [6]. U/G has a rational double point at 0 if and only if G is conjugate to a nontrivial finite subgroup of $SU(2)$.

(3) **Elliptic singularities**. The singularity (X, x) is called **minimally elliptic** if $p_g(X, x) = 1$ and (X, x) is Gorenstein [23]. The following are examples of minimally elliptic singularities.

A singular point $x \in X$ is called **simply elliptic** if the exceptional set A of the minimal resolution is a smooth †elliptic curve [33]. When $A^2 = -1, -2, -3$, (X, x) is a hypersurface singularity given by the following weighted homogeneous polynomials:

$\tilde{E}_6: x^3+y^3+z^3+axyz$,

weights $(1/3,\ 1/3,\ 1/3)$, $A^2 = -3$;

$\tilde{E}_7: x^4+y^4+z^2+axyz$,

weights $(1/4,\ 1/4,\ 1/2)$, $A^2 = -2$;

$\tilde{E}_8: x^6+y^3+z^2+axyz$,

weights $(1/6,\ 1/3,\ 1/2)$, $A^2 = -1$,

(4) **Cusp singularities**. A singular point $x \in X$ is called a **cusp singularity** if the exceptional set of the minimal resolution is either a single rational curve with a †node or a cycle of smooth rational curves. Cusp singularities appear as the boundary of †Hilbert modular surfaces [18]. The hypersurface cusp singularities are given by the polynomials

$T_{p,q,r}: x^p+y^q+z^r+axyz$,

where $1/p+1/q+1/r<1$ and $a \neq 0$.

D. The Milnor Fibration for Hypersurface Singularities

Let V be an analytic set in $\mathbf{C}^N$, and take a point $z_0 \in V$. Let $S_\varepsilon = S(z_0, \varepsilon)$ be a $(2N-1)$-dimensional sphere in $\mathbf{C}^N$ with center z_0 and radius $\varepsilon > 0$, and let $K_\varepsilon = V \cap S_\varepsilon$. If ε is sufficiently small, the topological type of the pair $(S_\varepsilon, K_\varepsilon)$ is independent of ε [27]. By virtue of this fact, the study of singular points constitutes an important aspect of the application of topology to the theory of functions of several complex variables.

A singular point z_0 of V is said to be **isolated** if, for some open neighborhood W of z_0 in $\mathbf{C}^N$, $W \cap V - \{z_0\}$ is a smooth submanifold of $W - \{z_0\}$. In that case, K_ε is a closed smooth submanifold of S_ε, and the diffeomorphism type of $(S_\varepsilon, K_\varepsilon)$ is independent of (sufficiently small) $\varepsilon > 0$. So far, the topological study of such singular points has been primarily focused on isolated singularities. When V is a plane curve, that is, $N = 2$ and $r = 1$, all the singular points of V are isolated, and the submanifold K_ε of the 3-sphere S_ε can be described as an iterated torus link, where type numbers are completely determined by the †Puiseaux expansion of the defining equation f of V at the point z_0 [5]. In 1961, D. Mumford, using a resolution argument, showed that if an algebraic surface V is †normal at z_0 and if the closed 3-manifold K_ε is simply connected, then K_ε is diffeomorphic to the 3-sphere and z_0 is nonsingular [29]. The following theorem in the higher-dimensional case is due to E. Brieskorn [8] (1966):

Every †homotopy $(2n-1)$-sphere $(n \neq 2)$ that is a boundary of a †π-manifold is diffeomorphic to the K_ε of some complex hypersurface defined by an equation of the form $f(z)=z_1^{a_1}+\ldots+z_{n+1}^{a_{n+1}}=0$ at the origin in $\mathbf{C}^{n+1}$, provided that $n \neq 2$. The hypersurface of this type is called the **Brieskorn variety**. Inspired by Brieskorn's method, J. W. Milnor developed topological techniques for the study of hypersurface singularities and obtained results such as the **Milnor fibering theorem**, which can be briefly stated as follows:

Suppose that V is defined by a single equation $f(z)=0$ in the neighborhood of $z_0 \in \mathbf{C}^{n+1}$. Then there is an associated smooth †fiber bundle $\varphi: S_\varepsilon - K_\varepsilon \to S^1$, where $\varphi(z)=f(z)/|f(z)|$ for $z \in S_\varepsilon - K_\varepsilon$. The fiber $F=\varphi^{-1}(p)$ $(p \in S^1)$ has the homotopy type of a finite CW-complex of dimension n, and K_ε is $(n-2)$-connected.

Suppose that z_0 is an isolated critical point of f. Then F has the homotopy type of a †bouquet of spheres of dimension n [27]. **The Milnor number** $\mu(f)$ of f is defined by the nth Betti number of F, and it is equal to $\dim_{\mathbf{C}} \mathcal{O}_{\mathbf{C}^{n+1},z_0}/(\partial f/\partial z_1, \ldots, \partial f/\partial z_{n+1})$, where $\mathcal{O}_{\mathbf{C}^{n+1},z_0}$ is the ring of the germs of analytic functions of $n+1$ variables at $z=z_0$. **The Milnor monodromy** h_* is the automorphism of $H_n(F)$ that is induced by the action of the canonical generator of the fundamental group of the base space S^1. The †Lefschetz number of h_* is zero if z^0 is a singular point of V. Let $\Delta(t)$ be the characteristic polynomial of h_*. Then K_ε is a homology sphere if and only if $\Delta(1)=\pm 1$ [27]. It is known that $\Delta(t)$ is a product of †cyclotomic polynomials.

The diffeomorphism class of $(S_\varepsilon, K_\varepsilon)$ is completely determined by the congruence class of the linking matrix $L(e_i, e_j)$ $(1 \leqslant i,j \leqslant \mu(f))$, where $e_1, \ldots, e_{\mu(f)}$ is an integral basis of $H_n(F)$ and $L(e_i, e_j)$ is the †linking number [21, 10].

The Milnor fibration is also described in the following way. Let $E(\varepsilon, \delta)$ be the intersection of $f^{-1}(D_\delta^*)$ and $B(\varepsilon)$, the open disk of radius ε and center z_0, where D_δ^* is $\{\eta \in \mathbf{C} \mid 0<|\eta|<\delta\}$. The restriction of f to $E(\varepsilon, \delta)$ is a †locally trivial fibration over D_δ^* if δ is sufficiently smaller than ε [27].

Let $f(z)$ be an analytic function; suppose that $f(0)=0$ and let $\sum_{p \in \mathbf{N}^{n+1}} a_p z^p$ be the Taylor expansion of f at $z=0$. Let $\Gamma_+(f)$ be the convex hull of the union of $\{p+(\mathbf{R}^+)^{n+1}\}$ for $p \in \mathbf{N}^{n+1} \subset \mathbf{R}^{n+1}$ with $a_p \neq 0$, where $\mathbf{R}^+=\{x \in \mathbf{R} \mid x \geqslant 0\}$, and let $\Gamma(f)$ be the union of compact faces of $\Gamma_+(f)$. We call $\Gamma(f)$ the **Newton boundary** of f in the **coordinates** $z_1, \ldots, z_{n+1}$. For a closed face Δ of $\Gamma(f)$ of any dimension, let $f_\Delta(z)=\sum_{p \in \Delta} a_p z^p$. We say that f has a **nondegenerate Newton boundary** if $(\partial f_\Delta/\partial z_1, \ldots, \partial f_\Delta/\partial z_{n+1})$ is a nonzero vector for any $z \in (\mathbf{C}^*)^{n+1}$ and any $\Delta \in \Gamma(f)$. Suppose that f has a nondegenerate Newton boundary and 0 is an isolated critical point of f. Then the Milnor fibration of f is determined by $\Gamma(f)$ and $\mu(f)$, and the characteristic polynomial can be explicitly computed by $\Gamma(f)$ [22, 38].

$f(z)$ is called **weighted homogeneous** if there exist positive rational numbers $r_1, \ldots, r_{n+1}$, which are called **weights**, such that $a_p=0$ if $\sum_{i=1}^{n+1} p_i r_i \neq 1$. An analytic function $f(z)$ with an isolated critical point at 0 is weighted homogeneous in suitable coordinates if and only if f belongs to the ideal $(\partial f/\partial z_1, \ldots, \partial f/\partial z_{n+1})$ (K. Saito [32]). Suppose that $f(z)$ is a weighted homogeneous polynomial with an isolated critical point at 0. Then the Milnor fibration of f is uniquely determined by the weights, and $\mu(f)=\prod_{i=1}^{n+1}\left(\frac{1}{r_i}-1\right)$. The surface $f^{-1}(0)$ for $n=2$ is a rational double point if and only if $\sum_{i=1}^{3} r_i>1$.

E. Unfolding Theory

An **unfolding** of a germ of an analytic function $f(z)$ at 0 is a germ of an analytic function $F(z,t)$, where $t \in \mathbf{C}^m$ (m is finite) such that $F(z,0)=f(z)$. We assume that f has an isolated critical point at 0. Among all the unfoldings of f, there exists a universal one, in a suitable sense, that is unique up to a local analytic isomorphism. It is called the **universal unfolding** of f [36, 37, 26] (→ 51 Catastrophe Theory). Explicitly it can be given by $F(z,t)=f(z)+t_1\varphi_1(z)+\ldots+t_\mu\varphi_\mu(z)$, where $\varphi_i(z)$ $(i=1, \ldots, \mu)$ are holomorphic functions which form a $\mathbf{C}$-basis of the Jacobi ring $\mathcal{O}_{\mathbf{C}^{n+1},0}/(\partial f/\partial z_1, \ldots, \partial f/\partial z_{n+1})$ $(\mu=\mu(f))$.

In the universal unfolding $F(z,t)$ of f, the set of points (z_0, t_0) such that $F(z, t_0)$ has an isolated critical point at z_0 with the Milnor number $\mu(f)$ and $F(z_0, t_0)=0$ forms an analytic set at $(z,t)=0$. **The modulus number** of f is the dimension of this set at 0. This set is sometimes called the **μ-constant stratum**. Let g be a germ of an analytic function. g is said to be **adjacent** to f (denoted by $f \to g$), if there exists a sequence of points $(z(m), t(m))$ in $\mathbf{C}^{n+1} \times \mathbf{C}^\mu$ that converges to the origin such that the term of $F(z, t(m))$ at $z(m)$ is equivalent to g. Adjacency relations are important for the understanding of the degeneration phenomena of functions. The unfolding theory can be considered in exactly the same way as that for the germ of a real-valued smooth function that is finitely determined [36, 26].

The germs of analytic functions with modulus number 0, 1, and 2 are called **simple**, **unimodular**, and **bimodular**, respectively. They were classified by V. I. Arnold [1] (→ Appendix A, Table 5.V). Simple germs correspond to the equations for the rational double points, and unimodular germs define simply elliptic singularities or cusp singularities. Every unimodular or bimodular germ defines a singularity with $p_g = 1$.

F. Picard-Lefschetz Theory

Let $f(z)$ be a holomorphic function such that $f(0)=0$ and 0 is an isolated critical point with the Milnor number μ. Let $F(z,t)$ be a universal unfolding of f at 0. Let $f: E(\varepsilon,\delta) \to D_\delta^*$ be the Milnor fibration of f by the second description in Section D. There exists a positive number r and a codimension 1 analytic subset Δ (called the **bifurcation set**) of $B'(r)$, the open disk of radius r with the center 0 in the parameter space $\mathbf{C}^\mu$, such that for any $t_0 \in B'(r)-\Delta$, $f_{t_0} = F|_{B(\varepsilon)\times t_0}$ has μ different nondegenerate critical points in $B(\varepsilon)$. Let $p_1, \dots, p_\mu$ be the critical points of f_{t_0}. For each p_i, one can choose local coordinates $(y_1, \dots, y_{n+1})$ so that $f_{t_0}(y) = f_{t_0}(p_i) + y_1^2 + \dots + y_{n+1}^2$. Such an f_{t_0} is called a **Morsification** of f.

Let B_i be a small disk with center p_i in $\mathbf{C}^{n+1}$. Then for any q_i which is near enough to $f_{t_0}(p_i)$, the intersection $f_{t_0}^{-1}(q_i) \cap B_i$ is diffeomorphic to the tangent disk bundle of the sphere S^n. The **vanishing cycle** e_i is the corresponding n-dimensional homology class of $f_{t_0}^{-1}(q_i) \cap B_i$. (We fix q_i.) The self-intersection number of e_i is given by

$$\langle e_i, e_i \rangle = \begin{cases} 2(-1)^{n(n-1)/2}, & n \text{ even}, \\ 0, & n \text{ odd}. \end{cases}$$

For a sufficiently small $t_0 \in B'(r)-\Delta$, one has the following: (i) $|f_{t_0}(p_i)| < \delta$; (ii) the restriction of f_{t_0} to E is a fiber bundle over D', where $D' = \{w \in \mathbf{C} \,|\, |w| \leqslant \delta$, and $w \neq f_{t_0}(p_i)$ for $i=1, \dots, \mu\}$ and $E = f_{t_0}^{-1}(D') \cap B(\varepsilon)$; (iii) the restriction of the above fibration to $\{w \,|\, |w| = \delta\}$ is equivalent to the restriction of the Milnor fibration of f to $\{w \,|\, |w|=\delta\}$. Let w_0 be a fixed point of D', and let $F = f_{t_0}^{-1}(w_0) \cap E$. Then F is diffeomorphic to the Milnor fiber of f. Let l_i be a simple path from w_0 to q_i, and let γ_i be the loop $|w - f_{t_0}(p_i)| = |q_i - f_{t_0}(p_i)|$. We suppose that the union of the l_i is contractible to w_0. By parallel translation of the vanishing cycle e_i along l_i, we consider $e_i \in H_n(F)$. The collection $\{e_i \,|\, i = 1, \dots, \mu\}$ is an integral basis of $H_n(F)$, which is called a **strongly distinguished basis** (Fig. 1).

Now let h_i be the linear transformation of $H_n(F)$ that is induced by the parallel translation along $l_i \gamma_i l_i^{-1}$. The **Picard-Lefschetz formula** says that

$$h_i(e) = e - (-1)^{n(n-1)/2} \langle e, e_i \rangle \cdot e_i \quad \text{for} \quad e \in H_n(F).$$

Here $\langle\ ,\ \rangle$ is the intersection number in $H_n(F)$. For n even, h_i is a †reflection.

The Milnor monodromy h_* of f is equal to the composition $h_1 \dots h_\mu$ under a suitable ordering of the h_i. The subgroup of the group of linear isomorphisms of $H_n(F)$ generated by $h_1, \dots, h_\mu$ is called the **total monodromy group**.

When f is a simple germ and $n \equiv 2 \bmod 4$, the total monodromy group is isomorphic to the †Weyl group of the corresponding Dynkin diagram. Even-dimensional simple singularities are the only ones for which the monodromy group is finite. These are also characterized as the singularities with definite intersection forms.

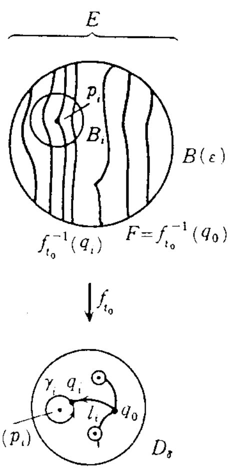


Fig. 1

G. Stratification Theory

The notion of **Whitney stratification** was first introduced by H. Whitney to study the singularities of analytic varieties [39] and was developed by R. Thom for the general case [37].

Let X and Y be submanifolds of the space $\mathbf{R}^n$. We say that the pair (X, Y) satisfies the **Whitney condition (b) at a point** $y \in Y$ if the following holds: Let x_i $(i=1,2,\dots)$ and y_i $(i=1,2,\dots)$ be sequences in X and Y, respectively, that converge to y. Suppose that the tangent space $T_{x_i}X$ converges to a plane T in the corresponding Grassmannian space and the secant $\overline{x_i y_i}$ converges to a line L. Then $L \subset T$. We say that (X, Y) satisfies the **Whitney condition (b)** if it satisfies the Whitney condition (b) at any point $y \in Y$. Let h be a local diffeomorphism of a neighborhood of y. One can see that $(h(X), h(Y))$ satisfies the Whitney condition (b) at $h(y)$ if (X, Y) satisfies it at y. Thus the Whitney condition can be considered for a pair of submanifolds X and Y of a manifold M using a local coordinate system. Let S be a subset of a manifold M, and let $\mathcal{S}$ be a family of submanifolds of M. $\mathcal{S}$ is called a **Whitney prestratification** of S if $\mathcal{S}$ is a locally finite disjoint cover of S satisfying the following: (i) For any $X \in \mathcal{S}$, the frontier $\bar{X} - X$ is a union of $Y \in \mathcal{S}$; (ii) for any pair (X, Y) $(X, Y \in \mathcal{S})$, the Whitney condition (b) is satisfied. A submanifold X in $\mathcal{S}$ is called a **stratum**. There exists a canonical partial order in $\mathcal{S}$ that is defined by $X < Y$ if and only if $X \subset \bar{Y} - Y$.

Let V be an analytic variety, and let $\mathcal{S}$ be an analytic stratification of V that satisfies the frontier condition (i). Then there exists a Whitney prestratification $\mathcal{S}'$ that is finer than $\mathcal{S}$ (Whitney [39]).

For a given Whitney prestratification $\mathcal{S}$, one can construct the following **controlled tubular neighborhood system**: For each $X \in \mathcal{S}$, a †tubular neighborhood $|T_X|$ of X in M and the projection $\pi_X : |T_X| \to X$ and a tubular function $\rho_X : |T_X| \to \mathbf{R}^+$ (= the square of a norm under the identification of $|T_X|$ with the †normal disk bundle of X) are given such that the commutation relations

$$\pi_X \cdot \pi_Y(m) = \pi_X(m), \ \rho_X \pi_Y(m) = \rho_X(m) \qquad \text{for } m \in M, \ X < Y.$$

are satisfied whenever both sides are defined.

By virtue of this, the notions of vector fields and their integral curves can be defined on a Whitney prestratified set so that several important results on a differentiable manifold can be generalized to the case of stratified sets. For example, the following is **Thom's first isotopy lemma**: Let M and P be differentiable manifolds, and let $(S, \mathcal{S})$ be a Whitney prestratified subset of M. Let $f: S \to P$ be a continuous mapping that is the restriction of a differentiable mapping from M to P. Suppose that the restriction of f to each stratum X of $\mathcal{S}$ is a proper submersion onto P. Then $f: S \to P$ is a fiber bundle [37].

H. b-Functions

Let $f(z)$ be a germ of an analytic function in $\mathbf{C}^{n+1}$ with $f(0)=0$. The **b-function** of f at 0 is the monic polynomial $b_f(s)$ of lowest degree among all polynomials $b(s)$ with the following property [4, 20]: There exists a differential operator $P(z, \partial/\partial z, s)$, which is a polynomial in s, such that $b(s) f^s(z) = P(z, \partial/\partial z, s) f^{s+1}(z)$. Since $b_f(s)$ is always divisible by $s+1$, we define $\tilde{b}_f(s) = b_f(s)/(s+1)$. All the roots of $\tilde{b}_f(s)=0$ are negative rational numbers (M. Kashiwara [20]. When f has an isolated critical point at 0, the set $\{\exp(2\pi i \alpha) \mid \alpha$ is a root of $b_f(s)=0\}$ coincides with the set of eigenvalues of the Milnor monodromy [25].

The name "b-function" is due to M. Sato. He first introduced it in the study of †prehomogeneous vector spaces. Some authors call it the **Bernstein (Bernshteĭn) polynomial**.

I. Hyperplane Sections

Let V be an algebraic variety of complex dimension k in the complex projective space $\mathbf{P}^n$. Let L be a hyperplane that contains the singular points of V. Then the †relative homotopy group $\pi_i(V, V \cap L)$ is zero for $i < k$. Thus the same assertion is true for the †relative homology groups (S. Lefschetz [24]; [28]).

Let f be a holomorphic function defined in the neighborhood of $0 \in \mathbf{C}^{n+1}$ and $f(0)=0$. Let H be the hypersurface $f^{-1}(0)$. There exists a †Zariski open subset U of the space ($=\mathbf{P}^n$) of hyperplanes such that for each $L \in U$, there exists a positive number ε such that $\pi_i(B(r) - H, (B(r) - H) \cap L) = 0$ for $i < n$ and $0 < r \leqslant \varepsilon$, where $B(r)$ is a disk of radius r (D. T. Lê and H. Hamm [15]). This implies the following theorem of Zariski: Let V be a hypersurface of $\mathbf{P}^n$, and let $\mathbf{P}^2$ be a general plane in $\mathbf{P}^n$. Then the fundamental group of $\mathbf{P}^n - V$ is isomorphic to the fundamental group of $\mathbf{P}^2 - C$, where $C = V \cap \mathbf{P}^2$. The fundamental group of $\mathbf{P}^2 - C$ is an Abelian group if C is a nodal curve [9, 13].

Suppose that f has an isolated critical point at 0. Let $\mu^{(n+1)}$ be the Milnor number $\mu(f)$. Take a generic hyperplane L. The Milnor number of $f|_L$ is well defined, and we let $\mu^{(n)} = \mu(f|_L)$. Similarly one can define $\mu^{(i)}$ of f and let $\mu^* = (\mu^{(n+1)}, \mu^{(n)}, \ldots, \mu^{(1)})$. Let $f_t(z)$ be a deformation of f. Each f_t has an isolated critical point at 0, and t is a point of a disk D of the complex plane. Let $W = \{(z, t) \mid f_t(z) = 0\}$ and $D' = \{0\} \times D$. $W - D'$ and D' satisfy the Whitney condition (b) if and only if $\mu^*(f_t)$ is invariant under the deformation [35]. The Whitney condition (b) implies topological triviality of the deformation.

References

[1] V. I. Arnol'd, Local normal forms of functions, Inventiones Math., 35 (1976), 87–109.
[2] V. I. Arnol'd, Normal forms for functions near degenerate critical points, the Weyl group of A_k, D_k, E_k and Lagrangian singularities, Functional Anal. Appl., 6 (1972), 254–272. (Original in Russian, 1972.)

[3] M. Artin, On isolated rational singularities of surfaces, Amer. J. Math., 88 (1966), 129–136.
[4] I. N. Bernshteĭn, The analytic continuation of generalized functions with respect to parameter, Functional Anal. Appl., 6 (1972), 273–285. (Original in Russian, 1972.)
[5] K. Brauer, Zur Geometrie der Funktionen zweier komplexen Veränderlichen III, Abh. Math. Sem. Hamburg, 6 (1928), 8–54.
[6] E. Brieskorn, Rationale Singularitäten Komplexer Flächen, Inventiones Math., 4 (1968), 336–358.
[7] E. Brieskorn, Singular elements of semi-simple algebraic groups, Actes Congrès Int. Math., 1970, vol. 2, 279–284.
[8] E. Brieskorn, Beispiele zur Differentialtopologie, Inventiones Math., 2 (1966), 1–14.
[9] P. Deligne, Le groupe fondamental du complément d'une courbe plane n'ayant que des points doubles ordinaires est abelian, Sém. Bourbaki, no. 543, 1979–1980.
[10] H. Durfee, Fibered knots and algebraic singularities, Topology, 13 (1974), 47–59.
[11] A. H. Durfee, Fifteen characterizations of rational double points and simple critical points, Enseignement Math., 25 (1979), 131–163.
[12] P. Du Val, On isolated singularities of surfaces which do not affect the condition of adjunction I–III, Proc. Cambridge Philos. Soc., 30 (1933/34), 453–465, 483–491.
[13] W. Fulton, On the fundamental group of the complement of a node curve, Ann. Math., (2) 111, (1980), 407–409.
[14] H. Grauert, Über Modifikationen und exzeptionelle anaytische Mengen, Math. Ann., 146 (1962), 331–368.
[15] H. Hamm and D. T. Lê, Un théorème de Zariski du type de Lefschetz, Ann. Sci. Ecole Norm. Sup., 3, 1973.
[16] H. Hironaka, Resolution of singularities of an algebraic variety over a field of characteristic zero, Ann. Math., (2) 79 (1964), 109–326.
[17] H. Hironaka, Bimeromorphic smoothing of a complex analytic space, Acta Math. Vietnamica, 2, no. 2 (1977), 106–168.
[18] F. Hirzebruch, Hilbert modular surfaces, Enseignement Math., 19 (1973), 183–281.
[19] P. Holm, Real and complex singularities, Sijthoff and Noordhoff, 1976.
[20] M. Kashiwara, B-functions and holonomic systems, Inventiones Math., 38 (1976), 33–54.
[21] M. Kato, A classification of simple spinnable structure on a 1-connected Alexander manifold, J. Math. Soc. Japan, 26 (1974), 454–463.
[22] A. G. Kouchnirenko, Polyèdre de Newton et nombres de Milnor, Inventiones Math., 32 (1976), 1–31.
[23] H. Laufer, On minimally elliptic singularities, Amer. J. Math., 99 (1977), 1257–1295.
[24] S. Lefschetz, L'analysis situs et la géométrie algébrique, Gauthiers-Villars, 1924.
[25] B. Malgrange, Le polynome de Bernstein d'une singularité isolée, Lecture notes in math. 459, Springer, 1976, 98–119.
[26] J. N. Mather, Stability of C^∞-mappings. I, Ann. Math., (2) 87 (1968), 89–104; II, (2) 89 (1969), 254–291; III, Publ. Math. Inst. HES, 35 (1970), 301–336; IV, Lecture notes in math. 192, Springer, 1971, 207–253.
[27] J. W. Milnor, Singular points of complex hypersurfaces, Ann. math. studies 61, Princeton Univ. Press, 1968.
[28] J. W. Milnor, Morse theory, Ann. math. studies 51, Princeton Univ. Press, 1963.
[29] D. Mumford, The topology of normal singularities of an algebraic surface and a criterion for simplicity, Publ. Math. Inst. HES, 9 (1961), 5–22.
[30] F. Pham, Formules de Picard-Lefschetz généralisées et ramification des intégrales, Bull. Soc. Math. France, 93 (1965), 333–367.
[31] E. Picard and S. Simart, Traité des fonctions algébriques de deux variables I, Gauthier-Villars, 1897.
[32] K. Saito, Quasihomogene isolierte Singularitäten von Hyperflächen, Inventiones Math., 14 (1971), 123–142.
[33] K. Saito, Einfach elliptische Singularitäten, Inventiones Math., 23 (1974), 289–325.
[34] P. Slodowy, Simple singularities and simple algebraic groups, Lecture notes in math. 815, Springer, 1980.
[35] B. Teissier, Cycles evanescents, sections planes et conditions de Whitney, Astérisque, 78 (1973), 285–362.
[36] R. Thom, La stabilité topologique des applications polynomiales, Enseignement Math., 8 (1962), 24–33.
[37] R. Thom, Ensembles et morphismes stratifiés, Bull. Amer. Math. Soc., 75 (1969), 240–284.
[38] A. N. Varchenko, Zeta function of the monodromy and Newton's diagram, Inventiones Math., 37 (1976), 253–262.
[39] H. Whitney, Tangents to an analytic variety, Ann. Math., (2) 81 (1964), 496–549.
[40] O. Zariski, Algebraic surfaces, Erg. math. 61, Springer, 1971.

419 (XX.18) Thermodynamics

A. Basic Concepts and Postulates

Thermodynamics traditionally focuses its attention on a particular class of states of a given system called (thermal) equilibrium states, although a more recent extension, called the thermodynamics of irreversible processes, deals with certain nonequilibrium states. In a simple system, an **equilibrium state** is completely specified (up to the shape of the volume it occupies) by the volume V (a positive real number), the **mole numbers** $N_1, \ldots, N_r$ (nonnegative reals) of its chemical components, and the **internal energy** U (real). (More variables might be needed if the system were, e.g., inhomogeneous, anisotropic, electrically charged, magnetized, chemically not inert, or acted on by electric, magnetic, or gravitational fields.) This means that any of the quantities associated with equilibrium states (called **thermodynamical quantities**) of a simple system under consideration is a function of V, $N_1, \ldots, N_r$, and U.

When n copies of the same state are put next to each other and the dividing walls are removed, V, $N_1, \ldots, N_r$, and U for the new state will be n times the old values of these variables under the assumptions that each volume is sufficiently large and that the effects of the boundary walls can be neglected. Thermodynamical quantities behaving in this manner are called **extensive**. Those that are invariant under the foregoing procedure are called **intensive**. More precisely, the thermodynamic variables are defined by homogeneity of degree 1 and 0 as functions of V, $N_1, \ldots, N_r$, and U.

By a shift of the position of the boundary (called an **adiabatic wall** if energy and chemical substances do not move through it) or by transport of energy through the boundary (called a **diathermal wall** if this is allowed) or by transport of chemical components through the boundary (called a **permeable membrane**) (in short, by thermodynamical processes), these variables can change their values. If these shifts or transports are not permitted (especially for a composite system consisting of several simple systems, at its boundary with the outside), the system is called **closed**. Otherwise it is called **open**.

Those equilibrium states that do not undergo any change when brought into contact with each other across an immovable and impermeable diathermal wall (called a **thermal contact**) form an equivalence class. This is sometimes called the **0th law of thermodynamics**. The equivalence class, called the **temperature** of states belonging to it, is an intensive quantity.

The force needed to keep a movable wall at rest, divided by the area of the wall, is called the **pressure**. It is another intensive quantity. For a (slow) change of the volume by an amount dV under a constant pressure P, mechanical work of amount $-P\,dV$ is done on the system. Together with a possible change of the internal energy, say of amount dU, the amount

$$\delta Q = dU - P\,dV \tag{1}$$

of energy is somehow gained (if it is positive) or lost (if it is negative) by the system. This amount of energy is actually transported from or to a neighboring system through diathermal walls so that the total energy for a bigger closed composite system is conserved. This is called the **first law of thermodynamics**, and δQ is called the **heat gain** or **loss** by the system.

If two states of different temperatures T_1 and T_2 are brought into thermal contact, energy is transferred from one, say T_1, to the other (called heat transfer). This defines a binary class relation denoted by $T_1 > T_2$. The Clausius formulation of the **second law of thermodynamics** says that it is impossible to make a positive heat transfer from a state of lower temperature to another state of higher temperature without another change elsewhere. By considering a certain composite system, one reaches the conclusion that there exists a labeling of temperatures by positive real numbers T, called the **absolute temperature**, for which the following is an exact differential:

$$\delta Q/T = (dU - P\,dV)/T = dS. \tag{2}$$

The integral S is an extensive quantity, called the **entropy**. Furthermore, the sum of the entropies of component simple systems in an isolated composite system is nondecreasing during any thermodynamic process, and the following **entropy maximum principle** holds: An isolated composite system reaches an equilibrium at those values of extensive parameters that maximize the sum of the entropies of component simple systems (for constant total energy and volume and within the set of allowed states under a given constraint).

A relation expressing the entropy of a given system as a function of the extensive parameters (specifying equilibrium states) is known as the **fundamental relation** of the system. If it is given as a continuous and differentiable homogeneous function of V, $N_1, \ldots, N_r$, and U and is monotone increasing in U for fixed V, $N_1, \ldots, N_r$, then one can develop the thermodynamics of the system based on the above entropy maximum principle. A relation expressing an intensive parameter as a function of some other independent variables is called an **equation of state**.

Another postulate, which is much less frequently used, is the **Nernst postulate** or the **third law of thermodynamics**, which says that the entropy vanishes at the vanishing absolute temperature.

B. Various Coefficients and Relationships

The partial derivative $\partial/\partial x$ of a function $f(x, y, \ldots)$ with respect to the variable x with the variables $y, \ldots$ fixed is denoted by $(\partial f/\partial x)_{y,\ldots}$. We abbreviate $N_1, \ldots, N_r$ as N in the following.

If the fundamental relation is written as $U = U(V, N_1, \ldots, N_r, S)$ (instead of S being represented as a function of the other quantities), then (2) implies

$$(\partial U/\partial S)_{V,N} = T, \quad (\partial U/\partial V)_{N,S} = -P.$$

The other first-order partial derivatives of U are

$$\mu_j = (\partial U/\partial N_j)_{V,N_1,\ldots,\hat{j},\ldots,N_r,S},$$

with μ_j called the **chemical potential** (or electrochemical potential) of the jth component.

If a system is surrounded by an adiabatic wall (i.e., the system is thermally isolated) and goes through a gradual reversible change (**quasistatic adiabatic process**), then the entropy has to stay constant. If a system is in thermal contact through a diathermal wall with a large system (called the **heat bath**) whose temperature is assumed to remain unchanged during the thermal contact, then the temperature of the system itself remains constant (an **isothermal process**). The decrease of the volume per unit increase of pressure under the latter circumstance is called the **isothermal compressibility** and is given by

$$\kappa_T = -V^{-1}(\partial V/\partial P)_{T,N}.$$

Under constant pressure, the increase of the volume per unit increase of the temperature is called the **coefficient of thermal expansion** and is given by

$$\alpha = V^{-1}(\partial V/\partial T)_{P,N}.$$

Under constant pressure, the amount of (quasistatic) heat transfer into the system per mole required to produce a unit increase of temperature is called the **specific heat at constant pressure** and is given by

$$c_P = N^{-1}T(\partial S/\partial T)_{P,N},$$

where $N = N_1 + \ldots + N_r$. The same quantity under constant volume is called the **specific heat at constant volume** and is given by

$$c_v = N^{-1}T(\partial S/\partial T)_{V,N}.$$

The positivity of c_v is equivalent to the convexity of energy as a function of entropy for fixed values of V and N.

Because of the first-order homogeneity of an extensive quantity as a function of other extensive variables, one can derive an **Euler relation**, such as

$$U = TS - PV + \mu_1 N_1 + \ldots + \mu_r N_r,$$

for a simple system. Its differential form implies the following **Gibbs-Duhem relation**:

$$S\,dT - V\,dP + N_1\,d\mu_1 + \ldots + N_r\,d\mu_r = 0.$$

Because of the identity

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

there arise relationships among second derivatives, known as the **Maxwell relations**:

$$(\partial T/\partial V)_{S,N} = -(\partial P/\partial S)_{V,N},$$

$$(\partial V/\partial S)_{P,N} = (\partial T/\partial P)_{S,N},$$

$$(\partial S/\partial V)_{T,N} = (\partial P/\partial T)_{V,N},$$

$$(\partial S/\partial P)_{T,N} = -(\partial V/\partial T)_{P,N}.$$

By computing the Jacobian of transformations of variables, further relations can be obtained. For example,

$$c_P = c_v + N^{-1}TV\alpha^2/\kappa_T.$$

C. Legendre Transform and Variational Principles

The **Legendre transform** of a function $f(x_1, \ldots, y_1, \ldots)$ relative to the variables x is given by

$$g(p_1, \ldots, y_1, \ldots) = f - \sum_j x_j p_j$$

as a function of the variables $p_j = \partial f/\partial x_j$ and y. The original variables x can be recovered as $-x_j = \partial g/\partial p_j$.

In terms of Legendre transforms, the entropy maximum principle can be reformulated in various forms:

Energy minimum principle: For given values of the total entropy and volume, the equilibrium is reached at those values of unconstrained parameters that minimize the total energy. This principle is applicable in reversible processes where the total entropy stays constant.

Helmholtz free energy minimum principle: For given values of the temperature (equal to that of a heat bath in thermal contact with the system) and the total volume, the equilibrium is reached at those values of the unconstrained parameters that minimize the total Helmholtz free energy, where the **Helmholtz free energy** for a simple system is defined as a function of $T, V, N_1, \ldots, N_r$ by

$$F = U - TS,$$

$$dF = -S\,dT - P\,dV + \mu_1\,dN_1 + \mu_r\,dN_r.$$

Enthalpy minimum principle: For given values of the pressure and the total entropy, the equilibrium is reached at those values of unconstrained parameters that minimize the total enthalpy, where the **enthalpy** for a simple system is defined as a function of S, P, $N_1, \ldots, N_r$ by

$$H = U + PV,$$

$$dH = T\,dS + V\,dP + \mu_1\,dN_1 + \ldots + \mu_r\,dN_r.$$

Gibbs free energy minimum principle: For constant temperature and pressure, the equilibrium is reached at those values of unconstrained parameters that minimize the total Gibbs free energy, where the **Gibbs free energy** for a simple system is given as a function of T, P, $N_1, \ldots, N_r$ by

$$G = U - TS + PV,$$

$$dG = -S\,dT + V\,dP + \mu_1\,dN_1 + \ldots + \mu_r\,dN_r.$$

References

[1] M. W. Zemansky, Heat and thermodynamics, McGraw-Hill, 1951.
[2] H. B. Callen, Thermodynamics, Wiley, 1960.
[3] J. W. Gibbs, The collected works of J. Willard Gibbs, vol. 1, Thermodynamics, Yale Univ. Press, 1948.
[4] M. Planck, Treatise on thermodynamics, Dover, 1945. (Original in German, seventh edition, 1922.)
[5] E. A. Guggenheim, Thermodynamics, North-Holland, 1949.
[6] F. C. Andrews, Thermodynamics: Principles and applications, Wiley, 1971.

420 (XX.8) Three-Body Problem

A. n-Body Problem and Classical Integrals

In the **n-body problem**, we study the motions of n particles $P_i(x_i, y_i, z_i)$ $(i = 1, 2, \ldots, n)$ with arbitrary masses $m_i(>0)$ following †Newton's law of motion,

$$m_i \frac{d^2 w_i}{dt^2} = \frac{\partial U}{\partial w_i}, \qquad i = 1, 2, \ldots, n, \tag{1}$$

where w_i is any one of x_i, y_i, or z_i,

$$U = \sum_{i \neq j} k^2 m_i m_j / r_{ij},$$

with k^2 the gravitation constant, and

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}.$$

Although the one-body and two-body problems have been completely solved, the problem has not been solved for $n > 2$. The **three-body problem** is well known and is important both in celestial mechanics and in mathematics. For $n > 3$ the problem is called the **many-body problem**.

The equations (1) have the so-called ten classical integrals, that is, the **energy integral** $\sum_i (m_i/2)((\dot{x}_i)^2 + (\dot{y}_i)^2 + (\dot{z}_i)^2) - U = \text{constant}$ $(\dot{w} = dw/dt)$, six **integrals of the center of mass** $\sum_i m_i \dot{w}_i = \text{constant}$, $\sum_i m_i w_i = (\sum_i m_i \dot{w}_i)t + \text{constant}$, and three **integrals of angular momentum** $\sum_i m_i (u_i \dot{w}_i - w_i \dot{u}_i) = \text{constant}$ $(u \neq w)$. Using these integrals and eliminating the time t and the ascending node by applying Jacobi's method, the order of the equations (1) can be reduced to $6n - 12$. H. Bruns proved that algebraic integrals cannot be found except for the classical integrals, and H. Poincaré showed that there is no other single-valued integral (Bruns, *Acta Math.*, 11 (1887); Poincaré [2, I, ch. 5]). These results are called **Poincaré-Bruns theorems**. Therefore we cannot hope to obtain general solutions for the equations (1) by †quadrature. General solutions for $n \geqslant 3$ have not been discovered except for certain specific cases.

B. Particular Solutions

Let r_i be the position vector of the particle P_i with respect to the center of mass of the n-body system. A configuration $r \equiv \{r_1, \ldots, r_n\}$ of the system is said to form a **central figure** (or **central configuration**) if the resultant force acting on each particle P_i is proportional to $m_i r_i$, where each proportionality constant is independent of i. The proportionality constant is uniquely determined as $-U/\sum_{i=1}^n m_i r_i^2$ by the configuration of the system. A configuration r is a central figure if and only if r is a †critical point of the mapping $r \mapsto U^2(r) \sum_{i=1}^n m_i r_i^2$ [5, 6]. A rotation of the system, in planar central figure, with appropriate angular velocity is a particular solution of the planar n-body problem.

Particular solutions known for the three-body problem are the **equilateral triangle solution** of Lagrange and the **straight line solution** of Euler. They are the only solutions known for the case of arbitrary masses, and their configuration stays in the central figure throughout the motion.

C. Domain of Existence of Solutions

The solutions for the three-body problem are analytic, except for the collison case, i.e., the case where $\min r_{ij} = 0$, in a strip domain enclosing the real axis of the t-plane (Poincaré, P.

Painlevé). K. F. Sundman proved that when two bodies collide at $t=t_0$, the solution is expressed as a power series in $(t-t_0)^{1/3}$ in a neighborhood of t_0, and the solution which is real on the real axis can be uniquely and analytically continued across $t=t_0$ along the real axis. When all three particles collide, the total angular momentum f with respect to the center of mass must vanish (and the motion is planar) (**Sundman's theorem**); so under the assumption $f \neq 0$, introducing $s=\int^t (U+1)dt$ as a new independent variable and taking it for granted that any binary collision is analytically continued, we see that the solution of the three-body problem is analytic on a strip domain $|\operatorname{Im} s|<\delta$ containing the real axis of the s-plane. The conformal mapping

$$\omega=(\exp(\pi s/2\delta)-1)/(\exp(\pi s/2\delta)+1)$$

maps the strip domain onto the unit disk $|\omega|<1$, where the coordinates of the three particles w_κ, their mutual distances r_{kl}, and the time t are all analytic functions of ω and give a complete description of the motion for all real time (Sundman, *Acta Math.*, 36 (1913); Siegel and Moser [7]).

When a triple collision occurs at $t=t_0$, G. Bisconcini, Sundman, H. Block, and C. L. Siegel showed that as $t \to t_0$, (i) the configuration of the three particles approaches asymptotically the Lagrange equilateral triangle configuration or the Euler straight line configuration, (ii) the collision of the three particles takes place in definite directions, and (iii) in general the triple-collision solution cannot be analytically continued beyond $t=t_0$.

D. Final Behavior of Solutions

Suppose that the center of mass of the three-body system is at rest. The motion of the system was classified by J. Chazy into seven types according to the asymptotic behavior when $t \to +\infty$, provided that the angular momentum f of the system is different from zero. In terms of the †order of the three mutual distances r_{ij} (for large t) these types are defined as follows:
(i) H^+: **Hyperbolic motion.** $r_{ij} \sim t$.
(ii) HP^+: **Hyperbolic-parabolic motion.** r_{13}, $r_{23} \sim t$ and $r_{12} \sim t^{2/3}$.
(iii) HE^+: **Hyperbolic-elliptic motion.** $r_{13}, r_{23} \sim t$ and $r_{12}<a$ (a = finite).
(iv) P^+: **Parabolic motion.** $r_{ij} \sim t^{2/3}$.
(v) PE^+: **Parabolic-elliptic motion.** $r_{13}, r_{23} \sim t^{2/3}$ and $r_{12}<a$.
(vi) L^+: **Lagrange-stable motion** or **bounded motion.** $r_{ij}<a$.
(vii) OS^+: **Oscillating motion.** $\overline{\lim}_{t\to\infty} \sup r_{ij}=\infty$, $\underline{\lim}_{t\to\infty} \sup r_{ij}<\infty$.

Define H^-, HE^-, etc. analogously but with $t \to -\infty$. There are three classes for each of the motions HP, HE, and PE, depending on which of the three bodies separates from the other two bodies and recedes to infinity, denoted by HP_i, HE_i, PE_i ($i=1,2,3$), respectively. The energy constant h is positive for H- and HP-motion, zero for P-motion, and negative for PE-, L-, and OS-motion. For HE-motion, h may be positive, zero, or negative.

We say that a **partial capture** takes place when the motion is H^- for $t \to -\infty$ and HE_i^+ for $t \to +\infty$ (for $h>0$), and a **complete capture** when the motion is HE_i^- for $t \to -\infty$ and L^+ for $t \to +\infty$ (for $h<0$). We say also that an **exchange** takes place when HE_i^- for $t \to -\infty$ and HE_j^+ for $t \to +\infty$ $(t \neq j)$. The probability of complete capture in the domain $h<0$ is zero (J. Chazy, G. A. Merman).

E. Perturbation Theories

The radius of convergence in the s-plane for Sundman's solution is too small and the convergence is too slow in the ω-plane to make it possible to compute orbits of celestial bodies, and for that purpose a perturbation method is usually adopted. When the masses $m_2, \ldots, m_n$ are negligibly small compared with m_1 for the n-body problem, the motion of the nth body is derived as the solution of the two-body problem for m_1 and m_n by assuming $m_2=\ldots=m_{n-1}=0$ as a first approximation, and then the deviations of the true orbit from the ellipse are derived as †perturbations. In the **general theory of perturbations** the deviations are derived theoretically by developing a disturbing function, whereas in the **special theory of perturbations** they are computed by numerical integration. In general perturbation theory, problems concerning convergence of the solution are important, and it becomes necessary to simplify the disturbing function in dealing with the actual relations among celestial bodies. Specific techniques have to be developed in order to compute perturbations for lunar motion, motions of characteristic asteroids, and motions of satellites (e.g., the system of the Sun, Jupiter, and Jovian satellites).

F. The Restricted Three-Body Problem

Since the three-body problem is very difficult to handle mathematically, mathematical interest has been concentrated on the **restricted three-body problem** (in particular, the planar problem) since Hill studied lunar theory in the 19th century. For the restricted three-body problem, the third body, of zero mass, cannot have any influence on the motion of the other two bodies, which are of finite masses and which move uniformly on a circle around the center of mass. In the planar case, let us choose units so that the total mass, the angular velocity of the two bodies about their center of mass, and the gravitation constant are all equal to 1, and let (q_1, q_2) be the coordinates of the third body with respect to a rotating coordinate system chosen in such a way that the origin is at the center of mass and the two bodies of finite masses μ and $1-\mu$ are always fixed on the q_1-axis. Then the equations of motion for the third body are given by a Hamiltonian system:

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i}, \quad i=1,2, \tag{2}$$

with

$$H=\frac{1}{2}(p_1^2+p_2^2)+q_2p_1-q_1p_2-U(q_1,q_2),$$

$$U=\frac{1-\mu}{\sqrt{(q_1+\mu)^2+q_2^2}}+\frac{\mu}{\sqrt{(q_1+\mu-1)^2+q_2^2}}.$$

The equations (2) have the energy integral $H(p,q)=$constant, called **Jacobi's integral**. Siegel showed that there is no other algebraic integral, and it can be proved by applying Poincaré's theorem that there is no other single-valued integral. Regularization of the two singular points for the equations (2) and solutions passing through the singular points were studied by T. Levi-Civita, and solutions tending to infinity were studied by B. O. Koopman.

After reducing the number of variables by means of the Jacobi integral, the equations (2) give rise to a flow in a 3-dimensional manifold of which the topological type was clarified by G. D. Birkhoff (*Rend. Circ. Mat. Palermo*, 39 (1915)). Since this flow has an †invariant measure, the equations have been studied topologically, and important results for the restricted three-body problem, particularly on periodic solutions, have been obtained.

G. Stability of Equilateral Triangular Solutions

Suppose that the origin $q_i=p_i=0$ is an †equilibrium point for an autonomous Hamiltonian system with two degrees of freedom:

$$\frac{dq_i}{dt}=\frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt}=-\frac{\partial H}{\partial q_i}, \quad i=1,2,$$

with the Hamiltonian H being analytic at the origin. When the †eigenvalues of the corresponding linearized system are purely imaginary and distinct, denoted by $\pm\lambda_1$, $\pm\lambda_2$, and $\lambda_1k_1+\lambda_2k_2\neq0$ for $0<|k_1|+|k_2|\leqslant4$ (where k_i is an integer), we can find suitable coordinates ξ_i, η_i so that the Hamiltonian H takes the form

$$H=\lambda_1\zeta_1+\lambda_2\zeta_2+\frac{1}{2}(c_{11}\zeta_1^2+2c_{12}\zeta_1\zeta_2+c_{22}\zeta_2^2)$$
$$+H_5+\ldots$$

with $\zeta_i\equiv\xi_i\eta_i$ and real c_{ij}. It is necessary that $\eta_i=\sqrt{-1}\,\bar{\xi}_i$ for the solutions to be real. In addition, if the condition

$$D\equiv c_{11}\lambda_2^2-2c_{12}\lambda_1\lambda_2+c_{22}\lambda_1^2\neq0$$

is satisfied, then the origin is a †stable equilibrium point of the original system (V. I. Arnol'd, J. Moser) [7].

For Lagrange equilateral triangular solutions of the planar restricted three-body problem, the eigenvalues λ of the linearized system derived from (2) are given as roots of the equation $\lambda^4+\lambda^2+(27/4)\mu(1-\mu)=0$ and are purely imaginary if $\mu(1-\mu)<1/27$. Applying the Arnol'd-Moser result, A. M. Leontovich and A. Deprit and Bartholomé showed that the Lagrange equilibrium points are stable for μ such that $0<\mu<\mu_0$, where μ_0 is the smaller root of $27\mu(1-\mu)=1$, excluding three values: μ_1, μ_2 at which $\lambda_1k_1+\lambda_2k_2=0$ $|k_1|+|k_2|\leqslant4$ and μ_3 at which $D=0$.

Arnol'd proved that if the masses $m_2,\ldots,m_n$ are negligibly small in comparison with m_1, the motion of the n-body system is †quasi-periodic for the majority of initial conditions for which the eccentricities and inclinations of the osculating ellipses are small.

References

[1] Y. Hagihara, Celestial mechanics. I, Dynamical principles and transformation theory, MIT Press, 1970; II, Perturbation theory, pts. 1, 2, MIT 1972; III, Differential equations in celestial mechanics, pts. 1, 2, Japan Society for the Promotion of Science, 1974; IV, Periodic and quasi-periodic solutions, pts. 1, 2, Japan Society, 1975; V, Topology of the three-body problem, pts. 1, 2, Japan Society, 1976.
[2] H. Poincaré, Les méthodes nouvelles de la mécanique céleste I–III, Gauthier-Villars, 1892–1899.
[3] E. T. Whittaker, A treatise on the analytical dynamics of particles and rigid bodies, Cambridge Univ. Press, fourth edition, 1937.
[4] H. Happel, Das Dreikörperproblem, Koehler, 1941.
[5] A. Wintner, The analytical foundations of celestial mechanics, Princeton Univ. Press, 1947.
[6] R. Abraham and J. E. Marsden, Foundations of mechanics, Benjamin, second edition, 1978.
[7] C. L. Siegel and J. Moser, Lectures on celestial mechanics, Springer, 1971.

[8] V. M. Alekseev (Alexeyev), Sur l'allure finale de mouvement dans le problème des trois corps, Actes Congrès Intern. Math., 1970, vol. 2, 893–907.
[9] V. Szebehely, Theory of orbits, Academic Press, 1967.

421 (XVIII.11) Time Series Analysis

A. Time Series

A time series is a sequence of observations ordered in time. Here we assume that measurements are quantitative and the times of measurements are equally spaced. We consider this sequence to be a realization of a stochastic process X_t (→ 407 Stochastic Processes). Usually time series analysis means a statistical analysis based on samples drawn from a stationary process (→ 395 Stationary Processes) or a related process. In what follows we denote the sample by $\mathbf{X}=(X_1, X_2, \ldots, X_T)'$.

B. Statistical Inference of the Autocorrelation

Let us assume X_t (t an integer) to be real-valued and weakly stationary (→ 395 Stationary Processes) and for simplicity $EX_t=0$ and consider the estimation of the **autocorrelation** $\rho_h=R_h/R_0$ of time lag h, where $R_l=EX_tX_{t+l}$. We denote the **sample autocovariance** of time lag h as

$$\tilde{R}_h=\frac{1}{T-|h|}\sum_{t=1}^{t-|h|}X_tX_{t+|h|},$$

and define the **serial correlation coefficient** of time lag h by $\tilde{\rho}_h=\tilde{R}_h/\tilde{R}_0$. It can be shown that the joint distribution of $\{\sqrt{T}(\tilde{\rho}_h-\rho_h) \mid 1\leqslant h\leqslant H\}$ tends to an H-dimensional †normal (Gaussian) distribution with mean vector $\mathbf{0}$, if one assumes that X_t is expressed as $X_t=\sum_{j=-\infty}^{\infty}b_j\xi_{t-j}$, where $\sum_{j=-\infty}^{\infty}|b_j|<+\infty$, $\sum_{j=-\infty}^{\infty}|j|^{1/2}b_j^2<+\infty$, and the ξ_t are independently and identically distributed random variables with $E\xi_t=0$ and $E\xi_t^4<+\infty$.

When X_t is an autoregressive process of order K (→ Section D) and also a †Gaussian process, it can be shown that the asymptotic distribution of $\{\sqrt{T}(\tilde{\rho}_h-\rho_h) \mid 1\leqslant h\leqslant K\}$ as $T\to\infty$ is equal to the asymptotic distribution of $\{\sqrt{T}(\hat{\rho}_h-\rho_h) \mid 1\leqslant h\leqslant K\}$, where $\hat{\rho}_h$ is the †maximum likelihood estimator of ρ_h. In general, it is difficult to obtain the maximum likelihood estimator of ρ_h. The statistical properties of other estimators of ρ_h, e.g., an estimator constructed by using $\mathrm{sgn}(X_t)$ ($\mathrm{sgn}(y)$ means 1 ($y>0$), 0 ($y=0$), -1 ($y<0$)) have also been investigated.

Testing hypotheses concerning autocorrelation can be carried out by using the above results. Let us now consider the problem of testing the hypothesis that X_t is a †white noise. Assume that X_t is a Gaussian process and that a white noise with $EX_t^2=\sigma^2$ exists, and define $\tilde{C}_h=\sum_{t=1}^{T}(X_t-\bar{X})(X_{t+h}-\bar{X})$ and $\tilde{\gamma}_h=\tilde{C}_h/\tilde{C}_0$ for $h\geqslant 0$, where $X_{T+j}=X_j$ and $\bar{X}=\sum_{t=1}^{T}X_t/T$. Then the probability density function of $\tilde{\gamma}_1$ can be obtained and it can be shown that

$$P(\tilde{\gamma}_1>\gamma)=\sum_{j=1}^{m}(\lambda_j-\gamma)^{(T-3)/2}\frac{1}{\Lambda_j},\qquad \lambda_{m+1}\leqslant\gamma\leqslant\lambda_m,$$

where $\lambda_j=\cos 2\pi j/T$ and

$$\Lambda_j=\prod_{\substack{k=1\\(k\neq j)}}^{(T-1)/2}(\lambda_j-\lambda_k),\qquad T=3,5,\ldots,$$

$$\Lambda_j=\prod_{\substack{k=1\\(k\neq j)}}^{T/2-1}(\lambda_j-\lambda_k)\sqrt{1+\lambda_j},\qquad T=4,6,\ldots,$$

$$1\leqslant m\leqslant(T-3)/2\qquad\text{if } T \text{ is odd},$$

$$1\leqslant m\leqslant T/2-1\qquad\text{if } T \text{ is even}.$$

This can be used to obtain a test of significance.

C. Statistical Inference of the Spectrum

To find the periodicities of a real-valued †weakly stationary process X_t with mean 0, the statistic, called the **periodogram**,

$$I_T(\lambda)=\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_te^{-2\pi it\lambda}\right|^2$$

is used. If X_t is expressed as

$$X_t=\sum_{l=1}^{L}\{m_l\cos 2\pi\lambda_lt+m_l'\sin 2\pi\lambda_lt\}+Y_t,$$

where $\{m_l\}$, $\{m_l'\}$, and $\{Y_t\}$ are mutually independent random variables with $Em_l=Em_l'=0$ and $V(m_l)=V(m_l')=\sigma_l^2$ and $\{Y_t\}$ is independent and identically distributed with means 0 and finite variances σ^2, the distribution of $I_T(\lambda)$ converges to a distribution with finite mean and finite variance at $\lambda\neq\pm\lambda_l$ for $1\leqslant l\leqslant L$ when T tends to infinity. On the other hand, the magnitude of $I_T(\lambda)$ is of the order of T at $\lambda=\pm\lambda_l$, $1\leqslant l\leqslant L$. This means that we can find the periodicities of X_t by using $I_T(\lambda)$. When $X_t=Y_t$, we find that the distribution of $2I_T(\lambda)/\sigma^2$ (when $\lambda\neq 0$, $\pm 1/2$) or $I_T(\lambda)/\sigma^2$ (when $\lambda=0$ or $\pm 1/2$) tends to the †χ^2 distribution with degrees of freedom 2 or 1, respectively, and $I(\mu_1), I(\mu_2), \ldots, I(\mu_M)$ are asymptotically independent random variables for $0\leqslant|\mu_1|<|\mu_2|<\ldots<|\mu_M|\leqslant 1/2$ when $T\to\infty$. Applying this result, we can test for periods in the data.

Let $f(\lambda)$ be the spectral density function of a real-valued weakly stationary process X_t. In general, the variance of $|\sum_{t=1}^{T} X_t e^{-2\pi i t\lambda}/\sqrt{T}|$ does not tend to 0 as T tends to infinity; hence $I_T(\lambda)$ cannot be used as a good estimator for the spectral density. To obtain an estimate of $f(\lambda)$, several estimators defined by using weight functions have been proposed by several authors. Let $W_T(\lambda)$ be a weight function defined on $(-\infty, \infty)$, and construct a statistic $\tilde{f}(\lambda)=\int_{-1/2}^{1/2} I_T(\mu)W_T(\lambda-\mu)d\mu$. Let us use $\tilde{f}(\mu)$ for the estimation of $f(\lambda)$. $W_T(\lambda)$ is called a **window**. An important class of $W_T(\lambda)$ is as follows. Let $W(\lambda)$ be continuous, $W(\lambda)=W(-\lambda)$, $W(0)=1$, $|W(\lambda)|<1$, and $\int_{-\infty}^{\infty} W(\lambda)^2 d\lambda < +\infty$, and let H be a positive integer depending on T such that $H\to\infty$ and $H/T\to 0$ as $T\to\infty$. Putting $w_j=W(j/H)$, we define $W_T(\lambda)$ by $W_T(\lambda)=\sum_{j=-T+1}^{T-1} w_j e^{-2\pi i j\lambda}$. Then $\tilde{f}(\lambda)$ can be expressed as $\tilde{f}(\lambda)=\sum_{h=-T+1}^{T-1}\tilde{\tilde{R}}_h w_h e^{-2\pi i h\lambda}$, where $\tilde{\tilde{R}}_h=\sum_{t=1}^{T-h} X_{t+h}X_t/T$ for $h\geqslant 0$ and $\tilde{\tilde{R}}_h=\sum_{t=|h|+1}^{T} X_{t+h}X_t/T$ for $h<0$. Let X_t be stationary to the fourth order (→ 395 Stationary Processes) and satisfy

$$\sum_{h=-\infty}^{\infty} |R_h| < +\infty,$$

$$\sum_{h,l,p=-\infty}^{\infty} |C_{o,h,l,p}| < +\infty,$$

where $C_{o,h,l,p}$ is the fourth-order joint †cumulant of X_t, X_{t+h}, X_{t+l}, and X_{t+p}. Then we have

$$\lim_{T\to\infty}\frac{T}{H}V(\tilde{f}(0))=2f(0)^2\int_{-\infty}^{\infty} W(\lambda)^2 d\lambda,$$

$$\lim_{T\to\infty}\frac{T}{H}V(\tilde{f}(\pm 1/2))=2f(1/2)^2\int_{-\infty}^{\infty} W(\lambda)^2 d\lambda,$$

$$\lim_{T\to\infty}\frac{T}{H}V(\tilde{f}(\lambda))=f(\lambda)^2\int_{-\infty}^{\infty} W(\lambda)^2 d\lambda, \qquad \lambda\neq 0, \quad \pm 1/2,$$

$$\lim_{T\to\infty}\frac{T}{H}\mathrm{Cov}(\tilde{f}(\lambda),\ \tilde{f}(\mu))=0, \quad \lambda\neq\mu. \tag{1}$$

$\{w_h\}$ or $W_T(\lambda)$ should have an optimality, e.g., to minimize the mean square error of $\tilde{f}(\lambda)$. But, generally, it is difficult to obtain such a $\{w_h\}$ or $W_T(\lambda)$.

Several authors have proposed specific types of windows. The following are some examples: (i) (Bartlett) $w_h=(1-|h|/H)$ for $|h|\leqslant H$ and $w_h=0$ for $|h|>H$; (ii) (Tukey) $w_h=\sum_{l=-\infty}^{\infty} a_l\cos(\pi lh/H)$ for $|h|\leqslant H$ and $w_h=0$ for $|h|>H$, where the a_l are constants such that $\sum_{l=-\infty}^{\infty}|a_l|<+\infty$, $\sum_{l=-\infty}^{\infty} a_l=1$ and $a_l=a_{-l}$. The Hanning and Hamming windows are $a_0=0.50$, $a_1=a_{-1}=0.25$, and $a_l=0$ for $|l|\geqslant 2$ and $a_0=0.54$, $a_1=a_{-1}=0.23$, and $a_l=0$ for $|l|\geqslant 2$, respectively [2]. Let $X_t=\sum_{j=-\infty}^{\infty} b_j\varepsilon_{t-j}$, where $\sum_{j=-\infty}^{\infty}|b_j|<+\infty$ and the ε_t are independently and identically distributed random variables with $E\varepsilon_t=0$ and $E\varepsilon_t^4<+\infty$. Let $\{\lambda_j \mid 1\leqslant j\leqslant M\}$ be arbitrary real numbers such that $0\leqslant\lambda_1<\lambda_2<\ldots<\lambda_M\leqslant 1/2$, where M is an arbitrary positive integer. Then the joint distribution of $\{\sqrt{T/H}(\tilde{f}(\lambda_\nu)-E\tilde{f}(\lambda_\nu)) \mid 1\leqslant\nu\leqslant M\}$ tends to the normal distribution with means 0 and covariance matrix Σ, which is defined by (1). Let us assume, furthermore, that $\lim_{x\to 0}(1-w(x))/|x|^q=C$ and $\sum_{h=-\infty}^{\infty}|h|^p|R_h|<+\infty$, where C, q, and p are some positive constants satisfying the following conditions: (i) when $p\geqslant q$, $H^q/T\to 0$ $(p\geqslant 1)$ and $H^{q+1-p}/T\to 0$ $(p\geqslant 1)$ as $T\to\infty$ and $\lim_{T\to\infty} T/H^{2q+1}$ is finite; (ii) when $p<q$, $H^p/T\to 0$ $(p\geqslant 1)$ and $H/T\to 0$ $(p\leqslant 1)$ as $T\to\infty$ and $\lim_{T\to\infty} T/H^{2p+1}=0$. Then $\sqrt{T/H}(\tilde{f}(\lambda_\nu)-E\tilde{f}(\lambda_\nu))$ in the results above can be replaced by $\sqrt{T/H}(\tilde{f}(\lambda_\nu)-f(\lambda_\nu))$.

Estimation of higher-order spectra, particularly the bispectrum, has also been discussed. Let X_t be a weakly stationary process with mean 0, and let its spectral decomposition be given by $X_t=\int_{-1/2}^{1/2} e^{2\pi i t\lambda}dZ(\lambda)$ (→ 395 Stationary Processes). We assume that X_t is a weakly stationary process of degree 3 and put $R_{h_1,h_2}=EX_tX_{t+h_1}X_{t+h_2}$ for any integers h_1 and h_2. Then we have

$$R_{h_1,h_2}=\int_{1/2}^{1/2}\int_{-1/2}^{1/2} e^{2\pi i(h_1\lambda_1+h_2\lambda_2)}dF(\lambda_1,\lambda_2).$$

Symbolically, $dF(\lambda_1,\lambda_2)=EdZ(\lambda_1)dZ(\lambda_2)dZ(-\lambda_1-\lambda_2)$. If $F(\lambda_1,\lambda_2)$ is absolutely continuous with respect to the Lebesgue measure of $\mathbf{R}^2$ and $\partial F(\lambda_1,\lambda_2)/\partial\lambda_1\partial\lambda_2=f(\lambda_1,\lambda_2)$, we call $f(\lambda_1,\lambda_2)$ the **bispectral density function**. When X_t is Gaussian, $R_{h_1,h_2}=0$ and $f(\lambda_1,\lambda_2)=0$ for any h_1, h_2 and any λ_1, λ_2. $f(\lambda_1,\lambda_2)$ can be considered to give a kind of measure of the departure from a Gaussian process or a kind of nonlinear relationship among waves of different frequencies. We can construct an estimator for $f(\lambda_1,\lambda_2)$ by using windows as in the estimation of a spectral density [3].

D. Statistical Analysis of Parametric Models

When we assume merely that X_t is a stationary process and nothing further, then X_t contains infinite-dimensional unknown parameters. In this case, it may be difficult to develop a satisfactory general theory for statistical inference about X_t. But in most practical applications of time series analysis, we can safely assume at least some of the time dependences to be known. For this reason, we can often use a model with finite-dimensional parameters. This means, mainly, that the moments (usually, second-order moments) or the spectral density are assumed to be expressible in terms of finite-dimensional parameters. As examples of such models, autoregressive models, moving average models, and autoregressive moving average models are widely used.

A process X_t is called an **autoregressive process** of order K if X_t satisfies a difference equation $\sum_{k=0}^{K} a_k X_{t-k} = \xi_t$, where the a_k are constants, $a_0 = 1$, $a_K \neq 0$, and the ξ_t are mutually uncorrelated with $E\xi_t = 0$ and $V(\xi_t) = \sigma_\xi^2 > 0$. We usually assume that X_t is a weakly stationary process with $EX_t = 0$. We sometimes use the notation AR(K) to express a weakly stationary and autoregressive process of order K. Let $\{\xi_t\}$ be as above. If X_t is expressed as $X_t = \sum_{l=0}^{L} b_l \xi_{t-l}$, where the b_l are constants, $b_0 = 1$ and $b_L \neq 0$, X_t is called a **moving average process** of order L (MA(L) process). Furthermore, if X_t is weakly stationary with $EX_t = 0$ and expressed as $\sum_{k=0}^{K} a_k X_{t-k} = \sum_{l=0}^{L} b_l \xi_{t-l}$ with $a_0 = 1$, $b_0 = 1$, and $a_K b_L \neq 0$, then X_t is called an **autoregressive moving average process** of order (K, L) (ARMA(K, L) process). Let $A(Z)$ and $B(Z)$ be two polynomials of Z such that $A(Z) = \sum_{k=0}^{K} a_k Z^{K-k}$ and $B(Z) = \sum_{l=0}^{L} b_l Z^{L-l}$, and let $\{\alpha_k \mid 1 \leqslant k \leqslant K\}$ and $\{\beta_l \mid 1 \leqslant l \leqslant L\}$ be the solutions of the associated polynomial equations $A(Z) = 0$ and $B(Z) = 0$, respectively, we assume that $|\alpha_k| < 1$ for $1 \leqslant k \leqslant K$ and $|\beta_l| < 1$ for $1 \leqslant l \leqslant L$. This condition implies that X_t is purely nondeterministic. Let the observed sample be $\{X_t \mid 1 \leqslant t \leqslant T\}$. If we assume that X_t is Gaussian and an ARMA(K, L) process, we can show that the †maximum likelihood estimators $\{\hat{a}_k\}$ and $\{\hat{b}_l\}$ of $\{a_k\}$ and $\{b_l\}$ are †consistent and asymptotically efficient when $T \to \infty$ ("asymptotically efficient" means that the covariance matrix of the distribution of the estimators is asymptotically equal to the inverse of the information matrix) [5] (→ 399 Statistical Estimation D). Furthermore, if X_t is an AR(K) process, the joint distribution of $\{\sqrt{T}(\hat{a}_k - a_k) \mid 1 \leqslant k \leqslant K\}$ tends to a K-dimensional normal distribution with means 0, and this distribution is the same as the one to which the distribution of the †least-square estimators $\{\hat{\hat{a}}_k\}$ minimizing $Q = \sum_{t=K+1}^{T} (X_t + \sum_{k=1}^{K} a_k X_{t-k})^2$ tends when $T \to \infty$. If X_t is a MA(L) or ARMA(K, L) process $(L \geqslant 1)$, the likelihood equations are complicated and cannot be solved directly. Many approximation methods have been proposed to obtain the estimates.

When X_t is an AR(K) process with $|\alpha_k| < 1$ for $1 \leqslant k \leqslant K$, R_t satisfies $\sum_{k=0}^{K} a_k R_{h-k} = 0$ for $h \geqslant 1$. These are often called the **Yule-Walker equations**. R_h can be expressed as $R_h = \sum_{j=1}^{K} C_j \alpha_j^h$ if the α_k are distinct and $a_K \neq 0$, where $\{C_j\}$ are constants and determined by R_h for $0 \leqslant h \leqslant K-1$. When X_t is an ARMA(K, L) process, $\sum_{k=0}^{K} a_k R_{h-k} = 0$ for $h \geqslant L+1$, and the C_j of $R_h = \sum_{j=1}^{k} C_j \alpha_j^h$ are determined by $\{R_h \mid 0 \leqslant h \leqslant \max(K, L)\}$.

The spectral density is expressed as $f(\lambda) = \sigma_\xi^2 |B(e^{2\pi i\lambda})|^2 / |A(e^{2\pi i\lambda})|^2$. If X_t is Gaussian, the maximum likelihood estimator of $f(\lambda)$ is asymptotically equal to the statistic obtained by replacing σ_ξ^2, $\{b_l\}$, and $\{a_k\}$ in $f(\lambda)$ with $\hat{\sigma}_\xi^2$, $\{\hat{b}_l\}$, and $\{\hat{a}_k\}$, respectively, where $\hat{\sigma}_\xi^2$ is the maximum likelihood estimator of σ_ξ^2, when $T \to \infty$.

When we analyze a time series and intend to fit an ARMA(K, L) model, we have to determine the values of K and L. For AR(K) models, many methods have been proposed to determine the value of K. Some examples are: (i) (Quenouille) Let $(Z^K A(1/Z))^2 = \sum_{j=0}^{2K} A_j Z^j$, and $G_K = \sum_{K-j=0}^{2K} \hat{A}_j(\tilde{R}_j/\tilde{R}_0)$, where $\hat{A}_j$ is obtained by replacing $\{a_k\}$ in A_j by $\{\hat{a}_k\}$, and we construct the statistic $\chi_f^2 = \sum_{l=1}^{f} G_{K+l}$. Then χ_f^2 has a †χ^2 distribution asymptotically with f degrees of freedom under the assumption that $K \geqslant K_0$, where K_0 is the true order, as $T \to \infty$. Using this fact, we can determine the order of an AR model. (ii) (Akaike) We consider choosing an order K satisfying $K_L \leqslant K \leqslant K_M$, where K_L and K_M are minimum order and maximum order, respectively, specified a priori. Then we construct the statistic AIC$(K) = (T-K)\log \hat{\hat{\sigma}}_\xi^2(K) + 2K$, where

$$\hat{\hat{\sigma}}_\xi^2(K) = \sum_{t=K+1}^{T} (X_t + \hat{a}_1 X_{t-1} + \ldots + \hat{a}_K X_{t-K})^2 / T$$

and $\{\hat{a}_k \mid 1 \leqslant k \leqslant K\}$ are the least square estimators of the autoregressive coefficients of an AR(K) model fitting X_t. Calculate AIC(K) for $K = K_L, K_L+1, \ldots, K_M$. If AIC$(K)$ has the minimum value at $K = \hat{K}$, we determine the order to be $\hat{K}$ [6] (→ 403 Statistical Models F). Parzen proposed another method by using the criterion autoregressive transfer function (CAT). Here CAT$(K) = 1 - \tilde{\sigma}^2(\infty)/\tilde{\sigma}_\xi^2(K) + K/T$, where $\tilde{\sigma}_\xi^2(K) = (T/(T-K))\hat{\hat{\sigma}}_\xi^2(K)$ and $\tilde{\sigma}^2(\infty)$ is an estimator of $\sigma^2(\infty) = \exp(\int_{-1/2}^{1/2} \log f(\lambda)\, d\lambda)$ [7]. (iii) We can construct a test statistic for the null hypothesis AR(K) against the alternative hypothesis AR$(K+1)$ (Jenkins) or use a multiple decision procedure (T. W. Anderson [8]).

Not much is known about the statistical properties of the above methods, and few comparisons have been made among them.

Another parametric model is an exponential model for the spectrum. The spectral density is expressed by $f(\lambda) = C^2 \exp\{2\sum_{k=1}^{K} \theta_k \cos(2\pi k\lambda)\}$, where the θ_k and C are constants.

We now discuss some general theories of estimation for finite-dimensional-parameter models. Let X_t be a real-valued Gaussian process of mean 0 and of spectral density $f(\lambda)$ which is continous and positive in $[-1/2, 1/2]$, and let the moving average representation of X_t be $X_t = \sum_{l=0}^{\infty} b_l \xi_{t-l}$, where ξ_t is a white noise and $\sigma_\xi^2 = E\xi_t^2$. We assume that $f(\lambda)/\sigma_\xi^2 = g(\lambda)$

depends only on M parameters $\boldsymbol{\theta}=(\theta_1,\theta_2,\ldots,\theta_M)'$ which are independent of σ_ξ^2. Then the logarithm of the †likelihood function can be approximated by $-(1/2)\{T\log 2\pi\sigma_\xi^2+\mathbf{X}'\Sigma_T^{-1}(\boldsymbol{\theta})\mathbf{X}/\sigma_\xi^2\}$ by ignoring the lower-order terms in T, where $\sigma_\xi^2\Sigma_T(\boldsymbol{\theta})$ is the covariance matrix of $\mathbf{X}$. Usually, it is difficult to find an explicit expression for each element of $\Sigma_T^{-1}(\boldsymbol{\theta})$. Another approximation for the logarithm of the likelihood function is given by

$$-\frac{T}{2}\int_{-1/2}^{1/2}\left[\log f(\lambda)+\frac{I_T(\lambda)}{f(\lambda)}\right]d\lambda.$$

Under mild conditions on the regularity of $g(\lambda)$, the estimators $\hat{\boldsymbol{\theta}}=(\hat\theta_1,\hat\theta_2,\ldots,\hat\theta_M)$ and $\hat\sigma_\xi^2$, obtained as the solutions of the likelihood equations, are †consistent and asymptotically normal as T tends to infinity. This means that the distribution of $\sqrt{T}(\hat\sigma_\xi^2-\sigma_\xi^2)$ is asymptotically normal and $\sqrt{T}(\hat\sigma_\xi^2-\sigma_\xi^2)$ and $\sqrt{T}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ are asymptotically independent. The asymptotic distribution of $\sqrt{T}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ is the normal distribution $N(\mathbf{0},\Gamma^{-1})$, where the (k,l)-component Γ_{kl} of Γ is given by

$$\Gamma_{kl}=\frac{1}{2}\int_{-1/2}^{1/2}\left(\frac{\partial\log g(\lambda)}{\partial\theta_k}\cdot\frac{\partial\log g(\lambda)}{\partial\theta_l}\right)_{\boldsymbol{\theta}}d\lambda.$$

E. Statistical Analysis of Multiple Time Series

Let $\mathbf{X}_t=(X_t^{(1)},X_t^{(2)},\ldots,X_t^{(p)})'$ be a complex-valued weakly stationary process with $E\mathbf{X}_t=\mathbf{0}$ and $E\mathbf{X}_t\bar{\mathbf{X}}_s'=R_{t-s}$. R_{t-s} is the $p\times p$ matrix whose (k,l)-component is $R_{t-s}^{(k,l)}=EX_t^{(k)}\bar X_s^{(l)}$. We discuss the case when t is an integer. R_h has the spectral representation

$$R_h=\int_{-1/2}^{1/2}e^{2\pi ih\lambda}dF(\lambda),$$

where $F(\lambda)$ is a $p\times p$ matrix and $F(\lambda_1)-F(\lambda_2)$, $\lambda_1\geqslant\lambda_2$, is Hermitian nonnegative. Let $f^{k,l}(\lambda)$ be the (k,l)-component of the spectral density matrix $f(\lambda)$, i.e., $F_a(\lambda)=\int_{-1/2}^{\lambda}f(\mu)d\mu$, of the absolutely continuous part in the Lebesgue decomposition of $F(\lambda)$. The function $f^{k,l}(\lambda)$ for $k\neq l$ is called the **cross spectral density function**. $f^{k,l}(\lambda)$ represents a kind of correlation between the wave of frequency λ included in $X_t^{(k)}$ and the one included in $X_t^{(l)}$.

Let $\mathbf{X}_t=(X_t^{(1)},X_t^{(2)},\ldots,X_t^{(p)})'$ and $\mathbf{Y}_t=(Y_t^{(1)},Y_t^{(2)},\ldots,Y_t^{(q)})'$ be two complex-valued weakly stationary processes with $E\mathbf{X}_t=0$, $E\mathbf{Y}_t=0$, $E\mathbf{X}_t\bar{\mathbf{X}}_s'=R_{t-s}^{\mathbf{X}}$, $E\mathbf{Y}_t\bar{\mathbf{Y}}_s'=R_{t-s}^{\mathbf{Y}}$ and $E\mathbf{X}_t\bar{\mathbf{Y}}_s'=R_{t-s}^{\mathbf{XY}}$. We assume $\mathbf{Y}_t=\sum_{s=-\infty}^{\infty}A_s\mathbf{X}_{t-s}$, where A_s is a $q\times p$ matrix whose components are constants depending on s. Put $A(\lambda)=\sum_{s=-\infty}^{\infty}A_se^{-2\pi is\lambda}$. $A(\lambda)$ should exist in the sense of mean square convergence with respect to the spectral distribution function F for $\mathbf{X}_t$. The function $A(\lambda)$ is called the matrix **frequency response function**.

As a measure of the strength of association between $X_t^{(k)}$ and $X_t^{(l)}$ at frequency λ, we introduce the quantity $\gamma^{k,l}(\lambda)=|f^{k,l}(\lambda)|^2/f^{k,k}(\lambda)f^{l,l}(\lambda)$. $\gamma^{k,l}(\lambda)$ is called the **coherence**. Let $X_t^{(k)}=\sum_{s=-\infty}^{\infty}a_s^{k,l}X_{t-s}^{(l)}+\eta_t$, where η_t is a weakly stationary process with mean 0 and uncorrelated with $X_s^{(l)}$, $-\infty<s<\infty$. If $E|\eta_t|^2=0$, $\gamma^{k,l}(\lambda)=1$. If $E|\sum_{-\infty}^{\infty}a_s^{k,l}X_{t-s}^{(l)}|^2=0$, $\gamma^{k,l}(\lambda)=0$. Generally, we have $0\leqslant\gamma^{k,l}(\lambda)\leqslant 1$.

For the estimation of $F(\lambda)$, $A(\lambda)$, and $\gamma^{k,l}(\lambda)$, the theories have been similar to those for the estimation of the spectral density of a scalar time series. For example, an estimator of $f(\lambda)$ is given [11] in the form

$$\hat f(\lambda)=\sum_{h=-(T-1)}^{T-1}\tilde{\bar{R}}_hw_he^{-2\pi ih\lambda},$$

where

$$\tilde{\bar{R}}_h=\sum_{t=1}^{T-|h|}\mathbf{X}_{t+|h|}\bar{\mathbf{X}}_t'/T$$

and the w_h are the same as in Section C.

We can define an autoregressive, moving average, or autoregressive moving average process in a similar way as for a scalar time series. The a_k and b_l in Section D should be replaced by $p\times p$ matrices and the associated polynomial equations $A(Z)=0$ and $B(Z)=0$ should be understood in the vector sense [11]. There are problems with determining the coefficients uniquely or identifying an ARMA(K,L) model, and these problems have been discussed to some extent.

F. Statistical Inference of the Mean Function

Let X_t be expressed as $X_t=m_t+Y_t$, where m_t is a real-valued deterministic function of t and Y_t is a real-valued weakly stationary process with mean 0 and spectral distribution function $F(\lambda)$. This means that $EX_t=m_t$. We consider the case when $m_t=\sum_{j=1}^{M}C_j\varphi_t^{(j)}$, where $\mathbf{C}=(C_1,C_2,\ldots,C_M)'$ is a vector of unknown coefficients and $\boldsymbol{\varphi}_t=(\varphi_t^{(1)},\varphi_t^{(2)},\ldots,\varphi_t^{(M)})'$ is a set of known (regression) functions.

Let us construct †linear unbiased estimators $\{\tilde C_j=\sum_{t=1}^{T}\gamma_{jt}X_t\,|\,1\leqslant j\leqslant M\}$ for the coefficients C_j, where the γ_{jt} are known constants. Put $\Phi=(\boldsymbol{\varphi}_1,\boldsymbol{\varphi}_2,\ldots,\boldsymbol{\varphi}_T)'$. Then the †least squares estimator of $\mathbf{C}$ is given by $\hat{\mathbf{C}}=(\Phi'\Phi)^{-1}\Phi'\mathbf{X}$ when $\Phi'\Phi$ is nonsingular. Let Σ be the covariance matrix of $\mathbf{X}$. Then the †best linear unbiased estimator is $\hat{\mathbf{C}}^*=(\Phi'\Sigma^{-1}\Phi)^{-1}\Phi'\Sigma^{-1}\mathbf{X}$. We put $\|\varphi^{(j)}\|_T^2=\sum_{t=1}^{T}(\varphi_t^{(j)})^2$ and assume that $\lim_{T\to\infty}\|\varphi^{(j)}\|_T^2=\infty$, $\lim_{T\to\infty}\|\varphi^{(j)}\|_{T+h}^2/\|\varphi^{(j)}\|_T^2=1$ for $1\leqslant j\leqslant M$ and any fixed h and assume the existence of $\psi_h^{(j,k)}=\lim_{T\to\infty}\sum_{t=1}^{T}\varphi_{t+h}^{(j)}\varphi_t^{(k)}/\|\varphi^{(j)}\|_T\|\varphi^{(k)}\|_T$ for $1\leqslant j,k\leqslant M$. We also assume that $F(\lambda)$ is absolutely continuous and $F'(\lambda)=f(\lambda)$ is positive and piecewise continuous. Let ψ_h be the $M\times M$ matrix whose (j,k)-component is $\psi_h^{(j,k)}$. Then ψ_h can be represented by

$$\psi_h=\int_{-1/2}^{1/2} e^{2\pi ih\lambda}\,dG(\lambda),$$

where $G(\lambda)-G(\mu)$ is a nonnegative definite matrix for $\lambda>\mu$. Assume that $\psi_0=G(1/2)-G(-1/2)$ is nonsingular and put $H(\lambda)=\psi_0^{-1/2}G(\lambda)\psi_0^{-1/2}$, and for any set S, $H(S)=\int_S H(d\lambda)$. Suppose further that $S_1, S_2, \ldots, S_q$ are q sets such that $H(S_j)>0$, $\sum_{j=1}^q H(S_j)=I$, $H(S_j)H(S_k)=0, j\neq k$, and for any j there is no subset $S_j'\subset S_j$ such that $H(S_j')>0$, $H(S_j-S_j')>0$ and $H(S_j')H(S_j-S_j')=0$. We have $q\leqslant M$. It can be shown that the spectrum of the regression can be decomposed into such disjoint sets $S_1,\ldots,S_q$. Then we can show that $\hat{\mathbf{C}}$ is asymptotically efficient in the sense that the asymptotic covariance matrix of $\hat{\mathbf{C}}$ is equivalent to that of $\hat{\mathbf{C}}^*$ if and only if $f(\lambda)$ is constant on each of the elements S_j. Especially, if $\psi_t^{(j)}=t^je^{2\pi it\mu_j}$, $\hat{\mathbf{C}}$ is asymptotically efficient.

G. Nonstationary Models

It is difficult to develop a statistical theory for a general class of nonstationary time series, but some special types of nonstationary processes have been investigated more or less in detail. Let X_t (t an integer) be a real-valued stochastic process and ∇ be the backward difference operator defined by $\nabla X_t=X_t-X_{t-1}$ and $\nabla^d X_t=\nabla(\nabla^{d-1}X_t)$ for $d\geqslant 2$. We assume that X_t is defined for $t\geqslant t_0$ (t_0 a finite integer), and $EX_t^2<+\infty$. For analyzing a nonstationary time series, Box and Jenkins introduced the following model: For a positive integer d, $Y_t=\nabla^d X_t$, $t\geqslant t_0+d$, is stationary and is an autoregressive moving average process of order (K,L) for $t\geqslant t_0+d+\max(K,L)$. They called such an X_t an **autoregressive integrated moving average process** of order (K,d,L) and denoted it by ARIMA(K,d,L). The word "integrated" means a kind of summation; in fact, X_t can be expressed as a sum of the weakly stationary process Y_t, i.e.,

$$X_t=X_0+(\nabla X_0)t+(\nabla^2X_0)\left(\sum_{s_2=1}^{t}\sum_{s_1=1}^{s_2}\right)+\ldots$$
$$+(\nabla^{d-1}X_0)\left(\sum_{s_{d-1}=1}^{t}\cdots\sum_{s_1=1}^{s_2}\right)$$
$$+\sum_{s_d=1}^{t}\sum_{s_{d-1}=1}^{s_d}\cdots\sum_{s_1=1}^{s_2}Y_{s_1}$$

when $t_0=-d+1$. Using this model, methods of forecasting and of model identification and estimation can be discussed [13].

Another nonstationary model is based on the concept of evolutionary spectra [14]. In this approach, spectral distribution functions are taken to be time-dependent. Let X_t be a complex-valued stochastic process (t an integer) with $EX_t=0$ and $R_{t,s}=EX_t\bar{X}_s$. In the following, we write simply $\int$ for $\int_{-1/2}^{1/2}$. We now restrict our attention to the class of X_t for which there exist functions $\{u_t(\lambda)\}$ defined on $[-1/2,1/2]$ such that $R_{t,s}$ can be expressed as $R_{t,s}=\int u_t(\lambda)\bar{u}_s(\lambda)\,d\mu(\lambda)$, where $\mu(\lambda)$ is a measure. $u_t(\lambda)$ should satisfy $\int|u_t(\lambda)|^2\,d\mu(\lambda)<+\infty$. Then X_t admits a representation of the form $X_t=\int u_t(\lambda)\,dZ(\lambda)$, where $Z(\lambda)$ is a process with orthogonal increments and $E|dZ(\lambda)|^2=d\mu(\lambda)$. If $u_t(\lambda)$ is expressed as $u_t(\lambda)=\gamma_t(\lambda)e^{2\pi i\theta(\lambda)t}$ and $\gamma_t(\lambda)$ is of the form $\gamma_t(\lambda)=\int e^{2\pi itw}\,d\Gamma_\lambda(w)$ with $|d\Gamma_\lambda(w)|$ having the absolute maximum at $w=0$, we call $u_t(\lambda)$ an oscillatory function and X_t an oscillatory process. The evolutionary power spectrum $dF_t(\lambda)$ is defined by $dF_t(\lambda)=|\gamma_t(\lambda)|^2\,d\mu(\lambda)$.

Other models, such as an autoregressive model whose coefficients vary with time or whose associated polynomial has roots outside the unit circle, have also been discussed.

References

[1] E. J. Hannan, Time series analysis, Methuen, 1960.
[2] R. B. Blackman and J. W. Tukey, The measurement of power spectra from the point of view of communications engineering, Dover, 1959.
[3] D. R. Brillinger, Time series: Data analysis and theory, Holt, Rinehart and Winston, 1975.
[4] M. S. Bartlett, An introduction to stochastic processes with special reference to methods and applications, Cambridge Univ. Press, second edition, 1966.
[5] P. Whittle, Estimation and information in stationary time series, Ark. Mat.. 2 (1953), 423–434.
[6] H. Akaike, Information theory and an extension of the maximum likelihood principle, 2nd Int. Symp. Information Theory, B. N. Petrov and F. Csáki (eds.), Akadémiai Kiadó, 1973, 267–281.
[7] E. Parzen, Some recent advances in time series modeling, IEEE Trans. Automatic Control, AC-19 (1974), 723–730.
[8] T. W. Anderson, The statistical analysis of time series, Wiley, 1971.
[9] P. Whittle, Hypothesis testing in time series analysis, Almqvist and Wiksell, 1951.
[10] H. B. Mann and A. Wald, On the statistical treatment of linear stochastic difference equations, Econometrica, 11 (1943), 173–220.

[11] E. J. Hannan, Multiple time series, Wiley, 1970.
[12] U. Grenander and M. Rosenblatt, Statistical analysis of stationary time series, Wiley, 1957.
[13] G. E. P. Box and G. M. Jenkins, Time series analysis: Forecasting and control, Holden-Day, revised edition, 1976.
[14] M. B. Priestley, Evolutionary spectra and nonstationary processes, J. Roy. Statist. Soc., ser. B, 27 (1965), 204–237.

422 (IV.7) Topological Abelian Groups

A. Introduction

A commutative topological group is called a **topological Abelian group**. Throughout this article, except in Section L, all topological groups under consideration are locally compact Hausdorff topological Abelian groups and are simply called groups (→ 423 Topological Groups).

B. Characters

A **character** of a group is a continuous function $\chi(x)$ ($x \in G$) that takes on as values complex numbers of absolute value 1 and satisfies $\chi(xy)=\chi(x)\chi(y)$. Equivalently, χ is a 1-dimensional and therefore an irreducible †unitary representation of G. Conversely any irreducible unitary representation of G is 1-dimensional. Indeed, for a topological Abelian group, the set of its characters coincides with the set of its irreducible unitary representations. If the product of two characters χ, χ' is defined by $\chi\chi'(x)=\chi(x)\chi'(x)$, then the set of all characters forms the **character group** $C(G)$ of G. With †compact-open topology, $C(G)$ itself becomes a locally compact topological Abelian group.

C. The Duality Theorem

For a fixed element x of G, $\chi(x)$ ($\chi \in C(G)$) is a character of $C(G)$, namely, an element of $CC(G)$. Denote this character of $C(G)$ by $x(\chi)$, and consider the correspondence $G \ni x \rightarrow x(\chi)$. That this correspondence is one-to-one follows from the fact that any locally compact G has †sufficiently many irreducible unitary representations (→ 437 Unitary Representations) and the fact that if G is an Abelian group, then any irreducible unitary representation of G is a character of G. Furthermore, any character of $C(G)$ is given as one of the $x(\chi)$; indeed, by this correspondence, we have $G \cong CC(G)$ (**Pontryagin's duality theorem**).

By the duality theorem, each of G and $C(G)$ is isomorphic to the character group of the other. In this sense, G and $C(G)$ are said to be **dual** to each other.

D. Correspondence between Subgroups

Let $G, G'=C(G)$ be groups that are dual to each other. Given a closed subgroup g of G, the set of all χ' such that $\chi'(x)=1$ for all x in g forms a closed subgroup of G', usually denoted by (G', g). The definition of (G, g') is similar. Then $g \leftrightarrow (G', g)=g'$ gives a one-to-one correspondence between the closed subgroups of G and those of G'. If $g_1 \supset g_2$, then g_1/g_2 and $(G', g_2)/(G', g_1)$ are dual to each other. If the group operations of G, G' are written in additive form, with 0 for the identity, then $x(\chi')=1$ is written as $x(\chi')=0$. In this sense, (G', g) is called the **annihilator** (or **annulator**) of g.

E. The Structure Theorem

Let $\mathfrak{A}$ be the set of all groups (more precisely, of all locally compact Hausdorff topological Abelian groups). If $G_1, G_2 \in \mathfrak{A}$, then the direct product $G_1 \times G_2 \in \mathfrak{A}$, and if $G \in \mathfrak{A}$ and H is a closed subgroup of G, then $H \in \mathfrak{A}$ and $G/H \in \mathfrak{A}$. In addition, if H is a closed subgroup of a group G such that $H \in \mathfrak{A}$ and $G/H \in \mathfrak{A}$, then $G \in \mathfrak{A}$. In other words, $\mathfrak{A}$ is closed under the operations of forming direct products, closed subgroups, quotient groups, and †extensions by members of $\mathfrak{A}$. Furthermore, the operation C that assigns to each element of $\mathfrak{A}$ its dual element is a reflexive correspondence of $\mathfrak{A}$ onto $\mathfrak{A}$, and if $G \supset H$, the annihilator $(C(G), H)$ of H is a closed subgroup of $C(G)$. Also, $C(G/H) \cong (C(G), H)$, $C(H) \cong C(G)/(C(G), H)$. Furthermore, $C(G_1 \times G_2) \cong C(G_1) \times C(G_2)$. Finally, $H=(G,(C(G), H))$ (**reciprocity of annihilators**).

Typical examples of groups in $\mathfrak{A}$ are the additive group $\mathbf{R}$ of real numbers, the additive group $\mathbf{Z}$ of rational integers, the 1-dimensional †torus group $\mathbf{T}=\mathbf{R}/\mathbf{Z}$, and finite Abelian groups $\mathbf{F}$. The torus group $\mathbf{T}$ is also isomorphic to the multiplicative group $U(1)$ of complex numbers of absolute value 1. The direct product $\mathbf{R}^n$ of n copies of $\mathbf{R}$ is the **vector group** of dimension n, and the direct product $\mathbf{T}^n$ of n copies of $\mathbf{T}$ is the **torus** (or **torus group**) of dimension n (or n-**torus**). Both $\mathbf{T}^n$ and $\mathbf{F}$ are compact, while $\mathbf{R}^n$ and $\mathbf{Z}^n$ are not. We have $C(\mathbf{R})=\mathbf{R}$, $C(\mathbf{T})=\mathbf{Z}$, $C(\mathbf{Z})=\mathbf{T}$. Any finite Abelian group $\mathbf{F}$ is isomorphic to its character group $C(\mathbf{F})$. The direct product of a finite number of copies of **R**, **T**, **Z**, and a finite Abelian group **F**, namely, a group of the form $\mathbf{R}^l \times \mathbf{T}^m \times \mathbf{Z}^n \times \mathbf{F}$, is called an **elementary topological Abelian group**.

Any group in $\mathfrak{A}$ is isomorphic to the direct product of a vector group of some dimension and the extension of a compact group by a discrete group (the **structure theorem**). Hence, if the effect of the operation C is explicitly known, then the problem of finding the structure of groups in $\mathfrak{A}$ is reduced to the problem concerning discrete groups alone. For the structure of groups in $\mathfrak{A}$, the following theorem is known: If $G \in \mathfrak{A}$ is generated by a compact neighborhood of the identity e, then G is isomorphic to the direct product of a compact subgroup K and a group of the form $\mathbf{R}^n \times \mathbf{Z}^m$ (n, m are nonnegative integers). Then any compact subgroup of G is contained in K, which is the unique maximal compact subgroup of G. A group $G \in \mathfrak{A}$ generated by a compact neighborhood of e is the †projective limit of elementary topological Abelian groups. L. S. Pontryagin first proved a structure theorem of this type and then the duality theorem.

F. Compact Elements

An element a of a group $G \in \mathfrak{A}$ is called a **compact element** if the cyclic group $\{a^n | n \in \mathbf{Z}\}$ generated by a is contained in a compact subset of G. The set C_0 of all compact elements of G is a closed subgroup of G, and the quotient group G/C_0 does not contain any compact element other than the identity. In particular, if G is generated by a compact neighborhood of the identity, then C_0 coincides with the maximal compact subgroup K of G. Let C_0 be the set of all compact elements of a group $G \in \mathfrak{A}$. The annihilator $(C(G), C_0)$ is a connected component of the character group $C(G)$ of G. If G is a discrete group, then a compact element of G is an element of G of finite order.

G. Compact Groups and Discrete Groups

Suppose that two groups $G, X \in \mathfrak{A}$ are dual to each other. Then one group is compact if and only if the other group is discrete. By the duality theorem, the properties of a compact Abelian group G can be stated, in principle, through the properties of the discrete Abelian group $C(G)$. The following are a few such examples. Let G be a compact Abelian group. Then its †dimension is equal to the †rank of the discrete Abelian group $C(G)$. A subgroup Y of a discrete Abelian group X is called a **divisible subgroup** if the quotient group X/Y contains no element of finite order other than the identity. A compact Abelian group G is locally connected if and only if any finite subset of the character group $C(G)$ is contained in some divisible subgroup of $C(G)$ generated by a finite number of elements. Hence if a compact locally connected Abelian group G has an †open basis consisting of a countable number of open sets, then G is of the form $\mathbf{T}^a \times \mathbf{F}$, where $\mathbf{F}$ is a finite Abelian group and $\mathbf{T}^a$ is the direct product of an at most countable number of 1-dimensional torus groups **T**.

H. Dual Decomposition into Direct Products

Let G be a compact or discrete Abelian group, and let $\mathfrak{M} = \{H_\alpha | \alpha \in A\}$ be a family of closed subgroups of G. Let $\Delta(\mathfrak{M}) = \bigcap_{\alpha \in A} H_\alpha$, and denote by $\Sigma(\mathfrak{M})$ the smallest closed subgroup of G containing $\bigcup_{\alpha \in A} H_\alpha$. Then, with $\Omega = \{(C(G), H_\alpha) | \alpha \in A\}$, the relations $\Delta(\Omega) = (C(G), \Sigma(\mathfrak{M}))$ and $\Sigma(\Omega) = (C(G), \Delta(\mathfrak{M}))$ hold. Furthermore, suppose that G is decomposed into the direct product $G = \prod_{\alpha \in A} H_\alpha$, and for each $\alpha \in A$ put $K_\alpha = \Sigma(\mathfrak{M} - \{H_\alpha\})$, $X_\alpha = (C(G), K_\alpha)$. Then X_α is the character group of H_α, and $C(G)$ can be decomposed into the direct product $C(G) = \prod_{\alpha \in A} X_\alpha$. This decomposition of $C(G)$ into a direct product is called the **dual direct product decomposition** corresponding to the decomposition $G = \prod_{\alpha \in A} H_\alpha$.

I. Orthogonal Group Pairs

Suppose that for two groups G, G' there exists a mapping $(x, x') \rightarrow xx'$ of the Cartesian product $G \times G'$ into the set $U(1)$ of all complex numbers of absolute value 1 such that

$$(x_1 x_2) x' = (x_1 x')(x_2 x'),$$

$$x(x'_1 x'_2) = (x x'_1)(x x'_2).$$

Then G, G' are said to form a **group pair**. Suppose that G, G' form a group pair, and consider xx' to be a function $x(x')$ in x'. If two functions $x_1(x')$ and $x_2(x')$ coincide only when $x_1 = x_2$ and the same is true when the roles of G and G' are interchanged, then G, G' are said to form an **orthogonal group pair**. If G is a compact Abelian group, G' is a discrete Abelian group, and G, G' form an orthogonal group pair, then G, G' are dual to each other.

J. Commutative Lie Groups

An elementary topological Abelian group $\mathbf{R}^l \times \mathbf{T}^m \times \mathbf{Z}^n \times \mathbf{F}$ is a commutative †Lie group. Conversely, any commutative Lie group G

generated by a compact neighborhood of the identity is isomorphic to an elementary topological Abelian group. In particular, any connected commutative Lie group G is isomorphic to $\mathbf{R}^l \times \mathbf{T}^m$ for some l and m. A closed subgroup H of the vector group $\mathbf{R}^n$ of dimension n is isomorphic to $\mathbf{R}^p \times \mathbf{Z}^q$ $(0 \leqslant p+q \leqslant n)$. More precisely, there exists a basis $a_1, \ldots, a_n$ of the vector group $\mathbf{R}^n$ such that $H = \{\sum_{i=1}^{p} x_i a_i + \sum_{j=p+1}^{r} n_j a_j \mid x_i \in \mathbf{R}, n_j \in \mathbf{Z}\}$. Hence the quotient groups of $\mathbf{R}^n$ that are †separated topological groups are all isomorphic to groups of the form $\mathbf{R}^l \times \mathbf{T}^m$ $(0 \leqslant l+m \leqslant n)$. Any closed subgroup of the torus group $\mathbf{T}^n$ of dimension n is isomorphic to a group of the form $\mathbf{T}^p \times \mathbf{F}$ $(0 \leqslant p \leqslant n)$, where $\mathbf{F}$ is a finite Abelian group. Hence the quotient groups of $\mathbf{T}^n$ that are separated topological groups are all isomorphic to $\mathbf{T}^m$ $(0 \leqslant m \leqslant n)$. A †regular linear transformation of the linear space $\mathbf{R}^n$ is a continuous automorphism of the vector group $\mathbf{R}^n$, and in fact, any continuous automorphism of $\mathbf{R}^n$ is given by a regular linear transformation. Indeed, the group of all continuous automorphisms of $\mathbf{R}^n$ is isomorphic to the †general linear group $GL(n, \mathbf{R})$ of degree n. Any continuous automorphism of the torus group $\mathbf{T}^n = \mathbf{R}^n/\mathbf{Z}^n$ of dimension n is given by a regular linear transformation φ of $\mathbf{R}^n$ such that $\varphi(\mathbf{Z}^n) = \mathbf{Z}^n$. Hence the group of continuous automorphisms of $\mathbf{T}^n$ is isomorphic to the multiplicative group of all $n \times n$ matrices, with determinant ± 1 and with entries in the set of rational integers.

K. Kronecker's Approximation Theorem

Let H be a subgroup of a group $G \in \mathfrak{A}$ (not necessarily closed). Then $(G,(C(G), H))$ coincides with the closure $\bar{H}$ of H. In particular, H is †dense in G if and only if the annihilator $(C(G), H)$ consists of the identity alone. Now let $G = \mathbf{R}^n$ and let H be the subgroup of $\mathbf{R}^n$ generated by $\theta = (\theta_1, \ldots, \theta_n) \in \mathbf{R}^n$ and the natural †basis $e_1 = (1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$ of $\mathbf{R}^n$. Then H is dense in $\mathbf{R}^n$ if and only if $(\mathbf{R}^n, H) = \{0\}$; that is, $\theta_1, \ldots, \theta_n, 1$ are linearly independent over the rational number field $\mathbf{Q}$ (**Kronecker's approximation theorem**). This theorem implies that the torus group $\mathbf{T}^n$ of dimension n has a cyclic subgroup and a 1-parameter subgroup that are both dense in $\mathbf{T}^n$.

L. Linear Topology

Consider the discrete topology in a field Ω. Suppose that an Ω-module G has a topology that satisfies †Hausdorff's separation axiom and is such that a base for the neighborhood system of the zero element 0 consists of Ω-submodules, and suppose that G together with this topology constitutes a topological Abelian group. Then this topology is called a **linear topology**. If a linear topology is restricted to a Ω-submodule, then it is also a linear topology. If G is of finite rank, then any linear topology is the discrete topology. The discrete topology on G is a linear topology. Let H be a Ω-submodule. Then the subset $V = H + g$ of G obtained by translating H by an element g of G is called a **linear variety** in G. If V is a linear variety, then $\bar{V}$ is also a linear variety. If Ω-modules G, G' have linear topologies, a homomorphism of G into G' is always assumed to be open and continuous with respect to these topologies. A linear variety V in G is said to be **linearly compact** if, for any system $\{V_\alpha\}$ of linear varieties closed in V with the †finite intersection property, we have $\bigcap_\alpha V_\alpha \neq \varnothing$. In this case V is closed in G. If linearly compact Ω-submodules can be chosen as a base for the neighborhood system of the zero element of G, we say that G is **locally linearly compact**. The set $C_\Omega(G)$ of homomorphisms of an Ω-module G with linear topology into Ω is also an Ω-module. For any linearly compact Ω-submodule H of G, let $U(H) = \{\chi \mid \chi(g) = 0, g \in H\}$. Then, with $\{U(H)\}$ as a base for the neighborhood system, a linear topology can be introduced in $C_\Omega(G)$. According as G is discrete, linearly compact, or locally linearly compact, $C_\Omega(G)$ is linearly compact, discrete, or locally linearly compact. Let G, H be Ω-modules each of which has a linear topology, and let $\varphi: G \ni g \to \varphi_g \in C_\Omega(H)$, $\psi: H \ni h \to \psi_h \in C_\Omega(G)$ be homomorphisms such that $\varphi_g(h) = \psi_h(g)$. Then if one of φ, ψ is an isomorphism, so is the other. This is an analog of the Pontryagin duality theorem and is called the **duality theorem for Ω-modules**. In particular, a linearly compact Ω-module is the direct sum of 1-dimensional spaces (S. Lefschetz [3]).

References

[1] L. S. Pontryagin, Topological groups, first edition, Princeton Univ. Press, 1939; second edition, Gordon & Breach, 1966 (translation of the 1954 Russian edition).
[2] A. Weil, L'intégration dans les groupes topologiques et ses applications, Actualités Sci. Ind., Hermann, 1940.
[3] S. Lefschetz, Algebraic topology, Amer. Math. Soc. Colloq. Publ., 1942.
[4] W. Rudin, Fourier analysis on groups, Interscience, 1962.
[5] E. Hewitt and K. A. Ross, Abstract harmonic analysis, Springer, I, 1963; II, 1970.

423 (IV.6) Topological Groups

A. Definitions

If a †group G has the structure of a †topological space such that the mapping $(x, y) \rightarrow xy$ (product) of the Cartesian product $G \times G$ into G and the mapping $x \rightarrow x^{-1}$ (inverse) of G into G are both continuous, then G is called a **topological group**. The group G without a topological structure is called the **underlying group** of the topological group G, and the topological space G is called the **underlying topological space** of the topological group G. Let G, G' be topological groups. A mapping f of G into G' is called an **isomorphism** of the topological group G onto the topological group G' if f is an †isomorphism of the underlying group G onto the underlying group G' and also a †homeomorphism of the underlying topological space G onto the underlying topological space G'. Two topological groups are said to be **isomorphic** if there exists an isomorphism of one onto the other.

B. Neighborhood Systems

Let $\mathfrak{N}$ be the †neighborhood system of the identity e of a topological group G. Namely, $\mathfrak{N}$ consists of all subsets of G each of which contains an open set containing the element e. Then $\mathfrak{N}$ satisfies the following six conditions: (i) If $U \in \mathfrak{N}$ and $U \subset V$, then $V \in \mathfrak{N}$. (ii) If U, $V \in \mathfrak{N}$, then $U \cap V \in \mathfrak{N}$. (iii) If $U \in \mathfrak{N}$, then $e \in U$. (iv) For any $U \in \mathfrak{N}$, there exists a $W \in \mathfrak{N}$ such that $WW = \{xy \mid x, y \in W\} \subset U$. (v) If $U \in \mathfrak{N}$, then $U^{-1} \in \mathfrak{N}$. (vi) If $U \in \mathfrak{N}$ and $a \in G$, then $aUa^{-1} \in \mathfrak{N}$. Conversely, if a nonempty family $\mathfrak{N}$ of subsets of a group G satisfies conditions (i)–(vi), then there exists a †topology $\mathfrak{O}$ of G such that $\mathfrak{N}$ is the neighborhood system of e and G is a topological group with this topology. Moreover, such a topology is uniquely determined by $\mathfrak{N}$. †Left translation $x \rightarrow ax$ and †right translation $x \rightarrow xa$ in a topological group G are homeomorphisms of G onto G; thus if $\mathfrak{N}$ is the neighborhood system of the identity e, then $a\mathfrak{N} = \mathfrak{N}a$ is the neighborhood system of a, where $a\mathfrak{N} = \{aU \mid U \in \mathfrak{N}\}$.

If the underlying topological space of a topological group G is a †Hausdorff space, G is called a **T_2-topological group (Hausdorff topological group** or **separated topological group)**. If the underlying topological space of a topological group G is a †T_0-topological space, then, as is easily seen, it is a †T_1-topological space. If it is a T_1-topological space, then by the fact that the topology may be defined by a †uniformity, it is a †completely regular space, hence, in particular, a Hausdorff space (→ Section G). Thus a topological group whose underlying topological space is a T_0-topological space is a T_2-topological group.

C. Direct Product of Topological Groups

Consider a family $\{G_\alpha\}_{\alpha \in A}$ of topological groups. The Cartesian product $G = \prod_{\alpha \in A} G_\alpha$ of the underlying groups of G_α is a topological group with the †product topology of the underlying topological spaces of G_α. This topological group $G = \prod_{\alpha \in A} G_\alpha$ is called the **direct product** of topological groups G_α $(\alpha \in A)$.

D. Subgroups

Let H be a subgroup of the underlying group of a topological group G. Then H is a topological group with the topology of a †topological subspace of G (†relative topology). This topological group H is called a **subgroup** of G. A subgroup that is a closed (open) set is called a **closed (open) subgroup**. Any open subgroup is also a closed subgroup. For any subgroup H of a topological group G, the closure $\bar{H}$ of H is also a subgroup. If H is a normal subgroup, so is $\bar{H}$. If H is commutative, so is $\bar{H}$. In a T_2-topological group G, the †centralizer $C(M) = \{x \in G \mid xm = mx \ (m \in M)\}$ of a subset M of G is a closed subgroup of G. In particular, the †center $C = C(G)$ of a T_2-topological group is a closed normal subgroup.

E. Quotient Spaces

Given a subgroup H of a topological group G, let $G/H = \{aH \mid a \in G\}$ be the set of †left cosets, and let p be the canonical surjection $p(a) = aH$ of G onto G/H. Consider the †quotient topology on G/H, namely, the strongest topology such that p is a continuous mapping. Since a subset A of G/H is open when $p^{-1}(A)$ is an open set of G, p is also an †open mapping. The set G/H with this topology is called the **left quotient space** (or **left coset space**) of G by H. The **right quotient space** (or **right coset space)** $H \backslash G = \{Ha \mid a \in G\}$ is defined similarly. The quotient space G/H is discrete if and only if H is an open subgroup of G. The quotient space is a Hausdorff space if and only if H is a closed subgroup. If G/H and H are both †connected, then G itself is connected. If G/H and H are both †compact, then G is compact. If H is a closed subgroup of G and G/H, H are both †locally compact, then G is locally compact.

Suppose that H is a normal subgroup of a topological group G. Then the quotient group

G/H is a topological group with the topology of the quotient space G/H. This topological group is called the **quotient group** of the topological group G by the normal subgroup H.

F. Connectivity

The †connected component G_0 containing the identity e of a topological group G is a closed normal subgroup of G. The connected component that contains an element $a \in G$ is the coset $aG_0 = G_0 a$. G_0 is called the **identity component** of G. The quotient group G/G_0 is †totally disconnected. A connected topological group G is generated by any neighborhood U of the identity. Namely, any element of G can be expressed as the product of a finite number of elements in U. Totally disconnected (in particular, discrete) normal subgroups of a connected topological group G are contained in the center of G.

G. Uniformity

Let $\mathfrak{N}_0$ be the neighborhood system of the identity of a topological group G, and let $U_l = \{(x, y) \in G \times G \mid y \in xU\}$ for $U \in \mathfrak{N}_0$. Then a †uniformity having $\{U_l \mid U \in \mathfrak{N}_0\}$ as a base is defined on G. This uniformity is called the **left uniformity** of G. Left translation $x \rightarrow ax$ of G is †uniformly continuous with respect to the left uniformity. The **right uniformity** is defined similarly by $U_r = \{(x, y) \mid y \in Ux\}$. These two uniformities do not necessarily coincide. The mapping $x \rightarrow x^{-1}$ is a †uniform isomorphism of G considered as a uniform space with respect to the left uniformity onto the same group G considered as a uniform space with respect to the right uniformity. A topological group G is thus a †uniform space under a uniformity †compatible with its topology, and hence it is a completely regular space if the underlying topological space is a T_1-space.

H. Completeness

If a topological group G is †complete with respect to the left uniformity, then it is also complete with respect to the right uniformity, and conversely. In this case the topological group G is said to be **complete**. A locally compact T_2-topological group is complete. If a T_2-topological group G is isomorphic to a dense subgroup of a complete T_2-topological group $\hat{G}$, then $\hat{G}$ is called the **completion** of G, and G is said to be **completable**. A T_2-topological group G is not always completable. For a T_2-topological group G to be completable it is necessary and sufficient that any †Cauchy filter of G considered as a uniform space with respect to the left uniformity is mapped to a Cauchy filter of the same uniform space G under the mapping $x \rightarrow x^{-1}$. Then the completion $\hat{G}$ of G is uniquely determined up to isomorphism. A commutative T_2-topological group always has a completion $\hat{G}$, and $\hat{G}$ is also commutative. If each point of a T_2-topological group G has a †totally bounded neighborhood, there exists a completion $\hat{G}$, and $\hat{G}$ is locally compact.

I. Metrization

If a †metric can be introduced in a T_2-topological group G so that the metric gives the topology of G, then G is said to be **metrizable**. For a T_2-topological group G to be metrizable it is necessary and sufficient that G satisfy the †first axiom of countability. Then the metric can be chosen so that it is **left invariant**, i.e., invariant under left translation. Similarly, it can be chosen so that it is right invariant. In particular, the topology of a compact T_2-topological group that satisfies the first axiom of countability can be given by a metric that is both left and right invariant.

J. Isomorphism Theorems

Let G and G' be topological groups. If a homomorphism f of the underlying group of G into the underlying group of G' is a continuous mapping of the underlying topological space of G into that of G', f is called a **continuous homomorphism**. If f is a continuous open mapping, f is called a **strict morphism** (or **open continuous homomorphism**). A continuous homomorphism of a †paracompact locally compact topological group onto a locally compact T_2-topological group is an open continuous homomorphism.

A topological group G' is said to be **homomorphic** to a topological group G if there exists an open continuous homomorphism f of G onto G'. Let N denote the kernel $f^{-1}(e)$ of f. Then the quotient group G/N is isomorphic to G', with G/N and G' both considered as topological groups (**homomorphism theorem**). Let f be an open continuous homomorphism of a topological group G onto a topological group G', and let H' be a subgroup of G'. Then $H = f^{-1}(H')$ is a subgroup of G, and the mapping φ defined by $\varphi(gH) = f(g)H'$ is a homeomorphism of the quotient space G/H onto G'/H'. In particular, if H' is a normal subgroup, then H is also a normal subgroup and φ is an isomorphism of the quotient group G/H onto G'/H' as topological groups (**first isomorphism theorem**). Let H and N be subgroups of a topological group G such that $HN=NH$. Then the canonical mapping $f: h(H \cap N) \to hN$ of the quotient space $H/H \cap N$ to HN/N is a continuous bijection but not necessarily an open mapping. In particular, if N is a normal subgroup of the group HN, then f is a continuous homomorphism. In addition, if f is an open mapping, the quotient groups $H/H \cap N$ and HN/N are isomorphic as topological groups (**second isomorphism theorem**). For example, f is an open mapping (1) if N is compact or (2) if G is locally compact, HN and N are closed subgroups of G, and H is the union of a countable number of compact subsets. Let H be a subgroup of a topological group G and N be a normal subgroup of G such that $H \supset N$. Then the canonical mapping of the quotient space $(G/N)/(H/N)$ onto G/H is a homeomorphism. In particular, if H is also a normal subgroup, the quotient groups $(G/N)/(H/N)$ and G/H are isomorphic as topological groups (**third isomorphism theorem**).

K. The Projective Limit

Let $\{G_\alpha\}_{\alpha \in A}$ be a family of topological groups indexed by a †directed set A, and suppose that if $\alpha \leqslant \beta$, there exists a continuous homomorphism $f_{\alpha\beta}: G_\beta \to G_\alpha$ such that $f_{\alpha\gamma} = f_{\alpha\beta} \circ f_{\beta\gamma}$ if $\alpha \leqslant \beta \leqslant \gamma$. Then the collection $\{G_\alpha, f_{\alpha\beta}\}$ of the family $\{G_\alpha\}_{\alpha \in A}$ of topological groups together with the family $\{f_{\alpha\beta}\}$ of mappings is called a **projective system** of topological groups. Consider the direct product $\prod_{\alpha \in A} G_\alpha$ of topological groups $\{G_\alpha\}$, and denote by G the set of all elements $x = \{x_\alpha\}_{\alpha \in A}$ of $\prod G_\alpha$ that satisfy $x_\alpha = f_{\alpha\beta}(x_\beta)$ for $\alpha \leqslant \beta$. Then G is a subgroup of $\prod G_\alpha$. The topological group G obtained in this way is called the **projective limit** of the projective system $\{G_\alpha, f_{\alpha\beta}\}$ of topological groups and is denoted by $G = \varprojlim G_\alpha$. If each G_α is a T_2-topological (resp. complete) group, then G is also a T_2-topological (complete) group.

Now consider another projective system $\{G'_\alpha, f'_{\alpha\beta}\}$ of topological groups indexed by the same A, and consider continuous homomorphisms $u_\alpha: G_\alpha \to G'_\alpha$ such that $u_\alpha \circ f_{\alpha\beta} = f'_{\alpha\beta} \circ u_\beta$ for $\alpha \leqslant \beta$. Then there exists a unique continuous homomorphism u of $G = \varprojlim G_\alpha$ into $G' = \varprojlim G'_\alpha$ such that for any $\alpha \in A$, $u_\alpha \circ f_\alpha = f'_\alpha \circ u$ holds, where $f_\alpha(f'_\alpha)$ is the restriction to $G(G')$ of the projection of $\prod G_\alpha$ ($\prod G'_\alpha$) onto $G_\alpha(G'_\alpha)$. The homomorphism u is called the **projective limit** of the family $\{u_\alpha\}$ of continuous homomorphisms and is denoted by $u = \varprojlim u_\alpha$. Let G be a T_2-topological group, and let $\{H_\alpha\}_{\alpha \in A}$ be a decreasing sequence ($H_\alpha \supset H_\beta$ for $\alpha \leqslant \beta$) of closed normal subgroups of G. Consider the quotient group G/H_α, and let $f_{\alpha\beta}$ be the canonical mapping $gH_\beta \to gH_\alpha$ of G_β to G_α for $\alpha \leqslant \beta$. Then $\{G_\alpha, f_{\alpha\beta}\}$ is a projective system of topological groups. Let f_α be the projection of G onto $G_\alpha = G/H_\alpha$, and let $f = \varprojlim f_\alpha$. Now assume that any neighborhood of the identity of G contains some H_α and that some H_γ is complete. Then $f = \varprojlim f_\alpha$ is an isomorphism of G onto $\varprojlim G/H_\alpha$ as topological groups. (For a general discussion of the topological groups already discussed → [1, 4].)

L. Locally Compact Groups

For the rest of this article, all topological groups under consideration are assumed to be T_2-topological groups. The identity component G_0 of a locally compact group G is the intersection of all open subgroups of G. In particular, any neighborhood of the identity of a totally disconnected locally compact group contains an open subgroup. A totally disconnected compact group is a projective limit of finite groups with discrete topology.

A T_1-topological space L is called a **local Lie group** if it satisfies the following six conditions: (i) There exist a nonempty subset M of $L \times L$ and a continuous mapping $\mu: M \to L$, called **multiplication** ($\mu(a,b)$ is written as ab). (ii) If (a,b), (ab,c), (b,c), (a,bc) are all in M, then $(ab)c = a(bc)$. (iii) L contains an element e, called the **identity**, such that $L \times \{e\} \subset M$ and $ae = a$ for all $a \in L$. (iv) There exists a nonempty open subset N of L and a continuous mapping $v: N \to L$ such that $av(a) = e$ for all $a \in N$. (v) There exist a neighborhood U of e in L and a homeomorphism f of U into a neighborhood V of the origin in the Euclidean space $\mathbf{R}^n$. (vi) Let D be the open subset of $V \times V$ defined by $D = \{(x,y) \in V \times V \mid (f^{-1}(x), f^{-1}(y)) \in M, f^{-1}(x), f^{-1}(y) \in U\}$. Then the function $F: D \to V$ defined by $F(x,y) = f\mu(f^{-1}(x), f^{-1}(y))$ is of †class C^ω.

For any neighborhood U of the identity e of a connected locally compact group G, there exist a compact normal subgroup K and a subset L that is a local Lie group under the †induced topology and the group operations of G such that the product LK is a neighborhood of e contained in U. Furthermore, under $(l,k) \to lk$, LK is homeomorphic to the product space $L \times K$. Any compact subgroup of a connected locally compact group G is contained in a maximal compact subgroup, and maximal compact subgroups of G are †conjugate. For a maximal compact subgroup K of G, there exists a finite number of subgroups $H_1, \ldots, H_r$ of G, each of which is isomorphic to the additive group of real numbers such that $G = KH_1 \ldots H_r$, and the mapping $(k, h_1, \ldots, h_r) \to kh_1 \ldots h_r$ is a homeomorphism of the direct product $K \times H_1 \times \ldots \times H_r$ onto G. Any locally compact group has a left-invariant positive measure and a right-invariant positive measure, which are uniquely determined up to constant multiples (→ 225 Invariant Measures). Using these measures, the theory of harmonic analysis on the additive group **R** of real numbers can be extended to that on G (→ 69 Compact Groups; 192 Harmonic Analysis; 422 Topological Abelian Groups; 437 Unitary Representations).

M. Locally Euclidean Groups

Suppose that each point of a topological group G has a neighborhood homeomorphic to an open set of a given Euclidean space. Then G is called a **locally Euclidean group**. If the underlying topological space of a topological group has the structure of a †real analytic manifold such that the group operation $(x, y) \rightarrow xy^{-1}$ is a real analytic mapping, then G is called a †**Lie group**. A Lie group is a locally Euclidean group.

N. Hilbert's Fifth Problem

Hilbert's fifth problem asks if every locally Euclidean group is a Lie group (→ 196 Hilbert). This problem was solved affirmatively in 1952; it was proved that any †locally connected finite-dimensional locally compact group is a Lie group (D. Montgomery and L. Zippin [3]). In connection with this, the relation between Lie groups and general locally compact groups has been studied, and the following results have been obtained: A necessary and sufficient condition for a locally compact group to be a Lie group is that there exist a neighborhood of the identity e that does not contain any subgroup (or any normal subgroup) other than $\{e\}$. A locally compact group has an open subgroup that is the projective limit of Lie groups. Hilbert's fifth problem is closely related to the following problem: Find the conditions for a †topological transformation group operating †effectively on a manifold to be a Lie group (→ 431 Transformation Groups).

O. Covering Groups

Let $\mathfrak{G}$ be the collection of all †arcwise connected and †locally arcwise connected T_2-topological groups. Suppose that $G^* \in \mathfrak{G}$ is a †covering space of $G \in \mathfrak{G}$ and the †covering mapping $f: G^* \rightarrow G$ is an open continuous homomorphism, with G^* and G considered as topological groups. Then G^* (or, more precisely, (G^*, f)) is called a **covering group** of G. Then the kernel $f^{-1}(e) = D$ of f is a discrete subgroup contained in the center of G^*, and G^*/D and G, considered as topological groups, are isomorphic to each other. Let $\pi_1(G)$ be the †fundamental group of G. The natural homomorphism $f^*: \pi_1(G^*) \rightarrow \pi_1(G)$ induced by f is an injective homomorphism, and if we identify $\pi_1(G^*)$ with the subgroup $f^*(\pi_1(G^*))$ of $\pi_1(G)$, we have $D \cong \pi_1(G)/\pi_1(G^*)$. Conversely, if D is any discrete subgroup contained in the center of $G^* \in \mathfrak{G}$, then G^* is a covering group of $G = G^*/D$. For any covering space (G^*, f) of $G \in \mathfrak{G}$, a multiplication law can be introduced in G^* so that G^* is a topological group belonging to $\mathfrak{G}$ and (G^*, f) is a covering group of G. In particular, any $G \in \mathfrak{G}$ has a †simply connected covering group $(\tilde{G}, \varphi)$. Then for any covering group (G^*, f) of G, there exists a homomorphism $f^*: \tilde{G} \rightarrow G^*$, and $(\tilde{G}, f^*)$ is a covering group of G^*. Furthermore, $\varphi = f \circ f^*$. Hence, in particular, any simply connected covering group of G is isomorphic to $\tilde{G}$, with G and $\tilde{G}$ considered as topological groups. This simply connected covering group $(\tilde{G}, \varphi)$ is called the **universal covering group**.

Let G and G' be topological groups, and let e and e' be their identities. A homeomorphism f of a neighborhood U of e onto a neighborhood U' of e' is called a **local isomorphism** of G to G' if it satisfies the following two conditions: (i) If a, b, ab are all contained in U, then $f(ab) = f(a)f(b)$. (ii) Let $f^{-1} = g$, then if $a', b', a'b' \in U'$, $g(a'b') = g(a')g(b')$ holds. If there exists a local isomorphism of G to G', we say that G and G' are **locally isomorphic**. If G^* is a covering group of G, then G^* and G are locally isomorphic. For two topological groups G and G' to be locally isomorphic it is necessary and sufficient that the universal covering groups of G and G' be isomorphic. For two connected Lie groups to be locally isomorphic it is necessary and sufficient that their †Lie algebras be isomorphic.

Let f be a mapping of a neighborhood U of the identity of a topological group G into a group H such that if a, b, ab are all contained in U, then $f(ab) = f(a)f(b)$. Then f is called a **local homomorphism** of G into H and U is called its **domain**. A local homomorphism of a simply connected group $G \in \mathfrak{G}$ into a group H can be extended to a homomorphism of G into H if the domain is connected [2, 4].

P. Topological Rings and Fields

If a ring R has the structure of a topological group such that $(x, y) \rightarrow x + y$ (sum) and $(x, y) \rightarrow xy$ (product) are both continuous mappings of $R \times R$ into R, then R is called a **topological ring**. If a topological ring K is a field (not necessarily commutative) such that $x \rightarrow x^{-1}$

(inverse element) is a continuous mapping of $K^* = K - \{0\}$ into K^*, then K is called a **topological field**. Let us assume that K is a topological field that is a locally compact Hausdorff space and is not discrete. If K is connected, then K is a †division algebra of finite rank over the field $\mathbf{R}$ of real numbers; hence it is isomorphic to the field $\mathbf{R}$ of real numbers, the field $\mathbf{C}$ of complex numbers, or the †quaternion field $\mathbf{H}$. If K is not connected, then K is totally disconnected and is isomorphic to a division algebra of finite rank over the †p-adic number field $\mathbf{Q}_p$ or a division algebra of finite rank over the †formal power series field with coefficients in a finite field [4].

For various important classes of topological groups → 69 Compact Groups; 249 Lie Groups; 422 Topological Abelian Groups; 424 Topological Linear Spaces.

References

[1] N. Bourbaki, Eléments de mathématique, Topologie générale, ch. 3, Actualités Sci. Ind., 1143c, Hermann, third edition, 1960; English translation, General topology, pt. 1, Addison-Wesley, 1966.
[2] C. Chevalley, Theory of Lie groups I, Princeton Univ. Press, 1946.
[3] D. Montgomery and L. Zippin, Topological transformation groups, Interscience, 1955.
[4] L. S. Pontryagin, Topological groups, first edition, Princeton Univ. Press, 1939, second edition, Gordon & Breach, 1966. (Second English edition translated from the second Russian edition, 1954.)

424 (XII.5) Topological Linear Spaces

A. Definition

A †linear space E over the real or complex number field K is said to be a **topological linear space**, **topological vector space**, or **linear topological space** if E is a †topological space and the basic operations $x+y$ and αx ($x, y \in E$, $\alpha \in K$) in the linear space are continuous as mappings of $E \times E$ and $K \times E$, respectively, into E. The coefficient field K may be a general †topological field, although it is usually assumed to be the real number field $\mathbf{R}$ or the complex number field $\mathbf{C}$, and accordingly E is called a **real topological linear space** or a **complex topological linear space**. Topological linear spaces are generalizations of †normed linear spaces and play an important role in the study of †function spaces, such as the †space of distributions, that are not †Banach spaces.

Each topological linear space E is equipped with a †uniform topology in which translations of the neighborhoods of zero form a †uniform family of neighborhoods, and the addition $x+y$ and the multiplication αx by a scalar α are uniformly continuous relative to this uniform topology. In particular, if for each $x \neq 0$ there is a neighborhood of the origin that does not contain x, then E satisfies the †separation axiom T_1 and hence is a †completely regular space. The †completion $\hat{E}$ of E is also a topological linear space.

We assume in this article that K is the real or complex number field and E is a topological linear space over K satisfying the axiom of T_1-spaces. Then E is finite-dimensional if and only if E has a †totally bounded neighborhood of zero. The topology of E is †metrizable if and only if it satisfies the †first countability axiom.

B. Linear Functional

A K-valued function $f(x)$ on E is said to be a **linear functional** if it satisfies (i) $f(x+y)=f(x)+f(y)$ and (ii) $f(\alpha x)=\alpha f(x)$. A linear functional that is continuous relative to the topologies of E and K is said to be a continuous linear functional. (Sometimes continuous linear functionals are simply called linear functionals, while abstract linear functionals are called **algebraic linear functionals**.) The following three statements are equivalent for linear functionals $f(x)$: (i) $f(x)$ is continuous; (ii) the half-space $\{x \in E \mid \operatorname{Re} f(x) > 0\}$ is open; (iii) the hyperplane $\{x \in E \mid f(x)=0\}$ is closed.

C. The Hahn-Banach Theorem

A linear functional $f(x)$ defined on a linear subspace F of E can be extended to a continuous linear functional on E if and only if there exists an open †convex neighborhood V of the origin in E that is disjoint with $\{x \in F \mid f(x)=1\}$. Furthermore, if $f(x)$ can be extended, at least one extension $f(x)$ never takes the value 1 on V (**Hahn-Banach theorem**).

D. Dual Spaces

The set E' of all continuous linear functionals on E is called the **dual space** of E. It is often denoted by E^* and is also called the **conjugate space** or **adjoint space**. It forms a linear space when $f+g$ and αf ($f, g \in E'$, $\alpha \in K$) are defined by $(f+g)(x)=f(x)+g(x)$ and $(\alpha f)(x)=\alpha(f(x))$ for $x \in E$.

E. Locally Convex Spaces

A topological linear space is said to be **locally convex** if it has a family of convex sets as a †base of the neighborhood system of 0. It follows from the Hahn-Banach theorem that for each $x \neq 0$ in a locally convex space E there is a continuous linear functional f such that $f(x) \neq 0$. A subset M of E is said to be **circled** if M contains $\alpha M = \{\alpha x \mid x \in M\}$ whenever $|\alpha| \leqslant 1$. A set that is both circled and convex is called **absolutely convex**. In a locally convex space, a family of absolutely convex and closed sets can be chosen as a base of the neighborhood system of the origin. Let A and B be subsets of E. A is said to **absorb** B if there is an $\alpha > 0$ such that $\alpha A \supset B$. A set V that absorbs every point $x \in E$ is called **absorbing**. Neighborhoods of 0 are absorbing.

F. Seminorms

A real-valued function $p(x)$ on E is said to be a **seminorm** (or **pseudonorm**) if it satisfies (i) $0 \leqslant p(x) < +\infty$ $(x \in E)$; (ii) $p(x+y) \leqslant p(x) + p(y)$; and (iii) $p(\alpha x) = |\alpha| p(x)$. The relation $V = \{x \mid p(x) \leqslant 1\}$ gives a one-to-one correspondence between seminorms $p(x)$ and absolutely convex absorbing sets V whose intersection with any line through the origin is closed. In terms of seminorms, the Hahn-Banach theorem states: Let E be a linear space on which a seminorm $p(x)$ is given. If a linear functional $f(x)$ defined on a linear subspace F of E satisfies $|f(x)| \leqslant p(x)$ on F, then $f(x)$ can be extended to the whole space E in such a way that the inequality holds on E.

The topology of a locally convex space is determined by the family of continuous seminorms on it. Conversely, if there is a family of seminorms $\{p_\lambda(x)\}$ $(\lambda \in \Lambda)$ on a linear space E over K that satisfies (iv) $p_\lambda(x) = 0$ for all λ implies $x = 0$, then there exists on E the weakest locally convex topology that renders the seminorms continuous. This topology is called the locally convex topology determined by $\{p_\lambda(x)\}$.

We assume that E is a locally convex space whose topology is determined by the family of seminorms $\{p_\lambda(x)\}$ $(\lambda \in \Lambda)$. Then a †net x_ν of E converges to x if and only if $p_\lambda(x_\nu - x) \to 0$ for all $\lambda \in \Lambda$. If F is a locally convex space whose topology is determined by the family of seminorms $\{q_\mu(y)\}$, then a necessary and sufficient condition for a linear mapping $u: E \to F$ to be continuous is that for every $q_\mu(y)$ there exist a finite number of $\lambda_1, \ldots, \lambda_n \in \Lambda$ and a constant C such that $q_\mu(u(x)) \leqslant C(p_{\lambda_1}(x) + \ldots + p_{\lambda_n}(x))$ $(x \in E)$.

A set is said to be **bounded** if it is absorbed by every neighborhood of zero. When the topology of E is determined by the family $\{p_\lambda(x)\}$ of seminorms a set B is bounded if and only if every p_λ is bounded on B. Totally bounded sets are bounded. The unit ball in a normed space is bounded. Conversely, a locally convex space is normable if it has a bounded neighborhood of 0. A locally convex space is called **quasicomplete** if every bounded closed set is complete. Since Cauchy sequences $\{x_n\}$ are totally bounded, all Cauchy sequences converge in a quasicomplete space (i.e., the space is sequentially complete).

G. Pairing of Linear Spaces

Let E and F be linear spaces over the same field K. A K-valued function $B(x,y)$ $(x \in E, y \in F)$ on $E \times F$ is called a **bilinear functional** or **bilinear form** if for each fixed $y \in F$ (resp. $x \in E$), it is a linear functional of x (resp. y). When a bilinear functional $\langle x, y \rangle$ on $E \times F$ is given so that $\langle x, y \rangle = 0$ for all $y \in F$ (all $x \in E$) implies $x = 0$ $(y = 0)$, then E and F are said to form a (separated) **pairing** relative to the **inner product** $\langle x, y \rangle$. A locally convex space E and its dual space E' form a pairing relative to the natural inner product $\langle x, x' \rangle = x'(x)$ $(x \in E, x' \in E')$.

H. Weak Topologies

When E and F form a pairing relative to an inner product $\langle x, y \rangle$, the locally convex topology on E determined by the family of seminorms $\{|\langle x, y \rangle| \mid y \in F\}$ is called the **weak topology (relative to the pairing** $\langle E, F \rangle$) and is denoted by $\sigma(E, F)$. A net x_ν in E is said to **converge weakly** if it converges in the weak topology. When E and E' are a locally convex space and its dual space, $\sigma(E, E')$ is called the **weak topology** of E, and $\sigma(E', E)$ the **weak* topology** of E'. The weak topology on a locally convex space E is weaker than the original topology on E. Consequently, a weakly closed set is closed. If the set is convex, the converse holds, and hence a convex closed set is weakly closed. Also, boundedness is preserved if we replace the original topology by the weak topology. Thus a weakly bounded set is bounded.

Let E and F form a pairing relative to $\langle x, y \rangle$, and let A be a subset of E. Then the set A° of points $y \in F$ satisfying $\operatorname{Re}\langle x, y \rangle \geqslant -1$ for all $x \in A$ is called the **polar** of A (relative to the pairing). If A is absolutely convex, A° is also absolutely convex and is the set of points y such that $|\langle x, y \rangle| \leqslant 1$ for all $x \in A$. If A is a convex set containing zero, its (weak) closure is equal to the **bipolar** $A^{\circ\circ} = (A^\circ)^\circ$ (**bipolar theorem**). In general, let A be a subset of a topological linear space E. We call the smallest closed convex set containing A the **closed convex hull** of A. If E is locally convex, the bipolar $A^{\circ\circ}$ relative to E' coincides with the closed convex hull of $A \cup \{0\}$.

A subset B of the dual space E' is †equicontinuous on E if and only if it is contained in the polar V° of a neighborhood V of 0 in E. Also, V° is weak*- compact in E' (**Banach-Alaoglu theorem**).

I. Barreled Spaces and Bornological Spaces

An absorbing absolutely convex closed set in a locally convex space E is called a **barrel**. In a sequentially complete space (hence in a quasi-complete space also), a barrel absorbs every bounded set. A locally convex space is said to be **barreled** if each barrel is a neighborhood of 0. A locally convex space is said to be **quasi-barreled** (or **evaluable**) if each barrel that absorbs every bounded set is a neighborhood of 0. Furthermore, a locally convex space is said to be **bornological** if each absolutely convex set that absorbs every bounded set is a neighborhood of 0. Bornological spaces are quasi-barreled. However, they are not necessarily barreled. Furthermore, barreled spaces are not necessarily bornological. A metrizable locally convex space, i.e., a space whose topology is determined by a countable number of seminorms, is bornological. A complete metrizable locally convex space is called a **locally convex Fréchet space** (**(F)-space** or simply **Fréchet space**). To distinguish it from Fréchet space as in 37 Banach Spaces, it is sometimes called a **Fréchet space in the sense of Bourbaki**. (F)-spaces are bornological and barreled.

A continuous linear mapping $u: E \to F$ of one locally convex space into another maps each bounded set of E to a bounded set in F. Conversely, if E is bornological, then each linear mapping that maps every bounded sequence to a bounded set is continuous.

J. The Banach-Steinhaus Theorem

In the dual space of a barreled space E, each (weak*-)bounded set is equicontinuous. Thus if a sequence of continuous linear mappings u_n of E into a locally convex space F converges at each point of E, then u_n converges uniformly on each totally bounded set of E, and the limit linear mapping is continuous (**Banach-Steinhaus theorem**).

K. The S-Topology

Let E and F be paired linear spaces relative to the inner product $\langle x, y \rangle$. When a family S of (weakly) bounded sets of F generates a dense subspace of F, the family of seminorms $\{\sup_{y \in B} |\langle x, y \rangle| \mid B \in S\}$ determines a locally convex topology on E. This is called the **S-topology** or **topology of uniform convergence on members of** S, because $x_\nu \to x$ in the S-topology is equivalent to the uniform convergence of $\langle x_\nu, y \rangle \to \langle x, y \rangle$ on each $B \in S$. The space E with the S-topology is denoted by E_S. The weak topology is the same as the topology of pointwise convergence. The S-topology in which S is the family of all bounded sets in F is called the **strong topology** and is denoted by $\beta(E, F)$. The dual space E' of a locally convex space E is usually regarded as a locally convex space with the strong topology $\beta(E', E)$. It is called the **strong dual space**. The topology of a locally convex space E is that of uniform convergence on equicontinuous sets of E'. The topology of a barreled space E coincides with the strong topology $\beta(E, E')$.

L. Grothendieck's Criterion of Completeness

Let E and F be paired spaces, and let S be a family of absolutely convex bounded sets of F such that: (i) the sets of S generate F; (ii) if B_1, $B_2 \in S$, then there is a $B_3 \in S$ such that $B_3 \supset B_1$ and $B_3 \supset B_2$. Then E_S is complete if and only if each algebraic linear functional $f(y)$ on F that is weakly continuous on every $B \in S$ is expressed as $f(y) = \langle x, y \rangle$ for some $x \in E$. When E_S is not complete, the space of all linear functionals satisfying this condition gives the completion $\hat{E}_S$ of E_S.

M. Mackey's Theorem

Let E, F, and S satisfy the same conditions as in Section L. Then the dual space of E_S is equal to the union of the weak completions of λB, where $\lambda > 0$ and $B \in S$ (**Mackey's theorem**).

N. The Mackey Topology

When E and F form a pairing, the topology on E of uniform convergence on convex weakly compact sets of F is called the **Mackey topology** and is denoted by $\tau(E, F)$. The dual space of E endowed with a locally convex topology T coincides with F if and only if T is stronger than the weak topology $\sigma(E, F)$ and weaker than the Mackey topology $\tau(E, F)$ (**Mackey-Arens theorem**). A locally convex space is said to be a **Mackey space** if the topology is equal to the Mackey topology $\tau(E, E')$. Every quasi-barreled space is a Mackey space.

O. Reflexivity

Let E be a locally convex space. The dual space E'' of the dual space E' equipped with the strong topology contains the original space E. We call E **semireflexive** if $E''=E$, and **reflexive** if in addition the topology of E coincides with the strong topology $\beta(E, E')$. E is semireflexive if and only if every bounded weakly closed set of E is weakly compact. E is reflexive if and only if E is semireflexive and (quasi)barreled.

A barreled space in which every bounded closed set is compact is called a **Montel space** or **(M)-space**. (M)-spaces are reflexive, and their strong dual spaces are also (M)-spaces.

Many of the function spaces that appear in applications are (F)-spaces or their dual spaces. For these spaces detailed consequences of the countability axiom are known [7, 8]. A convex set C in the dual space E' of an (F)-space E is weak*-closed if and only if for every neighborhood V of 0 in E, $C \cap V^{\circ}$ is weak*-closed (**Kreĭn-Shmul'yan theorem**). The strong dual space E' of an (F)-space E is (quasi)barreled if and only if it is bornological. In particular, the dual space of a reflexive (F)-space is bornological.

P. (DF)-Spaces

A locally convex space is called a **(DF)-space** if it satisfies: (i) There is a countable base of bounded sets (i.e., every bounded set is included in one of them); (ii) if the intersection V of a countable number of absolutely convex closed neighborhoods of 0 absorbs every bounded set, then V is also a neighborhood of 0. The dual space of an (F)-space is a (DF)-space, and the dual space of a (DF)-space is an (F)-space. A linear mapping of a (DF)-space E into a locally convex space F is continuous if and only if its restriction to every bounded set of E is continuous. A quasicomplete (DF)-space is complete.

Q. Bilinear Mappings

A bilinear mapping $b(x, y)$ on locally convex spaces E and F ($x \in E, y \in F$) to a locally convex space G is said to be **separately continuous** if for each fixed $y \in F$ ($x \in E$) it is continuous as a function of x (y). The linear mappings obtained from $b(x, y)$ by fixing x (y) are denoted by $b_x(y)$ ($b_y(x)$). We call $b(x, y)$ **hypocontinuous** if for each bounded set B of E and B' of F, $\{b_x(y) \mid x \in B\}$ and $\{b_y(x) \mid y \in B'\}$ are equicontinuous. A continuous bilinear mapping is hypocontinuous. However, the converse is not always true. A separately continuous bilinear mapping is not necessarily hypocontinuous. If both E and F are barreled, however, then every separately continuous mapping is hypocontinuous. If E is an (F)-space and F is metrizable, then every separately continuous bilinear mapping is continuous. Similarly, if both E and F are (DF)-spaces, then every hypocontinuous bilinear mapping is continuous.

R. Tensor Products

It is possible to introduce many topologies in the †tensor product $E \otimes F$ of locally convex spaces E and F. The **projective topology** (or **topology** π) is defined to be the strongest topology such that the natural bilinear mapping $E \times F \rightarrow E \otimes F$ is continuous. The dual space of $E \otimes_{\pi} F$ is identified with the space $B(E, F)$ of all continuous bilinear functionals on $E \times F$. The completion of $E \otimes_{\pi} F$ is denoted by $E \hat{\otimes} F$. The **topology of biequicontinuous convergence** (or **topology** ε) is defined to be the topology of uniform convergence on sets $V^{\circ} \times U^{\circ}$, where V and U are neighborhoods of 0 in E and F, respectively, considering the elements of $E \otimes F$ as linear functionals on $E' \otimes F'$ by the natural pairing of $E \otimes F$ and $E' \otimes F'$. The completion of $E \otimes_{\varepsilon} F$ is denoted by $E \hat{\hat{\otimes}} F$. The dual space of $E \otimes_{\varepsilon} F$ coincides with the subspace $J(E, F)$ of $B(E, F)$ composed of the union of the absolute convex hulls of the products $V^{\circ} \otimes U^{\circ}$ of equicontinuous sets. The elements of $J(E, F)$ are called **integral bilinear functionals**.

Closely related to $E \hat{\hat{\otimes}} F$ is L. Schwartz's ε **tensor product** $E \varepsilon F$ [12]. (They coincide if E and F are complete and if E or F has the †approximation property.) $E \varepsilon F$ can be regarded as (i) a space of vector-valued functions if E is a space of functions and F is an abstract locally convex space, especially a space of functions of two variables if E and F are, respectively, spaces of functions of one variable, and (ii) a space of operators $G \rightarrow F$ if E is the dual space G' of a locally convex space G.

S. Nuclear Spaces

Let E be a locally convex space, V be an absolutely convex closed neighborhood of the origin, and $p(x)$ be the seminorm corresponding to V. Then we denote by E_V the normed space with norm $p(x)$ obtained from E by identifying the two elements x and y with $p(x-y)=0$. If $U \subset V$, then a natural linear mapping $\varphi_{U,V}: E_U \rightarrow E_V$ is defined.

A locally convex space E is said to be a

nuclear space (resp. **Schwartz space** or simply **(S)-space**) if for each absolutely convex closed neighborhood V of 0 there is another U such that $\varphi_{U,V}$ is a †nuclear operator (resp. †compact operator) as an operator of E_U into the completion of E_V. A nuclear space or (S)-space is an (M)-space if it is quasicomplete and quasibarreled. A locally convex space E is a nuclear space if and only if the topologies π and ε coincide on the tensor product $E \otimes F$ with any locally convex space F. Accordingly, it follows that $B(E, F) = J(E, F)$. This can be regarded as a generalization of Schwartz's **kernel theorem**, which says that every separately continuous bilinear functional on $\mathscr{D}_x \times \mathscr{D}_y$ is represented by an integral with kernel in $\mathscr{D}'_{xy}$. The theory of topological tensor products and nuclear spaces is due to Grothendieck [9].

A locally convex space E is a nuclear (F)-space if and only if E is isomorphic to a closed subspace of $C^\infty(-\infty, \infty)$ (T. Kōmura and Y. Kōmura, 1966). An example of a nuclear (F)-space without basis is known (B. S. Mityagin and N. M. Zobin, 1974).

T. Gel'fand Triplet

Let H and L be Hilbert spaces. If L is a dense subspace of H and the injection $L \to H$ is a †Hilbert-Schmidt operator, then $H = H'$ is regarded as a dense subspace of L' and the injection $H' \to L'$ is a Hilbert-Schmidt operator. In this case, (L, H, L') is called a **Gel'fand triplet** (or a **rigged Hilbert space**).

A subset of H is called a cylindrical set if it is expressed in the form $P_F^{-1}(B)$ by the orthogonal projection P_F onto a finite-dimensional subspace F and a Borel subset B of F. If a finitely additive positive measure μ with $\|\mu\|_1 = 1$ defined on the cylindrical sets of H satisfies (i) μ is countably additive on cylindrical sets for a fixed F and (ii) for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|x\| < \delta$ implies $\mu\{y \in H \mid |\langle x, y\rangle| \geq 1\} < \varepsilon$, then μ is the restriction of a countably additive measure $\tilde{\mu}$ defined on the Borel subsets of L' (**Minlos's theorem**, 1959).

Let T be a self-adjoint operator in H. Then T has a natural extension $\tilde{T}$ in L' and almost every continuous spectrum λ of T has an associated eigenvector x_λ in L': $\tilde{T}x_\lambda = \lambda x_\lambda$, $x_\lambda \in L'$.

U. The Extreme Point Theorem

Let A be a subset of a linear space E. A point $x \in A$ is said to be an **extreme point** if x is an extreme point of any real segment containing x and contained in A. If A is a compact convex subset of a locally convex space E, A is the convex closed hull of (i.e., smallest convex closed set containing) the set of its extreme points (**Kreĭn-Milman theorem**). In applications it is important to know whether every point of A is represented uniquely as an integral of extreme points. For a metrizable convex compact subset A of a locally convex space E, the following two conditions are equivalent (**Choquet's theorem**): (i) A is a **simplex**, i.e., if we put $\tilde{A} = \{(\lambda x, \lambda) \mid x \in A, \lambda > 0\} \subset E \times \mathbf{R}^1$, the vector space $\tilde{A} - \tilde{A}$ becomes a †lattice with positive cone $\tilde{A}$; (ii) for any $x \in A$ there exists a unique positive measure μ on A with $\|\mu\|_1 = 1$ such that $l(x) = \int_A l(y)\, d\mu(y)$ $(l \in E')$ and the support of μ is contained in the set of extreme points of A.

V. Weakly Compact Set

A subset of a quasicomplete locally convex space is relatively weakly compact if and only if every sequence in the set has a weak accumulation point (**Eberlein's theorem**). If E is a metrizable locally convex space, every weakly compact set of E is weakly sequentially compact (**Shmul'yan's theorem**). If E is a quasicomplete locally convex space, the convex closed hull of any weakly compact subset is weakly compact (**Kreĭn's theorem**). If E is not quasicomplete, this is not necessarily true.

W. Permanence

Each subspace, quotient space, direct product, direct sum, projective limit, and inductive limit (of a family) of locally convex spaces has a unique natural locally convex topology. These spaces, except for quotient spaces and inductive limits, are separated, and a quotient space E/A is separated if and only if the subspace A is closed. The limit of a sequence $E_1 \subset E_2 \subset \ldots$ is said to be a **strictly inductive limit** if E_n has the induced topology as a subspace of E_{n+1}. If E is a strictly inductive limit of a sequence E_n such that E_n is closed in E_{n+1} or if E is the inductive limit of a sequence $E_1 \subset E_2 \subset \ldots$ such that the mapping $E_n \to E_{n+1}$ maps a neighborhood of 0 to a relatively weakly compact set, then E is separated and each bounded set of E is the image of a bounded set in some E_n. If $E = \bigcup E_n$ is the strictly inductive limit of the sequence $\{E_n\}$, then the topology of E_n coincides with the relative topology of $E_n \subset E$. The strictly inductive limit of a sequence of (F)-spaces is called an **(LF)-space**.

Any complete locally convex space (resp. any locally convex space) is (resp. a dense linear subspace of) the projective limit of Banach spaces. Every (F)-space E is the projective limit of a sequence of Banach spaces $E_1 \leftarrow E_2 \leftarrow \ldots$. In particular, E is said to be a **count-**

ably normed space if the mappings $E \to E_n$ are one-to-one and $\|x\|_n \leqslant \|x\|_{n+1}$ for all $x \in E$ with E considered as a subspace E_n. We call E a **countably Hilbertian space** if, in particular, the E_n are †Hilbert spaces. An (F)-space with at least one continuous norm is a nuclear space if and only if it is a countably Hilbertian space such that the mappings $E_{n+1} \to E_n$ are Hilbert-Schmidt operators or nuclear operators.

A locally convex space is bornological if and only if it is the inductive limit of normed spaces. A locally convex space is said to be **ultrabornological** if it is the inductive limit of Banach spaces, or in particular, if it is quasi-complete and bornological.

Properties of spaces, such as being complete, quasicomplete, semireflexive, or having every bounded closed set compact, are inherited by closed subspaces, direct products, projective limits, direct sums, and strictly inductive limits formed from the original spaces, and properties of spaces, such as being Mackey, quasibarreled, barreled, and bornological, are inherited by quotient spaces, direct sums, inductive limits, and direct products formed from the spaces. (For direct products of high power of bornological spaces, unsolved problems still exist concerning the inheritance of properties.) Quotient spaces of (F)-spaces are (F)-spaces, but quotient spaces of general complete spaces are not necessarily complete. There are examples of a Montel (F)-space whose quotient space is not reflexive and a Montel (DF)-space whose closed subspace is neither a Mackey space nor a (DF)-space. The property of being a Schwartz space or a nuclear space is inherited by the completions, subspaces, quotient spaces of closed subspaces, direct products, projective limits, direct sums of countable families, and inductive limits of countable families formed from such spaces. Tensor products of nuclear spaces are nuclear spaces. Y. Kōmura gave an example of a non-complete space that is quasicomplete, bornological, and nuclear (and hence a Montel space).

X. The Open Mapping Theorem and the Closed Graph Theorem

Let E and F be topological linear spaces. The statement that every continuous linear mapping of E onto F is open is called the **open mapping theorem** (or **homomorphism theorem**), and the statement that every linear mapping of F into E is continuous if its graph is closed in $F \times E$ is called the **closed graph theorem**. These theorems hold if both E and F are complete and metrizable (S. Banach).

A locally convex space is said to be **B-complete** (or **fully complete**) if a subspace C of E' is weak*-closed whenever $C \cap V^{\circ}$ is weak*-closed for every neighborhood V of 0 in E. (F)-spaces and the dual spaces of reflexive (F)-spaces are B-complete. B-complete spaces are complete, and closed subspaces and quotient spaces by closed subspaces of B-complete spaces are B-complete. If E is B-complete and F is barreled, then the open mapping theorem and the closed graph theorem hold (V. Pták).

Both theorems hold also if F is ultrabornological and E is a locally convex space obtained from a family of (F)-spaces after a finite number of operations of taking closed subspaces, quotient spaces by closed spaces, direct products of countable families, projective limits of countable families, direct sums of countable families, and inductive limits of countable families. This was conjectured by Grothendieck and proved by W. Słowikowski (1961) and D. A. Raikov. Later, L. Schwartz, A. Martineau, M. De Wilde, W. Robertson, and M. Nakamura simplified the proof and enlarged the class of spaces E [15].

(LF)-spaces, the dual spaces of Schwartz (F)-spaces, and the space $\mathscr{D}'$ of distributions are examples of spaces E described in the previous paragraph.

Y. Nonlocally Convex Spaces

The space L_p for $0<p<1$ shows that nonlocally convex spaces are meaningful in functional analysis. Recently, the Banach-Steinhaus theorem, closed graph theorems, etc. have been investigated for nonlocally convex topological linear spaces [13].

Z. Diagram of Topological Linear Spaces

The spaces in Fig. 1 are all locally convex spaces over the real number field or the complex number field and satisfy the separation axiom T_1. The notation $A \to B$ means that spaces with property A have property B. Main properties of dual spaces are listed in Table 1.

References

[1] N. Bourbaki, Eléments de mathématique, Espaces vectoriels topologiques, Actualités Sci. Ind., Hermann, 1189a, 1966; 1229b, 1967; 1230a, 1955.
[2] A. Grothendieck, Espaces vectoriels topologiques, Lecture notes, São Paulo, 1954.
[3] G. Köthe, Topologische lineare Räume I, Springer, second edition, 1966; English translation, Topological vector spaces I, II, Springer, 1969, 1979.

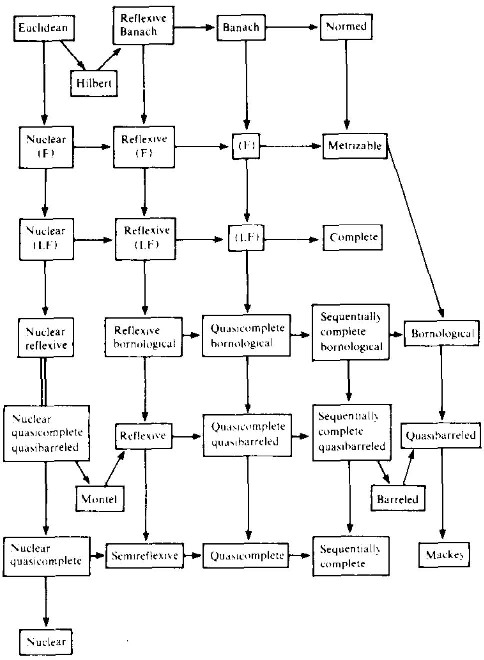


Fig. 1
Topological linear spaces.

Table 1

E	E'
Semireflexive	Barreled
Reflexive	Reflexive
Quasibarreled	Quasicomplete
Bornological	Complete
Reflexive, (F)	Bornological
(F)	(DF)
(DF)	(F)
(M)	(M)
Nuclear, (LF) or (DF)	Nuclear, reflexive
Complete, (S)	Ultrabornological

[4] I. M. Gel'fand et al., Generalized functions. II, Functions and generalized function spaces (with G. E. Shilov), Academic Press, 1968; IV, Applications of harmonic analysis (with N. Ya. Vilenkin), Academic, 1964. (Original in Russian, 1958–1966.)
[5] A. Pietsch, Nukleare lokalkonvexe Räume, Berlin, 1965.
[6] A. P. Robertson and W. Robertson, Topological vector spaces, Cambridge Univ. Press, 1964.
[7] J. Dieudonné and L. Schwartz, La dualité dans les espaces ($\mathfrak{F}$) et ($\mathfrak{L}\mathfrak{F}$), Ann. Inst. Fourier, 1 (1949), 61–101.
[8] A. Grothendieck, Sur les espaces (F) et (DF), Summa Brasil. Math., 3 (1954), 57–123.
[9] A. Grothendieck, Produits tensoriels topologiques et espaces nucléaires, Mem. Amer. Math. Soc., 1955.
[10] G. Choquet, Le théorème de représentation intégrale dans les ensembles convexes compacts, Ann. Inst. Fourier, 10 (1960), 333–344.
[11] J. L. Kelley, I. Namioka, et al., Linear topological spaces, Van Nostrand, 1963.
[12] L. Schwartz, Théorie des distributions à valeurs vectorielles I, Ann. Inst. Fourier, 7 (1957), 1–141.
[13] N. Adasch, B. Ernst, and D. Keim, Topological vector spaces: The theory without convexity conditions, Lecture notes in math. 639, Springer, 1978.
[14] F. Treves, Topological vector spaces, distributions and kernels, Academic Press, 1967.
[15] M. De Wilde, Closed graph theorems and webbed spaces, Pitman, 1978.

425 (II.16) Topological Spaces

A. Introduction

Convergence and continuity, as well as the algebraic operations on real numbers, are fundamental notions in analysis. In an abstract space too, it is possible to provide an additional structure so that convergence and continuity can be defined and a theory analogous to classical analysis can be developed. Such a structure is called a **topological structure** (for a precise definition, → Section B). There are several ways of giving a topology to a space. One method is to axiomatize the notion of convergence (M. Fréchet [1], 1906; → 87 Convergence). However, defining a topology in terms of either a neighborhood system (due to F. Hausdorff [3], 1914), a closure operation (due to C. Kuratowski, *Fund. Math.*, 3 (1922)), or a family of open sets is more common.

B. Definition of a Topology

Let X be a set. A **neighborhood system** for X is a function $\mathfrak{U}$ that assigns to each point x of X, a family $\mathfrak{U}(x)$ of subsets of X subject to the following axioms (**U**):
(1) $x \in U$ for each U in $\mathfrak{U}(x)$.
(2) If $U_1, U_2 \in \mathfrak{U}(x)$, then $U_1 \cap U_2 \in \mathfrak{U}(x)$.
(3) If $U \in \mathfrak{U}(x)$ and $U \subset V$, then $V \in \mathfrak{U}(x)$.
(4) For each U in $\mathfrak{U}(x)$, there is a member W of $\mathfrak{U}(x)$ such that $U \in \mathfrak{U}(y)$ for each y in W.

A **system of open sets** for a set X is a family $\mathfrak{O}$ of subsets of X satisfying the following axioms (**O**):

(1) $X, \varnothing \in \mathfrak{O}$.
(2) If $O_1, O_2 \in \mathfrak{O}$, then $O_1 \cap O_2 \in \mathfrak{O}$.
(3) If $O_\lambda \in \mathfrak{O}$ $(\lambda \in \Lambda)$, then $\bigcup_{\lambda\in\Lambda} O_\lambda \in \mathfrak{O}$.

A **system of closed sets** for a space X is a family $\mathfrak{F}$ of subsets of X satisfying the following axioms **(F)**:
(1) $X, \varnothing \in \mathfrak{F}$.
(2) If $F_1, F_2 \in \mathfrak{F}$, then $F_1 \cup F_2 \in \mathfrak{F}$.
(3) If $F_\lambda \in \mathfrak{F}$ $(\lambda \in \Lambda)$, then $\bigcap_{\lambda\in\Lambda} F_\lambda \in \mathfrak{F}$.

A **closure operator** for a space X is a function that assigns to each subset A of X, a subset A^a of X satisfying the following axioms **(C)**:
(1) $\varnothing^a = \varnothing$.
(2) $(A \cup B)^a = A^a \cup B^a$.
(3) $A \subset A^a$.
(4) $A^a = A^{aa}$.

An **interior operator** for a space X is a function that assigns to each subset A of X a subset A^i of X satisfying the following axioms **(I)**:
(1) $X^i = X$.
(2) $(A \cap B)^i = A^i \cap B^i$.
(3) $A^i \subset A$.
(4) $A^{ii} = A^i$.

Any one of these five structures for a set X, i.e., a structure satisfying any one of **(U)**, **(O)**, **(F)**, **(C)**, or **(I)**, determines the four other structures in a natural way. For instance, assume that a system of open sets $\mathfrak{O}$ satisfying **(O)** is given. In this case, each member of $\mathfrak{O}$ is called an **open set**. A subset U of X is called a **neighborhood** of a point x in X provided that there is an open set O such that $x \in O \subset U$. If $\mathfrak{U}(x)$ is the family of all neighborhoods of x, the function $x \to \mathfrak{U}(x)$ satisfies **(U)**. The complement of an open set in X is called a **closed set**. The family $\mathfrak{F}$ of all closed sets satisfies **(F)**. Given a subset A of X, the intersection A^a of the family of all closed sets containing A is called the **closure** of A, and each point of A^a is called an **adherent point** of A. The closure A^a is the smallest closed set containing A, and the function $A \to A^a$ satisfies **(A)**. The closure A^a is also denoted by $\bar{A}$ or $\operatorname{Cl} A$. Dually, there is a largest open subset A^i of A. The set A^i (also denoted by A° or $\operatorname{Int} A$) is called the **interior** of A, and each point of A° is called an **interior point** of A. The closure and interior are related by $A^\circ = X - \overline{(X-A)}$ and $\bar{A} = X - (X-A)^\circ$. The correspondence $A \to A^\circ$ satisfies **(I)**. Conversely, open sets can be characterized variously as follows:

$$\begin{aligned} A \text{ is open} &\Leftrightarrow A \in \mathfrak{U}(x) \text{ for each } x \text{ in } A \\ &\Leftrightarrow X - A \in \mathfrak{F} \\ &\Leftrightarrow \overline{(X-A)} = X - A \\ &\Leftrightarrow A^\circ = A. \end{aligned}$$

When a structure satisfying **(U)**, **(F)**, **(C)**, or **(I)** is given, one of the four characterizations of open sets can be used to define a system of open sets satisfying **(O)** and hence the other structure.

A **topological structure** or simply a **topology** for a space X is any of these five structures for X. If two topologies τ_1 and τ_2 for X give rise to identical systems of open sets, then τ_1 and τ_2 are considered to be identical. For this reason "topology" frequently means simply "system of open sets" in the literature. A **topological space** is a set X provided with a topology τ and is denoted by (X, τ) or simply X when there is no ambiguity.

C. Examples

(1) Discrete Topology. Let X be a set, and let the system $\mathfrak{O}$ of open sets be the family of all subsets of X. The resulting topology is called the **discrete topology**, and X with the discrete topology is a **discrete topological space**. In this space, $\bar{A} = A^\circ = A$ for each subset A, and A is a neighborhood of each of its points.

(2) Trivial Topology. The **trivial** (or **indiscrete**) **topology** for a set X is defined by the system of open sets which consists of X and $\varnothing$ only. If $A \subsetneqq X$, then $A^\circ = \varnothing$, and if $A \neq \varnothing$, then $\bar{A} = X$. Each point of X has only one neighborhood, X itself.

(3) Metric Topology. Let (X, ρ) be a †metric space, i.e., a set X provided with a †metric ρ. For a positive number ε, the ε-neighborhood of a point x is defined to be the set $U_\varepsilon(x) = \{y \mid y \in X, \rho(x, y) < \varepsilon\}$. Let $\mathfrak{U}(x)$ be the family of all sets V such that $U_\varepsilon(x) \subset V$ for some ε; then the assignment $x \to \mathfrak{U}(x)$ satisfies **(U)** and hence defines a topology. This topology is the **metric topology** for the metric space (X, ρ).

(4) Order Topology. Let X be a set †linearly ordered by $\leqslant$. For each point x in X, let $\mathfrak{U}(x)$ be the family of all subsets U such that $x \in \{y \mid a < y < b\} \subset U$ for some a, b. The function $x \to \mathfrak{U}(x)$ satisfies **(U)** and defines the **order topology** for the linearly ordered set X.

(5) Convergence and Topology. We can define the notion of convergence in a topological space, and conversely we can define a topology using convergence as a primitive notion (→ 87 Convergence). In particular, for a metric space, the metric topology can be defined in terms of convergent sequences (→ 273 Metric Spaces).

D. Generalized Topological Spaces

When a space X is equipped with a closure operator that does not satisfy all of **(C)**, the space is called a **generalized topological space** by some authors. Topological implications of each axiom in (C) have been investigated for such spaces.

E. Local Bases

Let X be a topological space, and let x be a point of X. A collection $\mathfrak{U}_0(x)$ of neighborhoods of x is called a **base for the neighborhood system (fundamental system of neighborhoods** of a point x or **local base** at x) if each neighborhood of x contains a member of $\mathfrak{U}_0(x)$. Let $\{\mathfrak{U}_0(x) | x \in X\}$ be a system of local bases; then the system has the following properties ($\mathbf{U}_0$):
(1) For each V in $\mathfrak{U}_0(x)$, $x \in V \subset X$.
(2) If $V_1, V_2 \in \mathfrak{U}_0(x)$, then there is a V_3 in $\mathfrak{U}_0(x)$ such that $V_3 \subset V_1 \cap V_2$.
(3) For each V in $\mathfrak{U}_0(x)$, there exists a $W \subset V$ in $\mathfrak{U}_0(x)$ such that for each y in W, V contains some member of $\mathfrak{U}_0(y)$.

Conversely, suppose that $\{\mathfrak{U}_0(x) | x \in X\}$ is a system satisfying ($\mathbf{U}_0$). For each x in X, let $\mathfrak{U}(x)$ consist of all subsets V of X such that $V \supset U$ for some U in $\mathfrak{U}_0(x)$. Then the system $\{\mathfrak{U}(x) | x \in X\}$ satisfies (U) and therefore defines a topology for X. This topology is called the topology determined by the system $\{\mathfrak{U}_0(x) | x \in X\}$.

For instance, in a metric space X, the set of ε-neighborhoods of $x (\varepsilon > 0)$ is a local base at x with respect to the metric topology. In an arbitrary topological space, the collection of all open sets containing x, i.e., the **open neighborhoods** of x, is a local base at x.

Two systems satisfying ($\mathbf{U}_0$) are called **equivalent** if they determine the same topology. For systems $\{\mathfrak{U}_0(x) | x \in X\}$ and $\{\mathfrak{V}_0(x) | x \in X\}$ to be equivalent it is necessary and sufficient that for each x in X each member of $\mathfrak{U}_0(x)$ contain a member of $\mathfrak{V}_0(x)$ and each member of $\mathfrak{V}_0(x)$ contain a member of $\mathfrak{U}_0(x)$.

Sometimes the word "neighborhood" stands for a member of a local base or for an open neighborhood. However, this convention is not used here.

F. Bases and Subbases

A family $\mathfrak{O}_0$ of open sets of a topological space X is called a **base for the topology (base for the space**, or **open base**) if each open set is the union of a subfamily of $\mathfrak{O}_0$. A base $\mathfrak{O}_0$ for the topology of a topological space X has the following properties ($\mathbf{O}_0$):
(1) $\bigcup \mathfrak{O}_0 = X$.
(2) If $W_1, W_2 \in \mathfrak{O}_0$ and $x \in W_1 \cap W_2$, then there is a W_3 in $\mathfrak{O}_0$ such that $x \in W_3 \subset W_1 \cap W_2$.

Conversely, if a family $\mathfrak{O}_0$ of subsets of a set X satisfies ($\mathbf{O}_0$), then $\mathfrak{O}_0$ is a base for a unique topology. A member of $\mathfrak{O}_0$ is called a **basic open set**.

A family $\mathfrak{O}_{00}$ of open sets of a topological space X is a **subbase for the topology** (or **subbase for the space**) if the family of all finite intersections of members of $\mathfrak{O}_{00}$ is a base for the topology. If $\mathfrak{O}_{00}$ a subbase for the topology of a topological space X, then $\bigcup \mathfrak{O}_{00} = X$. Conversely, if $\mathfrak{O}_{00}$ is a family of subsets of a set X such that $\bigcup \mathfrak{O}_{00} = X$, then the family of all finite intersections of members of $\mathfrak{O}_{00}$ is a base for a unique topology τ. A subset of X is open for τ if and only if it is the union of a family of finite intersections of members of $\mathfrak{O}_{00}$. The system of open sets relative to τ is said to be **generated** by the family $\mathfrak{O}_{00}$. Thus any family of sets defines a topology for its union.

A set $\mathfrak{F}$ of subsets of a topological space is called a **network** if for each point x and its neighborhood U there is a member $F \in \mathfrak{F}$ such that $x \in F \subset U$ (A. V. Arkhangel'skiĭ, 1959). If all $F \in \mathfrak{F}$ are required to be open, the network $\mathfrak{F}$ is exactly an open base.

G. Continuous Mappings

A mapping f on a topological space X into a topological space Y is called **continuous** at a point a of X if it satisfies one of the following equivalent conditions:
(1) For each neighborhood V of $f(a)$, there is a neighborhood U of a such that $f(U) \subset V$. (1′) For each neighborhood V of $f(a)$, the inverse image $f^{-1}(V)$ is a neighborhood of a.
(2) For an arbitrary subset A of X such that $a \in \bar{A}$, $f(a) \in \overline{f(A)}$.

Continuity can also be defined in terms of convergence (→ 87 Convergence).

If f is continuous at each point of X, f is said to be **continuous**. Continuity of f is equivalent to each of the following conditions:
(1) For each open subset O of Y, the inverse image $f^{-1}(O)$ is open in X.
(1′) The inverse image under f of each member of a subbase for the topology of Y is open in X.
(2) For each closed subset F of Y, the inverse image $f^{-1}(F)$ is closed.
(3) For each subset A of X, $f(\bar{A}) \subset \overline{f(A)}$.

The image $f(X)$ of X under a continuous mapping f is called a **continuous image** of X. Let X, Y, and Z be topological spaces, and let $f: X \to Y$ and $g: Y \to Z$ be mappings. If f is continuous at a point a of X and g is continuous at $f(a)$, then the composite mapping $g \circ f: X \to Z$ is continuous at the point a. Hence if f and g are continuous, so is $g \circ f$.

When a continuous mapping $f: X \to Y$ is

†bijective and f^{-1} is continuous, the mapping f is called a **homeomorphism** (named by H. Poincaré, 1895) or **topological mapping**. Two topological spaces X and Y are **homeomorphic**, $X \approx Y$, if there is a homeomorphism $f: X \to Y$.

The relation of being homeomorphic is an †equivalence relation. A property which, when held by a topological space, is also held by each space homeomorphic to it is a **topological property** or **topological invariant**. The problem of deciding whether or not given spaces are homeomorphic is called the **homeomorphism problem**.

A mapping $f: X \to Y$ is called **open** (resp. **closed**) if the image under f of each open (resp. closed) subset of X is open (closed) in Y. A continuous bijection that is either open or closed is a homeomorphism.

A continuous surjection $f: X \to Y$ is called a **quotient mapping** if $U \subset Y$ is open whenever $f^{-1}(U)$ is open (→ Section L). If moreover $f \mid f^{-1}(S)$ is quotient for each $S \subset Y$ as a mapping from the subspace (→ Section J) $f^{-1}(S)$ onto the subspace S, then f is called a **hereditarily quotient mapping**. Open or closed continuous mappings are hereditarily quotient mappings.

H. Comparison of Topologies

When a set X is provided with two topologies τ_1 and τ_2 and the identity mapping: $(X, \tau_1) \to (X, \tau_2)$ is continuous, the topology τ_1 is said to be **stronger (larger** or **finer)** than the topology τ_2, τ_2 is said to be **weaker (smaller** or **coarser)** than τ_1, and the notation $\tau_1 \geqslant \tau_2$ or $\tau_2 \leqslant \tau_1$ is used. Let $\mathfrak{O}_i$, $\mathfrak{F}_i$, $\mathfrak{U}_i$, and a_i be the system of open sets, system of closed sets, neighborhood system, and closure operation for X relative to the topology τ_i $(i = 1, 2)$, respectively. Then each of the following is equivalent to the statement $\tau_1 \geqslant \tau_2$:

(1) $\mathfrak{O}_1 \supset \mathfrak{O}_2$.
(2) $\mathfrak{F}_1 \supset \mathfrak{F}_2$.
(3) For each x in X, $\mathfrak{U}_1(x) \supset \mathfrak{U}_2(x)$.
(4) $A^{a_1} \subset A^{a_2}$ for each subset A of X.

Let S be the family of all topologies for X. Then S is ordered by the relation $\geqslant$. The discrete topology is the strongest topology for X. If $\{\tau_\lambda \mid \lambda \in \Lambda\}$ is a subfamily of S, then among the topologies stronger than each τ_λ, there is a weakest one $\tau_1 = \sup\{\tau_\lambda \mid \lambda \in \Lambda\}$. Similarly, among the topologies weaker than each τ_λ, there is a strongest one $\tau_2 = \inf\{\tau_\lambda \mid \lambda \in \Lambda\}$. In fact, let $\mathfrak{O}_\lambda$ be the family of all open sets relative to τ_λ; then the system of open sets for τ_1 is generated by $\bigcup_{\lambda \in \Lambda} \mathfrak{O}_\lambda$, and the system of open sets for τ_2 is precisely $\bigcap_{\lambda \in \Lambda} \mathfrak{O}_\lambda$. The family S is therefore a †complete lattice.

I. Induced Topology

Let f be a mapping from a set X into a topological space Y. Then the family $\{f^{-1}(O) \mid O$ is open in $Y\}$ satisfies axioms **(O)** and defines a topology for X. This topology is called the **topology induced by** f (or simply **induced topology**), and it is characterized as the weakest one among the topologies for X relative to which the mapping f is continuous.

J. Subspaces

Let (X, τ) be a topological space and M be a subset of X. The topology for M induced by the inclusion mapping $f: M \to X$, i.e., the mapping f defined by $f(x) = x$ for each x in M, is called the **relativization** of τ to M or the **relative topology**. The set M provided with the relative topology is called a **subspace** of the topological space (X, τ). Topological terms, when applied to a subspace, are frequently preceded by the adjective "relative" to avoid ambiguity. Thus a **relative neighborhood** of a point x in M is a set of the form $M \cap U$, where U is a neighborhood of x in X. A **relatively open (relatively closed)** set in M is a set of the form $M \cap A$, where A is open (closed) in X. For a subset A of M, the relative closure of A in M is $M \cap \bar{A}$, where $\bar{A}$ is the closure of A in X. A mapping $f: X \to Y$ is called an **embedding** if f is a homeomorphism from X to the subspace $f(X)$, and in this case X is said to be **embedded** into Y. A property P is said to hold **locally** on a topological space X if each point x of X has a neighborhood U such that the property P holds on the subspace U. A subset A of X is **locally closed** if for each point x of X, there exists a neighborhood V of x such that $V \cap A$ is relatively closed in V. A subset of X is locally closed if and only if it can be represented as $O \cap F$, where O is open and F is closed in X.

K. Product Spaces

Let X be a set, and for each member λ of an index set Λ, let f_λ be a mapping of X into a topological space X_λ. Then there is a weakest topology for X that makes each f_λ continuous. In fact, this topology is $\sup\{\tau_\lambda\}$, where τ_λ is the topology for X induced by f_λ. In particular, let $\{X_\lambda \mid \lambda \in \Lambda\}$ be a family of topological spaces, and let X be the Cartesian product $\prod_{\lambda \in \Lambda} X_\lambda$. Then the weakest topology for X such that each projection $\mathrm{pr}_\lambda: X \to X_\lambda$ is continuous is called the **product topology** or **weak topology**. The Cartesian product $\prod_{\lambda \in \Lambda} X_\lambda$ equipped with the product topology is called the **product topological space** or simply the **product space** or **direct product** of the family $\{X_\lambda | \lambda \in \Lambda\}$ of topological spaces. If $\mathfrak{O}$ is the family of all open subsets of X_λ, the union $\bigcup_\lambda \mathrm{pr}_\lambda^{-1}(\mathfrak{O}_\lambda)$ is a subbase for the product topology. If $x = \{x_\lambda\}$ is a point of X, then sets of the type $\bigcap_{j=1}^n \mathrm{pr}_j^{-1}(U_j) = \prod_{\lambda \neq \lambda_1, \ldots, \lambda_n} X_\lambda \times U_1 \times \ldots \times U_n$ form a local base at x for the product topology, where $\lambda_1, \ldots, \lambda_n \in \Lambda$ and U_j is a neighborhood of x_{λ_j}. Each projection $\mathrm{pr}_\lambda : X \to X_\lambda$ is continuous and open, and a mapping f from a topological space Y into the product space $\prod_\lambda X_\lambda$ is continuous if and only if $\mathrm{pr}_\lambda \circ f : Y \to X_\lambda$ is continuous for each λ. Given a family $\{f_\lambda\}$ of continuous mappings $f_\lambda : X_\lambda \to Y_\lambda$, the product mapping $\prod_\lambda f_\lambda : \prod_\lambda X_\lambda \to \prod_\lambda Y_\lambda$ is continuous with respect to the product topologies.

For the Cartesian product $\prod_\lambda X_\lambda$ of a family $\{X_\lambda | \lambda \in \Lambda\}$ of topological spaces, there is another topology called the **box topology** (or **strong topology**). A base for the box topology is the family of all sets $\prod_\lambda O_\lambda$, where O_λ is open in X_λ for each λ. For a point $x = \{x_\lambda\}$, the family of all sets of the form $\prod_\lambda U_\lambda$ is a local base at x relative to the box topology, where U_λ is a neighborhood of x_λ for each λ. With respect to the box topology, each projection $\mathrm{pr}_\lambda : \prod_\lambda X_\lambda \to X_\lambda$ is continuous and open, and the product mapping $\prod f_\lambda : \prod_\lambda X_\lambda \to \prod_\lambda Y_\lambda$ of a family $\{f_\lambda\}$ of continuous mappings $f_\lambda : X_\lambda \to Y_\lambda$ is continuous. For a finite product of topological spaces, the product topology agrees with the box topology, but for an arbitrary product the product topology is weaker than the box topology. For the Cartesian product of topological spaces the usual topology considered is the product topology rather than the box topology.

L. Quotient Spaces

Let f be a mapping of a topological space X onto a set Y. The **quotient topology** for Y (relative to the mapping f) is the strongest topology for Y such that f is continuous. A subset O of Y is open relative to the quotient topology if and only if $f^{-1}(O)$ is open. Given an equivalence relation $\sim$ on a topological space X, the †quotient set $Y = X/\sim$ provided with the quotient topology relative to the projection $\varphi : X \to Y$ is called the **quotient topological space** (or simply **quotient space**). A mapping f from the quotient space $Y = X/\sim$ into a topological space is continuous if and only if $f \circ \varphi$ is continuous.

A **partition** of a space X is a family $\{A_\lambda | \lambda \in \Lambda\}$ of pairwise disjoint subsets of X such that $\bigcup_\lambda A_\lambda = X$. A partition $\{A_\lambda\}$ of a topological space X determines an equivalence relation $\sim$ on X such that the family $\{A_\lambda\}$ is precisely the family of all equivalence classes under $\sim$, and therefore the partition determines the quotient space $Y = X/\sim$. This space is called the **identification space** of X by the given partition. Each member A_λ of the partition can be regarded as a point of Y, and the projection $\varphi : X \to Y$ satisfies $\varphi(x) = A_\lambda$ whenever $x \in A_\lambda$. A partition $\{A_\lambda | \lambda \in \Lambda\}$ of a topological space is called **upper semicontinuous** if for each A_λ and each open set U containing A_λ, there is an open set V such that $A_\lambda \subset V \subset U$, and V is the union of members of $\{A_\lambda | \lambda \in \Lambda\}$. A partition $\{A_\lambda | \lambda \in \Lambda\}$ is upper semicontinuous if and only if the projection $\varphi : X \to Y = \{A_\lambda | \lambda \in \Lambda\}$ is a closed mapping.

M. Topological Sums

Let X be a set, and for each member λ of an index set Λ, let f_λ be a mapping of a topological space X_λ to X. Then the family $\{O \subset X | f_\lambda^{-1}(O)$ is open for any $\lambda\}$ satisfies the axioms of the open sets. This topology τ is characterized as the strongest one for X that makes each f_λ continuous. A mapping g on X with τ to a topological space Y is continuous if and only if $g \circ f_\lambda : X_\lambda \to Y$ is continuous for each $\lambda \in \Lambda$. The simplest is the case where X is the disjoint union of X_λ and f_λ is the inclusion mapping. Then we call the topological space X the **direct sum** or the **topological sum** of $\{X_\lambda\}$ and denote it by $\oplus X_\lambda$ or $\coprod X_\lambda$. More generally let the set X be the union of topological spaces $\{X_\lambda\}_{\lambda \in \Lambda}$ such that for each λ and $\mu \in \Lambda$ the relative topologies of $X_\lambda \cap X_\mu$ from X_λ and X_μ coincide. Then we call the topology τ the **weak topology** with respect to $\{X_\lambda\}$. If $X_\lambda \cap X_\mu$ is closed (resp. open) in X_μ for any μ, then X_λ is closed (resp. open) in X and the original topology of X_λ coincides with the relative topology. If, moreover, for each subset Γ of Λ, $F = \bigcup_{\lambda \in \Gamma} X_\lambda$ is closed and the weak topology of F with respect to $\{X_\lambda\}_{\lambda \in \Gamma}$ coincides with the relative topology induced by τ, then X with τ is said to have the **hereditarily weak topology** with respect to $\{X_\lambda\}$ (or to be **dominated** by $\{X_\lambda\}$). A topological space has the hereditarily weak topology with respect to any locally finite closed covering, and every CW-complex (→ 70 Complexes) has the hereditarily weak topology with respect to the covering of all finite subcomplexes.

When $\{X_n\}$ is an increasing sequence of topological spaces such that each X_n is a subspace of X_{n+1}, then the union $X = \bigcup X_n$ with the weak topology is called the **inductive limit** of $\{X_n\}$ and is denoted by $\varinjlim X_n$. Each X_n may again be regarded as a subspace of X.

N. Baire Spaces

For a subset A of a topological space X, the set $X-\bar{A}$ is called the **exterior** of A, and the set $\bar{A}\cap\overline{X-A}$ is called the **boundary** of A, denoted by $\mathrm{Bd}\,A$, $\mathrm{Fr}\,A$, or ∂A. A point belonging to the exterior (boundary) of A is an **exterior point** (**boundary point** or **frontier point**) of A. If the closure of A is X, then A is said to be **dense** in X. When $X-A$ is dense in X, i.e., when the interior of A is empty, A is called a **boundary set** (or **border set**), and if the closure $\bar{A}$ is a boundary set, A is said to be **nowhere dense**. The union of a countable family of nowhere dense sets is called a **set of the first category** (or **meager set**). A set that is not of the first category is called a **set of the second category** (or **nonmeager set**). The complement of a set of the first category is called a **residual set**. In the space $\mathbf{R}$ of real numbers, the set $\mathbf{Q}$ of all rational numbers is of the first category, and the set $\mathbf{R}-\mathbf{Q}$ of all irrational numbers is of the second category. Both $\mathbf{Q}$ and $\mathbf{R}-\mathbf{Q}$ are dense in X and hence are boundary sets. The union of a finite family of nowhere dense sets is nowhere dense, and the union of a countable family of sets of the first category is also of the first category. A subset A of X is nowhere dense in X if and only if for each open set O, $O\cap A$ is not dense in O.

A topological space X is called a **Baire space** (Baire, 1899) if each subset of X of the first category has an empty interior. Each of the following conditions is necessary and sufficient for a space X to be a Baire space:

(1) Each nonempty open subset of X is of the second category.

(2) If $F_1, F_2, \ldots$ is a sequence of closed subsets of X such that the union $\bigcup_{n=1}^{\infty} F_n$ has an interior point, then at least one F_n has an interior point.

(3) If $O_1, O_2, \ldots$ is a sequence of dense open subsets of X, then the intersection $\bigcap_{n=1}^{\infty} O_n$ is dense in X.

An open subset of a Baire space is a Baire space for the relative topology. A topological space that is homeomorphic to a complete metric space (→ 436 Uniform Spaces I) is a Baire space (**Baire-Hausdorff theorem**). A locally compact Hausdorff space (→ Section V) is also a Baire space. The class of Čech-complete completely regular spaces (→ Section T) includes both of these spaces, but there are also Baire spaces that are not in the class. A subset A of a topological space is said to satisfy **Baire's condition** or to have the **Baire property** if there exist an open set O and sets P_1, P_2 of the first category such that $A=(O\cup P_1)-P_2$. A †Borel set satisfies Baire's condition.

O. Accumulation Points

A point x is called an **accumulation point**, or a **cluster point** of a subset A of a topological space X if $x\in\overline{A-\{x\}}$. The set of all accumulation points of a set A is called the **derived set** of A and is denoted by A' or A^d. A point x belongs to A' if and only if each neighborhood of x contains a point of A other than x itself. A point belonging to the set $A^s=A-A'$ is called an **isolated point** of A, and a set A consisting of isolated points only, i.e., $A=A^s$, is said to be **discrete**. If each nonempty subset of A contains an isolated point, then A is said to be **scattered**; and if A does not possess an isolated point, i.e., $A\subset A'$, then A is said to be **dense in itself**. The largest subset of A which is dense in itself is called the **kernel** of A. If $A=A'$, then A is called a **perfect set**.

If x is an accumulation point of A, then for each neighborhood U of x, $U\cap(A-\{x\})\neq\varnothing$. Furthermore, it is possible to classify an accumulation point of A according to the †cardinality of $U\cap(A-\{x\})$. A point x is called a **condensation point** of a set A if for each neighborhood U of x, the set $U\cap A$ is uncountable. A point x is a **complete accumulation point** of A if for each neighborhood U of x, the set $U\cap A$ has the same cardinality as A.

P. Countability Axioms

A topological space X satisfies the **first countability axiom** if each point x of X has a countable local base (F. Hausdorff [3]). Metric spaces satisfy the first countability axiom. In fact, the family of $(1/n)$-neighborhoods ($n=1,2,\ldots$) of a point is a local base of the point. The topology of a topological space that satisfies the first countability axiom is completely determined by convergent sequences. For instance, the closure of a subset A of such a space consists of all limits of sequences in A (→ 87 Convergence). A topological space X is said to satisfy the **second countability axiom** or to be **perfectly separable** if there is a countable base for the topology. †Euclidean spaces satisfy the second countability axiom. If X contains a countable dense subset, X is said to be **separable**. A space that satisfies the second countability axiom satisfies the first and is also a separable Lindelöf space (→ Section S). However, the converse is not true. Each of the following properties is independent of the others: separability, the first countability axiom, and the Lindelöf property. If a metric space is separable, then it satisfies the second countability axiom. There are metric spaces that are not separable.

Q. Separation Axioms

Topological spaces that are commonly encountered usually satisfy some of the following separation axioms.

(T_0) **Kolmogorov's axiom**. For each pair of distinct points, there is a neighborhood of one point of the pair that does not contain the other.

(T_1) **The first separation axiom** or **Fréchet's axiom**. For each pair x, y of distinct points, there are neighborhoods U of x and V of y such that $x \notin V$ and $y \notin U$.

Axiom (T_1) can be restated as follows:

(T'_1) For each point x of the space, the singleton $\{x\}$ is closed.

(T_2) **The second separation axiom** or **Hausdorff's axiom** [3]. For each pair x, y of distinct points of the space X, there exist disjoint neighborhoods of x and y.

Axiom (T_2) is equivalent to the following:

(T'_2) In the product space $X \times X$ the diagonal set Δ is closed.

(T_3) **The third separation axiom** or **Vietoris's axiom** (*Monatsh. Math. Phys.*, 31 (1921)). Given a point x and a subset A such that $x \notin \bar{A}$, there exist disjoint open sets O_1 and O_2 such that $x \in O_1$ and $A \subset O_2$. (In this case, the sets $\{x\}$ and A are said to be **separated** by open sets.)

Axiom (T_3) can be restated as (T'_3) or (T''_3):

(T'_3) For each point x of the space, there is a local base at x consisting of closed neighborhoods of x.

(T''_3) An arbitrary closed set and a point not belonging to it can be separated by open sets.

(T_4) **The fourth separation axiom** or **Tietze's first axiom** (*Math. Ann.*, 88 (1923)). Two disjoint closed sets F_1 and F_2 can be separated by open sets, i.e., there exist disjoint open sets O_1 and O_2 such that $F_1 \subset O_1$ and $F_2 \subset O_2$.

(T_5) **Tietze's second axiom**. Whenever two subsets A_1 and A_2 satisfy $A_1 \cap \bar{A}_2 = \bar{A}_1 \cap A_2 = \varnothing$, A_1 and A_2 can be separated by open sets.

It is easily seen that (T_5) $\Rightarrow$ (T_4), (T_0) and (T_3) $\Rightarrow$ (T_2), (T_4) and (T_1) $\Rightarrow$ (T_3). Axiom (T_4) is equivalent to each of (T'_4) and (T''_4):

(T'_4) Whenever F_1 and F_2 are disjoint closed subsets, there exists a continuous function f on the space into the interval $[0, 1]$ such that f is identically 0 on F_1 and 1 on F_2.

(T''_4) Each real-valued continuous function defined on a closed subspace can be extended to a real-valued continuous function on the entire space.

The implications (T_4) $\Rightarrow$ (T'_4) and (T_4) $\Rightarrow$ (T''_4) are known as **Uryson's lemma** (*Math. Ann.*, 94 (1925)) and the **Tietze extension theorem** (*J. Reine Angew. Math.*, 145 (1915)), respectively. In addition, there are two more related axioms:

($T_{3\frac{1}{2}}$) **Tikhonov's separation axiom**. For each closed subset F and each point x not in F, there is a real-valued continuous function f on the space such that $f(x) = 0$ and f is identically 1 on F.

(T_6) (N. Vedenisov). For each closed subset F, there is a real-valued continuous function f on the space such that $F = \{x \mid f(x) = 0\}$.

Axioms (T_5) and (T_6) are equivalent to the following (T'_5) and (T'_6), respectively:

(T'_5) Each subspace satisfies (T_4)

(T'_6) X satisfies (T_4) and each closed set is a †G_δ-set.

The following implications are valid: ($T_{3\frac{1}{2}}$) $\Rightarrow$ (T_3), (T_6) $\Rightarrow$ (T_5), (T_4) and (T_1) $\Rightarrow$ ($T_{3\frac{1}{2}}$).

Table 1 gives a classification of topological spaces by the separation axioms. Each line represents a special case of the preceding line.

A †metrizable space is perfectly normal, but the converse is false (for metrization theorems → 273 Metric Spaces). Among the spaces satisfying the second countability axiom, regular spaces are normal (**Tikhonov's theorem**, *Math. Ann.*, 95 (1925)) and metrizable (**Tikhonov-Uryson theorem**; P. Uryson, *Math. Ann.*, 94 (1925)).

Table 2 shows whether various topological properties are preserved in subspaces, product spaces, and quotient spaces. The topological properties considered are T_1, T_2 = Hausdorff, T_3 = regular, CR = completely regular, T_4 = normal, T_5 = completely normal, M = metrizable, C_I = first axiom of countability, C_{II} = second axiom of countability, C = compact, S = separable, and L = Lindelöf. Each position is filled with ○ or × according as the property (say, P) listed at the head of the column is preserved or not in the sort of space listed on the left obtained from space(s) all having property P.

R. Coverings

A family $\mathfrak{M} = \{M_\lambda\}_{\lambda \in \Lambda}$ of subsets of a set X is called a **covering** of a subset A of X if $A \subset \bigcup_\lambda M_\lambda$. If $\mathfrak{M}$ is finite (countable), it is called a **finite covering (countable covering)**. An **open (closed) covering** is a covering consisting of open (closed) sets.

A family $\mathfrak{M}$ of subsets of a topological space X is said to be **locally finite** if for each point x of X, there is a neighborhood of x which intersects only a finite number of members of $\mathfrak{M}$. If moreover $\{\bar{M}_\lambda\}_{\lambda \in \Lambda}$ is disjoint, then $\mathfrak{M}$ is called **discrete**. $\mathfrak{M}$ is called **star-finite** if each member of $\mathfrak{M}$ intersects only a finite number of members of $\mathfrak{M}$. A **σ-locally finite** or **σ-discrete** family of subsets of X is respectively the union of a countable number of locally finite or discrete families of subsets of X. A covering $\mathfrak{M}$

Table 1. Separation Axioms

Axioms	Spaces Satisfying the Axioms
(T_0)	**T_0-space (Kolmogorov space)**
(T_1)	**T_1-space (Kuratowski space)**
(T_2)	**T_2-space (Hausdorff space, separated space)**
(T_0) and (T_3)	**T_3-space (regular space)**
(T_1) and ($T_{3\frac{1}{2}}$)	**Completely regular space (Tikhonov space)**
(T_1) and (T_4)	**T_4-space (normal space)**
(T_1) and (T_5)	**T_5-space (completely normal space, hereditarily normal space)**
(T_1) and (T_6)	**T_6-space (perfectly normal space)**

Table 2. Topological Properties and Spaces

Space	T_1	T_2	T_3	CR	T_4	T_5	M	C_I	C_{II}	C	S	L
Subspace	○	○	○	○	×	○	○	○	○	×	×	×
Closed subspace	○	○	○	○	○	○	○	○	○	○	×	○
Open subspace	○	○	○	○	×	○	○	○	○	×	○	×
Product	○	○	○	○	×	×	×	×	×	○	×	×
Countable product	○	○	○	○	×	×	○	○	○	○	○	×
Quotient space	×	×	×	×	×	×	×	×	×	○	○	○

is called **point-finite** if each infinite number of members of $\mathfrak{M}$ has an empty intersection. A covering $\mathfrak{M}$ is a **refinement** of a covering $\mathfrak{N}$ (written $\mathfrak{M} \prec \mathfrak{N}$) if each member of $\mathfrak{M}$ is contained in a member of $\mathfrak{N}$. The **order** of the covering $\mathfrak{M}$ is the least integer r such that any subfamily of $\mathfrak{M}$ consisting of $r+1$ members has an empty intersection.

Let $\mathfrak{M}$ be a covering of X, and let A be a subset of X. The **star** of A relative to $\mathfrak{M}$, denoted by $S(A, \mathfrak{M})$, is the union of all members of $\mathfrak{M}$ whose intersection with A is nonempty. Let $\mathfrak{M}^\Delta$ denote the family $\{S(\{x\}, \mathfrak{M})\}_{x \in X}$ and $\mathfrak{M}^*$ the family $\{S(M, \mathfrak{M})\}_{M \in \mathfrak{M}}$. Then $\mathfrak{M}^\Delta$ and $\mathfrak{M}^*$ are coverings of X, and $\mathfrak{M} \prec \mathfrak{M}^\Delta \prec \mathfrak{M}^* \prec \mathfrak{M}^{\Delta\Delta}$. A covering $\mathfrak{M}$ is a **star refinement** of a covering $\mathfrak{N}$ if $\mathfrak{M}^* \prec \mathfrak{N}$, and $\mathfrak{M}$ is a **barycentric refinement** (or **Δ-refinement**) of $\mathfrak{N}$ if $\mathfrak{M}^\Delta \prec \mathfrak{N}$.

A sequence $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ of open coverings of a topological space is called a **normal sequence** if $\mathfrak{M}_{n+1}^\Delta \prec \mathfrak{M}_n$ for $n = 1, 2, \ldots$, and an open covering $\mathfrak{M}$ is said to be a **normal covering** if there is a normal sequence $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ such that $\mathfrak{M}_1 \prec \mathfrak{M}$. The **support** (or **carrier**) of a real-valued function f on a topological space X is defined to be the closure of the set $\{x \mid f(x) \neq 0\}$. Let $\{f_\alpha\}_{\alpha \in A}$ be a family of continuous nonnegative real-valued functions on a topological space X, and for each α in A, let C_α be the support of f_α. The family $\{f_\alpha\}_{\alpha \in A}$ is called a **partition of unity** if the family $\{C_\alpha\}_{\alpha \in A}$ is locally finite and $\sum_\alpha f_\alpha(x) = 1$ for each x in X. If the covering $\{C_\alpha\}_{\alpha \in A}$ is a refinement of a covering $\mathfrak{M}$, the family $\{f_\alpha\}_{\alpha \in A}$ is called a **partition of unity subordinate to the covering** $\mathfrak{M}$. A partition of unity subordinate to a covering $\mathfrak{M}$ exists only if $\mathfrak{M}$ is a normal covering (→ Section X). If ρ is a continuous †pseudometric on a T_1-space X, then define a covering M_n for each natural number n by $M_n = \{U(x; 2^{-n})\}_{x \in X}$, where $U(x; \varepsilon) = \{y \mid \rho(x, y) < \varepsilon\}$. Then the sequence $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ is a normal sequence of open coverings. Conversely, given a normal sequence $\mathfrak{M}_1, \mathfrak{M}_2, \ldots$ of open coverings of X, there exists a continuous pseudometric ρ such that $\rho(x, y) \leqslant 2^{-n}$ whenever $x \in S(y, \mathfrak{M}_n)$, and $\rho(x, y) \geqslant 2^{-n-1}$ whenever $x \notin S(y, \mathfrak{M}_n)$. If in addition for each x the family $\{S(x, \mathfrak{M}_n) \mid n = 1, 2, \ldots\}$ is a local base at x, then the metric topology of ρ agrees with the topology of X.

S. Compactness

If each open covering of a topological space X admits a finite open covering as its refinement, the space X is called **compact**; if each open covering of X admits a countable open refinement, X is said to be a **Lindelöf space** (P. Aleksandrov and P. Uryson, *Verh. Akad. Wetensch.*, Amsterdam, 19 (1929)); if each open covering of X admits a locally finite open refinement, X is called **paracompact** (J. Dieudonné, *J. Math. Pures Appl.*, 23 (1944)); and if each open covering of X admits a star-finite open refinement, X is said to be **strongly paracompact** (C. H. Dowker, *Amer. J. Math.*, 69 (1947)) or to have the **star-finite property** (K. Morita, *Math. Japonicae*, 1 (1948)). The space X is compact (Lindelöf) if for each open covering $\mathfrak{M}$ of X, there is a finite (countable) subfamily of $\mathfrak{M}$ whose union is X.

The following properties for a topological space X are equivalent: (1) The space X is compact. (2) If a family $\{F_\lambda\}_{\lambda \in \Lambda}$ of closed subsets of X has the **finite intersection property**, i.e., each finite subfamily of $\{F_\lambda\}_{\lambda \in \Lambda}$ has nonempty intersection, then $\bigcap_\lambda F_\lambda \neq \varnothing$. (3) Each infinite subset of X has a complete accumulation point. (4) Each †net has a convergent †subnet. (5) Each †universal net and each †ultrafilter converge.

If a subset A of X is compact for the relative topology, A is called a **compact** subset. A subset A of X is said to be **relatively compact** if the closure of A in X is a compact subset. A closed subset of a compact topological space is compact, and a compact subset of a Hausdorff space is closed. A continuous image of a compact space is compact, each continuous mapping of a compact space into a Hausdorff space is a closed mapping, and a continuous bijection of a compact space onto a Hausdorff space is a homeomorphism. The product space of a family $\{X_\lambda\}_{\lambda \in \Lambda}$ of topological spaces is compact if and only if each factor space is compact (**Tikhonov's product theorem**, *Math. Ann.*, 102 (1930)). A compact Hausdorff space is normal. A compact Hausdorff space is metrizable if and only if it satisfies the second countability axiom. A metric space or a †uniform space is compact if and only if it is †totally bounded and †complete. A subset of a Euclidean space is compact if and only if it is closed and bounded. In a discrete space only finite subsets are compact. The cardinality of a compact Hausdorff space with the first countability axiom cannot exceed the power of the continuum (Arkhangel'skiĭ).

There are a number of conditions related to compactness. A topological space is **sequentially compact** if each sequence in X has a convergent subsequence. A space X is **countably compact** (M. Fréchet [1]) if each countable open covering of X contains a finite subfamily that covers X. A space X is **pseudocompact** (E. Hewitt, 1948) if each continuous real-valued function on X is bounded. Some authors use *compact* and *bicompact* for what we call countably compact and compact, respectively. N. Bourbaki [9] uses *compact* and *quasicompact* instead of compact Hausdorff and compact, respectively. A T_1-space is countably compact if and only if each infinite set possesses an accumulation point. If X is countably compact, then X is pseudocompact, and if X is normal, the converse also holds. If a †complete uniform space is pseudocompact, then it is compact. A space satisfying the second countability axiom is compact if and only if it is sequentially compact. If X is sequentially compact, then X is countably compact, and if X satisfies the first countability axiom, the converse is true.

T. Compactification

A **compactification** of a topological space X consists of a compact space Y and a homeomorphism of X onto a dense subspace X_1 of Y. We can always regard X as a dense subspace of a compactification Y. If X is completely regular, then there is a Hausdorff compactification Y such that each bounded real-valued continuous function on X can be extended continuously to Y. Such a compactification is unique up to homeomorphism; it is called the **Stone-Čech compactification** of X (E. Čech, *Ann. Math.*, 38 (1937); M. H. Stone, *Trans. Amer. Math. Soc.*, 41 (1937)) and is denoted by $\beta(X)$. Let $\{f_\lambda\}_{\lambda \in \Lambda}$ be the set of all continuous functions on a completely regular space X into the closed interval $I=[0,1]$. Then a continuous mapping φ of X into a **parallelotope** $I^\Lambda = \prod_\lambda I_\lambda$ $(I_\lambda = I)$ is defined by $\varphi(x) = \{f_\lambda(x)\}_{\lambda \in \Lambda}$, and the mapping φ is a homeomorphism of X onto the subspace $\varphi(X)$ of I^Λ (**Tikhonov's embedding theorem**, *Math. Ann.*, 102 (1930)). The closure $\overline{\varphi(X)}$ of $\varphi(X)$ in I^Λ is the Stone-Čech compactification of X. The natural mapping $\beta(X_1 \times X_2) \rightarrow \beta(X_1) \times \beta(X_2)$ is a homeomorphism if and only if $X_1 \times X_2$ is pseudocompact (I. Glicksberg, 1959).

For a topological space X, let ∞ be a point not in X, and define a topology on the union $X \cup \{\infty\}$ as follows: A subset U of $X \cup \{\infty\}$ is open if and only if either $\infty \neq U$ and U is open in X, or $\infty \in U$ and $X - U$ is a compact closed subset of X. The topological space $X \cup \{\infty\}$ thus obtained is compact, and if X is not already compact, the space $X \cup \{\infty\}$ is a compactification of X called the **one-point compactification** of X (P. S. Aleksandrov, *C. R. Acad. Sci. Paris*, 178 (1924)). The one-point compactification of a Hausdorff space is not necessarily Hausdorff. The one-point compactification of the n-dimensional Euclidean space $\mathbf{R}^n$ is homeomorphic to the n-dimensional sphere S^n.

A completely regular space X is a †G_δ-set in the Stone-Čech compactification $\beta(X)$ if and only if it is a G_δ-set in any Hausdorff space Y which contains X as a dense subspace. Then X is said to be **Čech-complete**.

U. Absolutely Closed Spaces

A Hausdorff space X is said to be **absolutely closed** (or **H-closed**; P. Aleksandrov and P. Uryson, 1929) if X is closed in each Hausdorff space containing it. A compact Hausdorff space is absolutely closed. A Hausdorff space is absolutely closed if and only if for each open covering $\{N_\lambda\}_{\lambda \in \Lambda}$ of X, there is a finite subfamily of $\{\overline{N}_\lambda\}_{\lambda \in \Lambda}$ that covers X. The product space of a family of absolutely closed spaces is absolutely closed. Each Hausdorff space is a dense subset of an absolutely closed space. Similarly, a regular space X is said to be **r-**

closed if X is closed in each regular space containing it (N. Weinberg, 1941).

V. Locally Compact Spaces

A topological space X is said to be **locally compact** if each point of X has a compact neighborhood (P. Aleksandrov and P. Uryson, 1929). A †uniform space X is said to be **uniformly locally compact** if there is a member U of the †uniformity such that $U(x)$ is compact for each x in X (→ 436 Uniform Spaces). A noncompact space X is locally compact and Hausdorff if and only if the one-point compactification of X is Hausdorff, and this is the case if and only if X is homeomorphic to an open subset of a compact Hausdorff space. A locally compact Hausdorff space is completely regular, and for each point of the space, the family of all of its compact neighborhoods forms a local base at the point. A locally closed, hence open or closed, subset of a locally compact Hausdorff space is also locally compact for the relative topology. If a subspace A of a Hausdorff space X is locally compact, then A is a locally closed subset of X. The Euclidean space $\mathbf{R}^n$ is locally compact, and hence each **locally Euclidean space**, i.e., a space such that each point admits a neighborhood homeomorphic to a Euclidean space, is locally compact. A topological space is called **σ-compact** if it can be expressed as the union of at most countably many compact subsets.

W. Proper (Perfect) Mappings

A mapping f of a topological space X into a topological space Y is said to be **proper** (N. Bourbaki [9]) (or **perfect** [14]) if it is continuous and for each topological space Z, the mapping $f\times 1: X\times Z\to Y\times Z$ is closed, where $(f\times 1)(x,z)=(f(x),z)$. A continuous mapping $f: X\to Y$ is proper if and only if it is closed and $f^{-1}(\{y\})$ is compact for each y in Y. Another necessary and sufficient condition is that if $\{x_\alpha\}_{\mathfrak{A}}$ is a †net in X such that its image $\{f(x_\alpha)\}$ converges to $y\in Y$, then a subnet of $\{x_\alpha\}$ converges to an $x\in f^{-1}(y)$ in X. A continuous mapping of a compact space into a Hausdorff space is always proper. For a compact Hausdorff space X, a quotient space Y is Hausdorff if and only if the canonical projection $\varphi: X\to Y$ is proper.

For a continuous mapping f of a locally compact Hausdorff space X into a locally compact Hausdorff space Y, the following three conditions are equivalent: (1) f is proper. (2) For each compact subset K of Y, the inverse image $f^{-1}(K)$ is compact. (3) If $X\cup\{x_\infty\}$ and $Y\cup\{y_\infty\}$ are the one-point compactifications of X and Y, then the extension f_1 of f such that $f_1(x_\infty)=y_\infty$ is continuous.

The composition of two proper mappings is proper and the direct product of an arbitrary number of proper mappings is proper.

X. Paracompact Hausdorff Spaces

A paracompact Hausdorff space (often called simply a paracompact space) is normal. For a Hausdorff space X, the following five conditions are equivalent: (1) X is paracompact. (2) X is **fully normal** (J. W. Tukey [8]), i.e., each open covering of X admits an open barycentric refinement. (3) Each open covering has a partition of unity subordinate to it. (4) Each open covering is refined by a closed covering $\{F_\alpha\mid\alpha\in A\}$ that is **closure-preserving**, i.e., $\bigcup\{F_\beta\mid\beta\in B\}$ is closed for each $B\subset A$. (5) Each open covering $\{U_\alpha\mid\alpha\in A\}$ has a **cushioned refinement** $\{V_\alpha\mid\alpha\in A\}$, i.e., $\mathrm{Cl}(\bigcup\{V_\beta\mid\beta\in B\})\subset\bigcup\{U_\beta\mid\beta\in B\}$ for each $B\subset A$. The implication (1)–(2) is **Dieudonné's theorem**. The implication (2)→(1) is A. H. **Stone's theorem** (1948), from which it follows that each metric space is paracompact. The implications (5)→(4)→(1) is **Michael's theorem** (1959, 1957).

For normal spaces, the following weaker versions of (2) and (3) hold: A T_1-space X is normal if and only if each finite open covering of X admits a finite open star refinement (or finite open barycentric refinement). For each locally finite open covering of a normal space, there is a partition of unity subordinate to it.

For a regular space X the following three conditions are equivalent: (1) X is paracompact. (2) Each open covering of X is refined by a σ-discrete open covering. (3) Each open covering of X is refined by a σ-locally finite open covering. **Tamano's product theorem**: For a completely regular space X to be paracompact it is necessary and sufficient that $X\times\beta(X)$ be normal (1960).

For a †connected locally compact space X, the following conditions are equivalent: (1) X is paracompact. (2) X is σ-compact. (3) In the one-point compactification $X\cup\{\infty\}$, the point ∞ admits a countable local base. (4) There is a locally finite open covering $\{U_\lambda\}_{\lambda\in\Lambda}$ of X such that $\bar{U}_\lambda$ is compact for each λ. (5) X is the union of a sequence $\{U_n\}$ of open sets such that $\bar{U}_n$ is compact and $\bar{U}_n\subset U_{n+1}$ $(n=1,2,\dots)$. (6) X is strongly paracompact.

Every †F_σ-set of a paracompact Hausdorff space is paracompact (Michael, 1953). When a T_1-space X has the hereditarily weak topology with respect to a closed covering $\{F_\lambda\}$, then X is paracompact Hausdorff (normal, completely normal or perfectly normal) if and only if each

F_λ is (Morita, 1954; Michael, 1956). In particular, every CW-complex is paracompact and perfectly normal (Morita, 1953).

Y. Normality and Paracompactness of Direct Products

A topological space X is discrete if $X \times Y$ is normal for any normal space Y (M. Atsuji and M. Rudin, 1978). There are a paracompact Lindelöf space X and a separable metric space Y such that the product $X \times Y$ is not normal (Michael, 1963). The following are conditions under which the products are normal or paracompact. Let $\mathfrak{m}$ be an infinite †cardinal number. A topological space X is called $\mathfrak{m}$-paracompact if every open covering consisting of at most $\mathfrak{m}$ open sets admits a locally finite open covering as its refinement. When $\mathfrak{m}$ is countable, it is called **countably paracompact**. If X has an open base of at most $\mathfrak{m}$ members, $\mathfrak{m}$-paracompact means paracompact. The following conditions are equivalent for a topological space X: (1) X is normal and countably paracompact; (2) The product $X \times Y$ is normal and countably paracompact for any compact metric space Y; (3) $X \times I$ is normal, where $I = [0, 1]$ (C. H. Dowker, 1951). Rudin (1971) constructed an example of a collectionwise normal space (→ Section AA) that is not countably paracompact. When $\mathfrak{m}$ is general the following conditions are equivalent: (1) X is normal and $\mathfrak{m}$-paracompact; (2) If Y is a compact Hausdorff space with an open base consisting of at most $\mathfrak{m}$ sets, then $X \times Y$ is normal and $\mathfrak{m}$-paracompact; (3) $X \times I^{\mathfrak{m}}$ is normal; (4) $X \times \{0, 1\}^{\mathfrak{m}}$ is normal (Morita, 1961). In particular, the product $X \times Y$ of a paracompact Hausdorff space X and a compact Hausdorff space Y is paracompact (Dieudonné, 1944).

A topological space X is called a **P-space** if it satisfies the following conditions: Let Ω be an arbitrary set and $\{G(\alpha_1, \ldots, \alpha_i) \mid \alpha_1, \ldots \alpha_i \in \Omega, i = 1, 2, \ldots\}$ be a family of open sets such that $G(\alpha_1, \ldots, \alpha_i) \subset G(\alpha_1, \ldots, \alpha_i, \alpha_{i+1})$. Then there is a family of closed sets $\{F(\alpha_1, \ldots, \alpha_i) \mid \alpha_1, \ldots, \alpha_i \in \Omega, i = 1, 2, \ldots\}$ such that $F(\alpha_1, \ldots, \alpha_i) \subset G(\alpha_1, \ldots, \alpha_i)$ and that if $\bigcup_{i=1}^{\infty} G(\alpha_1, \ldots, \alpha_i) = X$ for a sequence $\{\alpha_i\}$, then $\bigcup_{i=1}^{\infty} F(\alpha_1, \ldots, \alpha_i) = X$. Perfectly normal spaces, countably compact spaces, Čech-complete paracompact spaces and σ-compact regular spaces are P-spaces. Normal P-spaces are countably paracompact. A Hausdorff space X is a normal (resp. paracompact) P-space if and only if the product $X \times Y$ is normal (resp. paracompact) for any metric space Y (Morita, *Math. Ann.*, 154 1964).

The product $X \times Y$ of locally compact Hausdorff spaces X and Y is a locally compact Hausdorff space. If, in this case, X and Y are paracompact, then so is the product. If the direct product space $\prod_\lambda X_\lambda$ of metric spaces is normal, then X_λ are compact except for at most countably many λ, and hence the product space is paracompact (A. H. Stone, 1948).

A class $\mathscr{C}$ of topological spaces is called **countably productive** if for a sequence X_i of members of $\mathscr{C}$ their product $\prod X_i$ is again a member of $\mathscr{C}$. The classes of (complete) (separable) metric spaces form such examples. The class of paracompact and Čech-complete spaces is countably productive (Z. Frolík, 1960). A topological space X is called a **p-space** if it is completely regular and there is a sequence $\mathfrak{M}_i$ of families of open sets in the Stone-Čech compactification $\beta(X)$ such that, for each point $x \in X$, $x \in \bigcap S(x, \mathfrak{M}_i) \subset X$ (Arkhangel'skiĭ, 1963). X is called an **M-space** if there is a normal sequence $\mathfrak{M}_i$ of open coverings of X such that if $K_1 \supset K_2 \supset \ldots$ is a sequence of nonempty closed sets and $K_i \subset S(x, \mathfrak{M}_i)$, $i = 1, 2, \ldots$, for an $x \in X$, then $\bigcap K_i \neq \varnothing$ (Morita, 1963). The class of paracompact p-spaces and that of paracompact Hausdorff M-spaces are the same and are countably productive. For a covering $\mathfrak{F}$ of X and an $x \in X$ we set $C(x, \mathfrak{F}) = \bigcap \{F \mid x \in F \in \mathfrak{F}\}$. X is called a **Σ-space** if X admits a sequence $\mathfrak{F}_i$ of locally finite closed coverings such that if $K_1 \supset K_2 \supset \ldots$ is a sequence of nonempty closed sets and $K_i \subset C(x, \mathfrak{F}_i)$, $i = 1, 2, \ldots$, for an $x \in X$, then $\bigcap K_i \neq \varnothing$ (K. Nagami, 1969). Σ-spaces are P-spaces. The class of all paracompact Σ-spaces is also countably productive. Among the above classes each one is always wider than its predecessors. Yet the product $X \times Y$ of a paracompact Hausdorff P-space X and a paracompact Hausdorff Σ-space Y is paracompact. Other examples of countably productive classes are the Suslin spaces and the Luzin spaces (→ Section CC) introduced by Bourbaki (1958), the **stratifiable spaces** by J. G. Ceder (1961) and C. J. R. Borges (1966), the **$\aleph_0$-spaces** by Michael (1966) and the **σ-spaces** by A. Okuyama (1967).

Z. Strongly Paracompact Spaces

Regular Lindelöf spaces are strongly paracompact. Conversely, if a connected regular space is strongly paracompact, then it is a Lindelöf space (Morita, 1948). Hence a connected nonseparable metric space is not strongly paracompact. Paracompact locally compact Hausdorff spaces and uniformly locally compact Hausdorff spaces are strongly paracompact.

These classes of spaces coincide under suitable †uniform structures.

AA. Collectionwise Normal Spaces

A Hausdorff space X is called a **collectionwise normal space** if for each discrete collection $\{F_\alpha \mid \alpha\in A\}$ of closed sets of X there exists a disjoint collection $\{U_\alpha \mid \alpha\in A\}$ of open sets with $F_\alpha\subset U_\alpha$ $(\alpha\in A)$ (R. H. Bing, 1951). If X satisfies an analogous condition for the case where each F_α is a singleton, X is called a **collectionwise Hausdorff space**. Paracompact Hausdorff spaces are collectionwise normal (Bing). Every point-finite open covering of a collectionwise normal space has a locally finite open refinement (Michael, Nagami).

A topological space X is called a **developable space** if it admits a sequence $\mathfrak{U}_i$, $i=1,2,\ldots$, of open coverings such that, for each point $x\in X$, $\{S(x,\mathfrak{U}_i) \mid i=1,2,\ldots\}$ forms a base for the neighborhood system of x (R. L. Moore, 1916). A regular developable space is called a **Moore space**. The question of whether or not every normal Moore space is metrizable is known as the **normal Moore space problem** (→ 273 Metric Spaces K). Collectionwise normal Moore spaces are metrizable (Bing).

BB. Real-Compact Spaces

A completely regular space X is called **real-compact** if X is complete under the smallest †uniformity such that each continuous real-valued function on X is uniformly continuous (→ 422 Uniform Spaces). This notion was introduced by E. Hewitt (*Trans. Amer. Math. Soc.*, 64 (1948)) under the name of **Q-space**, and independently by L. Nachbin (*Proc. International Congress of Mathematicians*, Cambridge, Mass., 1950).

A Lindelöf space is real-compact. If X_1 and X_2 are real-compact spaces such that the rings $C(X_1)$ and $C(X_2)$ of continuous real-valued functions on X_1 and X_2 are isomorphic, then X_1 and X_2 are homeomorphic (Hewitt). If X is real-compact, then X is homeomorphic to a closed subspace of the product space of copies of the space of real numbers, and conversely.

CC. Images and Inverse Images of Topological Spaces

Each continuous mapping $f:X\to Y$ is decomposed into the product $i\circ h\circ p$ of continuous mappings $p:X\to X/\sim$, $h:X/\sim\,\to f(X)$ and i: $f(X)\to Y$, where $\sim$ is the equivalence relation such that $x_1\sim x_2$ if and only if $f(x_1)=f(x_2)$. The mapping f is open (resp. closed) if and only if these mappings are all open (resp. closed). Then h is a homeomorphism. The image of a paracompact Hausdorff space under a closed continuous mapping is paracompact (Michael, 1957).

Let $f:X\to Y$ be a perfect surjection. Then Y is called a **perfect image** of X and X a **perfect inverse image** of Y. If, in this case, one of X and Y satisfies a property such as being compact, locally compact, σ-compact, Lindelöf, or countably compact, then the other also satisfies the property. When X and Y are completely regular, the same is true with regard to Čech completeness. Properties such as regularity, normality, complete normality, perfect normality, and the second countability axiom are preserved in perfect images; but complete regularity and strong paracompactness are not. Perfect images of metric spaces are also metrizable (S. Hanai and Morita, A. H. Stone, 1956). Conversely, perfect inverse images of paracompact spaces are paracompact. If a Hausdorff space is a perfect inverse image of a regular space (resp. k-space; → below), then it is a regular space (resp. k-space). Every paracompact Čech-complete space is a perfect inverse image of a †complete metric space (Z. Frolík, 1961). A completely regular space is a paracompact p-space if and only if it is a perfect inverse image of a metric space (Arkhangel'skiĭ, 1963). A mapping $f:X\to Y$ is called **quasi-perfect** if it is closed and continuous and the inverse image $f^{-1}(y)$ of each point $y\in Y$ is countably compact. A topological space X is an M-space if and only if there is a quasi-perfect mapping from X onto a metric space Y (Morita, 1964). Let $f:X\to Y$ be a quasi-perfect surjection. If one of X and Y is a Σ-space, then the other is also a Σ-space (Nagami, 1969).

A topological space X is called a **Fréchet-Uryson space** (or a **Fréchet space**) if the closure of an arbitrary set $A\subset X$ is the set of all limits of sequences in A (Arkhangel'skiĭ, 1963). X is called a **sequential space** if $A\subset X$ is closed whenever A contains all the limits of sequences in A (S. P. Franklin, 1965). X is called a **k′-space** if the closure of an arbitrary set A is the set of all points adherent to the intersection $A\cap K$ for a compact set K in X (Arkhangel'skiĭ, 1963). X is called a **k-space** if $A\subset X$ is closed whenever $A\cap K$ is closed in K for any compact set K (→ Arkhangel'skiĭ, *Trudy Moskov. Mat. Obshch.*, 13 (1965)). Spaces satisfying the first countability axiom are Fréchet-Uryson spaces. The Fréchet-Uryson spaces (resp. sequential spaces) are characterized as the images under hereditarily quotient (resp. quotient) mappings of metric spaces or locally compact metric spaces. Similarly the k′-spaces (resp. k-spaces) coincide with the images under hereditarily quotient (resp. quotient) mappings of locally compact spaces. The image of a metric space under a closed continuous mapping is called a **Lashnev space**. Any subspace of a Fréchet-Uryson space is a Fréchet-Uryson space. Conversely, a Hausdorff space is a Fréchet-Uryson space if any of its subspaces is a k-space. Čech-complete spaces are k-spaces. A Hausdorff space is called a **Suslin space** (resp. **Luzin space**) if it is the image under a continuous surjection (resp. continuous bijection) of a complete separable metric space (Bourbaki [9]; also → 22 Analytic Sets).

In Figs. 1, 2, and 3, the relationships between the various properties are indicated by the arrows.

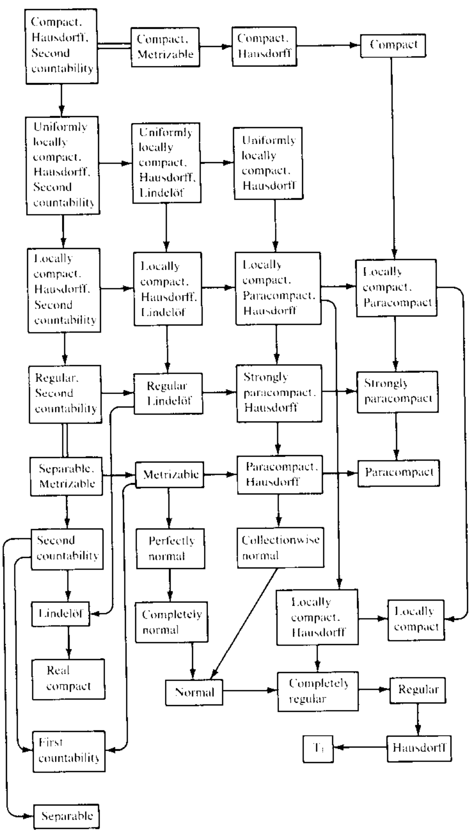


Fig. 1

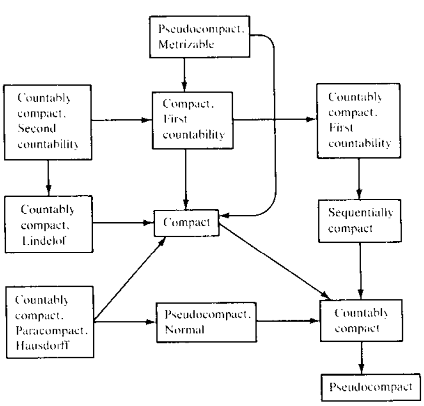


Fig. 2

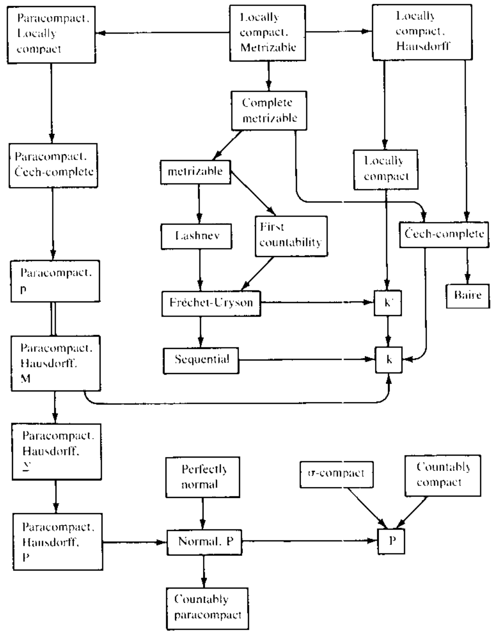


Fig. 3

References

[1] M. Fréchet, Sur quelques points du calcul fonctionnel, Rend. Circ. Mat. Palermo, 22 (1906), 1–74.

[2] F. Riesz, Stetigkeitsbegriff und abstrakte Mengenlehre, Atti del IV Congr. Intern. Mat. Roma 1908, II, 18–24, Rome, 1909.

[3] F. Hausdorff, Grundzüge der Mengenlehre, Teubner, 1914.

[4] M. Fréchet, Les espaces abstraits, Gauthier-Villars, 1928.

[5] C. Kuratowski, Topologie, I, Monograf. Mat., I, 1933, revised edition, 1948; II, 1950.

[6] W. Sierpiński, Introduction to general topology, University of Toronto Press, 1934.

[7] P. S. Aleksandrov (Alexandroff) and H. Hopf, Topologie I, Springer, 1935.

[8] J. W. Tukey, Convergence and uniformity in topology, Ann. Math. Studies, Princeton Univ. Press, 1940.

[9] N. Bourbaki, Eléments de mathématique, III. Topologie générale, 1940–1949; new edition, 1970–1974; Hermann, English translation, General topology, Addison-Wesley, 1966.

[10] J. L. Kelley, General topology, Van Nostrand, 1955.

[11] J. Dugundji, Topology, Allyn and Bacon, 1966.

[12] J. Nagata, Modern general topology, North-Holland, 1968.

[13] R. Engelking, General topology, Polish Scientific Publ., 1977.

[14] P. S. Aleksandrov, Some results in the theory of topological spaces, obtained within the last twenty-five years, Russian Math. Surveys, 16, no. 2 (1960), 23–83.
[15] A. V. Arkhangel'skiĭ, Mappings and spaces, Russian Math. Surveys, 21, no. 4 (1966), 115–162.
[16] Y. Kodama and K. Nagami, Theory of topological spaces (in Japanese), Iwanami, 1974.
[17] L. Steen and J. Seebach, Jr., Counterexamples in topology, Holt, Rinehart and Winston, 1970.
[18] L. Gillman and M. Jerison, Rings of continuous functions, Van Nostrand, 1960.
[19] S. A. Gaal, Point set topology, Academic Press, 1964.
[20] E. Čech, Topological spaces, Interscience, 1966.
[21] R. C. Walker, The Stone-Čech compactification, Springer, 1974.
[22] R. A. Alò and H. L. Shapiro, Normal topological spaces, Cambridge Univ. Press, 1974.
[23] M. D. Weir, Hewitt-Nachbin spaces, North-Holland, 1975.

426 (IX.1) Topology

The term **topology** means a branch of mathematics that deals with topological properties of geometric figures or point sets. A classical result in topology is the Euler relation on polyhedra: Let α_0, α_1, and α_2 be the numbers of vertices, edges, and faces of a polyhedron homeomorphic to the 2-dimensional sphere; then $\alpha_0-\alpha_1+\alpha_2=2$ (†Euler-Poincaré formula for the 2-dimensional case; actually, the formula was known to Descartes). It is one of the earliest results in topology. In 1833, C. F. Gauss used integrals to define the notion of †linking numbers of two closed curves in a space (→ 99 Degree of Mapping). It was in J. B. Listing's classical work *Vorstudien zur Topologie* (1847) that the term *topology* first appeared in print.

In the 19th century, B. Riemann published many works on function theory in which topological methods played an essential role. He solved the homeomorphism problem for compact surfaces (→ 410 Surfaces); his result is basic in the theory of algebraic functions. In the same period, mathematicians began to study topological properties of n-dimensional polyhedra. E. Betti considered the notion of †homology. H. Poincaré, however, was the first to recognize the importance of a topological approach to analysis in general; he defined the †homology groups of a complex [1]. He obtained the famous †Poincaré duality theorem and defined the †fundamental group. He considered †polyhedra as the basic objects in topology, and deduced topological properties utilizing †complexes obtained from polyhedra by †simplicial decompositions. He thus constructed a branch of topology known as **combinatorial topology**.

In its beginning stages combinatorial topology dealt only with polyhedra. In the late 1920s, however, it became possible to apply combinatorial methods to general †compact spaces. P. S. Alexandrov introduced the concept of approximation of a †compact metric space by an inverse sequence of complexes and the definition of homology groups for these spaces. His idea had a precursor in the notion of †simplicial approximations of continuous mappings, which was introduced by L. E. J. Brouwer in 1911. In 1932, E. Čech defined homology groups for arbitrary spaces utilizing the †inductive limit of the homology groups of polyhedra; and †Čech cohomology groups for arbitrary spaces were also defined. S. Eilenberg established †singular (co)homology theory using †singular chain complexes (1944). The axiomatic approach to (co)homology theory is due to Eilenberg and Steenrod, who gave axioms for (co)homology theory in a most comprehensive way and unified various (co)homology theories (1945) (→ 201 Homology Theory.

The approach using algebraic methods has progressed extensively in connection with the development of homology theory. This branch is called **algebraic topology**. In the 1920s and 1930s, a number of remarkable results in algebraic topology, such as the †Alexander duality theorem, the †Lefschetz fixed-point theorem, and the †Hopf invariant, were obtained. In the late 1930s, W. Hurewicz developed the theory of higher-dimensional †homotopy groups (→ 153 Fixed-Point Theorems, 201 Homology Theory, 202 Homotopy Theory). J. H. C. Whitehead introduced the concept of †CW complexes and proved an algebraic characterization of the homotopy equivalence of CW complexes. N. Steenrod developed †obstruction theory utilizing †squaring operations in the cohomology ring (1947). Subsequently, the theory of †cohomology operations was introduced (→ 64 Cohomology Operations, 305 Obstructions). The theory of †spectral sequences for †fiber spaces was originated by J. Leray (1945) and J.-P. Serre (1951) and was successfully applied to cohomology operations and homotopy theory by H. Cartan and Serre (1954) (→ 148 Fiber Spaces, 200 Homological Algebra). The study of the combinatorial structures of polyhedra and †piecewise linear mappings has flourished since 1940 in the works of Whitehead, S. S. Cairns, and others. S. Smale and, independently, J. Stallings solved the †generalized Poincaré conjecture in 1960. The †Hauptvermutung in combinatorial topology was solved negatively in 1961 by B. Mazur and J. Milnor. E. C. Zeeman proved the unknottedness of codimension 3 (1962). The recent development of the theory in conjunction with progress in †differential topology is notable. The Hauptvermutung for combinatorial manifolds was solved in 1969 by Kirby, Siebenmann, and Wall. In particular, there exist different combinatorial structures on tori of dimension $\geqslant 5$, and there are topological manifolds that do not admit any combinatorial structure (→ 65 Combinatorial Manifolds, 114 Differential Topology, 235 Knot Theory).

The global theory of differentiable manifolds started from the algebraic-topological study of †fiber bundles and †characteristic classes in the 1940s. R. Thom's fundamental theorem of †cobordism (1954) was obtained through extensive use of cohomology operations and homotopy groups. Milnor (1956) showed that the sphere S^7 may have differentiable structures that are essentially distinct from each other by using †Morse theory and the †index theorem of Thom and Hirzebruch. These results led to the creation of a new field, †differential topology (→ 56 Characteristic Classes, 114 Differential Topology).

Since 1959, A. Grothendieck, M. F. Atiyah, F. Hirzebruch, and J. F. Adams have developed †K-theory, which is a generalized cohomology theory constructed using stable classes of †vector bundles (→ 237 K-Theory).

†Knot theory, an interesting branch of topology, was one of the classical branches of topology and is now studied in connection with the theory of low-dimensional manifolds (→ 235 Knot Theory).

On the other hand, G. Cantor established general set theory in the 1870s and introduced such notions as †accumulation points, †open sets, and †closed sets in Euclidean space. The first important generalization of this theory was the concept of †topological space, which was proposed by M. Fréchet and developed by F. Hausdorff at the beginning of the 20th century. The theory subsequently became a new field of study, called **general topology** or **set-theoretic topology**. It deals with the topological properties of point sets in a Euclidean or topological space without reference to polyhedra. There has been a remarkable development of the theory since abount 1920, notably by Polish mathematicians S. Janiszewski, W. Sierpiński, S. Mazurkiewicz, C. Kuratowski, and others. The contributions of R. L. Moore, G. T. Whyburn, and K. Menger are also important (→ 382 Shape Theory, 425 Topological Spaces).

Topology is not only a foundation of various theories, but is also itself one of the most important branches of mathematics. It consists of †homology theory, †homotopy theory, †differential topology, †combinatorial manifolds, †K-theory, †transformation groups, †theory of singularities, †foliations, †dynamical systems, †catastrophe theory, etc. It continues to develop in interaction with other branches of mathematics (→ 51 Catastrophe Theory, 126 Dynamical Systems, 154 Foliations, 418 Theory of Singularities, 431 Transformation Groups).

References

[1] H. Poincaré, Analysis situs, J. Ecole Polytech., (2) 1 (1895), 1–121. (Oeuvres, Gauthier-Villars, 1953, vol. 6, 193–288.)
[2] P. S. Aleksandrov and H. Hopf, Topologie I, Springer, 1935 (Chelsea, 1965).
[3] S. Lefschetz, Algebraic topology, Amer. Math. Soc. Colloq. Publ., 1942.
[4] S. Eilenberg and N. E. Steenrod, Foundations of algebraic topology, Princeton Univ. Press, 1952.

427 (IX.12) Topology of Lie Groups and Homogeneous Spaces

A. General Remarks

Among various topological structures of †Lie groups and †homogeneous spaces, the structures of their †(co)homology groups and †homotopy groups are of special interest. Let G/H be a homogeneous space, where G is a Lie group and H is its closed subgroup. Then $(G, G/H, H)$ is a †fiber bundle, where G/H is the base space and H is the fiber. Thus homology and homotopy theory of fiber bundles (†spectral sequences and †homotopy exact sequences) can be applied. The †cellular decomposition of †Stiefel manifolds, †Grassmann manifolds, and †Kähler homogeneous spaces are known. Concerning †symmetric Riemannian spaces, we have various interesting methods, such as the use of invariant differential forms in connection with real cohomology rings and the use of †Morse theory in order to establish relations between the diagrams of symmetric Riemannian spaces G/H and homological properties of their †loop spaces and some related homogeneous spaces [4, 5]. Lie groups can be regarded as special cases of homogeneous spaces or symmetric spaces, although their group structures are of particular importance. A connected Lie group is homeomorphic to the product of one of its compact subgroups and a Euclidean space (†Cartan-Mal'tsev-Iwasawa theorem). Hence the topological structure of a connected Lie group is essentially determined by the topological structures of its compact subgroups.

B. Homology of Compact Lie Groups

Let G be a connected compact Lie group. Since G is an †H-space whose multiplication is given by its group multiplication h, $H^*(G;k)$ and $H_*(G;k)$ are dual †Hopf algebras for any coefficient field k. Also, $H^*(G;k)$ is isomorphic as a †graded algebra to the tensor product of †elementary Hopf algebras (→ 203 Hopf Algebras), but no factor of the tensor product is isomorphic to a polynomial ring because G is a finite †polyhedron. In particular, if $k=\mathbf{R}$ (the field of real numbers), then $H^*(G;\mathbf{R})\cong \bigwedge_{\mathbf{R}}(x_1,\dots,x_l)$ (the exterior (Grassmann) algebra over $\mathbf{R}$ with generators $x_1,\dots,x_l$ of odd degrees). Here we can choose generators x_i such that $h^*(x_i)=1\otimes x_i+x_i\otimes 1$, $1\leqslant i\leqslant l$. The x_i that satisfy this property are said to be **primitive**. Since in this case the †comultiplication h^* is commutative, the multiplication h_* is also commutative and the Hopf algebra $H_*(G;R)$ is an exterior algebra generated by elements y_i having the same degree as x_i $(i=1,\dots,l)$. When the characteristic of the coefficient field k is nonzero, h_* need not be commutative.

The dimension of a †maximal torus of a connected compact Lie group G is independent of the choice of the maximal torus and is called the **rank** of G. The rank of G coincides with the number l of generators of $H^*(G;\mathbf{R})$. E. Cartan studied $H^*(G;\mathbf{R})$ by utilizing invariant differential forms. The cohomology theory of Lie algebras originated from the method he used in his study. $H^*(G;\mathbf{R})$ is invariant under †local isomorphisms of groups G. For †classical compact simple Lie groups G, R. Brauer calculated $H^*(G;\mathbf{R})$, while C.-T. Yen and C. Chevalley calculated $H^*(G;\mathbf{R})$ for †exceptional compact simple Lie groups (→ Appendix A, Table 6.IV). The degrees of the generators have group-theoretic meaning. Suppose that the degree of the ith generator is $2m_i-1$, $1\leqslant i\leqslant l$, and that $m_1\leqslant m_2\leqslant\ldots\leqslant m_l$. When G is simple, there is a relation $m_i+m_{l-i+1}=$ constant (Chevalley's duality). We have a proof for this property that does not use classification.

The cohomology groups $H^*(G;\mathbf{Z}_p)$ (where p is a prime and $\mathbf{Z}_p=\mathbf{Z}/p\mathbf{Z}$) have been determined as graded algebras for all compact simple Lie groups by A. Borel, S. Araki, and P. Baum and W. Browder (→ Appendix A, Table 6.IV).

C. Cohomology of Classifying Spaces

Let (E_G, B_G, G) be a †universal bundle of a connected compact Lie group G and p a prime or zero. Suppose that the integral cohomology of G has no p-torsion (no torsion when $p=0$). Then $H^*(G;\mathbf{Z}_p)=\bigwedge_{\mathbf{Z}_p}(x'_1,\dots,x'_l)$ $(H^*(G;\mathbf{Z})=\bigwedge_{\mathbf{Z}}(x'_1,\dots,x'_l)$ when $p=0)$, an exterior algebra with $\deg x'_i=2m_i-1$, $1\leqslant i\leqslant l$, and the generators x'_i can be chosen to be †transgressive in the spectral sequence of the universal bundle. Let $y_1,\dots,y_l$ be their transgression images. Then $\deg y_i=2m_i$, $1\leqslant i\leqslant l$, and the cohomology of the †classifying space B_G over $\mathbf{Z}_p$ (resp. $\mathbf{Z}$) is the polynomial algebra with generators $y_1,\dots,y_l$. Let T be a maximal torus of G. Then $B_T=E_G/T$ is a classifying space of T, the †Weyl group $W=N(T)/T$ of G with respect to T operates on B_T by †right translations, and $H^*(T;\mathbf{Z})$ has no torsion and is an exterior algebra with l generators of degree 1. Thus $H^*(B_T;\mathbf{Z})=\mathbf{Z}[u_1,\dots,u_l]$, $\deg u_i=2$. Let I_W be the subalgebra of $H^*(B_T;\mathbf{Z})$ consisting of W-invariant polynomials, and let ρ be the projection of the bundle $(B_T, B_G, G/T)$. Then under the assumption that G has no p-torsion (no torsion), the cohomology mapping ρ^* over $\mathbf{Z}_p(\mathbf{Z})$ is monomorphic, and $\rho^*: H^*(B_G;\mathbf{Z}_p)\cong I_W\otimes\mathbf{Z}_p$ $(H^*(B_G;\mathbf{Z})\cong I_W)$ [1]. In the case of real coefficients, we have $H^*(B_G;\mathbf{R})\cong I_W\otimes\mathbf{R}$ for all G, and $m_1,\dots,m_l$ are the degrees of generators of the ring I_W of W-invariant polynomials.

Example (1) $G=U(n)$: $l=n$ and G has no torsion. W operates on $H^*(B_T;\mathbf{Z})$ as the group of all permutations of generators $u_1,\dots,u_n$. Thus generators of I_W are the †elementary symmetric polynomials $\sigma_1,\dots,\sigma_n$ of $u_1,\dots,u_n$. Let $c_1,\dots,c_n$ be the †universal Chern classes; then $\rho^*(c_i)=\sigma_i$ and $H^*(B_{U(n)};\mathbf{Z})=\mathbf{Z}[c_1,\dots,c_n]$.

Example (2) $G=SO(n)$: $l=[n/2]$ and G has no p-torsion for $p\neq 2$. W operates on $H^*(B_T;\mathbf{Z})$ as the group generated by the permutations of generators $u_1,\dots,u_l$ and by the transformations $\sigma(u_i)=e_iu_i$, $e_i=\pm 1$, where the number of u_i for which $e_i=-1$ is arbitrary for odd n and even for even n. Thus the generators of I_W are the elementary symmetric polynomials $\sigma'_1,\dots,\sigma'_l$ of $u_1^2,\dots,u_l^2$ for odd n and $\sigma'_1,\dots,\sigma'_{l-1}$ and $u_1\dots u_l$ for even n. Let $p_1,\dots,p_l$ be the †universal Pontryagin classes and χ be the †universal Euler-Poincaré class in the case of even n. Then $\rho^*(p_i)=\sigma'_i$ and $\rho^*(\chi)=u_1\dots u_l$ for integral cohomology. Denote the mod p

reduction of p_i and χ by $\bar{p}_i$ and $\bar{\chi}$, respectively. Then $H^*(B_{SO(2l+1)}; \mathbf{Z}_p) = \mathbf{Z}_p[\bar{p}_1, \ldots, \bar{p}_l]$ and $H^*(B_{SO(2l)}; \mathbf{Z}_p) = \mathbf{Z}_p[\bar{p}_1, \ldots, \bar{p}_{l-1}, \bar{\chi}]$ ($p=0$ or >2).

Example (3) $G = O(n)$: If we use the subgroup Q consisting of all diagonal matrices instead of T, then we can make a similar argument for $\mathbf{Z}_2$-cohomology. Since $Q \cong (\mathbf{Z}_2)^n$, $H^*(B_Q; \mathbf{Z}_2) = \mathbf{Z}_2[v_1, \ldots, v_n]$ ($\mathbf{Z}_2[v_1, \ldots, v_n]$ is a polynomial ring with $\deg v_i = 1$), and $W_2 = N(Q)/Q$ operates on B_Q by right translations and on $H^*(B_Q; \mathbf{Z}_2)$ as the group of all permutations of $v_1, \ldots, v_n$. Let I_{W_2} be the subalgebra of $H^*(B_Q; \mathbf{Z}_2)$ consisting of all W_2-invariant polynomials. Then I_{W_2} is a polynomial ring generated by the elementary symmetric polynomials $\sigma_1'', \ldots, \sigma_n''$ of $v_1, \ldots, v_n$. The projection $\rho_2 : B_Q \to B_{O(n)}$ induces a monomorphic cohomology mapping ρ_2^* over $\mathbf{Z}_2$, and ρ_2^*: $H^*(B_{O(n)}; \mathbf{Z}_2) \cong I_{W_2}$. Let $w_1, \ldots, w_n$ be the †universal Stiefel-Whitney classes. Then $\rho_2^*(w_i) = \sigma_i''$ and $H^*(B_{O(n)}; \mathbf{Z}_2) = \mathbf{Z}_2[w_1, \ldots, w_n]$ [2].

D. Grassmann Manifolds

The following manifolds are called **Grassmann manifolds**: The manifold $M_{n+m,n}(\mathbf{R})$ consisting of all n-subspaces of $\mathbf{R}^{n+m}$; the manifold $\tilde{M}_{n+m,n}(\mathbf{R})$ consisting of all oriented n-subspaces of $\mathbf{R}^{n+m}$; and the manifold $M_{n+m,n}(\mathbf{C})$ consisting of all complex n-subspaces of $\mathbf{C}^{n+m}$. These are expressed as quotient spaces as follows: $M_{n+m,n}(\mathbf{R}) = O(n+m)/O(n) \times O(m)$, $\tilde{M}_{n+m,n}(\mathbf{R}) = SO(n+m)/SO(n) \times SO(m)$, and $M_{n+m,n}(\mathbf{C}) = U(n+m)/U(n) \times U(m)$. They admit cellular decompositions by †Schubert varieties from which their cohomologies can be computed (→ 56 Characteristic Classes). $M_{n+m,n}(\mathbf{R})$ and $\tilde{M}_{n+m,n}(\mathbf{R})$ have no p-torsion for $p \neq 2$, and $M_{n+m,n}(\mathbf{C})$ has no torsion. These spaces are m-, m-, and $(2m+1)$-classifying spaces of $O(n)$, $SO(n)$, and $U(n)$, respectively. Hence their cohomologies are isomorphic to those of B_G ($G = O(n)$, $SO(n)$, $U(n)$) in dimensions $<m$, $<m$, and $\leqslant 2m$, respectively; and they are polynomial rings generated by suitable universal characteristic classes in low dimensions.

E. Cohomologies of Homogeneous Spaces G/U (Rank G = Rank U)

Let G be a compact connected Lie group and U a closed subgroup of G with the same rank as G. Denote the degrees of generators of $H^*(G; \mathbf{R})$ and $H^*(U; \mathbf{R})$ by $2m_1 - 1, \ldots, 2m_l - 1$, and $2n_1 - 1, \ldots, 2n_l - 1$, respectively. Then the real-coefficient †Poincaré polynomial P_0 of the homogeneous space G/U is given by $P_0(G/U, t) = \prod_i (1 - t^{2m_i})/(1 - t^{2n_i})$ (G. Hirsch). When G, U, and G/U have no p-torsion, the same formula is valid for the $\mathbf{Z}_p$-coefficient Poincaré polynomial [1]. When U is the †centralizer of a torus, G/U has a complex analytic cellular decomposition [3]. Hence G/U has no torsion in this case. This was proved by R. Bott and H. Samelson by utilizing Morse theory [5] (→ 279 Morse Theory). The case $U = T$ has also been studied.

F. Homotopy Groups of Compact Lie Groups

The †fundamental group $\pi_1(G)$ of a compact Lie group G is Abelian. Furthermore, $\pi_2(G) = 0$. If we apply Morse theory to G, the variational completeness of G can be utilized to show that the loop space ΩG has no torsion and that its odd-dimensional cohomologies vanish [4]. Consequently, when G is non-Abelian and simple, we have $\pi_3(G) \cong \mathbf{Z}$. A †periodicity theorem on †stable homotopy groups of classical groups proved by Bott is used in K-theory (→ 202 Homotopy Theory; 237 K-Theory). (For explicit forms of homotopy groups → Appendix A, Table 6.VI).

Homotopy groups of Stiefel manifolds are used to define characteristic classes by †obstruction cocycles (→ 147 Fiber Bundles; Appendix A, Table 6.VI).

References

[1] A. Borel, Sur la cohomologie des espaces fibrés principaux et des espaces homogènes de groupes de Lie compacts, Ann. Math., (2) 57 (1953), 115–207.
[2] A. Borel, La cohomologie mod 2 de certains espaces homogènes, Comment. Math. Helv., 27 (1953), 165–197.
[3] A. Borel, Kählerian coset spaces of semisimple Lie groups, Proc. Nat. Acad. Sci. US, 40 (1954), 1147–1151.
[4] R. Bott, An application of the Morse theory to the topology of Lie-groups, Bull. Soc. Math. France, 84 (1956), 251–281.
[5] R. Bott and H. Samelson, Applications of the theory of Morse to symmetric spaces, Amer. J. Math., 80 (1958), 964–1029.

428 (XIII.17) Total Differential Equations

A. Pfaff's Problem

A **total differential equation** is an equation of the form

$$\omega = 0, \tag{1}$$

where ω is a †differential 1-form $\sum_{i=1}^{n} a_i(x)dx_i$ on a manifold X. A submanifold M of X is called an **integral manifold** of (1) if each vector ξ of the †tangent vector space $T_x(M)$ of M at every point x on M satisfies $\omega(\xi)=0$. We denote the maximal dimension of integral manifolds of (1) by $m(\omega)$. J. F. Pfaff showed that $m(\omega)\geqslant(n-1)/2$ for any ω. The problem of determining $m(\omega)$ for a given form ω is called **Pfaff's problem**. This problem was solved by G. Frobenius, J. G. Darboux, and others as follows: Form an †alternating matrix

$$(a_{ij})_{1\leqslant i,j\leqslant n} \tag{2}$$

from the coefficients of the †exterior derivative of ω,

$$d\omega=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}(x)dx_i\wedge dx_j,$$

where $a_{ij}=\partial a_j/\partial x_i-\partial a_i/\partial x_j$. Suppose that the rank of (2) is $2t$. Then the rank of the matrix

$$\begin{pmatrix} a_{ij} & -a_i \\ a_j & 0 \end{pmatrix}_{1\leqslant i,j\leqslant n}$$

is $2t$ or $2t+2$. In the former case $m(\omega)=n-t$, and ω can be expressed in the form

$$\sum_{i=1}^{t} u_{2i-1}\,du_{2i}$$

by choosing a suitable coordinate system $(u_1,\dots,u_n)$. In the latter case $m(\omega)=n-t-1$, and ω can be expressed in the form

$$\sum_{i=1}^{t} u_{2i-1}\,du_{2i}+du_{2t+1}$$

by choosing a suitable coordinate system $(u_1,\dots,u_n)$. This theorem is called **Darboux's theorem.**

A 1-form ω is called a **Pfaffian form**, and equation (1) is called a **Pfaffian equation**. A system of equations $\omega_i=0$ $(1\leqslant i\leqslant s)$ for 1-form ω_i is called a **system of Pfaffian equations** or a **system of total differential equations** [6, 12, 26].

B. Systems of Differential Forms and Systems of Partial Differential Equations

Let Ω be a system of differential forms ω_i^p, $0\leqslant p\leqslant n$, $1\leqslant i\leqslant v_p$, on X, where ω_i^p is a p-form on X. A submanifold M of X is called an **integral manifold** of $\Omega=0$ if for each p $(0\leqslant p\leqslant \dim M)$, any p-dimensional subspace E_p of $T_x(M)$ satisfies $\omega_i^p(E_p)=0$ $(1\leqslant i\leqslant v_p)$ at every point x on M. Denote the maximal dimension of integral manifolds of $\Omega=0$ by $m(\Omega)$. The problem of determining $m(\Omega)$ for a given system Ω is called the **generalized Pfaff problem**, and will be explained in later sections. By fixing a local coordinate system of X and dividing it into two systems $(x_1,\dots,x_r)$ and $(y_1,\dots,y_m)$ $(m=n-r)$, we can consider the problem of finding an integral manifold of $\Omega=0$ defined by

$$y_\alpha=y_\alpha(x_1,\dots,x_r),\qquad 1\leqslant\alpha\leqslant m.$$

This problem can be reduced to solving a system of partial differential equations of the first order on the submanifold N with the local coordinate system $(x_1,\dots,x_r)$.

Consider a system of partial differential equations $\Phi=0$ of order l:

$$\varphi_\lambda(x_i,y_\alpha,p_\beta^{j_1\dots j_r})=0,\qquad 1\leqslant\lambda\leqslant s, \tag{3}$$

with $1\leqslant i\leqslant r$, $1\leqslant\alpha,\ \beta\leqslant m$, $j_1+\dots+j_r\leqslant l$, where

$$p_\beta^{j_1\dots j_r}=\frac{\partial^{j_1+\dots+j_r}y_\beta}{\partial x_1^{j_1}\dots\partial x_r^{j_r}}. \tag{4}$$

A submanifold defined by $y_\alpha=y_\alpha(x_1,\dots,x_r)$, $1\leqslant\alpha\leqslant m$, is called a **solution** of $\Phi=0$ if it satisfies (3) identically. The problem of determining whether a given system $\Phi=0$ has a solution was solved by C. Riquier, who showed that any system can be prolonged either to a passive orthonomic system or to an incompatible system by a finite number of steps. A system of partial differential equations is called a **prolongation** of another system if the former contains the latter and they have the same solution. A **passive orthonomic system** is one whose general solution can be parametrized by an infinite number of arbitrary constants. A solution containing parameters is called a **general solution** if by specifying the parameters we can obtain a solution of the †Cauchy problem for any initial data. A system (3) is said to be **incompatible** if it implies a nontrivial relation $f(x_1,\dots,x_r)=0$ among the x_i.

The problem of solving a system $\Phi=0$ of partial differential equations can be reduced to that of finding integral manifolds of a system of differential forms Σ as follows: Let J^l be a manifold with the local coordinate system

$$(x_i,y_\alpha,p_\beta^{j_1\dots j_r};\ 1\leqslant i\leqslant r,\ 1\leqslant\alpha,\ \beta\leqslant m,\ j_1+\dots+j_r\leqslant l),$$

and Σ be a system of 0-forms φ_λ $(1\leqslant\lambda\leqslant s)$ and 1-forms

$$dy_\alpha-\sum_{i=1}^{r}p_\alpha^i dx_i,$$

$$dp_\beta^{j_1\dots j_r}-\sum_{k=1}^{r}p_\beta^{j_1\dots j_k+1\dots j_r}dx_k$$

$(1\leqslant\alpha,\beta\leqslant m, j_1+\dots+j_r<l)$. Then an integral manifold of $\Sigma=0$ of the form

$$y_\alpha=y_\alpha(x_1,\dots,x_r),\qquad 1\leqslant\alpha\leqslant m,$$

$$p_\beta^{j_1\dots j_r}=p_\beta^{j_1\dots j_r}(x_1,\dots,x_r),$$

$$1\leqslant\beta\leqslant m,\qquad j_1+\dots+j_r\leqslant l,$$

gives a solution $y_\alpha = y_\alpha(x_1, \dots, x_r)$, $1 \leqslant \alpha \leqslant m$, of $\Phi = 0$, and y_β and $p_\beta^{j_1 \dots j_r}$ satisfy (4).

Conversely, a solution $y_\alpha = y_\alpha(x_1, \dots, x_r)$, $1 \leqslant \alpha \leqslant m$, of $\Phi = 0$ gives an integral manifold of $\Sigma = 0$ if we define $p_\beta^{j_1 \dots j_r}(x_1, \dots, x_r)$ by (4) [23, 24, 26].

C. Systems of Partial Differential Equations of First Order with One Unknown Function

Consider a system of independent †vector fields on N:

$$L_\lambda = \sum_{i=1}^{r} b_{\lambda i}(x) \frac{\partial}{\partial x_i}, \qquad 1 \leqslant \lambda \leqslant s.$$

We solve a system of inhomogeneous equations

$$L_\lambda y - f_\lambda(x) y - g_\lambda(x) = 0, \qquad 1 \leqslant \lambda \leqslant s, \qquad (5)$$

for a given system of $f_\lambda(x)$ and $g_\lambda(x)$. The system (5) is called a **complete system** if each of the expressions

$$[L_\lambda, L_\mu] y - (L_\lambda f_\mu - L_\mu f_\lambda) y - (f_\mu g_\lambda - f_\lambda g_\mu)$$
$$- (L_\lambda g_\mu - L_\mu g_\lambda), \qquad 1 \leqslant \lambda < \mu \leqslant s, \quad (6)$$

is a linear combination of the left-hand sides of (5), where $[L_\lambda L_\mu]$ means the †commutator of L_λ and L_μ. This condition is called the **complete integrability condition** for (5). Suppose that the homogeneous system

$$L_\lambda y = 0, \qquad 1 \leqslant \lambda \leqslant s, \qquad (7)$$

is complete. Then it has a system of †functionally independent solutions $y_1, \dots, y_{r-s}$, and any solution y of (8) is a function of them: $y = \psi(y_1, \dots, y_{r-s})$. If the inhomogeneous system (5) is complete, then the homogeneous system (7) is complete. This notion of a complete system is due to Lagrange and was extended to a system of nonlinear equations by Jacobi as follows (→ 324 Partial Differential Equations of First Order C).

Consider a system of nonlinear equations

$$F_\lambda(x_1, \dots, x_r, y, p_1, \dots, p_r) = 0, \qquad 1 \leqslant \lambda \leqslant s, \qquad (8)$$

where $p_i = \partial y / \partial x_i$. The system (8) is called an **involutory system** if each of $[F_\lambda, F_\mu]$, $1 \leqslant \lambda < \mu \leqslant s$, is a linear combination of $F_1, \dots, F_s$. Here †Lagrange's bracket $[F, G]$ is defined by

$$[F, G] = \sum_{i=1}^{r} \frac{\partial F}{\partial p_i} \left(\frac{\partial G}{\partial x_i} + p_i \frac{\partial G}{\partial y} \right)$$
$$- \sum_{i=1}^{r} \frac{\partial G}{\partial p_i} \left(\frac{\partial F}{\partial x_i} + p_i \frac{\partial F}{\partial y} \right).$$

Suppose that the system (8) is involutory and $F_1, \dots, F_s$ are functionally independent. Then, in general, we can solve the following †Cauchy problem for an $(r-s)$-dimensional submanifold N_{r-s} of N: Given a function f on N_{r-s}, find a solution y of (8) satisfying $y = f$ on N_{r-s}. We can construct a solution by integrating a system of ordinary differential equations called a †characteristic system of differential equations. Hence the solution of these problems may be carried out in the C^∞-category (→ 322 Partial Differential Equations (Methods of Integration) B) [7, 11].

D. Frobenius's Theorem

Let X be a †differentiable manifold of class C^∞ and Ω be a system of independent 1-forms ω_i, $1 \leqslant i \leqslant s$, on X. Then the system of Pfaffian equations $\Omega = 0$ is called a **completely integrable system** if at every point x of X,

$$d\omega_i = \sum_{j=1}^{s} \theta_{ij} \wedge \omega_j, \qquad 1 \leqslant i \leqslant s,$$

for 1-forms θ_{ij} on a neighborhood of x. Suppose that $\Omega = 0$ is completely integrable. Then at every point x of X, there exists a local coordinate system $(f_1, \dots, f_s, x_{s+1}, \dots, x_n)$ in a neighborhood U of x for which a tangent vector ξ of X at $z \in U$ satisfies $\omega_i(\xi) = 0$, $1 \leqslant i \leqslant s$, if and only if $\xi f_i = 0$, $1 \leqslant i \leqslant s$. In this case, each of the df_i is a linear combination of $\omega_1, \dots, \omega_s$, and conversely, each of the ω_i is a linear combination of $df_1, \dots, df_s$. In general, a function f for which df is a linear combination of $\omega_1, \dots, \omega_s$ is called a **first integral** of $\Omega = 0$.

The theorem of the previous paragraph is called **Frobenius's theorem**, which can be stated in the dual form as follows: Let $D(X)$ be a †subbundle of the †tangent bundle $T(X)$ over X. The mapping $X \ni x \to D_x(X)$ is called a **distribution** on X. It is said to be an **involutive distribution** if at every point x of X we can find a system of independent vector fields L_i $(1 \leqslant i \leqslant s)$ on a neighborhood U of x such that the $L_i(z)$ $(1 \leqslant i \leqslant s)$ form a basis of $D_z(X)$ at every $z \in U$ and satisfy $[L_i, L_j] \equiv 0 \; (L_1, \dots, L_s)$, $1 \leqslant i < j \leqslant s$, on U. A connected submanifold M of X is called an **integral manifold** of $D(X)$ if $T_x(M) = D_x(X)$ at every point x of M. Suppose that $D(X)$ gives an involutive distribution on X. Then every point x of X is in a maximal integral manifold M that contains any integral manifold including x as a submanifold.

E. Cartan-Kähler Existence Theorems

Let X be a †real analytic manifold. Denote the †sheaf of rings of differential forms on X by $\Lambda(X)$ and its subsheaf of $\mathcal{O}(X)$-modules of p-forms on X by $\Lambda_p(X)$, $1 \leqslant p \leqslant n$, where $\mathcal{O}(X)$ is the sheaf of rings of 0-forms on X. A subsheaf of ideals Σ is called a **differential ideal** if it is generated by Σ_p, $0 \leqslant p \leqslant n$, and contains $d\Sigma$, where $\Sigma_p = \Sigma \cap \Lambda_p(X)$. Consider a differential ideal Σ on X. Denote the †Grassmann manifold of p-dimensional subspaces of $T_x(X)$ with origin $x \in X$ by $G_p(x)$, and the Grassmann manifold $\bigcup_{x\in X} G(x)$ over X by $G_p(X)$. An element E_p of $G_p(x)$ is called a p-dimensional **contact element** with **origin** x. An element E_p of $G_p(x)$ is called an **integral element** of Σ_p if $\omega(E_p)=0$ at x for any p-form ω in Σ; furthermore, E_p is called an integral element of Σ if any element E_q contained in E_p, $0 \leqslant q \leqslant p$, is an integral element of Σ_q. In particular, 0-dimensional and 1-dimensional integral elements are called **integral points** and **integral vectors**, respectively. It can be proved that an element E_p is an integral element of Σ if and only if it is an integral element of Σ_p. The **polar element** $H(E_p)$ of an integral element E_p with origin x is defined as the subspace of $T_x(X)$ consisting of all vectors that generate with E_p an integral element of Σ. Let $(\Sigma_p)^0$, $0 \leqslant p \leqslant n$, be the subsheaf of $\mathcal{O}(X)$-modules in $\mathcal{O}(G_p(X))$ consisting of all 0-forms

$$\sum_{1 \leqslant i_1 < \ldots < i_p \leqslant n} a_{i_1 \ldots i_p} z_{i_1 \ldots i_p}$$

on $G_p(X)$ derived from a p-form

$$\sum_{1 \leqslant i_1 < \ldots < i_p \leqslant n} a_{i_1 \ldots i_p} dx_{i_1} \wedge \ldots \wedge dx_{i_p} \in \Sigma_p,$$

where $z_{i_1 \ldots i_p}$ is the †Grassmann coordinate of E_p. An integral element E_p^0 is called a **regular integral element** if the following two conditions are satisfied: (i) $(\Sigma_p)^0$ is a regular local equation of $I\Sigma_p$ at E_p^0, where $I\Sigma_p$ is the set of all integral elements of Σ_p; (ii) $\dim H(E_p)=$constant around E_p^0 on $I\Sigma_p$. This definition, due to E. Kähler, is different from that given by E. Cartan [4].

Here, in general, a subsheaf Φ of $\mathcal{O}(X)$ is called a **regular local equation** of $I\Phi$ at an integral point x_0 if there exists a neighborhood U of x_0 and †cross sections $\varphi_1, \ldots, \varphi_s$ of Φ on U that satisfy the following two conditions: (i) $d\varphi_1, \ldots, d\varphi_s$ are linearly independent at every x on U; (ii) a point x of U is an integral point of Φ if and only if $\varphi_1(x) = \ldots = \varphi_s(x) = 0$.

First existence theorem. Suppose that we are given a p-dimensional integral manifold M with a regular integral element $T_x(M)$ at a point x on M. Suppose further that there exists a submanifold F of X containing M such that $\dim F = n - t_{p+1}$, $\dim(T_x(F) \cap H(E_p)) = p+1$, where $E_p = T_x(M)$ and $t_{p+1} = \dim H(E_p) - p - 1$. Then around x there exists a unique integral manifold N such that $\dim N = p+1$ and $F \supset N \supset M$.

This theorem is proved by integrating a system of partial differential equations of Cauchy-Kovalevskaya type. E. Cartan [2–4] also tried to obtain an existence theorem by integrating a system of ordinary differential equations.

A chain of integral elements $E_0 \subset E_1 \subset \ldots \subset E_r$ is called a **regular chain** if each of E_p $(0 \leqslant p < r)$ is a regular integral element. For a regular chain $E_0 \subset E_1 \subset \ldots \subset E_r$, define t_{p+1} by $t_{p+1} = \dim H(E_p) - p - 1$, $0 \leqslant p < r$, and define s_p by $s_p = t_p - t_{p+1} - 1$ $(0 \leqslant p < r)$, $s_r = t_r$, where $t_0 = \dim I\Sigma_0$. Then we have $s_p \geqslant 0$ $(0 \leqslant p \leqslant r)$, $s_0 + \ldots + s_r = t_0 - r$, and we can take a local coordinate system $(x_1, \ldots, x_r, y_1, \ldots, y_m)$, $m = n - r$, around E_0 that satisfies the following four conditions:

(i) $I\Sigma_0$ is defined by $y_{t_0 - r + 1} = \ldots = y_m = 0$;

(ii) $$H(E_p) = \left\{ \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_r}, \frac{\partial}{\partial y_{s_0 + \ldots + s_{p-1} + 1}}, \ldots, \frac{\partial}{\partial y_{t_0 - r}} \right\}, \quad 0 \leqslant p < r;$$

(iii) $$E_p = \left\{ \frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_p} \right\}, \quad 1 \leqslant p \leqslant r;$$

(iv) $E_0 = (0, \ldots, 0, 0, \ldots, 0)$.

The integers $s_0, \ldots, s_r$ are called the **characters** of the regular chain $E_0 \subset \ldots \subset E_r$.

Second existence theorem. Suppose that a chain of integral elements $E_0 \subset \ldots \subset E_r$ is regular, and take a local coordinate system satisfying (i)–(iv). Consider a system of initial data

$f_1, \ldots, f_{s_0}$,

$f_{s_0+1}(x_1), \ldots, f_{s_0+s_1}(x_1)$,

$f_{s_0+s_1+1}(x_1, x_2), \ldots, f_{s_0+s_1+s_2}(x_1, x_2)$,

...

$f_{s_0+\ldots+s_{r-1}+1}(x_1, \ldots, x_r), \ldots, f_{t_0-r}(x_1, \ldots, x_r)$.

Then if their values and derivatives of the first order are sufficiently small, there exists a unique integral manifold defined by $y_\alpha = y_\alpha(x_1, \ldots, x_r)$, $y_\beta = 0$, $1 \leqslant \alpha \leqslant t_0 - r < \beta \leqslant m$, such that

$$y_\alpha(x_1, \ldots, x_p, 0, \ldots, 0) = f_\alpha(x_1, \ldots, x_p),$$

$$s_0 + \ldots + s_{p-1} < \alpha \leqslant s_0 + \ldots + s_p, \quad 0 \leqslant p \leqslant r.$$

This theorem is proved by successive application of the first existence theorem. These two theorems are called the **Cartan-Kähler existence theorems**. Σ is said to be **involutive** at an integral element E_r if there exists a regular chain $E_0 \subset \ldots \subset E_r$. An integral manifold possessing a tangent space at which Σ is involutive is called an **ordinary integral manifold** or **ordinary solution** of Σ. An integral manifold that does not possess such a tangent space is called a **singular integral manifold** or **singular solution** of Σ.

Cartan's definition of ordinary and regular integral elements is as follows: An integral point E_0^0 is an ordinary integral point if Σ_0 is a regular local equation of $I\Sigma_0$ at E_0^0. An ordinary integral point E_0^0 is a regular integral point if $\dim H(E_0)$ is constant on $I\Sigma_0$ around E_0^0. Inductively, an integral element E_p^0 is called an **ordinary integral element** if $(\Sigma_p)^0$ is a regular local equation of $I\Sigma_p$ at E_p^0 and E_p^0 contains a regular integral element E_{p-1}^0. An ordinary integral element E_p^0 is a regular integral element (in the sense of Cartan) if $\dim H(E_p)$ is constant on $I\Sigma_p$ around E_p^0. It can be proved that Σ is involutive at an integral element E_r if and only if E_r is an ordinary integral element of Σ. An integral manifold possessing a tangent space that is a regular integral element of Σ is called a **regular integral manifold** or **regular solution** of Σ. Let m_{p+1} be the minimal dimension of $H(E_p)$, where E_p varies over the set of p-dimensional ordinary integral elements, and g be an integer such that $m_p \geqslant p$ $(1 \leqslant p \leqslant g)$ and $m_{g+1} = p$. Then this integer g is called the **genus** of Σ. It is the maximal dimension of ordinary integral manifolds of Σ. However, in general, it is not the maximal dimension of integral manifolds of Σ.

D. C. Spencer and others have been trying to obtain an existence theorem in the C^∞-category analogous to that of Cartan and Kähler. (For a system of linear partial differential equations → [2, 4, 11, 13, 25, 27].)

F. Involutive Systems of Partial Differential Equations

To give a definition of an involutive system of partial differential equations, we define an involutive subspace of $\mathrm{Hom}(V, W)$, where V and W are finite-dimensional vector spaces over the real number field $\mathbf{R}$. Let A be a subspace of $\mathrm{Hom}(V, W)$. For a system of vectors $v_1, \dots, v_p$ in V, $A(v_1, \dots, v_p)$ denotes the subspace of A that annihilates $v_1, \dots, v_p$. Let g_p be the minimal dimension of $A(v_1, \dots, v_p)$ as $(v_1, \dots, v_p)$ varies, where $0 \leqslant p \leqslant r = \dim V$. A basis $(v_1, \dots, v_r)$ of V is called a generic basis if it satisfies $g_p = \dim A(v_1, \dots, v_p)$ for each p. There exists a generic basis for any A. Let $W \otimes S^2(V^*)$ be the subspace of $\mathrm{Hom}(V, \mathrm{Hom}(V, W))$ consisting of all elements ξ satisfying $\xi(u)v = \xi(v)u$ for any u and v in V. Then the prolongation pA of A is defined by $pA = \mathrm{Hom}(V, A) \cap W \otimes S^2(V^*)$. For any basis $(v_1, \dots, v_r)$ of V, we have the inequality

$$\dim pA \leqslant \sum_{p=0}^{r} \dim A(v_1, \dots, v_p).$$

The subspace A is called an **involutive subspace** of $\mathrm{Hom}(V, W)$ if $\dim pA = \sum_{p=0}^{r} g_p$. This notion of an involutive subspace was obtained by V. W. Guillemin and S. Sternberg [13].

A triple $(X, N; \pi)$ consisting of two manifolds X, N and a projection π from X onto N is called a **fibered manifold** if the †differential π_* is surjective at every point of X. Take the set of all mappings f from a domain in N to X satisfying $\pi \circ f = \text{identity}$ for a fibered manifold $(X, N; \pi)$. Then an †l-jet $j_x^l(f)$ is an equivalence class under the equivalence relation defined as follows: $j_x^l(f) = j_u^l(g)$ if and only if $x = u$, $f(x) = g(u)$, and

$$\frac{\partial^{i_1 + \dots + i_r} f}{\partial x_1^{i_1} \dots \partial x_r^{i_r}}(x) = \frac{\partial^{i_1 + \dots + i_r} g}{\partial x_1^{i_1} \dots \partial x_r^{i_r}}(u),$$

$i_1 + \dots + i_r \leqslant l$, where $(x_1, \dots, x_r)$ is a local coordinate system of N around $x = u$ (→ 105 Differentiable Manifolds X).

Denote the space of all l-jets of a fibered manifold $(X, N; \pi)$ by $J^l(X, N; \pi)$ or simply J^l. Then a subsheaf of ideals Φ in $\mathcal{O}(J^l)$ is called a **system of partial differential equations of order** l on N. A point z of J^l is called an integral point of Φ if $\varphi(z) = 0$ for all $\varphi \in \Phi$. The set of all **integral points** of Φ is denoted by $I\Phi$. Let π^l be the natural projection of J^l onto J^{l-1}. Then at a point z of J^l, we can identify $\mathrm{Ker}\, \pi_*^l$ with $\mathrm{Hom}(T_x(N), \mathrm{Ker}\, \pi_*^l)$, where $x = \pi\pi^1 \dots \pi^l z$. The principal part $C_z(\Phi)$ of Φ is defined as the subspace of $\mathrm{Ker}\, \pi_*^l$ that annihilates Φ. The **prolongation** $p\Phi$ of Φ is defined as the system of order $l+1$ on N generated by Φ and $\partial_k \Phi$, $1 \leqslant k \leqslant \dim N$, where ∂_k is the formal derivative with respect to a coordinate x_k of N:

$$(\partial_k \varphi)(j_x^{l+1}(f)) = \frac{\partial}{\partial x_k} \varphi(j_x^l(f)), \qquad \varphi \in \mathcal{O}(J^l).$$

Let w be an integral point of $p\Phi$ and z be $\pi^{l+1} w$. Then we have the identity

$$pC_z(\Phi) = C_w(p\Phi).$$

The following definition of an involutive system is due to M. Kuranishi [19]: Φ is involutive at an integral point z if the following two conditions are satisfied: (i) Φ is a regular local equation of $I\Phi$ at z; (ii) there exists a neighborhood U of z in J^l such that $(\pi^{l+1})^{-1} U \cap I(p\Phi)$ forms a fibered manifold with base $U \cap I\Phi$ and projection π^{l+1}.

A system of partial differential equations is said to be **involutive** (or **involutory**) if it has an integral point at which it is involutive. Fix a system of independent variables $(y_1, \dots, y_N)$ in X. Then a system of differential forms is said to be **involutive** (or **involutory**) if it has an integral element at which it is involutive and $dy_1 \wedge \dots \wedge dy_N \neq 0$. It can be proved that these two definitions of involutive system are equivalent [19, 25].

G. Prolongation Theorems

Cartan gave a method of prolongation by which we can obtain an involutive system from a given system with two independent variables, if it has a solution. He proposed the following problem: For any $r>2$, construct a method of prolongation by which we can obtain an involutive system from a given system with r independent variables, if it has a solution. To solve this problem, Kuranishi prolonged a given system Φ successively to $p^t\Phi$, $t=1,2,3,\ldots$, and proved the following theorem: Suppose that there exists a sequence of integral points z^t of $p^t\Phi$ with $\pi^{t+l}z^t=z^{t-l}$, $t=1,2,3,\ldots$, that satisfies the following two conditions for each t: (i) $p^t\Phi$ is a regular local equation of $I(p^t\Phi)$ at z^t; (ii) there exists a neighborhood V^t of z^t in $I(p^t\Phi)$ such that $\pi^{t+l}V^t$ contains a neighborhood of z^{t-1} in $I(p^{t-1}\Phi)$ and forms a fibered manifold $(V^t, \pi^{t+l}V^t; \pi^{t+l})$. Then $p^t\Phi$ is involutive at z^t for a sufficiently large integer t.

This prolongation theorem gives a powerful tool to the theory of †infinite Lie groups. However, if we consider a system of partial differential equations of general type, there exist examples of systems that cannot be prolonged to an involutive system by this prolongation, although they have a solution. To improve Kuranishi's prolongation theorem, M. Matsuda [22] defined the prolongation of the same order by $p_0\Phi=p\Phi\cap\mathcal{O}(J^l)$ for a system Φ of order l. This is a generalization of the classical method of completion given by Lagrange and Jacobi. Applying this prolongation successively to a given system Φ, we have $\Psi=\bigcup_{\sigma=1}^{\infty}p_0^{\sigma}\Phi$. Define the p_*-operation by $p_*=\bigcup_{\sigma=1}^{\infty}p_0^{\sigma}p$. Then applying this prolongation successively to Ψ, we have the following theorem: suppose that there exists a sequence of integral points z^t of $p_*^t\Psi$ with $\pi^{l+t}z^t=z^{t-1}$, $t=1,2,3,\ldots$, that satisfies the following two conditions for each t: (i) $p_*^t\Psi$ is a regular local equation of $I(p_*^t\Psi)$ at z^t; (ii) $\dim pC(p_*^t\Psi)$ is constant around z^t on $I(p_*^t\Psi)$. Then $p_*^t\Psi$ is involutive at z^t for a sufficiently large integer t.

To prove this theorem Matsuda applied the following theorem obtained by V. W. Guillemin, S. Sternberg, and J.-P. Serre [25, appendix]: suppose that we are given a subspace A_0 of $\mathrm{Hom}(V,W)$ and subspaces A_t of $\mathrm{Hom}(V,A_{t-1})$ satisfying $A_t\subset pA_{t-1}$, $t=1,2,3,\ldots$. Then A_t is an involutive subspace of $\mathrm{Hom}(V,A_{t-1})$ for a sufficiently large integer t. Thus Cartan's problem was solved affirmatively. To the generalized Pfaff problem these prolongation theorems give another solution, which differs from that obtained by Riquier.

H. Pfaffian Systems in the Complex Domain

Consider a linear system of Pfaffian equations

$$du_i=\sum_{k=1}^{n}\sum_{j=1}^{m}a_{ij}^k(x)u_j\,dx_k,\qquad i=1,\ldots,m,$$

where $x=(x_1,\ldots,x_n)$ is a local coordinate of a complex manifold X and a_{ij}^k are meromorphic functions on X. If we put $u={}^t(u_1,\ldots,u_m)$ and $A^k(x)=(a_{ij}^k(x))$, $k=1,\ldots,n$, the system is written as

$$du=\left(\sum_{k=1}^{n}A^k(x)\,dx_k\right)u. \tag{9}$$

System (9) is completely integrable if and only if

$$\frac{\partial A^j}{\partial x^l}-\frac{\partial A^l}{\partial x^j}=[A^l,A^j],\qquad j,l=1,\ldots,n.$$

Suppose that (9) is completely integrable. If the $A^k(x)$ are holomorphic at $x^0=(x_1^0,\ldots,x_n^0)\in X$, there exists for any $u^0\in\mathbf{C}^m$ one and only one solution of (9) that is holomorphic at x^0 and satisfies $u(x^0)=u^0$. This implies that the solution space of (9) is an m-dimensional vector space; the basis of this space is called a fundamental system of solutions. Therefore any solution is expressible as a linear combination of a fundamental system of solutions and can be continued analytically in a domain where the $A^k(x)$ are holomorphic. A subvariety of X that is the pole set of at least one of the $A^k(x)$ is called a singular locus of (9), and a point on a singular locus is called a singular point.

R. Gérard has given a definition of regular singular points and an analytic expression of a fundamental system of solutions around a regular singular point, and he studied systems of Fuchsian type [8; also 9, 30].

Let $\Omega=\sum_{k=1}^{n}A^k(x)\,dx_k$. Then the system (9) can be rewritten as

$$(d-\Omega)u=0.$$

If we consider a local coordinate (x,u) of a fiber bundle over X, the operator $d-\Omega$ induces a meromorphic linear connection ∇ over X. Starting from this point of view, P. Deligne [5] introduced several important concepts and obtained many results.

The first results for irregular singular points were obtained by Gérard and Y. Sibuya [10], and H. Majima [20] studied irregular singular points of mixed type.

The systems of partial differential equations that are satisfied by the hypergeometric functions of several variables are equivalent to linear systems of Pfaffian equations [1]. This means that such systems of partial differential equations are †holonomic systems. M. Kashiwara and T. Kawai [15] studied holonomic systems with regular singularities from the standpoint of microlocal analysis. Special types of holonomic systems were investigated by T. Terada [28] and M. Yoshida [29].

Consider a system of Pfaffian equations

$$\omega_j=0,\qquad j=1,\ldots,r, \tag{10}$$

where $\omega_j = \sum_{k=1}^{n} a_{jk}(x) dx_k$ and $x = (x_1, \ldots, x_n)$. Suppose that a_{jk} are holomorphic in a domain D of $\mathbf{C}^n$ and that $d\omega_j \wedge \omega_1 \wedge \ldots \wedge \omega_r = 0$ in D. Denote by S the zero set of $\omega_1 \wedge \ldots \wedge \omega_r = 0$. A point of S is called a singular point of (10). If the codimension of S is $\geqslant 1$, then system (10) is completely integrable in $D - S$. The following theorem was proved by B. Malgrange [21]: Let $x^0 \in S$, and suppose that the codimension of S is $\geqslant 3$ around x^0; then there exist functions $f_j, j = 1, \ldots, r$, and $g_{jk}, j, k = 1, \ldots, r$, that are holomorphic at x^0 and satisfy $\omega_j = \sum_{k=1}^{r} g_{jk} df_k$ and $\det(g_{jk}(x^0)) \neq 0$.

References

[1] P. Appell and J. Kampé de Fériet, Fonctions hypergéométriques et hypersphériques, Gauthier-Villars, 1926.
[2] E. Cartan, Sur l'intégration des systèmes d'équations aux différentielles totales, Ann. Sci. Ecole Norm. Sup., 18 (1901), 241–311.
[3] E. Cartan, Leçons sur les invariants intégraux, Hermann, 1922.
[4] E. Cartan, Les systèmes différentielles extérieures et leurs applications géométriques, Hermann, Actualités Sci. Ind., 1945.
[5] P. Deligne, Equations différentielles à points singuliers réguliers, Lecture notes in math. 163, Springer, 1970.
[6] A. R. Forsyth, Theory of differential equations, pt. I. Exact equations and Pfaff's problem, Cambridge Univ. Press, 1890.
[7] A. R. Forsyth, Theory of differential equations, pt. IV. Partial differential equations, Cambridge Univ. Press, 1906.
[8] R. Gérard, Théorie de Fuchs sur une variété analytique complexe, J. Math. Pures Appl., 47 (1968), 321–404.
[9] R. Gérard and A. H. M. Levelt, Sur les connexions à singularités régulières dans le cas de plusieurs variables, Funkcial. Ekvac., 19 (1976), 149–173.
[10] R. Gérard and Y. Sibuya, Etude de certains systèms de Pfaff avec singularités, Lecture notes in math. 712, Springer, 1979.
[11] E. Goursat, Leçons sur l'intégration des équations aux dérivées partielles du premier ordre, Hermann, second edition, 1920.
[12] E. Goursat, Leçons sur le problème de Pfaff, Hermann, 1922.
[13] V. W. Guillemin and S. Sternberg, An algebraic model of transitive differential geometry, Bull. Amer. Math. Soc., 70 (1964), 16–47.
[14] E. Kähler, Einführung in die Theorie der Systeme von Differentialgleichungen, Teubner, 1934.
[15] M. Kashiwara and T. Kawai, On holonomic systems of microdifferential equations IV, systems with regular singularities, Publ. Res. Inst. Math. Sci., 17 (1981), 813–979.
[16] M. Kuranishi, On E. Cartan's prolongation theorem of exterior differential systems, Amer. J. Math., 79 (1957), 1–47.
[17] M. Kuranishi, Lectures on exterior differential systems, Lecture notes, Tata Inst., 1962.
[18] M. Kuranishi, On the local theory of continuous infinite pseudo groups I, II, Nagoya Math. J., 15 (1959), 225–260; 19 (1961), 55–91.
[19] M. Kuranishi, Lectures on involutive systems of partial differential equations, Publ. Soc. Math. São Paulo, 1967.
[20] H. Majima, Asymptotic analysis for integrable connections with irregular singular points, Lecture notes in math. 1075, Springer, 1984.
[21] B. Malgrange, Frobenius avec singularités II, Le cas général, Inventiones Math., 39 (1976), 67–89.
[22] M. Matsuda, Cartan-Kuranishi's prolongation of differential systems combined with that of Lagrange and Jacobi, Publ. Res. Inst. Math. Sci., 3 (1967/68), 69–84.
[23] C. Riquier, Les systèmes d'équations aux dérivées partielles, Gauthier-Villars, 1910.
[24] J. F. Ritt, Differential algebra, Amer. Math. Soc. Colloq. Publ., 1950.
[25] I. M. Singer and S. Sternberg, The infinite groups of Lie and Cartan I. The transitive groups, J. Analyse Math., 15 (1965), 1–114.
[26] J. A. Schouten and W. v. d. Kulk, Pfaff's problem and its generalizations, Clarendon Press, 1949.
[27] D. C. Spencer, Overdetermined systems of linear partial differential equations, Bull. Amer. Math. Soc., 75 (1969), 179–239.
[28] T. Terada, Problème de Rieman et fonctions automorphes provenant des fonctions hypergéométriques de plusieurs variables, J. Math. Kyoto Univ., 13 (1973), 557–578.
[29] M. Yoshida, Local theory of Fuchsian systems with certain discrete monodromy groups I, II, Funkcial. Ekvac., 21 (1978), 105–137, 203–221.
[30] M. Yoshida and K. Takano, On a linear system of Pfaffian equations with regular singular point, Funkcial. Ekvac., 19 (1976), 175–189.

429 (XI.6) Transcendental Entire Functions

A. General Remarks

An **entire function** (or **integral function**) $f(z)$ is a complex-valued function of a complex variable

z that is holomorphic in the finite z-plane, $z \neq \infty$. If $f(z)$ has a pole at ∞, then $f(z)$ is a polynomial in z. A polynomial is called a **rational entire function**. If an entire function is bounded, it is constant (†Liouville's theorem). A **transcendental entire function** is an entire function that is not a polynomial, for example, $\exp z$, $\sin z$, $\cos z$. An entire function can be developed in a power series $\sum_{n=0}^{\infty} a_n z^n$ with infinite radius of convergence. If $f(z)$ is a transcendental entire function, this is actually an infinite series.

B. The Order of an Entire Function

If a transcendental entire function $f(z)$ has a zero of order m ($m \geqslant 0$) at $z=0$ and other zeros at $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ $(0<|\alpha_1| \leqslant |\alpha_2| \leqslant |\alpha_3| \leqslant \ldots \to \infty)$, multiple zeros being repeated, then $f(z)$ can be written in the form

$$f(z)=e^{g(z)} z^m \prod_{k=1}^{\infty}\left(1-\frac{z}{\alpha_k}\right) e^{g_k(z)},$$

where $g(z)$ is an entire function, $g_k(z)=(z/\alpha_k)+(1/2)(z/\alpha_k)^2+(1/3)(z/\alpha_k)^3+\ldots+(1/p_k)(z/\alpha_k)^{p_k}$, and $p_1, p_2, \ldots$ are integers with the property that $\sum_{k=1}^{\infty} |z/\alpha_k|^{p_k+1}$ converges for all z (**Weierstrass's canonical product**).

E. N. Laguerre introduced the concept of the genus of a transcendental entire function $f(z)$. Assume that there exists an integer p for which $\sum_{k=0}^{\infty} |\alpha_k|^{-(p+1)}$ converges, and take the smallest such p. Assume further that in the representation for $f(z)$ in the previous paragraph, when $p_1=p_2=\ldots=p$, the function $g(z)$ reduces to a polynomial of degree q; then $\max(p, q)$ is called the **genus** of $f(z)$. For transcendental entire functions, however, the order is more essential than the genus. The **order** ρ of a transcendental entire function $f(z)$ is defined by

$$\rho=\limsup_{r \to \infty} \log\log M(r)/\log r,$$

where $M(r)$ is the maximum value of $|f(z)|$ on $|z|=r$. By using the coefficients of $f(z)=\sum a_n z^n$, we can write

$$\rho=\limsup_{n \to \infty} n \log n/\log(1/|a_n|).$$

The entire functions of order 0, which were studied by Valiron and others, have properties similar to polynomials, and the entire functions of order less than 1/2 satisfy $\lim_{r_n \to \infty} \min_{|z|=r_n} |f(z)| = \infty$ for some increasing sequence $r_n \uparrow \infty$ (**Wiman's theorem**). Hence entire functions of order less than 1/2 cannot be bounded in any domain extending to infinity. Among the functions of order greater than 1/2 there exist functions bounded in a given angular domain $D: \alpha < \arg z < \alpha + \pi/\mu$. If $|f(z)| < \exp r^{\rho}$ $(\rho<\mu)$ and $f(z)$ is bounded on the boundary of D, then $f(z)$ is bounded in the angular domain (→ 272 Meromorphic Functions). In particular, if the order ρ of $f(z)$ is an integer p, then it is equal to the genus, and $g(z)$ reduces to a polynomial of degree $\leqslant p$ (J. Hadamard). These theorems originated in the study of the zeros of the †Riemann zeta function and constitute the beginning of the theory of entire functions.

There is some difference between the properties of functions of integral order and those of others. Generally, the point z at which $f(z) = w$ is called a **w-point** of $f(z)$. If $\{z_n\}$ consists of w-points different from the origin, the infimum $\rho_1(w)$ of k for which $\sum 1/|z_n|^k$ converges is called the **exponent of convergence** of $f-w$. If the order ρ of an entire function is integral, then $\rho_1(w)=\rho$ for each value w with one possible exception, and if ρ is not integral, then $\rho_1(w)=\rho$ for all w (É. Borel). Therefore any transcendental entire function has an infinite number of w-points for each value w except for at most one value, called an **exceptional value** of $f(z)$ (**Picard's theorem**). In particular, $f(z)$ has no exceptional values if ρ is not integral. For instance, $\sin z$ and $\cos z$ have no exceptional values, while e^z has 0 as an exceptional value. Since transcendental entire functions have no poles, ∞ can be counted as an exceptional value. Then we must change the statement in Picard's theorem to "except for at most two values." Since the theorem was obtained by E. Picard in 1879, problems of this type have been studied intensively (→ 62 Cluster Sets, 272 Meromorphic Functions).

After Picard proved the theorem by using the inverse of a †modular function, several alternative proofs were given. For instance, there is a proof using the Landau-Schottky theorem and †Bloch's theorem and one using †normal families. Picard's theorem was extended to meromorphic functions and has also been studied for analytic functions defined in more general domains. There are many fully quantitative results, too. For instance, Valiron [3] gave such results by performing some calculations on neighborhoods of points where entire functions attain their maximum absolute values.

Thereafter, the distribution of w-points in a neighborhood of an essential singularity was studied by many people, and in 1925 the Nevanlinna theory of meromorphic functions was established. The core of the theory consists of two fundamental theorems, †Nevanlinna's first and second fundamental theorems (→ 272 Meromorphic Functions). Concerning composite entire functions $F(z)=f(g(z))$, Pólya proved the following fact: The finiteness of the order of F implies that the order of f should be zero unless g is a polynomial. This gives the starting point of the factorization theory, on which several people have been working recently. Several theorems in the theory of meromorphic functions can be applied to the theory. One of the fundamental theorems is the following: Let $F(z)$ be an entire function, which admits the factorizations $F(z)=P_m(f_m(z))$ with a polynomial P_m of degree m and an entire function f_m for all integers m. Then $F(z)=A\cos\sqrt{H(z)}+B$ unless $F(z)=A\exp H(z)+B$. Here, H is a nonconstant entire function and A, B are constant, $A\neq 0$.

C. Julia Directions

Applying the theory of †normal families of holomorphic functions, G. Julia proved the existence of Julia directions as a precise form of Picard's theorem [5]. A transcendental entire function $f(z)$ has at least one direction $\arg z=\theta$ such that for any $\varepsilon>0$, $f(z)$ takes on every (finite) value with one possible exception infinitely often in the angular domain $\theta-\varepsilon<\arg z<\theta+\varepsilon$. This direction $\arg z=\theta$ is called a **Julia direction** of $f(z)$.

D. Asymptotic Values

†Asymptotic values, †asymptotic paths, etc., are defined for entire functions as for meromorphic functions. In relation to †Iversen's theorem and †Gross's theorem for inverse functions and results on †cluster sets, †ordinary singularities of inverse functions hold for entire functions in the same way as for meromorphic functions. Also, as for meromorphic functions, †transcendental singularities of inverse functions are divided into two classes, the †direct and the †indirect transcendental singularities.

The exceptional values in Picard's theorem are asymptotic values of the functions, and ∞ is an asymptotic value of any transcendental entire function. Therefore $f(z)\rightarrow\infty$ along some curve extending to infinity. Between the asymptotic paths corresponding to two distinct asymptotic values, there is always an asymptotic path with asymptotic value ∞. By †Bloch's theorem, A. Bloch showed that the †Riemann surface of the inverse function of a transcendental entire function contains a disk with arbitrarily large radius. Denjoy conjectured in 1907 that $\mu\leqslant 2\rho$, where ρ is the order of an entire function and μ is the number of distinct finite asymptotic values of the function, and L. V. Ahlfors gave the first proof (1929). This result contains Wiman's theorem. There are transcendental entire functions with $\mu=2\rho$. It was shown by W. Gross that among entire functions of infinite order there exists an entire function having every value as its asymptotic value.

References

[1] E. C. Titchmarsh, The theory of functions, Clarendon Press, second edition, 1939.
[2] R. P. Boas, Entire functions, Academic Press, 1954.
[3] G. Valiron, Lectures on the general theory of integral functions, Librairie de l'Université, Deighton, Bell and Co., 1923 (Chelsea, 1949).
[4] L. Bieberbach, Lehrbuch der Funktionentheorie II, Teubner, 1931 (Johnson Reprint Co., 1969).
[5] G. Julia, Sur quelques propriétés nouvelles des fonctions entières ou méromorphes, Ann. Sci. Ecole Norm. Sup., (3) 36 (1919), 93–125.

430 (V.11) Transcendental Numbers

A. History

A complex number α is called a **transcendental number** if α is not †algebraic over the field of rational numbers **Q**. C. Hermite showed in 1873 that e is a transcendental number. Following a similar line of thought as that taken by Hermite, C. L. F. Lindemann showed that π is also transcendental (1882). Among the 23 problems posed by D. Hilbert in 1900 (→ 196 Hilbert), the seventh was the problem of establishing the transcendence of certain numbers (e.g., $2^{\sqrt{2}}$). This stimulated fruitful investigations by A. O. Gel'fond, T. Schneider, C. L. Siegel, and others. The theory of transcendental numbers is, however, far from complete. There is no general criterion that can be utilized to characterize transcendental numbers. For example, neither the transcendence nor even the irrationality of the †Euler constant $C=\lim_{n\rightarrow\infty}(1+1/2+\ldots+1/n-\log n)$ has been established. A survey of the development of the theory of transcendental numbers can be found in [18], in which an extensive list of relevant publications up to 1966 is given.

B. Construction of Transcendental Numbers

Let $\bar{\mathbf{Q}}$ be the field of †algebraic numbers. Suppose that α is an element of $\bar{\mathbf{Q}}$ that satisfies the irreducible equation $f(x)=a_0x^n+a_1x^{n-1}+\ldots+a_n=0$, where the a_i are rational integers, $a_0\neq 0$, and $a_0,a_1,\ldots,a_n$ have no common factors. Then we define $H(\alpha)$ to be the maximum of $|a_i|$ $(i=0, \ldots, n)$ and call it the **height** of α. J. Liouville proved the following theorem (1844): Let ξ be a real number ($\xi \notin \mathbf{Q}$). If $\inf\{q^n|\xi - p/q| \mid p/q \in \mathbf{Q}\} = 0$ for any positive integer n, then ξ is transcendental.

Transcendental numbers having this property are called **Liouville numbers**. Examples are: (i) $\xi = \sum_{\nu=1}^{\infty} g^{-\nu!}$, where g is an integer not smaller than 2. (ii) Suppose that we are given a sequence $\{n_k\}$ of positive integers such that $n_k \to \infty$ $(k \to \infty)$. Let ξ be the real number expressed as an †infinite simple continued fraction $b_0 + 1/b_1 \dot{+} 1/b_2 \dot{+} \ldots$. Let B_l be the denominator of the lth †convergent of the continued fraction. If $b_{n_k+1} \geqslant B_{n_k}^{n_k-2}$ for $k \geqslant 1$, then ξ is a Liouville number.

On the other hand, K. Mahler [8, 9] proved the existence of transcendental numbers that are not Liouville numbers. For example, he showed that if $f(x)$ is a nonconstant integral polynomial function mapping the set of positive integers into itself, then a number ξ expressed, e.g., in the decimal system as $0y_1y_2y_3\ldots$ is such a number if we put $y_n = f(n)$, $n = 1, 2, 3, \ldots$. (In particular, from $f(x) = x$ we get the non-Liouville transcendental number $\xi = 0.123456789101112\ldots$.) Mahler proved this result by using †Roth's theorem (1955) (→ 182 Geometry of Numbers). Both Liouville and Mahler utilized the theory of †Diophantine approximation to construct transcendental numbers.

On the other hand, Schneider [10–12] and Siegel [3] constructed transcendental numbers using certain functions. Examples are: $\exp\alpha$ ($\alpha \in \bar{\mathbf{Q}}$, $\alpha \neq 0$); α^β ($\alpha \in \bar{\mathbf{Q}}$, $\alpha \neq 0, 1$; $\beta \in \bar{\mathbf{Q}} - \mathbf{Q}$); $J(\tau)$, where J is the †modular function and τ is an algebraic number that is not contained in any imaginary quadratic number field; $\wp(2\pi i/\alpha)$, where $\wp$ is the Weierstrass $\wp$-function, $\alpha \in \bar{\mathbf{Q}}$, and $\alpha \neq 0$; and $B(p, q)$, where B is the †Beta function and $p, q \in \mathbf{Q} - \mathbf{Z}$.

Since $e = \exp 1$ and $1 = \exp 2\pi i$, the transcendence of e and π is directly implied by the transcendence of $\exp\alpha$ ($\alpha \in \bar{\mathbf{Q}}$, $\alpha \neq 0$).

C. Classification of Transcendental Numbers

(1) Mahler's classification: Given a complex number ξ and positive integers n and H, we consider the following:

$$w_n(H, \xi) = \min\left\{\left|\sum_{\nu=0}^{n} a_\nu \xi^\nu\right| \;\middle|\; a_\nu \in \mathbf{Z},\ |a_\nu| \leqslant H,\ \sum_{\nu=0}^{n} a_\nu \xi^\nu \neq 0\right\},$$

$$w_n(\xi) = w_n = \limsup_{H \to \infty}(-\log w_n(H, \xi)/\log H),$$

$$w(\xi) = w = \limsup_{n \to \infty} w_n(\xi)/n,$$

and let μ = the first number n for which w_n is ∞. Then we have the following four cases: (i) $w = 0$, $\mu = \infty$; (ii) $0 < w < \infty$, $\mu = \infty$; (iii) $w = \mu = \infty$; (iv) $w = \infty$, $\mu < \infty$, corresponding to which we call ξ an **A-number, S-number, T-number**, or **U-number**. The set of A-numbers is denoted by **A**, and similarly we have the classes **S**, **T**, and **U**. It is known that $\mathbf{A} = \bar{\mathbf{Q}}$. If two numbers ξ and η are †algebraically dependent over $\mathbf{Q}$, then they belong to the same class. If ξ belongs to **S**, the quantity $\theta(\xi) = \sup\{w_n(\xi)/n \mid n = 1, 2, \ldots\}$ is called the **type** of ξ (in the sense of Mahler). Mahler conjectured that almost all transcendental numbers (except a set of Lebesgue measure zero) are S-numbers of the type 1 or 1/2 according as they belong to **R** or not. Various results were obtained concerning this conjecture (W. J. LeVeque, J. F. Koksma, B. Volkmann) until it was proved by V. G. Sprindzhuk in 1965 [14, 15]. The existence of T-numbers was proved by W. M. Schmidt (1968) [16]. All Liouville numbers are U-numbers [7]. On the other hand, $\log\alpha$ ($\alpha \in \mathbf{Q}$, $\alpha > 0$, $\alpha \neq 1$) and π are transcendental numbers that do not belong to **U**.

(2) Koksma's classification: For a given transcendental number ξ and positive numbers n and H, we consider the following:

$$w_n^*(H, \xi) = \min\{|\xi - \alpha| \mid \alpha \in \bar{\mathbf{Q}},\ H(\alpha) \leqslant H,\ [\mathbf{Q}(\alpha):\mathbf{Q}] \leqslant n\},$$

$$w_n^*(\xi) = w_n^* = \limsup_{H \to \infty}(-\log(H w_n^*(H, \xi))/\log H),$$

$$w^*(\xi) = w^* = \limsup_{n \to \infty} w_n^*(\xi)/n,$$

and let μ^* = the first number n for which w_n^* is ∞. Then we have the following three cases: (i) $w^* < \infty$, $\mu^* = \infty$; (ii) $w^* = \mu^* = \infty$; (iii) $w^* = \infty$, $\mu^* < \infty$. We call ξ an **S*-number, T*-number**, or **U*-number** according as (i), (ii), or (iii) holds and denote the set of S*-numbers by **S***, etc. If ξ belongs to **S***, we call $\theta^*(\xi) = \sup\{w_n^*(\xi)/n \mid n = 1, 2, \ldots\}$ the **type** of ξ (in the sense of Koksma). It can be shown that $\mathbf{S} = \mathbf{S}^*$, $\mathbf{T} = \mathbf{T}^*$, and $\mathbf{U} = \mathbf{U}^*$, and that if $\xi \in \mathbf{S}$, then $\theta^*(\xi) \leqslant \theta(\xi) \leqslant \theta^*(\xi) + 1$.

D. Algebraic Independence

Concerning the algebraic relations of transcendental numbers, we have the following three principal theorems:

(1) Let $\alpha_1, \ldots, \alpha_m$ be elements of $\bar{\mathbf{Q}}$ that are linearly independent over $\mathbf{Q}$. Then $\exp\alpha_1, \ldots, \exp\alpha_m$ are transcendental and algebraically independent over $\bar{\mathbf{Q}}$ (**Lindemann-Weierstrass theorem**).

(2) Let $J_0(x)$ be the †Bessel function and α a nonzero algebraic number. Then $J_0(\alpha)$ and $J_0'(\alpha)$ are transcendental and algebraically independent over $\mathbf{Q}$ (Siegel).

(3) Let $\alpha_1, \ldots, \alpha_n$ be nonzero elements of $\bar{\mathbf{Q}}$ such that $\log\alpha_1, \ldots, \log\alpha_n$ are linearly independent over $\mathbf{Q}$. Then $1, \log\alpha_1, \ldots, \log\alpha_n$ are linearly independent over $\bar{\mathbf{Q}}$ (A. Baker).

Besides these theorems, various related results have been obtained by A. B. Shidlovskiĭ, Gel'fond, N. I. Fel'dman, and others. A quantitative extension of theorem (3), also by Baker, will be discussed later.

First we give more detailed descriptions of theorems (1) and (2). Let $\alpha_1, \ldots, \alpha_m$ be as in theorem (1), $s=[\mathbf{Q}(\alpha_1, \ldots, \alpha_m):\mathbf{Q}]$, $P(X_1, \ldots, X_m)$ be an arbitrary polynomial in $\bar{\mathbf{Q}}[X_1, \ldots, X_m]$ of degree n, and $H(P)$ be the maximum of the absolute values of the coefficients of the polynomial P. Then there exists a positive number C determined only by the numbers $\alpha_1, \ldots, \alpha_m$ and $n(=\deg P)$ such that

$$|P(e^{\alpha_1}, \ldots, e^{\alpha_m})| > CH(P)^{-2s\left(2\left(\frac{2smn+m+n}{m}\right)-1\right)}.$$

In particular, if α is a nonzero algebraic number, then $\exp\alpha$ belongs to $\mathbf{S}$ and $\theta(\exp\alpha) \leqslant 8s^2+6s$.

(2′) Let α be a nonzero algebraic number, $s=[\mathbf{Q}(\alpha):\mathbf{Q}]$, $P\in\mathbf{Q}[X_1, X_2]$, $\deg P=n$. Then there exists a positive number C determined only by α and n such that $|P(J_0(\alpha), J_0'(\alpha))| > CH(P)^{-82s^3n^3}$.

Theorems (1) and (2) are actually special cases of a theorem obtained by Siegel. To state this theorem, the following terminology is used: An entire function $f(z)=\sum_{n=0}^{\infty} C_n \cdot z^n/n!$ is called an E-**function** defined over an †algebraic number field K of finite degree if the following three conditions are satisfied: (i) $C_n\in K$ $(n=0,1,2,\ldots)$. (ii) For any positive number ε, $C_n=O(n^{\varepsilon n})$. (iii) Let q_n be the least positive integer such that $C_k q_n$ belongs to the ring $\mathfrak{O}$ of algebraic integers in K $(0\leqslant n, 0\leqslant k\leqslant n)$. Then for an arbitrary positive number ε, $q_n=O(n^{\varepsilon n})$.

A system $\{f_1(z), \ldots, f_m(z)\}$ of E-functions defined over K is said to be **normal** if it satisfies the following two conditions: (i) None of the functions $f_i(z)$ is identically zero. (ii) If the functions $w_k=f_k(z)$ $(k=1, \ldots, m)$ satisfy a system of †homogeneous linear differential equations of the first order, then $w_k'=\sum_{l=1}^m Q_{kl}(z)w_l$, where the $Q_{kl}(z)$ are rational functions of z, with coefficients in the ring $\mathfrak{O}$. The matrix (Q_{kl}) can be decomposed by rearranging the order of the indices k, l if necessary into the form

$$\begin{pmatrix} W_1 & & 0 \\ & \ddots & \\ 0 & & W_r \end{pmatrix},$$

where

$$W_t=\begin{bmatrix} Q_{11,t} & \cdots & Q_{1m_t,t} \\ \cdots & \cdots & \cdots \\ Q_{m_t1,t} & \cdots & Q_{m_tm_t,t} \end{bmatrix}, \quad 1\leqslant t\leqslant r, \quad \sum_{t=1}^n m_t=m.$$

The decomposition is unique if we choose r as large as possible, in which case we call $W_1, \ldots, W_r$ the primitive parts of (Q_{kl}). The requirement is that the primitive parts W_i are independent in the following sense: If there are numbers $C_{st}\in K$ and polynomial functions $P_{kt}(z)\in K[z]$ such that

$$\sum_{t=1}^r (C_{1t} \ldots C_{m_tt})W_t \begin{bmatrix} P_{1t}(z) \\ \vdots \\ P_{m_tt}(z) \end{bmatrix}=0,$$

then $C_{st}=0$, $P_{kl}(z)=0$.

Let N be a positive integer. A normal system $\{f_1(z), \ldots, f_m(z)\}$ of E-functions is said to be of degree N if the system $\{F_{n_1, \ldots, n_m}(z)= f_1(z)^{n_1} \ldots f_m(z)^{n_m} \mid n_i\geqslant 0, \sum_{i=1}^m n_i\leqslant N\}$ is also a normal system of E-functions. Then the theorem obtained by Siegel [4] is: Let N be an arbitrary positive integer and $\{f_1(z), \ldots, f_m(z)\}$ be a normal system of E-functions of degree N defined over an algebraic number field of finite degree K satisfying the system of differential equations $f_k'(z)=\sum_{l=1}^m Q_{kl}(z)f_l(z)$, where $Q_{kl}(z)\in\mathfrak{O}(z)$, $1\leqslant k\leqslant m$. If α is a nonzero algebraic number that is not a †pole of any one of the functions $Q_{kl}(z)$, then $f_1(\alpha), \ldots, f_m(\alpha)$ are transcendental numbers that are algebraically independent over the field $\bar{\mathbf{Q}}$.

Theorem (3) at the beginning of this section implies, for example, the following: (i) If $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_n$ all belong to $\bar{\mathbf{Q}}$ and $\gamma=\alpha_1\log\beta_1+\ldots+\alpha_n\log\beta_n\neq 0$, then γ is transcendental. (ii) If $\alpha_1, \ldots, \alpha_n, \beta_0, \beta_1, \ldots, \beta_n$ are nonzero algebraic numbers, then $e^{\beta_0}\alpha_1^{\beta_1}\ldots\alpha_n^{\beta_n}$ is transcendental. (iii) If $\alpha_1, \ldots, \alpha_n$ are algebraic numbers other than 0 and 1, and $\beta_1, \ldots, \beta_n$ also belong to $\bar{\mathbf{Q}}$, with $1, \beta_1, \ldots, \beta_n$ linearly independent over $\mathbf{Q}$, then $\alpha_1^{\beta_1}\ldots\alpha_n^{\beta_n}$ is transcendental.

Baker [17] also obtained a quantitative extension of theorem (3): Suppose that we are given integers $A\geqslant 4$, $d\geqslant 4$ and nonzero algebraic numbers $\alpha_1, \ldots, \alpha_n$ $(n\geqslant 2)$ whose heights and degrees do not exceed A and d, respectively. Suppose further that $0<\delta\leqslant 1$, and let $\log\alpha_1, \ldots, \log\alpha_n$ be the principal values of the logarithms. If there exist rational integers $b_1, \ldots, b_n$ with absolute value at most H such that

$$0<|b_1\log\alpha_1+\ldots+b_n\log\alpha_n|<e^{-\delta H},$$

then

$$H<(4^{n^2}\delta^{-1}d^{2n}\log A)^{(2n+1)^2}.$$

This theorem has extensive applications in various problems of number theory, including a wide class of †Diophantine problems [19].

A number of new, interesting results on the algebraic independence of values of exponential functions, elliptic functions, and some other special functions have been obtained recently by D. Masser, G. V. Chudnovskiĭ, M. Waldschmidt, and other writers. In particular, Chudnovskiĭ (1975) obtained the remarkable result that Γ(1/3) and Γ(1/4) are transcendental numbers. See [20–24].

References

[1] A. O. Gel'fond, Transcendental and algebraic numbers, Dover, 1960. (Original in Russian, 1952.)
[2] T. Schneider, Einführung in die transzendenten Zahlen, Springer, 1957.
[3] C. L. Siegel, Über einige Anwendungen diophantischer Approximationen, Abh. Preuss. Akad. Wiss., 1929 (Gesammelte Abhandlungen, Springer, 1966, vol. 1, 209–266).
[4] C. L. Siegel, Transcendental numbers, Ann. Math. Studies, Princeton Univ. Press, 1949.
[5] N. I. Fel'dman, Approximation of certain transcendental numbers I, Amer. Math. Soc. Transl., 59 (1966), 224–245. (Original in Russian, 1951.)
[6] W. J. LeVeque, Note on S-numbers, Proc. Amer. Math. Soc., 4 (1953), 189–190.
[7] W. J. LeVeque, On Mahler's U-numbers, J. London Math. Soc., 28 (1953), 220–229.
[8] K. Mahler, Arithmetische Eigenschaften einer Klasse von Dezimalbrüchen, Proc. Acad. Amsterdam, 40 (1937), 421–428.
[9] K. Mahler, Über die Dezimalbruchentwicklung gewisser Irrationalzahlen, Mathematica B, Zutphen, 6 (1937), 22–36.
[10] T. Schneider, Transzendenzuntersuchungen periodischer Funktionen I, J. Reine Angew. Math., 172 (1935), 65–69.
[11] T. Schneider, Arithmetische Untersuchungen elliptischer Integrale, Math. Ann., 113 (1936), 1–13.
[12] T. Schneider, Zur Theorie der Abelschen Funktionen und Integrale, J. Reine Angew. Math., 183 (1941), 110–128.
[13] S. Lang, Introduction to transcendental numbers, Addison-Wesley, 1966.
[14] V. G. Sprindzhuk, A proof of Mahler's conjecture on the measure of the set of S-numbers, Amer. Math. Soc. Transl., 51 (1966), 215–272. (Original in Russian, 1965.)
[15] V. G. Sprindzhuk, Mahler's problem in metric number theory, Amer. Math. Soc. Transl. Math. Monographs, 1969. (Original in Russian, 1967.)
[16] W. M. Schmidt, T-numbers do exist, Rend. Convegno di Teoria dei Numeri, Roma, 1968.
[17] A. Baker, Linear forms in the logarithms of algebraic numbers I, II, III, IV, Mathematika, 13 (1966), 204–216; 14 (1967), 102–107; 14 (1967), 220–228; 15 (1968), 204–216.
[18] N. I. Fel'dman and A. B. Shidlovskiĭ, The development and present state of the theory of transcendental numbers, Russian Math. Surveys, 22 (1967), 1–79. (Original in Russian, 1967.)
[19] A. Baker, Transcendental number theory, Cambridge Univ. Press, 1975.
[20] D. Masser, Elliptic functions and transcendence, Lecture notes in math. 437, Springer, 1975.
[21] G. V. Chudnovskiĭ, Algebraic independence of values of exponential and elliptic functions, Proc. Int. Congr. Math., Helsinki, 1978.
[22] A. Baker and D. Masser (eds.), Transcendence theory: Advances and applications, Academic Press, 1977.
[23] M. Waldschmidt, Transcendence methods, Queen's papers in pure and appl. math. 52, Queen's Univ., 1979.
[24] M. Waldschmidt, Nombres transcendants et groups algébriques, Astérisque, 69–70 (1980).

431 (IX.19) Transformation Groups

A. Topological Transformation Groups

Let G be a group, M a set, and f a mapping from $G \times M$ into M. Put $f(g,x)=g(x)$ ($g \in G$, $x \in M$). Then the group G is said to be a **transformation group** of the set M if the following two conditions are satisfied: (i) $e(x)=x$ ($x \in M$), where e is the identity element of G; and (ii) $(gh)(x)=g(h(x))$ ($x \in M$) for any $g, h \in G$. In this case the mapping $x \to g(x)$ is a one-to-one mapping of M onto itself.

Let G be a transformation group of M. If G is a topological group, M a topological space, and the mapping $(g,x) \to g(x)$ a continuous mapping from $G \times M$ into M, then G is called a **topological transformation group** of M. In this case $x \to g(x)$ is a homeomorphism of M onto itself. The mapping $(g,x) \to g(x)$ is called an **action** of G on M. The space M, together with a given action of G, is called a G-**space**.

For a point x of M, the set $G(x)=\{g(x) \mid g \in G\}$ is called the **orbit** of G passing through the point x. Defining as equivalent two points x and y of M belonging to the same orbit, we get an equivalence relation in M. The quotient space of M by this equivalence relation, denoted by M/G, is called the **orbit space** of G-space M.

If $G(x)=\{x\}$, then x is called a **fixed point**. The set of all fixed points is denoted by M^G. For a point x of M, the set $G_x=\{g \in G \mid g(x)=$

$x\}$ is a subgroup of G called the **isotropy subgroup (stabilizer, stability subgroup)** of G at the point x. A conjugacy class of the subgroup G_x is called an **isotropy type** of the transformation group G on M.

The group G is said to act **nontrivially** (resp. **trivially**) on M if $M \neq M^G$ (resp. $M = M^G$). The group G is said to act **freely** on M if the isotropy subgroup G_x consists only of the identity element for any point x of M.

The group G is said to act **transitively** on M if for any two points x and y of M, there exists an element $g \in G$ such that $g(x) = y$.

Let N be the set of all elements $g \in G$ such that $g(x) = x$ for all points x of M. Then N is a normal subgroup of G. If N consists only of the identity element e, we say that G acts **effectively** on M, and if N is a discrete subgroup of G, we say that G acts **almost effectively** on M. When $N \neq \{e\}$, the quotient topological group G/N acts effectively on M in a natural fashion.

An **equivariant mapping (equivariant map)** (or a **G-mapping, G-map**) $h: X \to Y$ between G-spaces is a continuous mapping which commutes with the group actions, that is, $h(g(x)) = g(h(x))$ for all $g \in G$ and $x \in X$. An equivariant mapping which is also a homeomorphism is called an equivalence of G-spaces.

For a G-space M, an equivalence class of the G-spaces $G(x)$, $x \in M$, is called an **orbit type** of the G-space M.

B. Cohomological Properties

We consider only †paracompact G-spaces and †Čech cohomology theory in this section. We shall say that a topological space X is **finitistic** if every open covering has a finite-dimensional refinement. The following theorems are useful [1–3].

(1) If G is finite, X a finitistic paracompact G-space, and K a field of characteristic zero or prime to the order of G, then the induced homomorphism $\pi^*: H^*(X/G; K) \to H^*(X; K)^G$ is an isomorphism. Here, π is a natural projection of X onto X/G. The group G acts naturally on $H^*(X; K)$, and $H^*(X; K)^G$ denotes the fixed-point set of this G-action.

(2) Let X be a finitistic G-space and G cyclic of prime order p. Then, with coefficients in $\mathbf{Z}/p\mathbf{Z}$, we have

(a) for each n $\quad \sum_{i=n}^{\infty} \operatorname{rank} H^i(X^G) \leqslant \sum_{i=n}^{\infty} \operatorname{rank} H^i(X)$,

(b) $\quad \chi(X) + (p-1)\chi(X^G) = p\chi(X/G)$.

Here the †Euler-Poincaré characteristics $\chi(\ )$ are defined in terms of mod p cohomology.

(3) **Smith's theorem**: If G is a p-group (p prime) and if x is a finitistic G-space whose mod p cohomology is isomorphic to the n-sphere, then the mod p cohomology of the fixed-point set X^G is isomorphic to that of the r-sphere for some $-1 \leqslant r \leqslant n$, where (-1)-sphere means the empty set.

(4) Let T^k denote the k-dimensional toral group. Let X be a T^k-space whose rational cohomology is isomorphic to the n-sphere, and assume that there are only a finite number of orbit types and that the orbit spaces of all subtori are finitistic. Let H be a subtorus of T^k. Then by the above theorem the rational cohomology of X^H is isomorphic to that of the $r(H)$-sphere for some $-1 \leqslant r(H) \leqslant n$. Assume further that there is no fixed point of the T^k-action. Then, with H ranging over all subtori of dimension $k-1$, we have

$$n+1 = \sum_{H} (r(H)+1).$$

C. Differentiable Transformation Groups

Suppose that the group G is a transformation group of a †differentiable manifold M, G is a †Lie group, and the mapping $(g, x) \to g(x)$ of $G \times M$ into M is a differentiable mapping. Then G is called a **differentiable transformation group** (or **Lie transformation group**) of M, and M is called a differentiable **G-manifold.**

The following are basic facts about compact differentiable transformation groups [3, 4]:

(5) **Differentiable slice theorem**: Let G be a compact Lie group acting differentiably on a manifold M. Then, by averaging an arbitrary †Riemannian metric on M, we may have a G-invariant Riemannian metric on M. That is, the mapping $x \to g(x)$ is an †isometry of this Riemannian manifold M for each $g \in G$. For each point $x \in M$, the orbit $G(x)$ through x is a compact submanifold of M and the mapping $g \mapsto g(x)$ defines a G-equivariant diffeomorphism $G/G_x \cong G(x)$, where G/G_x is the left quotient space by the isotropy subgroup G_x. G_x acts orthogonally on the †tangent space T_xM at x (resp. the †normal vector space N_x of the orbit $G(x)$); we call it the **isotropy representation** (resp. **slice representation**) of G_x at x. Let E be the †normal vector bundle of the orbit $G(x)$. Since G acts naturally on E as a bundle mapping, the bundle E is equivalent to the bundle $(G \times N_x)/G_x$ over G/G_x as a †G-vector bundle, where G_x acts on N_x by means of the slice representation and G_x acts on G by the right translation. We can choose a small positive real number ε such that the †exponential mapping gives an equivariant †diffeomorphism of the ε-disk bundle of E onto an invariant †tubular neighborhood of $G(x)$.

(6) Assume that a compact Lie group G acts differentiably on M with the orbit space $M^* = M/G$ connected. Then there exists a maximum

orbit type G/H for G on M (i.e., H is an isotropy subgroup and H is conjugate to a subgroup of each isotropy group). The union $M_{(H)}$ of the orbits of type G/H is open and dense in M, and its image $M^*_{(H)}$ in M^* is connected.

The maximum orbit type for orbits in M guaranteed by the above theorem is called the **principal orbit type**, and orbits of this type are called **principal orbits**. The corresponding isotropy groups are called **principal isotropy groups**. Let P be a principal orbit and Q any orbit. If $\dim P > \dim Q$, then Q is called a **singular orbit**. If $\dim P = \dim Q$ but P and Q are not equivalent, then Q is called an **exceptional orbit**.

(7) Let G be a compact Lie group and M a compact G-manifold. Then the orbit types are finite in number.

By applying (5) and (6) we have that an isotropy group is principal if and only if its slice representation is trivial.

The situation is quite different in the case of noncompact transformation groups. For example, there exists an analytic action of $G = SL(4, \mathbf{R})$ on an analytic manifold M such that each orbit of G on M is closed and of codimension one and such that, for $x, y \in M$, G_x is not isomorphic to G_y unless x and y lie on the same G-orbit [5].

D. Compact Differentiable Transformation Groups

Many powerful techniques in †differential topology have been applied to the study of differentiable transformation groups. For example, using the techniques of †surgery, we can show that there are infinitely many free differentiable circle actions on †homotopy $(2n+1)$-spheres $(n \geqslant 3)$ that are differentiably inequivalent and distinguished by the rational †Pontryagin classes of the orbit manifolds (W. C. Hsiang [6]). Also, using †Brieskorn varieties, we can construct many examples of differentiable transformation groups on homotopy spheres [3, 4, 7]. Differentiable actions of compact connected Lie groups on homology spheres have been studied systematically (Hsiang and W. Y. Hsiang [4]).

The Atiyah-Singer †index theorem has many applications in the study of transformation groups. The following are notable applications:

(8) Let M be a compact connected †oriented differentiable manifold of dimension $4k$ with a †spin-structure. If a compact connected Lie group G acts differentiably and nontrivially on M, then the $\hat{A}$-genus $\langle \hat{\mathscr{A}}(M), [M] \rangle$ of M vanishes (where $\hat{\mathscr{A}}(M)$ denotes the †$\hat{A}$-characteristic class of M) (M. F. Atiyah and F. Hirzebruch [8], K. Kawakubo [9]). For further developments, see A. Hattori [10].

(9) Let M be a closed oriented manifold with a differentiable circle action. Then each connected component F_k of the fixed point set can be oriented canonically, and we have

$$I(M) = \sum_k I(F_k),$$

where $I(\ )$ denotes the †Thom-Hirzebruch index [8, 9].

Let G be a compact Lie group and $G \to EG \to BG$ the †universal G-bundle. Then the †singular cohomology $H^*(EG \times_G X)$ is called **equivariant cohomology** for a G-space X and is an $H^*(BG)$-module. Let $G = U(1)$, M a differentiable $U(1)$-manifold, $F = M^G$, and $i: F \to M$ the inclusion mapping. Then the †localization of the induced homomorphism

$$S^{-1} i^*: S^{-1} H^*(EG \times_G M) \to S^{-1} H^*(BG \times F)$$

is an isomorphism, where S^{-1} denotes the localization with respect to the multiplicative set $S = \{at^k\}$ with a, k ranging over all positive integers and t the generator of $H^2(BG)$. Theorems (8) and (9) can be proved by the above localization isomorphism.

Let M be a differentiable manifold. The upper bound $N(M)$ of the dimension of all the compact Lie groups that acts effectively and differentiably on M is called the **degree of symmetry** of M. It measures, in some crude sense, the symmetry of the differentiable manifold M. The number $N(M)$ depends heavily on the differentiable structure. For example, $N(S^m) = m(m+1)/2$ for the standard m-sphere, but $N(\Sigma^m) < (m+1)^2/16 + 5$ for a †homotopy m-sphere $(m \geqslant 300)$ that does not bound a †π-manifold [11]. Also, $N(P_n(\mathbf{C})) = n(n+2)$ for the complex projective n-space $P_n(\mathbf{C})$, but $N(hP_n(C)) < (n+1)(n+2)/2$ for any homotopy complex projective n-space $hP_n(\mathbf{C})$ $(n \geqslant 13)$ other than $P_n(\mathbf{C})$ (T. Watabe [12]).

Let X be a differentiable closed manifold and $h: X \to P_n(\mathbf{C})$ be an orientation-preserving †homotopy equivalence. There is a conjecture about the total $\hat{A}$-classes that states: If X admits a nontrivial differentiable circle action, then $\hat{\mathscr{A}}(X) = h^* \hat{\mathscr{A}}(P_n(\mathbf{C}))$ (T. Petrie [13]). It is known that if the action is free outside the fixed-point set, then the conjecture is true (T. Yoshida [14]).

E. Equivariant Bordism

Fix a compact Lie group G; a compact **oriented G-manifold** (ψ, M) consists of a compact †oriented differentiable manifold M and an orientation-preserving differentiable G-action $\psi: G \times M \to M$ on M.

Given families $F \supset F'$ of subgroups of G, a compact oriented G-manifold (ψ, M) is (F, F')-**free** if the following conditions are satisfied: (i) if $x \in M$, then the isotropy group G_x is conjugate to a member of F; (ii) if $x \in \partial M$, then G_x is conjugate to a member of F'.

If F' is the empty family, then necessarily ∂M is empty and M is closed. In this case we say that (ψ, M) is F-**free**.

Given (ψ, M), define $-(\psi, M) = (\psi, -M)$ with the structure precisely the same as (ψ, M) except for †orientation. Also define $\partial(\psi, M) = (\psi, \partial M)$. Note that if (ψ, M) is (F, F')-free, then $(\psi, \partial M)$ is F'-free. Define (ψ, M) and (ψ', M') to be isomorphic if there exists an equivariant orientation-preserving diffeomorphism of M onto M'.

An (F, F')-free compact oriented n-dimensional G-manifold (ψ, M) is said to **bord** if there exists an (F, F)-free compact oriented $(n+1)$-dimensional G-manifold (Φ, W) together with a regularly embedded compact n-dimensional manifold M_1 in ∂W with M_1 invariant under the G-action Φ such that (Φ, M_1) is isomorphic to (ψ, M) and G_x is conjugate to a member of F' for $x \in \partial W - M_1$. Also, M_1 is required to have its orientation induced by that of W.

We say that (ψ_1, M_1) is **bordant** to (ψ_2, M_2) if the disjoint union $(\psi_1, M_1) + (\psi_2, -M_2)$ bords. Bordism is an equivalence relation on the class of (F, F')-free compact oriented n-dimensional G-manifolds. The bordism classes constitute an Abelian group $\mathbf{O}_n^G(F, F')$ under the operation of disjoint union. If F' is empty, denote the above group by $\mathbf{O}_n^G(F)$. The direct sum

$$\mathbf{O}_*^G(F, F') = \bigoplus_n \mathbf{O}_n^G(F, F')$$

is naturally an Ω-module, where Ω is the †oriented cobordism ring. If F consists of all subgroups of G, then $\mathbf{O}_*^G(F)$ is denoted by $\mathbf{O}_*^G$.

Suppose now that $F \supset F'$ are fixed families of subgroups of G. Every F'-free G-manifold is also F-free, and so this inclusion induces a homomorphism $\alpha: \mathbf{O}_n^G(F') \to \mathbf{O}_n^G(F)$. Similarly every F-free G-manifold is also (F, F')-free, inducing a homomorphism $\beta: \mathbf{O}_n^G(F) \to \mathbf{O}_n^G(F, F')$. Finally, there is a homomorphism $\partial: \mathbf{O}_n^G(F, F') \to \mathbf{O}_{n-1}^G(F')$ given by $\partial(\psi, M) = (\psi, \partial M)$. Then the following sequence is exact [15]:

$$\ldots \xrightarrow{\partial} \mathbf{O}_n^G(F') \xrightarrow{\alpha} \mathbf{O}_n^G(F) \xrightarrow{\beta} \mathbf{O}_n^G(F, F') \xrightarrow{\partial} \mathbf{O}_{n-1}^G(F') \xrightarrow{\alpha} \ldots.$$

A weakly almost complex compact G-manifold (ψ, M) consists of a †weakly almost complex compact manifold M and a differentiable G-action $\psi: G \times M \to M$ that preserves the weakly almost complex structure on M. $\mathbf{U}_*^G(F, F')$, $\mathbf{U}_*^G$ are defined similarly, and they are $\mathbf{U}_*$-modules, where $\mathbf{U}_*$ is the †complex cobordism ring of compact weakly almost complex manifolds.

To study $\mathbf{O}_*^G$ and $\mathbf{U}_*^G$, (co)bordism theory is introduced (P. E. Conner and E. E. Floyd [16]), which is one of the †generalized (co)-homology theories. Miscellaneous results are known, in particular, for G a cyclic group of prime period. By means of the equivariant †Thom spectrum, equivariant cobordism theory can be developed (T. tom Dieck [17]); this is a multiplicative generalized cohomology theory with Thom classes (→ 114 Differential Topology; also → 201 Homology Theory, 56 Characteristic Classes).

F. Equivariant Homotopy

Let G be a compact Lie group. On the category of closed G-manifolds, we say that two objects M, N are χ-**equivalent** if $\chi(M^H) = \chi(N^H)$ for all closed subgroups H of G, where $\chi(\ )$ is the †Euler-Poincaré characteristic. On the set of equivalence classes $\mathbf{A}(G)$, a ring structure is imposed by disjoint union and the Cartesian product. We call $\mathbf{A}(G)$ the **Burnside ring** of G. If G is finite, $\mathbf{A}(G)$ is naturally isomorphic to the classical Burnside ring of G [18].

Denote by $S(V)$ the unit sphere of an orthogonal G-representation space V. Let V, W be orthogonal G-representation spaces. The equivariant stable homotopy group $[[S(V), S(W)]]$, which is defined as the direct limit of the equivariant homotopy sets $[S(V+U), S(W+U)]_G$ taken over orthogonal G-representation spaces U and suspensions, is denoted by ω_α for $\alpha = V - W \in RO(G)$. The †smash product of representatives induces a bilinear pairing $\omega_\alpha \times \omega_\beta \to \omega_{\alpha+\beta}$. Then ω_0 is a ring, and ω_α is an ω_0-module. The ring ω_0 is isomorphic to the Burnside ring of G, and ω_α is a †projective ω_0-module of rank one. The ω_0-module ω_α is free if and only if $S(V)$ and $S(W)$ are stably G-homotopy equivalent [18].

Let E be an orthogonal G-vector bundle over a compact G-space X. Denote by $S(E)$ the sphere bundle associated with E. Let E, F be orthogonal G-vector bundles over X. Then E and F have the **same spherical G-fiber homotopy type** if there exist fiber-preserving G-mappings $f: S(E) \to S(F)$, $f': S(F) \to S(E)$ and fiber-preserving G-homotopies $h_t: S(E) \to S(E)$, $h'_t: S(F) \to S(F)$ such that $h_0 = f' \circ f$, $h_1 = \text{identity}$, $h'_0 = f \circ f'$, $h'_1 = \text{identity}$. Let $KO_G(X)$ be the †equivariant K-group of real G-vector bundles over X. Let $T_G(X)$ be the additive subgroup of $KO_G(X)$ generated by elements of the form $[E] - [F]$, where E and F are orthogonal G-vector bundles having the same spherical G-fiber homotopy type. The factor group $J_G(X) = KO_G(X)/T_G(X)$ and the natural projection

$J_G: KO_G(X) \to J_G(X)$ are called an **equivariant J-group** and an **equivariant J-homomorphism**, respectively (→ 237 K-Theory).

In particular, $J_G(\{x_0\})$ is a factor group of the real representation ring $RO(G)$. †Adams operations on representation rings are the main tools for studying the group $J_G(\{x_0\})$ [18].

G. Infinitesimal Transformations

Let $f: G \times M \to M$ be a differentiable action of a Lie group G on a differentiable manifold M. Let X be a †left invariant vector field on G. Then we can define a differentiable vector field $f^+(X)$ on M as

$$f^+(X)_q h = \lim_{t \to 0} (h(f(\exp(-tX), q)) - h(q))/t$$

for each $q \in M$ and any differentiable function h defined on a neighborhood of q. It is easy to see that $f^+(X)_q = 0$ if and only if q is a fixed point of the one-parameter subgroup $\{\exp(tX)\}$. A vector field $f^+(X)$ is called an **infinitesimal transformation** of the differentiable transformation group G.

The set $\mathfrak{g}$ of all infinitesimal transformations of G forms a finite-dimensional †Lie algebra (the laws of addition and †bracket product are defined from those for the vector fields on M). If G acts effectively on M, $\mathfrak{g}$ is isomorphic to the Lie algebra of the Lie group G (→ 249 Lie Groups). In fact, the correspondence $X \to f^+(X)$ defines a Lie algebra homomorphism f^+ from the Lie algebra of all left invariant vector fields on G into the Lie algebra of all differentiable vector fields on M [19].

The following fact [20] is useful for the study of noncompact real analytic transformation groups. Let $\mathfrak{g}$ be a real †semisimple Lie algebra and $\rho: \mathfrak{g} \to L(M)$ be a Lie algebra homomorphism of $\mathfrak{g}$ into a Lie algebra of real analytic vector fields on a †real analytic manifold M. Let p be a point at which the vector fields in the image $\rho(\mathfrak{g})$ have common zero. Then there exists an analytic system of coordinates $(U; u_1, \ldots, u_m)$ with origin at p in which all the vector fields in $\rho(\mathfrak{g})$ are linear. Namely, there exists $a_{ij} \in \mathfrak{g}^* = \mathrm{Hom}_{\mathbf{R}}(\mathfrak{g}, \mathbf{R})$ such that

$$\rho(X)_q = -\sum_{i,j} a_{ij}(X) u_j(q) \frac{\partial}{\partial u_i}; \qquad X \in \mathfrak{g},\ q \in U.$$

The correspondence $X \to (a_{ij}(X))$ defines a Lie algebra homomorphism of $\mathfrak{g}$ into $\mathfrak{sl}(m, \mathbf{R})$.

For example, we can show that a real analytic $SL(n, \mathbf{R})$ action on the m-sphere is characterized by a certain real analytic vector field on $(m-n+1)$-sphere $(5 \leq n \leq m \leq 2n-2)$ [21]. In particular, there are infinitely many (at least the cardinality of the real numbers) inequivalent real analytic $SL(n, \mathbf{R})$ actions on the m-sphere $(3 \leq n \leq m)$.

Conversely, let $\mathfrak{g}$ be a finite-dimensional Lie algebra of vector fields on M. Although there is not always a differentiable transformation group G that admits $\mathfrak{g}$ as its Lie algebra of infinitesimal transformations, the following local result holds. Let $\tilde{G}$ be the †simply connected Lie group corresponding to the Lie algebra $\mathfrak{g}$. Then for each point x of M, there exist a neighborhood $\tilde{U}$ of the identity element e of $\tilde{G}$, neighborhoods V, W $(V \subset W)$ of x, and a differentiable mapping f of $\tilde{U} \times V$ into W with the following properties. Putting $f(g, y) = g(y)$ $(g \in \tilde{U}, y \in V)$, we have: (i) For all $y \in V$, $e(y) = y$. (ii) If g, $h \in \tilde{U}$, $y \in V$, then $(gh)(y) = g(h(y))$, provided that $gh \in \tilde{U}$, $h(y) \in V$. (iii) Let X be an arbitrary element of $\mathfrak{g}$. Put $g_t = \exp(-tX)$, the corresponding one-parameter subgroup of $\tilde{G}$. If $\varepsilon > 0$ is taken small enough, then we have $g_t \in \tilde{U}$ for $|t| < \varepsilon$ so that $g_t(y)$ $(|t| < \varepsilon, y \in V)$ is well defined. Therefore g_t determines a vector field $\tilde{X}$ on V by the formula

$$\tilde{X}_y h = \lim_{t \to 0} (h(g_t(y)) - h(y))/t.$$

The vector field $\tilde{X}$ coincides with the restriction of X to V. This local proposition is often expressed by the statement that $\mathfrak{g}$ generates a **local Lie group of local transformations**, which is called **Lie's fundamental theorem** on local Lie groups of local transformations.

H. Criteria

It is important to know whether a given transformation group is a topological or a Lie transformation group. The following theorems are useful for this purpose [22, 23]:

(10) Let G be a transformation group of a †locally compact Hausdorff space M. If we introduce the †compact-open topology in G, then G is a topological transformation group of M when M is locally connected or M is a †uniform topological space and G acts †equicontinuously on M.

(11) Suppose that M is a †C^1-manifold and G is a topological transformation group of M acting effectively on M. If G is locally compact and the mapping $x \to g(x)$ of M is of class C^1 for each element g of G, then G is a Lie transformation group of M.

(12) Assume that G is a transformation group of a differentiable manifold M and G acts effectively on M. Let $\mathfrak{g}$ be the set of all vector fields on M defined by one-parameter groups of transformations of M contained in G as subgroups. If $\mathfrak{g}$ is a finite-dimensional Lie algebra, then G has a Lie group structure with respect to which G is a Lie transformation group of M, and then $\mathfrak{g}$ coincides with the Lie

algebra formed by the infinitesimal transformations of G.

By applying theorems (10), (11), and (12) we can show that the following groups are Lie transformation groups: the group of all †isometries of a †Riemannian manifold; the group of all affine transformations of a differentiable manifold with a †linear connection (generally, the group of all transformations of a differentiable manifold that leave invariant a given †Cartan connection); the group of all analytic transformations of a compact complex manifold (this group is actually a complex Lie group); and the group of all analytic (holomorphic) transformations of a bounded domain in $\mathbf{C}^n$.

For related topics → 105 Differentiable Manifolds, 114 Differential Topology, 122 Discontinuous Groups, 153 Fixed-Point Theorems, 427 Topology of Lie Groups and Homogeneous Spaces, etc.

References

[1] A. Borel et al., Seminar on transformation groups, Ann. Math. Studies, Princeton Univ. Press, 1960.
[2] W. Y. Hsiang, Cohomology theory of topological transformation groups, Erg. math. 85, Springer, 1975.
[3] G. E. Bredon, Introduction to compact transformation groups, Academic Press, 1972.
[4] W. C. Hsiang and W. Y. Hsiang, Differentiable actions of compact connected classical groups I, II, III, Amer. J. Math. 89 (1967); Ann. Math., (2) 92 (1970); Ann. Math., (2) 99 (1974).
[5] R. W. Richardson, Deformations of Lie subgroups and the variation of isotropy subgroups, Acta Math., 129 (1972).
[6] W. C. Hsiang, A note on free differentiable actions of S^1 and S^3 on homotopy spheres, Ann. Math., (2) 83 (1966).
[7] F. Hirzebruch and K. H. Mayer, $O(n)$-Mannigfaltigkeiten, Exotische Sphären und Singularitäten, Lecture notes in math. 57, Springer, 1968.
[8] M. F. Atiyah and F. Hirzebruch, Spin-manifolds and group actions, Essays on Topology and Related Topics, Memoires dédiés à Georges de Rham, Springer, 1970.
[9] K. Kawakubo, Equivariant Riemann-Roch theorems, localization and formal group law, Osaka J. Math., 17 (1980).
[10] A. Hattori, Spin^c-structures and S^1-actions, Inventiones Math., 48 (1978).
[11] W. C. Hsiang and W. Y. Hsiang, The degree of symmetry of homotopy spheres, Ann. Math., (2) 89 (1969).
[12] T. Watabe, On the degree of symmetry of complex quadric and homotopy complex projective space, Sci. Rep. Niigata Univ., 11 (1974).
[13] T. Petrie, Smooth S^1 actions on homotopy complex projective spaces and related topics, Bull. Amer. Math. Soc., 78 (1972).
[14] T. Yoshida, On smooth semifree S^1 actions on cohomology complex projective spaces, Publ. Res. Inst. Math. Sci., 11 (1976).
[15] P. E. Conner and E. E. Floyd, Maps of odd period, Ann. Math., (2) 84 (1966).
[16] P. E. Conner and E. E. Floyd, Differentiable periodic maps, Erg. Math., 33 (1964).
[17] T. tom Dieck, Bordism of G-manifolds and integrality theorems, Topology, 9 (1970).
[18] T. tom Dieck, Transformation groups and representation theory, Lecture notes in math. 766, Springer, 1979.
[19] R. S. Palais, A global formulation of the Lie theory of transformation groups, Mem. Amer. Math. Soc., 22 (1957).
[20] V. Guillemin and S. Sternberg, Remarks on a paper of Hermann, Trans. Amer. Math. Soc., 130 (1968).
[21] F. Uchida, Real analytic $SL(n, R)$ actions on spheres, Tôhoku Math. J., 33 (1981).
[22] D. Montgomery and L. Zippin, Topological transformation groups, Interscience, 1955.
[23] S. Kobayashi, Transformation groups in differential geometry, Erg. math. 70, Springer, 1972.

432 (VI.8) Trigonometry

A. Plane Trigonometry

Fix an orthogonal frame O-XY in a plane, and take a point P on the plane such that the angle POX is α. Denote by (x, y) the coordinates of P, and put $OP = r$ (Fig. 1). We call the six ratios $\sin\alpha = y/r$, $\cos\alpha = x/r$, $\tan\alpha = y/x$, $\cot\alpha = x/y$, $\sec\alpha = r/x$, $\operatorname{cosec}\alpha = r/y$ the **sine**, **cosine**, **tangent**, **cotangent**, **secant**, and **cosecant** of α, respectively. These functions of the angle α are called **trigonometric functions** or **circular functions** (→ 131 Elementary Functions). They are periodic functions with the fundamental period π for the tangent and cotangent, and 2π for the others. The relation $\sin^2\alpha + \cos^2\alpha = 1$ and the **addition formulas** $\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta$, $\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$ follow from the definitions (→ Appendix A, Table 2). Given a plane triangle ABC (Fig. 2), we have the following three properties: (i) $a = b\cos C + c\cos B$ (**the first law of cosines**); (ii) $a^2 = b^2 + c^2 - 2bc\cos A$ (**the second law of cosines**); (iii) $a/\sin A = b/\sin B = c/\sin C = 2R$, where R is the radius of the circle circumscribed about $\triangle ABC$ (**laws of sines**) (→ Appendix A, Table 2). Thus we obtain relations among the six quantities a, b, c, $\angle A$, $\angle B$, and $\angle C$ associated with the triangle ABC. The study of plane figures by means of trigonometric functions is called **plane trigonometry**. For example, if a suitable combination of three of these six quantities (including a side) associated with a triangle is given, then the other three quantities are uniquely determined. The determination of unknown quantities associated with a triangle by means of these laws is called **solving a triangle**.

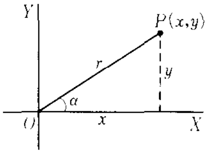


Fig. 1

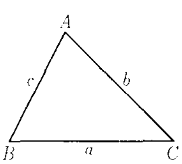


Fig. 2

B. Spherical Trigonometry

The part ABC of a spherical surface bounded by three arcs of great circles is called a **spherical triangle**. Points A, B, C are called the **vertices**; the three arcs a, b, c are called the **sides**; and the angles formed by lines tangent to the sides and intersecting at the vertices are called the **angles** of the spherical triangle (Fig. 3). If we denote the angles by A, B, C, we have the relation $A+B+C-\pi=E>0$, and E is called the **spherical excess**. Spherical triangles have properties similar to those of plane triangles: $\sin a/\sin A=\sin b/\sin B=\sin c/\sin C$ (**laws of sines**), and $\cos a=\cos b\cos c+\sin b\sin c\cos A$ (**law of cosines**). The study of spherical figures by means of trigonometric functions, called **spherical trigonometry**, is widely used in astronomy, geodesy, and navigation (→ Appendix A, Table 2).

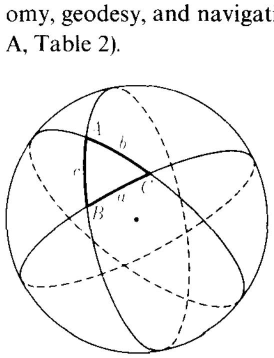


Fig. 3

C. History

Trigonometry originated from practical problems of determining a triangle from three of its elements. The development of spherical trigonometry, which was spurred on by its applications to astronomy, preceded the development of plane trigonometry. In Egypt, Babylon, and China, people had some knowledge of trigonometry, and the founder of trigonometry is believed to have been Hipparchus of Greece (fl. 150 B.C.). In the *Almagest* of Ptolemy (c. 150 A.D.) we find a table for $2\sin\alpha$ for $\alpha=0, 30', 1°, 1°30', \ldots$ that is exact to five decimal places, and the addition formulas. The Greeks calculated $2\sin\alpha$, which is the length of the chord corresponding to the double arc. Indian mathematicians, on the other hand, calculated half of the above quantities, that is, $\sin\alpha$ and $1-\cos\alpha$ for the arc α. In the book by Aryabhatta (c. 500 A.D.) we find laws of cosines. The Arabs, influenced by Indian mathematicians, expressed geometric computations algebraically, a technique also known to the Greeks. Abûl Wafâ (in the latter half of the 10th century A.D.) gave the correct sines of angles for every 30′ to 9 decimal places and studied with Al Battani the projection triangle of the sundial, thereby obtaining the concepts of sine, cosine, secant, and cosecant. Later, a table of sines and cosines for every minute was established by the Arabs. Regiomontanus (d. 1476), a German, elaborated on this table. The form he gave to trigonometry has been maintained nearly intact to the present day. Various theorems in trigonometry were established by G. J. Rhaeticus, J. Napier, J. Kepler, and L. Euler (1748). Euler treated trigonometry as a branch of analysis, generalized it to functions of complex variables, and introduced the abbreviated notations that are still in use (→ 131 Elementary Functions).

References

[1] H. Flanders and J. Price, Trigonometry, Academic Press, 1975.
[2] A. Washington and C. Edmond, Plane trigonometry, Addison-Wesley, 1977.

433 (XX.12) Turbulence and Chaos

Turbulent flow is the irregular motion of fluids, whereas relatively simple types of flows that are either stationary, slowly varying, or periodic in time are called **laminar flow**. When a laminar flow is stable against external disturbances, it remains laminar, but if the flow is unstable, it usually changes into either another type of laminar flow or a turbulent flow.

A. Stability and Bifurcation of Flows

The velocity field $\mathbf{u}(\mathbf{x}, t)$, $\mathbf{x}$ being the space coordinates and t the time, of a flow of an incompressible viscous fluid in a bounded domain G is determined by the †Navier-Stokes equation of motion,

$$\frac{\partial \mathbf{u}}{\partial t}+(\mathbf{u}\cdot \operatorname{grad})\mathbf{u}-\nu\Delta\mathbf{u}+\frac{1}{\rho}\operatorname{grad} p=0, \qquad (1)$$

and the equation of continuity,

$$\operatorname{div}\mathbf{u}=0, \qquad (2)$$

with the prescribed initial and boundary conditions, where Δ denotes the Laplacian, p the pressure, ρ the density, and ν the kinematic viscosity of the fluid. Suitable extensions must be made in the foregoing system of equations if other field variables, such as the temperature in thermal-convection problems, are to be considered.

The stability of a fluid flow is studied by examining the behavior of the solution of equations (1) and (2) against external disturbances, and, in particular, stability against infinitesimal disturbances constitutes the linear stability problem. The stability characteristics of the solution of equations (1) and (2) depend largely upon the value of the †Reynolds number $R=UL/\nu$, U and L being the representative velocity and length of the flow, respectively. Let a stationary solution of equations (1) and (2) be $\mathbf{u}_0(\mathbf{x}, R)$. If the perturbed flow is given by $\mathbf{u}_0(\mathbf{x}, R)+\mathbf{v}(\mathbf{x}, R)\exp(\sigma t)$, $\mathbf{v}$ being the perturbation velocity, and equation (1) is linearized with respect to $\mathbf{v}$, we obtain a †linear eigenvalue problem for σ. The flow is called linearly stable if $\max(\operatorname{Re}\sigma)$ is negative, and linearly unstable if it is positive. For small values of R, a flow is generally stable, but it becomes unstable if R exceeds a critical value R_c, which is called the critical Reynolds number [1].

The instability of a stationary solution gives rise to the †bifurcation to another solution at a †bifurcation point R_c of the parameter R. If $\operatorname{Im}\sigma=0$ for an eigenvalue σ at $R=R_c$, a stationary solution bifurcates from the solution u_0 at R_c, and if $\operatorname{Im}\sigma\neq 0$, a time-periodic solution bifurcates at R_c. The latter bifurcation is called the Hopf bifurcation. A typical example of stationary bifurcation is the generation of an axially periodic row of Taylor vortices in Couette flow between two rotating coaxial cylinders, which was studied by G. I. Taylor (1923), with excellent agreement between theory and experiment [2]. Hopf bifurcation is exemplified by the generation of Tollmien-Schlichting waves in the laminar †boundary layer along a flat plate, which was predicted theoretically by W. Tollmien (1929) and H. Schlichting (1933) and later confirmed experimentally by G. B. Schubauer and H. K. Skramstad (1947) [3].

In either type of bifurcation ($\operatorname{Im}\sigma=0$ or $\neq 0$) the bifurcation is called supercritical if the bifurcating solution exists only for $R>R_c$, subcritical if it exists only for $R<R_c$, and transcritical if it happens to exist on both sides of R_c. The amplitude of the departure of the bifurcating solution from the unperturbed solution u_0 tends to zero as $R\to R_c$. The behavior of the bifurcating solution around the bifurcation point R_c is dealt with systematically by means of bifurcation analysis. In supercritical bifurcation, the bifurcating solution is stable and represents an equilibrium state to which the perturbed flow approaches just as in the cases of Taylor vortices and Tollmien-Schlichting waves. On the other hand, for subcritical bifurcation the bifurcating solution is unstable and gives a critical amplitude of the disturbance above which the linearly stable basic flow ($R<R_c$) becomes unstable. In this case, the instability of the basic flow gives rise to a sudden change of the flow pattern resulting in either a stationary (or time-periodic) or even turbulent flow. The transition to turbulent flow that takes place in Hagen-Poiseuille flow through a circular tube and is linearly stable at all values of R ($R_c=\infty$) may be attributed to this type of bifurcation.

The concept of bifurcation can be extended to the case where the flow u_0 is nonstationary, but the bifurcation analysis then becomes much more difficult.

B. Onset of Turbulence

The fluctuating flow resulting from an instability does not itself necessarily constitute a turbulent flow. In order that a flow be turbulent, the fluctuations must take on some irregularity. The turbulent flow is usually defined in terms of the long-time behavior of the flow velocity $\mathbf{u}(\mathbf{x}, t)$ at a fixed point $\mathbf{x}$ in space. The flow is expected to be turbulent if the fluctuating velocity

$$\delta\mathbf{u}(\mathbf{x}, t)=\mathbf{u}(\mathbf{x}, t)-\lim_{T\to\infty}\frac{1}{T}\int_0^T \mathbf{u}(\mathbf{x}, t)\,dt \qquad (3)$$

satisfies the condition

$$\lim_{\tau\to\infty}\lim_{T\to\infty}\frac{1}{T}\int_0^T \delta u_i(\mathbf{x}, t)\delta u_i(\mathbf{x}, t+\tau)\,dt=0, \qquad (4)$$

where the subscripts label the components. Condition (4) implies that the †dynamical system of a fluid has the mixing property. This condition also states that the velocity fluctuation δu_i has a continuous frequency spectrum. In practical situations the frequency spectrum of a turbulent flow may contain both the line and continuous spectra, in which case the flow is said to be partially turbulent.

L. D. Landau (1959) and E. Hopf (1948) proposed a picture of turbulent flow as one composed of a †quasiperiodic motion, $\mathbf{u}(t) = \mathbf{f}(\omega_1 t, \omega_2 t, \ldots, \omega_n t)$, with a large number of rationally independent frequencies $\omega_1, \ldots, \omega_n$ produced by successive supercritical bifurcations of Hopf type. This picture of turbulence is not compatible with the foregoing definition of turbulence, since it does not satisfy the mixing property (4). The fact that the generation of real turbulence is not necessarily preceded by successive supercritical bifurcations casts another limitation on the validity of this picture.

The concept of turbulence is more clearly exhibited with respect to a dynamical system of finite dimension. Although we are without a general proof, it is expected that the Navier-Stokes equation with nonzero viscosity ν can be approximated within any degree of accuracy by a system of finite-dimensional †first-order ordinary differential equations

$$\frac{dX}{dt} = F(X). \tag{5}$$

Thus the onset and some general properties of turbulence are understood in the context of the theory of †dynamical systems. Turbulence is related to those solutions of equation (5) that tend to a †set in the †phase space that is neither a †fixed point, a †closed orbit, nor a †torus. A set of such complicated structure is called a nonperiodic †attractor or a strange attractor. Historically, the strange attractor originates from the strange Axiom A attractor that was found in a certain class of dynamical systems called the Axiom A systems. However, this term has come to be used in a broader sense, and it now represents a variety of nonperiodic motions exhibited by a system that is not necessarily of Axiom A type. The above-mentioned Landau-Hopf picture of turbulence was criticized by D. Ruelle and F. Takens (1971), who proved for the dynamical system (5) that an arbitrary small perturbation on a quasiperiodic †flow on a k-dimensional torus ($k \geqslant 4$) generically (in the sense of residual sets) produces a flow with a strange Axiom A attractor [4].

There exist a number of examples of first-order ordinary differential equations of relatively low dimension whose solutions exhibit nonperiodic behavior. An important model system related to turbulence is the Lorenz model (1963) of thermal convection in a horizontal fluid layer. This model is obtained by taking only three components out of an infinite number of spatial †Fourier components of the velocity and temperature fields. The model is written as

$$\frac{dX}{dt} = -\sigma X + \sigma Y,$$

$$\frac{dY}{dt} = -XZ + rX - Y,$$

$$\frac{dZ}{dt} = XY - bZ, \tag{6}$$

where σ ($> b+1$) and b are positive constants and r is a parameter proportional to the Rayleigh number. Obviously, equations (6) have a fixed point $X = Y = Z = 0$ representing the state of thermal convection without fluid flow. For $r < 1$, this fixed point is stable, but it becomes unstable for $r > 1$, and a pair of new fixed points $X = Y = \pm\sqrt{b(r-1)}$, $Z = r - 1$ emerges supercritically. This corresponds to the onset of stationary convection at $r = 1$. At a still higher value of $r = \sigma(\sigma + b + 3)/(\sigma - b - 1)$, a subcritical Hopf bifurcation occurs with respect to this pair of fixed points, and for a certain range of r above this threshold the solutions with almost any initial conditions exhibit nonperiodic behavior. This corresponds to the generation of turbulence. The property

$$\frac{\partial \dot{X}}{\partial X} + \frac{\partial \dot{Y}}{\partial Y} + \frac{\partial \dot{Z}}{\partial Z} = -(\sigma + b + 1) < 0, \tag{7}$$

where the dots denote time derivatives, shows that each volume element of the phase space shrinks asymptotically to zero as the time increases indefinitely. This property is characteristic of dynamical systems with energy dissipation, in sharp contrast to the †measure-preserving character of †Hamiltonian systems [5].

For a certain class of ordinary differential equations, the bifurcation to nonperiodic motion corresponds neither to the bifurcation of tori, just as in the Ruelle-Takens theory, nor to subcritical bifurcation, as in the Lorenz model. Such a bifurcation takes place when nonperiodic motion emerges as the consequence of an infinite sequence of supercritical bifurcations at each of which a periodic orbit of period T bifurcates into one of period $2T$. If we denote the nth bifurcation point by r_n, the distance $r_{n+1} - r_n$ between two successive bifurcation points decreases exponentially with increasing n, and eventually the bifurcation points accumulate at a point r_c, beyond which nonperiodic motion is expected to emerge. It is not yet clear if any of the above three types of bifurcation leading to nonperiodic behavior is actually responsible for the generation of real turbulence.

Some important properties of a dynamical system with a nonperiodic attractor, which may be either a flow or a †diffeomorphism, can be stated as follows:
(i) The distance between two points in the phase space that are initially close to each other grows exponentially in time, so that the solutions exhibit a sensitive dependence on the initial conditions.
(ii) The nonperiodic attractor has †Lebesgue measure zero, and such a system is expected to have many other †ergodic †invariant measures.

The irregular behavior of a deterministic dynamical system is also called **chaos**, but this concept is more abstract and general than that of turbulence, and covers phenomena exhibited by systems such as nonlinear electric circuits, chemical reactions, and ecological systems.

C. Statistical Theory of Turbulence

The statistical theory of turbulence deals with the statistical behavior of fully developed turbulence. The turbulent field is sometimes idealized for mathematical simplicity to be homogeneous or isotropic. In **homogeneous turbulence** the statistical laws are invariant under all parallel displacements of the coordinates, whereas in **isotropic turbulence** invariance under rotations and reflections of the coordinates is required in addition.

The instantaneous state of the fluid motion is completely determined by specifying the fluid velocity $\mathbf{u}$ at all space points $\mathbf{x}$ and can be expressed as a phase point in the infinite-dimensional †phase space spanned by these velocities. The phase point moves with time along a path uniquely determined by the solution of the Navier-Stokes equation. In the turbulent state the path is unstable to the initial disturbance and describes an irregular line in the phase space. In this situation the deterministic description is no longer useful and should be replaced by a statistical treatment. Abstractly speaking, turbulence is just a view of fluid motion as the random motion of the phase point $\mathbf{u}(\mathbf{x})$ (→ 407 Stochastic Processes). The equation for the †characteristic functional of the random velocity $\mathbf{u}(\mathbf{x})$ was first given by E. Hopf (1952). An exact solution obtained by Hopf represents a †normal distribution associated with a white energy spectrum, but so far no general solution has been obtained [6].

Besides the formulation in terms of the †probability distribution or the characteristic functional, there is another way of describing turbulence by †moments of lower orders. This is the conventional statistical theory originated by G. I. Taylor (1935) and T. von Kármán (1938), which made remarkable progress after World War II. The principal moments in this theory are the **correlation tensor**, whose (i, j)-component is the mean of the product of two velocity components u_i at a point $\mathbf{x}$ and u_j at another point $\mathbf{x}+\mathbf{r}$,

$$B_{ij}(\mathbf{r}) = \langle u_i(\mathbf{x}) u_j(\mathbf{x}+\mathbf{r}) \rangle, \tag{8}$$

and its †Fourier transform, or the **energy spectrum tensor**,

$$\Phi_{ij}(\mathbf{k}) = \frac{1}{(2\pi)^3} \int B_{ij}(\mathbf{r}) \exp(-\sqrt{-1}\, \mathbf{k}\cdot\mathbf{r})\, d\mathbf{r}. \tag{9}$$

In isotropic turbulence Φ_{ij} is expressed as

$$\Phi_{ij}(\mathbf{k}) = \frac{1}{4\pi k^2} E(k) \left(\delta_{ij} - \frac{k_i k_j}{k^2} \right), \qquad k = |\mathbf{k}|, \tag{10}$$

where $E(k)$ is the **energy spectrum function**, representing the amount of energy included in a spherical shell of radius k in the wave number space. The energy of turbulence $\mathscr{E}$ per unit mass is expressed as

$$\mathscr{E} = \frac{1}{2}\langle |\mathbf{u}|^2 \rangle = \frac{1}{2} B_{ii}(0) = \frac{1}{2}\int \Phi_{ii}(\mathbf{k})\, d\mathbf{k}$$
$$= \int_0^\infty E(k)\, dk. \tag{11}$$

The state of turbulence is characterized by the Reynolds number $R = E_0^{1/2}/(\nu k_0^{1/2})$, where E_0 and k_0 are representative values of $E(k)$ and k, respectively. For weak turbulence of small R, $E(k)$ is governed by a linear equation with the general solution

$$E(k,t) = E(k,0)\exp(-2\nu k^2 t), \tag{12}$$

$E(k,0)$ being an arbitrary function. Thus $E(k)$ decays in time due to the viscous dissipation. For strong turbulence of large R, it is difficult to obtain the precise form of $E(k)$, and this is usually done by way of some assumption that allows us to approximate the nonlinear effects [7].

Some of the similarity laws governing the energy spectrum and other statistical functions can be determined rigorously but not necessarily uniquely. For 3-dimensional incompressible turbulence, the energy spectrum satisfies an inviscid similarity law

$$E(k)/E_0 = F_e(k/k_0) \tag{13}$$

in the energy-containing region $k = O(k_0)$ characterized by a wave number k_0, and a viscous similarity law

$$E(k)/E_0 = R^{-5/4} F_d(k/(R^{3/4} k_0)), \tag{14}$$

in the energy dissipation region $k=O(R^{3/4}k_0)$, where F_e and F_d denote dimensionless functions generally dependent on the initial condition and the time [6].

If an assumption is made to the effect that the statistical state in the energy-dissipation region depends only upon the energy-dissipation rate $\varepsilon=-d\mathscr{E}/dt$ besides the viscosity ν (or R), then (14) becomes Kolmogorov's equilibrium similarity law (1941):

$$E(k)=\varepsilon^{1/4}\nu^{5/4}F(k/(\varepsilon^{1/4}\nu^{-3/4})), \tag{15}$$

where F is a dimensionless function. For extremely large R (or small ν) there exists an inertial subregion between the energy-containing and energy-dissipation regions such that the viscous effect vanishes and (15) takes the form

$$E(k)=K\varepsilon^{2/3}k^{-5/3}, \tag{16}$$

where K is an absolute constant. **Kolmogorov's spectrum** (16) has been observed experimentally several times, and now its consistency with experimental results at large Reynolds numbers is well established [8].

Kolmogorov (1962) and others modified (16) by taking account of the fluctuation of ε due to the **intermittent structure** of the energy-dissipation region as

$$E(k)=K'\varepsilon^{2/3}k^{-5/3}(Lk)^{-\mu/9}, \tag{17}$$

where ε is now the average of the fluctuating ε, μ is the covariance of the log-normal distribution of ε, and L is the length scale of the spatial domain in which the average of ε is taken [8]. A similar modification, with the exponent $-\mu/3$ in place of $-\mu/9$, is obtained using a fractal model of the energy-cascade process. These corrections to $E(k)$, based upon the experimentally estimated μ of 0.3–0.5, are too small to be detected experimentally, but the deviation is expected to appear more clearly in the higher-order moments [8–10]. It should be noted that Kolmogorov's spectrum (16) itself does not contradict the notion of intermittent turbulence and gives one of the possible asymptotic forms in the limit $R\to\infty$.

The 1-dimensional Burgers model of turbulence satisfies the same similarity laws as (13) and (14), but it has an inviscid spectrum $E(k)\propto k^{-2}$ instead of (16). Two-dimensional incompressible turbulence has no energy-dissipation region, and hence Kolmogorov's theory is not valid for this turbulence. It has an inviscid spectrum $E(k)\propto k^{-3}$, first derived by R. H. Kraichnan (1967), C. E. Leith (1968), and G. K. Batchelor (1969). These inviscid spectra for 1- and 2-dimensional turbulence have been confirmed by numerical simulation [11].

References

[1] C. C. Lin, The theory of hydrodynamical stability, Cambridge Univ. Press, 1955.
[2] S. Chandrasekhar, Hydrodynamic and hydromagnetic stability, Clarendon Press, 1961.
[3] H. L. Dryden, Recent advances in the mechanics of boundary layer flow, Adv. Appl. Mech., 1 (1948), 1–40.
[4] D. Ruelle and F. Takens, On the nature of turbulence, Comm. Math. Phys., 20 (1971), 167–192.
[5] E. N. Lorenz, Deterministic nonperiodic flow, J. Atmospheric Sci., 20 (1963), 130–141.
[6] T. Tatsumi, Theory of homogeneous turbulence, Adv. Appl. Mech., 20 (1980), 39–133.
[7] G. K. Batchelor, The theory of homogeneous turbulence, Cambridge Univ. Press, 1953.
[8] A. S. Monin and A. M. Yaglom, Statistical fluid mechanics II, MIT Press, 1975.
[9] B. Mandelbrot, Fractals and turbulence: Attractors and dispersion, Lecture notes in math. 615, Springer, 1977, 83–93.
[10] B. Mandelbrot, Intermittent turbulence and fractal dimension: Kurtosis and the spectral exponent $5/3+B$, Lecture notes in math. 565, Springer, 1977, 121–145.
[11] T. Tatsumi, Analytical theories and numerical experiments on two-dimensional turbulence, Theor. Appl. Mech., 29 (1979), 375–393.

U

434 (XX.22) Unified Field Theory

A. History

Unified field theory is a branch of theoretical physics that arose from the success of †general relativity theory. Its purpose is to discuss in a unified way the fields of gravitation, electromagnetism, and nuclear force from the standpoint of the geometric structure of space and time. Studies have continued since 1918, and many theories of mathematical interest have been published without attaining, however, any conclusive physical theory.

A characteristic feature of relativity theory is that it is based on a completely new concept of space and time. That is, in general relativity theory it is considered that when a gravitational field is generated by matter, the structure of space and time changes, and the flat †Minkowski world becomes a 4-dimensional †Riemannian manifold (with signature (1, 3)) having nonvanishing curvature. The †fundamental tensor g_{ij} of the manifold is interpreted as the gravitational potential, and the basic gravitational equation can be described as a geometric law of the manifold. It is characteristic of general relativity theory that gravitational phenomena are reduced to space-time structure (→ 359 Relativity). The introduction of the Minkowski world in †special relativity theory was a revolutionary advance over the 3-dimensional space of Newtonian mechanics. But the inner structure of the Minkowski world does not reflect gravitational phenomena. The latter shortcoming is overcome by introducing the concept of space-time represented by a Riemannian manifold into general relativity theory.

When a coexisting system of gravitational and electromagnetic fields is discussed in general relativity theory, simultaneous equations (Einstein-Maxwell equations) must be solved for the gravitational potential g_{ij} and the electromagnetic field tensor F_{ij}. Thus the gravitational potential g_{ij} is affected by the existence of an electromagnetic field. As the validity of general relativity began to be accepted, it came to be expected that all physical actions might be attributed to the gravitational and electromagnetic fields. Thus various extensions of general relativity theory have been proposed in order to devise a geometry in which the electromagnetic as well as the gravitational field directly contributes to the space-time structure, and to establish a unified theory of both fields on the basis of the geometry thus obtained. These attempts are illustrated in Fig. 1.

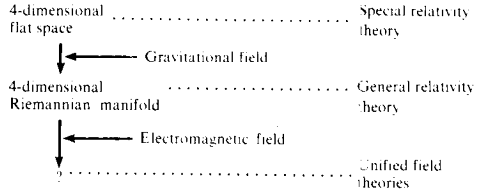


Fig. 1

B. Weyl's Theory

The first unified field theory was proposed by H. Weyl in 1918. In Riemannian geometry, which is the mathematical framework of general relativity theory, the †covariant derivative of the †fundamental tensor g_{ij} vanishes, i.e.,

$$\nabla_i g_{jk} \equiv \partial g_{jk}/\partial x^i - g_{ja}\Gamma^a_{ik} - g_{ak}\Gamma^a_{ij} = 0, \tag{1}$$

where Γ^i_{jk} is the †Christoffel symbol derived from g_{ij}. Conversely, if Γ^i_{jk} is considered as the coefficient of a general †affine connection and (1) is solved with respect to Γ^i_{jk} under the condition $\Gamma^i_{jk} = \Gamma^i_{kj}$, then the Christoffel symbol derived from g_{ij} coincides with Γ^i_{jk}. In this sense, (1) means that the space-time manifold has Riemannian structure. On the other hand, Weyl considered a space whose structure is given by an extension of (1),

$$\nabla_i g_{jk} = 2A_i g_{jk}, \tag{2}$$

and developed a unified field theory by regarding A_i as the electromagnetic potential. This theory has mathematical significance in that it motivated the discovery of Cartan's geometry of connection, but it has some unsatisfactory points concerning the derivation of the field equation and the equation of motion for a charged particle.

The scale transformation given by $\bar{g}_{ij} = \rho^2 g_{ij}$ is important in Weyl's theory. If in addition to this transformation, A_i is changed to

$$\bar{A}_i = A_i - \partial \log \rho / \partial x^i, \tag{3}$$

then (2) is left invariant and the space-time structure in Weyl's theory remains unchanged. We call (3) the **gauge transformation**, corresponding to the fact that the electromagnetic potential A_i is determined by the electromagnetic field tensor F_{ij} up to a gradient vector. In the †field theories known at present, the gauge transformation is generalized to various fields, and the law of charge conservation is derived from the invariance of field equations under generalized gauge transformation.

C. Further Developments

A unified field theory that appeared after Weyl's is **Kaluza's 5-dimensional theory** (Th.

Kaluza, 1921). This theory has been criticized as being artificial, but it is logically consistent, and therefore many of the later unified field theories are improved or generalized versions of it. The underlying space of Kaluza's theory is a 5-dimensional Riemannian manifold with the fundamental form

$$ds^2=(dx^4+A_a dx^a)^2+g_{ab}dx^a dx^b,$$

where A_i and g_{ij} are functions of x^i alone ($a, b, \ldots, i, j=0, 1, 2, 3$). The field equation and the equation of motion of a particle are derived from the variational principle in general relativity theory. The field equation is equivalent to the Einstein-Maxwell equations. The trajectory of a charged particle is given by a geodesic in the manifold, and its equation is reducible to the Lorentz equation in general relativity.

After the introduction of Kaluza's theory, various unifield field theories were proposed, and we give here the underlying manifolds or geometries of some mathematically interesting theories: a manifold with †affine connection admitting absolute parallelism (A. Einstein, 1928); a manifold with †projective connection (O. Veblen, B. Hoffman, 1930 [4]; J. A. Schouten, D. van Dantzig, 1932); **wave geometry** (a theory based on the linearization of the fundamental form; Y. Mimura, 1934 [3]); a nonholonomic geometry (G. Vranceanu, 1936); a manifold with †conformal connection (Hoffman, 1948).

The investigations since 1945 have been motivated by the problem of the representation of matter in general relativity theory. Einstein first represented matter by an energy-momentum tensor T_{ij} of class C^0, which must be determined by information obtained from outside relativity. Afterward he felt that this point was unsatisfactory and tried to develop a theory on the basis of field variables alone, without introducing such a quantity as T_{ij}. This theory is the so-called **unitary field theory**, and a solution without singularities is required from a physical point of view. His first attempt was to remove singularities from an exterior solution in general relativity by changing the topological structure of the space-time manifold. This idea was then extended to a unified field theory by J. A. Wheeler, and an interpretation was given to mass and charge by applying the theory of †harmonic integrals (1957) [2].

Einstein's second attempt was to propose a **nonsymmetric unified field theory** (1945) [1, Appendix II; 6]. The fundamental quantities in this theory are a nonsymmetric tensor g_{ij} and a nonsymmetric affine connection Γ^i_{jk}. The underlying space of the theory can be considered a direct extension of the Riemannian manifold, since (1) is contained in the field equations (notice the order of indices in this equation). E. Schrödinger obtained field equations of almost the same form by taking only Γ^i_{jk} as a fundamental quantity (1947) [5].

References

[1] A. Einstein, The meaning of relativity, Princeton Univ. Press, fifth edition, 1956.
[2] J. A. Wheeler, Geometrodynamics, Academic Press, 1962.
[3] Y. Mimura and H. Takeno, Wave geometry, Sci. Rep. Res. Inst. Theoret. Phys. Hiroshima Univ., no. 2 (1962).
[4] O. Veblen, Projektive Relativitätstheorie, Springer, 1933.
[5] E. Schrödinger, Space-time structure, Cambridge Univ. Press, 1950.
[6] A. Einstein, A generalization of the relativistic theory of gravitation, Ann. Math., (2) 46 (1945), 578–584.
[7] V. Hlavatý, Geometry of Einstein's unified field theory, Noordhoff, 1957.
[8] M. A. Tonnelat, Einstein's unified field theory, Gordon & Breach, 1966. (Original in French, 1955.)

435 (II.23) Uniform Convergence

A. Uniform Convergence of a Sequence of Real-Valued Functions

A sequence of real-valued functions $\{f_n(x)\}$ defined on a set B is said to be **uniformly convergent** (or to **converge uniformly**) to a function $f(x)$ on the set B if it converges with respect to the †norm $\|\varphi\|=\sup\{|\varphi(x)|\,|\,x\in B\}$, i.e., $\lim_{n\to\infty}\|f_n-f\|=0$ (→ 87 Convergence). In other words, $\{f_n(x)\}$ converges uniformly to $f(x)$ on B if for every positive constant ε we can select a number N independent of the point x such that $|f_n(x)-f(x)|<\varepsilon$ holds for all $n>N$ and $x\in B$. By the †completeness of the real numbers, a sequence of functions $\{f_n(x)\}$ converges uniformly on B if and only if we can select for every positive constant ε a number N independent of the point x such that $|f_m(x)-f_n(x)|<\varepsilon$ holds for all $m, n>N$ and $x\in B$.
The **uniform convergence** of a series $\sum_n f_n(x)$ or of an infinite product $\prod_n f_n(x)$ is defined by the uniform convergence of the sequence of its partial sums or products. If the series of the absolute values $\sum_n|f_n(x)|$ converges uniformly, then the series $\sum_n f_n(x)$ also converges uniformly. In this case the series $\sum_n f_n(x)$ is said to

be **uniformly absolutely convergent**. A sequence of (nonnegative) constants M_n satisfying $|f_n(x)| \leqslant M_n$ is called a **dominant** (or **majorant**) of the sequence of functions $\{f_n(x)\}$. A series of functions $\sum_n f_n(x)$ with converging **majorant series** $\sum_n M_n$ is uniformly absolutely convergent (**Weierstrass's criterion for uniform convergence**).

Let $\{\lambda_n(x)\}$ be another sequence of functions on B. The series $\sum_n \lambda_n(x) f_n(x)$ is uniformly convergent if either of the following conditions holds: (i) The series $\sum_n f_n(x)$ converges uniformly and the partial sums of the series $\sum_n |\lambda_n(x) - \lambda_{n+1}(x)|$ are uniformly bounded, i.e., bounded by a constant independent of $x \in B$ and of the number of terms; or (ii) the series $\sum_n |\lambda_n(x) - \lambda_{n+1}(x)|$ converges uniformly, the sequence $\{\lambda_n(x)\}$ converges uniformly to 0, and the partial sums of $\sum_n |f_n(x)|$ are uniformly bounded.

B. Uniform Convergence and Pointwise Convergence

Let $\{f_n(x)\}$ be a sequence of real-valued functions on B, and let $f(x)$ be a real-valued function also defined on B. If the sequence of numbers $\{f_n(x_0)\}$ converges to $f(x_0)$ for every point $x_0 \in B$, we say that $\{f_n(x)\}$ is **pointwise convergent** (or **simply convergent**) to the function $f(x)$. Pointwise convergence is, of course, weaker than uniform convergence. If we represent the function $f(x)$ by the point $\prod_{x \in B} f(x) = [f]$ of the †Cartesian product $\mathbf{R}^B = \prod_{x \in B} \mathbf{R}$, then the pointwise convergence of $\{f_n(x)\}$ to $f(x)$ is equivalent to the convergence of the sequence of points $\{[f_n]\}$ to $[f]$ in the †product topology of $\mathbf{R}^B$.

When B is a †topological space and every $f_n(x)$ is continuous, the pointwise limit $f(x)$ of the sequence $\{f_n(x)\}$ is not necessarily continuous. However, if the sequence of continuous functions $\{f_n(x)\}$ converges uniformly to $f(x)$, then the limit function $f(x)$ is continuous. On the other hand, the continuity of the limit does not imply that the convergence is uniform. If the set B is †compact and the sequence of continuous functions $\{f_n(x)\}$ is monotone (i.e., $f_n(x) \leqslant f_{n+1}(x)$ for all n or $f_n(x) \geqslant f_{n+1}(x)$ for all n) and pointwise convergent to a continuous function $f(x)$, then the convergence is uniform (**Dini's theorem**).

C. Uniform Convergence on a Family of Sets

Let B be a topological space. We say that a sequence of functions $\{f_n(x)\}$ is **uniformly convergent in the wider sense** to the function $f(x)$, depending on circumstances, in either of the following two cases: (i) Every point $x_0 \in B$ has a neighborhood U on which the sequence $\{f_n(x)\}$ converges uniformly to $f(x)$; or (ii) $\{f_n(x)\}$ converges uniformly to $f(x)$ on every compact subset K in B. If B is †locally compact, the two definitions coincide. The term **uniform convergence on compact sets** is also used for (ii).

In general, given a family $\mathscr{P}$ of subsets in B, we may introduce in the space $\mathscr{F}$ of real-valued functions on B a family of †seminorms $\|f\|_K = \sup\{|f(x)| \mid x \in K\}$ for every set $K \in \mathscr{P}$. Let T be the topology of $\mathscr{F}$ defined by this family of seminorms (→ 424 Topological Linear Spaces). A sequence $\{f_n(x)\}$ is called **uniformly convergent on** $\mathscr{P}$ if it is convergent with respect to T. In particular, when $\mathscr{P}$ coincides with $\{B\}$, $\{\{x\} \mid x \in B\}$, or the family of all compact sets in B, then uniform convergence on $\mathscr{P}$ coincides with the usual uniform convergence, pointwise convergence, or uniform convergence on compact sets, respectively. If $\mathscr{P}$ is a countable set, the topology T is †metrizable. Most of these definitions and results may be extended to the case of functions whose values are in the complex number field, in a †normed space, or in any †uniform space.

D. Topology of the Space of Mappings

Let X, Y be two topological spaces. Denote by $C(X, Y)$ the space of all continuous mappings $f: X \to Y$. This space $C(X, Y)$, or a subspace $\mathscr{F}$ of $C(X, Y)$, is called a **mapping space** (or **function space** or **space of continuous mappings**) from X to Y. A natural mapping $\Phi: \mathscr{F} \times X \to Y$ is defined by $\Phi(f, x) = f(x)$ $(f \in \mathscr{F}, x \in X)$. We define a topology in $\mathscr{F}$ as follows: for a compact set K in X and an open set U in Y, put $W(K, U) = \{f \in \mathscr{F} \mid f(K) \in U\}$, and introduce a topology in $\mathscr{F}$ such that the base for the topology consists of intersections of finite numbers of $W(K_i, U_i)$. This topology is called the **compact-open topology** (R. H. Fox, *Bull. Amer. Math. Soc.*, 51 (1945)). When X is a †locally compact Hausdorff space and Y is a †Hausdorff space, the compact-open topology is the †weakest topology on $\mathscr{F}$ for which the function Φ is continuous. If, in this case, $\mathscr{F}$ is compact with respect to the compact-open topology, then the compact-open topology coincides with the topology of pointwise convergence.

In particular, when Y is a †metric space (or, in general, a †uniform space with the uniformity $\mathfrak{U}$), the compact-open topology in $\mathscr{F}$ coincides with the topology of uniform convergence on compact sets. A family $\mathscr{F}$ is called

equicontinuous at a point $x \in X$ if for every positive number ε (in the case of uniform space, for every $V \in \mathfrak{U}$) there exists a neighborhood U of x such that $\rho(f(x), f(p)) < \varepsilon$ ($(f(x), f(p)) \in V$) for every point $p \in U$ and for every function $f \in \mathcal{F}$ (G. Ascoli, 1883–1884). If X is a †locally compact Hausdorff space, a necessary and sufficient condition for $\mathcal{F}$ to be relatively compact (i.e., for the closure of $\mathcal{F}$ to be compact) with respect to the compact-open topology (i.e., to the topology of uniform convergence on compact sets) is that $\mathcal{F}$ be equicontinuous at every point $x \in X$ and that the set $\{f(x) | f \in \mathcal{F}\}$ be relatively compact in Y for every point $x \in X$ (**Ascoli's theorem**). In particular, when X is a σ-compact locally compact Hausdorff space and Y is the space of real numbers, a family of functions $\mathcal{F}$ that are equicontinuous (at every point $x \in X$) and uniformly bounded is relatively compact. Hence, for any sequence of functions $\{f_n\}$ in $\mathcal{F}$, we can select a subsequence $\{f_{n(\nu)}\}$ which converges uniformly on compact sets (**Ascoli-Arzelà theorem**).

E. Normal Families

P. Montel (1912) gave the name **normal family** to the family of functions that is relatively compact with respect to the topology of uniform convergence on compact sets. This terminology is used mainly for the family of complex analytic functions. In that case, it is customary to compactify the range space and consider Y to be the †Riemann sphere. Using this notion, Montel succeeded in giving a unified treatment of various results in the theory of complex functions.

A family of analytic functions $\mathcal{F}$ on a finite-dimensional †complex manifold X is a normal family if it is uniformly bounded on each compact set (**Montel's theorem**). Another criterion is that there are three values on the Riemann sphere which no function $f \in \mathcal{F}$ takes. More generally, three exceptional values not taken by $f \in \mathcal{F}$ may depend on f, if there is a positive lower bound for the distances between these three values on the Riemann sphere. This gives an easy proof of the †Picard theorem stating that every †transcendental meromorphic function $f(z)$ in $|z| < \infty$ must take all values except possibly two values. In fact the family of functions $f_n(z) = f(z/2^n)$, $n = 1, 2, 3, \ldots$, in $\{1 < |z| < 2\}$ cannot be normal. Using a similar procedure, G. Julia obtained the results on †Julia's direction.

F. Marty introduced the notion of **spherical derivative** $|f'(z)|/(1 + |f(z)|^2)$ for the analytic or meromorphic function $f(z)$ and proved that for a family $\mathcal{F} = \{f(z)\}$ of analytic functions to be normal, it is necessary and sufficient that the spherical derivatives of $f \in \mathcal{F}$ be uniformly bounded. This theorem implies Montel's theorem and its various extensions, including, for example, quantitative results concerning †Borel's direction.

A family $\mathcal{F}$ of analytic functions of one variable defined on X is said to form a **quasi-normal family** if there exists a subset P of X consisting only of isolated points such that from any sequence $\{f_n\}$ ($f_n \in \mathcal{F}$) we can select a subsequence $\{f_{n(\nu)}\}$ converging uniformly on $X - P$. If P is finite and consists of p points, the family $\mathcal{F}$ is called a quasinormal family of order p. For example, the family of at most †p-valent functions is quasinormal of order p.

The theory of normal families of complex analytic functions is not only applied to †value distribution theory, as above, but also used to show the existence of a function that gives the extremal of functionals. The extremal function is usually obtained as a limit of a subsequence of a sequence in a normal family. A typical example of this method is seen in the proof of the †Riemann mapping theorem. This is perhaps the only general method known today in the study of the iteration of †holomorphic functions. By this method, Julia (1919) made an exhaustive study of the iteration of meromorphic functions; there are several other investigations on the iteration of elementary transcendental functions. On the other hand, A. Wintner (*Comm. Math. Helv.*, 23 (1949)) gave the implicit function theorem for analytic functions in a precise form using the theory of normal families of analytic functions of several variables.

References

[1] N. Bourbaki, Eléments de mathématique, III., Topologie générale, ch. 10, Espaces fonctionnels, Actualités Sci. Ind., 1084b, Hermann, second edition, 1967; English translation, Theory of sets, Addison-Wesley, 1966.
[2] J. L. Kelley, General topology, Van Nostrand, 1955, ch. 7.
[3] J. Dieudonné, Foundations of modern analysis, Academic Press, 1960, enlarged and corrected printing, 1969.
For normal families of complex functions,
[4] P. Montel, Leçons sur les familles normales de fonctions analytiques et leurs applications, Gauthier-Villars, 1927.
[5] G. Valiron, Familles normales et quasi-normales de fonctions méromorphes, Gauthier-Villars, 1929.

436 (II.22) Uniform Spaces

A. Introduction

There are certain properties defined on †metric spaces but not on general †topological spaces, for example, †completeness or †uniform continuity of functions. Generalizing metric spaces, A. Weil introduced the notion of uniform spaces. This notion can be defined in several ways [3,4]. The definition in Section B is that of Weil [1] without the †separation axiom for topology.

We denote by Δ_X the **diagonal** $\{(x,x) \mid x \in X\}$ of the Cartesian product $X \times X$ of a set X with itself. If U and V are subsets of $X \times X$, then the **composite** $V \circ U$ is defined to be the set of all pairs (x,y) such that for some element z of X, the pair (x,z) is in U and the pair (z,y) is in V. The inverse U^{-1} of U is defined to be the set of all pairs (x,y) such that $(y,x) \in U$.

B. Definitions

Let $\mathscr{U}$ be a nonempty family of subsets of $X \times X$ such that (i) if $U \in \mathscr{U}$ and $U \subset V$, then $V \in \mathscr{U}$; (ii) if U, $V \in \mathscr{U}$, then $U \cap V \in \mathscr{U}$; (iii) if $U \in \mathscr{U}$, then $\Delta_X \subset U$; (iv) if $U \in \mathscr{U}$, then $U^{-1} \in \mathscr{U}$; and (v) if $U \in \mathscr{U}$, then $V \circ V \subset U$ for some $V \in \mathscr{U}$. Then we say that a **uniform structure** (or simply a **uniformity**) is defined on X by $\mathscr{U}$. If a uniformity is defined on X by $\mathscr{U}$, then the pair $(X, \mathscr{U})$ or simply the set X itself is called a **uniform space**, and $\mathscr{U}$ is usually called a **uniformity** for X.

A subfamily $\mathscr{B}$ of the uniformity $\mathscr{U}$ is called a **base for the uniformity** $\mathscr{U}$ if every member of $\mathscr{U}$ contains a member of $\mathscr{B}$. If a family $\mathscr{B}$ of subsets of $X \times X$ is a base for a uniformity $\mathscr{U}$, then the following propositions hold: (ii′) if U, $V \in \mathscr{B}$, then there exists a $W \in \mathscr{B}$ such that $W \subset U \cap V$; (iii′) if $U \in \mathscr{B}$, then $\Delta_X \subset U$; (iv′) if $U \in \mathscr{B}$, then there exists a $V \in \mathscr{B}$ such that $V \subset U^{-1}$; (v′) if $U \in \mathscr{B}$, then there exists a $V \in \mathscr{B}$ such that $V \circ V \subset U$. Conversely, if a family $\mathscr{B}$ of subsets of a Cartesian product $X \times X$ satisfies (ii′)–(v′), then the family $\mathscr{U} = \{U \mid U \subset X \times X, V \subset U \text{ for some } V \in \mathscr{B}\}$ defines a uniformity on X and $\mathscr{B}$ is a base for $\mathscr{U}$. Given a uniform space $(X, \mathscr{U})$, a member V of $\mathscr{U}$ is said to be **symmetric** if $V = V^{-1}$. The family of all symmetric members of $\mathscr{U}$ is a base for $\mathscr{U}$.

C. Topology of Uniform Spaces

Given a uniform space $(X, \mathscr{U})$, an element $x \in X$, and $U \in \mathscr{U}$, we put $U(x) = \{y \mid y \in X, (x,y) \in U\}$. Then the family $\mathscr{U}(x) = \{U(x) \mid U \in \mathscr{U}\}$ forms a neighborhood system of $x \in X$, which gives rise to a topology of X (→ 425 Topological Spaces). This topology is called the **uniform topology** (or **topology of the uniformity**). When we refer to a topology of a uniform space $(X, \mathscr{U})$, it is understood to be the uniform topology; thus a uniform space is also called a **uniform topological space**. If $\mathscr{B}$ is a base for the uniformity of a uniform space $(X, \mathscr{U})$, then $\mathscr{B}(x) = \{U(x) \mid U \in \mathfrak{B}\}$ is a base for the neighborhood system at each point $x \in X$. Each member of $\mathscr{U}$ is a subset of the topological space $X \times X$, which is supplied with the product topology. The family of all open (closed) symmetric members of $\mathscr{U}$ forms a base for $\mathscr{U}$. A uniform space $(X, \mathscr{U})$ is a †T_1-topological space if and only if the intersection of all members of $\mathscr{U}$ is the diagonal Δ_X. In this case, the uniformity of $(X, \mathscr{U})$ is called a **T_1-uniformity**, and $(X, \mathscr{U})$ is called a **T_1-uniform space**. A T_1-uniform space is always †regular; a fortiori, it is a T_2-topological space. Hence a T_1-uniform space is also said to be a **Hausdorff uniform space** (or **separated uniform space**). Moreover, a uniform topology satisfies †Tikhonov's separation axiom; in particular, a T_1-uniform space is †completely regular.

D. Examples

(1) Discrete Uniformity. Let X be a nonempty set, and let $\mathscr{U} = \{U \mid \Delta_X \subset U \subset X \times X\}$. Then $(X, \mathscr{U})$ is a T_1-uniform space and $\mathscr{B} = \{\Delta_X\}$ is a base for $\mathscr{U}$. This uniformity is called the **discrete uniformity** for X.

(2) Uniform Family of Neighborhood System. A family $\{U_\alpha(x)\}_{\alpha \in A} (x \in X)$ of subsets of a set X is called a **uniform neighborhood system** in X if it satisfies the following four requirements: (i) $x \in U_\alpha(x)$ for each $\alpha \in A$ and each $x \in X$; (ii) if x and y are distinct elements of X, then $y \notin U_\alpha(x)$ for some $\alpha \in A$; (iii) if α and β are two elements of A, then there is another element $\gamma \in A$ such that $U_\gamma(x) \subset U_\alpha(x) \cap U_\beta(x)$ for all $x \in X$; (iv) if α is an arbitrary element in A, then there is an element β in A such that $y \in U_\alpha(x)$ whenever x, $y \in U_\beta(z)$ for some z in X. If we denote by $U_\alpha (\alpha \in A)$ the subset of $X \times X$ consisting of all elements (x,y) such that $x \in X$ and $y \in U_\alpha(x)$, then the family $\{U_\alpha \mid \alpha \in A\}$ satisfies all the conditions for a base for a uniformity. In particular, it follows from (ii) that $\bigcap_{\alpha \in A} U_\alpha = \Delta_X$, which is a stronger condition than (iii′) in Section B. For instance, if $\{U_\alpha \mid \alpha \in A\}$ is a base for the neighborhood system at the identity element of a T_1-topological group G, then we have two uniform neighborhood systems $\{U_\alpha^l(x)\}$ and $\{U_\alpha^r(x)\}$, where $U_\alpha^l(x) = xU_\alpha$ and

$U_\alpha^r(x)=U_\alpha x$. Two uniformities derived from these uniform neighborhood systems are called a †left uniformity and a †right uniformity, respectively. Generally, these two uniformities do not coincide (→ 423 Topological Groups).

(3) Uniform Covering System [4]. A family $\{\mathfrak{U}_\alpha\}_{\alpha\in A}$ of †coverings of a set X is called a **uniform covering system** if the following three conditions are satisfied: (i) if $\mathfrak{U}$ is a covering of X such that $\mathfrak{U}\prec\mathfrak{U}_\alpha$ for all $\alpha\in A$, then $\mathfrak{U}$ coincides with the covering $\Delta=\{\{x\}\}_{x\in X}$; (ii) if α, $\beta\in A$, then there is a $\gamma\in A$ such that $\mathfrak{U}_\gamma\prec\mathfrak{U}_\alpha$ and $\mathfrak{U}_\gamma\prec\mathfrak{U}_\beta$; (iii) if $\alpha\in A$, then there is a $\beta\in A$ such that $\mathfrak{U}_\beta$ is a †Δ-refinement of $\mathfrak{U}_\alpha$ $((\mathfrak{U}_\beta)^\Delta\prec\mathfrak{U}_\alpha)$. For an example of a uniform covering system of X, suppose that we are given a uniform neighborhood system $\{U_\alpha(x)\}_{\alpha\in A}$ $(x\in X)$. Let $\mathscr{U}_\alpha=\{U_\alpha(x)\}_{\alpha\in X}$ $(\alpha\in A)$. Then $\{\mathscr{U}_\alpha\}_{\alpha\in A}$ is a uniform covering system. On the other hand, for a covering $\mathfrak{U}=\{U_\lambda\}_{\lambda\in\Lambda}$, let $S(x,\mathfrak{U})$ be the union of all members of $\mathfrak{U}$ that contain x. If $\{\mathfrak{U}_\alpha\}_{\alpha\in A}$ is a uniform covering system and $U_\alpha(x)=S(x,\mathfrak{U}_\alpha)$, then $\{U_\alpha(x)\}_{\alpha\in A}$ $(x\in X)$ is a uniform neighborhood system. Hence defining a uniform covering system of X is equivalent to defining a T_1-uniformity on X.

(4). In a metric space (x,d) the subsets $\mathfrak{U}_r=\{(x,y)\mid d(x,Y)<r\}$, $r>0$, form a base of uniformity. The uniform topology defined by this coincides with the topology defined by the metric.

E. Some Notions on Uniform Spaces

Some of the terminology concerning topological spaces can be restated in the language of uniform structures. A mapping f from a uniform space $(X,\mathscr{U})$ into another $(X',\mathscr{U}')$ is said to be **uniformly continuous** if for each member U' in $\mathscr{U}'$ there is a member U in $\mathscr{U}$ such that $(f(x),f(y))\in U'$ for every $(x,y)\in U$. This condition implies that f is continuous with respect to the uniform topologies of the uniform spaces. Equivalently, the mapping is uniformly continuous with respect to the uniform neighborhood system $\{U_\alpha(x)\}_{\alpha\in A}$ if for any index β there is an index α such that $y\in U_\alpha(x)$ implies $f(y)\in U_\beta(f(x))$. If $f:X\to X'$ and $g:X'\to X''$ are uniformly continuous, then the composite $g\circ f:X\to X''$ is also uniformly continuous. A bijection f of a uniform space $(X,\mathscr{U})$ to another $(X',\mathscr{U}')$ is said to be a **uniform isomorphism** if both f and f^{-1} are uniformly continuous; in this case $(X,\mathscr{U})$ and $(X',\mathscr{U}')$ are said to be **uniformly equivalent**. A uniform isomorphism is a homeomorphism with respect to the uniform topologies, and a uniform equivalence defines an equivalence relation between uniform spaces.

If $\mathscr{U}_1$ and $\mathscr{U}_2$ are uniformities for a set X, we say that the uniformity $\mathscr{U}_1$ is **stronger** than the uniformity $\mathscr{U}_2$ and $\mathscr{U}_2$ is **weaker** than $\mathscr{U}_1$ if the identity mapping of $(X,\mathscr{U}_1)$ to $(X,\mathscr{U}_2)$ is uniformly continuous. The discrete uniformity is the strongest among the uniformities for a set X. The weakest uniformity for X is defined by the single member $X\times X$; this uniformity is not a T_1-uniformity unless X is a singleton. Generally, there is no weakest T_1-uniformity. A uniformity $\mathscr{U}_1$ for X is stronger than another $\mathscr{U}_2$ if and only if every member of $\mathscr{U}_2$ is also a member of $\mathscr{U}_1$.

If f is a mapping from a set X into a uniform space $(Y,\mathscr{V})$ and g is the mapping of $X\times X$ into $Y\times Y$ defined by $g(x,y)=(f(x),f(y))$, then $\mathscr{B}=\{g^{-1}(V)\mid V\in\mathscr{V}\}$ satisfies conditions (ii')–(v') in Section B for a base for a uniformity. The uniformity $\mathscr{U}$ for X determined by $\mathscr{B}$ is called the **inverse image** of the uniformity $\mathscr{V}$ for Y by f; $\mathscr{U}$ is the weakest uniformity for X such that f is uniformly continuous. Hence a mapping f from a uniform space $(X,\mathscr{U})$ into another $(Y,\mathscr{V})$ is uniformly continuous if and only if the inverse image of the uniformity $\mathscr{V}$ under f is weaker than the uniformity $\mathscr{U}$. If A is a subset of a uniform space $(X,\mathscr{U})$, then there is a uniformity $\mathscr{V}$ for A determined as the inverse image of $\mathscr{U}$ by the inclusion mapping of A into X. This uniformity $\mathscr{V}$ for A is called the **relative uniformity** for A induced by $\mathscr{U}$, or the **relativization** of $\mathscr{U}$ to A, and the uniform space $(A,\mathscr{V})$ is called a **uniform subspace** of $(X,\mathscr{U})$. The uniform topology for $(A,\mathscr{V})$ is the relative topology for A induced by the uniform topology for $(X,\mathscr{U})$.

If $\{(X_\lambda,\mathscr{U}_\lambda)\}_{\lambda\in\Lambda}$ is a family of uniform spaces, then the **product uniformity** for $X=\prod_{\lambda\in\Lambda}X_\lambda$ is defined to be the weakest uniformity $\mathscr{U}$ such that the projection of X onto each X_λ is uniformly continuous, and $(X,\mathscr{U})$ is called the **product uniform space** of $\{(X_\lambda,\mathscr{U}_\lambda)\}_{\lambda\in\Lambda}$. The topology for $(X,\mathscr{U})$ is the product of the topologies for $(X_\lambda,\mathscr{U}_\lambda)$ $(\lambda\in\Lambda)$.

F. Metrization

Each †pseudometric d for a set X generates a uniformity in the following way. For each positive number r, let $V_{d,r}=\{(x,y)\in X\times X\mid d(x,y)<r\}$. Then the family $\{V_{d,r}\mid r>0\}$ satisfies conditions (ii')–(v') in Section B for a base for a uniformity $\mathscr{U}$. This uniformity is called the **pseudometric uniformity** or **uniformity generated by** d. The uniform topology for $(X,\mathscr{U})$ is the pseudometric topology. A uniform space $(X,\mathscr{U})$ is said to be **pseudometrizable** (**metrizable**) if there is a pseudometric (metric)

d such that the uniformity $\mathscr{U}$ is identical with the uniformity generated by d. A uniform space is pseudometrizable if and only if its uniformity has a countable base. Consequently, a uniform space is metrizable if and only if its uniformity is a T_1-uniformity and has a countable base. For a family P of pseudometrics on a set X, let $V_{d,r} = \{(x, y) \in X \times X \mid d(x, y) < r\}$ for $d \in P$ and positive r. The weakest uniformity containing every $V_{d,r}$ $(d \in P, r > 0)$ is called the **uniformity generated by** P. This uniformity may also be described as the weakest one such that each pseudometric in P is uniformly continuous on $X \times X$ with respect to the product uniformity.

Each uniformity $\mathscr{U}$ on a set X coincides with the uniformity generated by the family P_X of all pseudometrics that are uniformly continuous on $X \times X$ with respect to the product uniformity of $\mathscr{U}$ with itself. It follows that each uniform space is uniformly isomorphic to a subspace of a product of pseudometric spaces (in which the number of components is equal to the cardinal number of P_X) and that each T_1-uniform space is uniformly isomorphic to a subspace of a product of metric spaces. A topology τ for a set X is the uniform topology for some uniformity for X if and only if the topological space (X, τ) satisfies †Tikhonov's separation axiom; in particular, the uniformity is a T_1-uniformity if and only if (X, τ) is †completely regular.

G. Completeness

If $(X, \mathscr{U})$ is a uniform space, a subset A of X is called a **small set of order** U $(U \in \mathscr{U})$ if $A \times A \subset U$. A †filter on X is called a **Cauchy filter** (with respect to the uniformity $\mathscr{U}$) if it contains a small set of order U for each U in $\mathscr{U}$. If a filter on X converges to some point in X, then it is a Cauchy filter. If f is a uniformly continuous mapping from a uniform space X into another X', then the image of a base for a Cauchy filter on X under f is a base for a Cauchy filter on X'. A point contained in the closure of every set in a Cauchy filter $\mathfrak{F}$ is the limit point of $\mathfrak{F}$. Hence if a filter converges to x, a Cauchy filter contained in the filter also converges to x.

A †net $x(\mathfrak{A}) = \{x_\alpha\}_{\alpha \in \mathfrak{A}}$ (where $\mathfrak{A}$ is a directed set with a preordering $\leqslant$) in a uniform space $(X, \mathscr{U})$ is called a **Cauchy net** if for each U in $\mathscr{U}$ there is a γ in $\mathfrak{A}$ such that $(x_\alpha, x_\beta) \in U$ for every α and β such that $\gamma \leqslant \alpha$, $\gamma \leqslant \beta$. If $\mathfrak{A}$ is the set $\mathbf{N}$ of all natural numbers, a Cauchy net $\{x_n\}_{n \in \mathbf{N}}$ is called a **Cauchy sequence** (or **fundamental sequence**). Given a Cauchy net $\{x_\alpha\}_{\alpha \in \mathfrak{A}}$, let $A_\alpha = \{x_\beta \mid \beta \geqslant \alpha\}$. Then $\mathfrak{B} = \{A_\alpha \mid \alpha \in \mathfrak{A}\}$ is a base for a filter, and the filter is a Cauchy filter. On the other hand, let $\mathfrak{B}$ be a base for a Cauchy filter $\mathfrak{F}$. For $U, V \in \mathfrak{B}$, we put $U \leqslant V$ if and only if $U \supset V$. Then $\mathfrak{B}$ is a directed set with respect to $\leqslant$. The net $\{x_U\}_{U \in \mathfrak{B}}$, where x_U is an arbitrary point in U, is a Cauchy net. A proposition concerning convergence of a Cauchy filter is always equivalent to a proposition concerning convergence of the corresponding Cauchy net.

A Cauchy filter (or Cauchy net) in a uniform space X does not always converge to a point of X. A uniform space is said to be **complete** (with respect to the uniformity) if every Cauchy filter (or Cauchy net) converges to a point of that space. A complete uniform space is called for brevity a **complete space**. A closed subspace of a complete space is complete with respect to the relative uniformity. A pseudometrizable uniform space is complete if and only if every Cauchy sequence in the space converges to a point. Hence in the case of a metric space, our definition of completeness coincides with the usual one (→ 273 Metric Spaces).

A mapping f from a uniform space X to another X' is said to be **uniformly continuous on a subset** A of X if the restriction of f to A is uniformly continuous with respect to the relative uniformity for A. If f is a uniformly continuous mapping from a subset A of a uniform space into a complete T_1-uniform space, then there is a unique uniformly continuous extension $\bar{f}$ of f on the closure $\bar{A}$.

Each T_1-uniform space is uniformly equivalent to a dense subspace of a complete T_1-uniform space; this property is a generalization of the fact that each metric space can be mapped by an isometry onto a dense subset of a complete metric space. **A completion** of a uniform space $(X, \mathscr{U})$ is a pair $(f, (X^*, \mathscr{U}^*))$, where $(X^*, \mathscr{U}^*)$ is a complete space and f is a uniform isomorphism of X onto a dense subspace of X^*. The T_1-completion of a T_1-uniform space is unique up to uniform equivalence.

H. Compact Spaces

A uniformity $\mathscr{U}$ for a topological space (X, τ) is said to be **compatible** with the topology τ if the uniform topology for $(X, \mathscr{U})$ coincides with τ. A topological space (X, τ) is said to be **uniformizable** if there is a uniformity compatible with τ. If (X, τ) is a compact Hausdorff space, then there is a unique uniformity $\mathscr{U}$ compatible with τ; in fact, $\mathscr{U}$ consists of all neighborhoods of the diagonal Δ_X in $X \times X$; and the compact Hausdorff space is complete with this uniformity. Hence every subspace of a compact Hausdorff space is uniformizable, and every †locally compact Hausdorff space is uniformizable. Any continuous mapping from a compact Hausdorff space to a uniform space is uniformly continuous. A uniform space $(X, \mathscr{U})$ is said to be **totally bounded** (or **precompact**) if for each $U \in \mathscr{U}$ there is a finite covering consisting of small sets of order U; a subset of a uniform space is called **totally bounded** if it is totally bounded with respect to the relative uniformity. A uniform space X is said to be **locally totally bounded** if for each point of X there is a base for a neighborhood system consisting of totally bounded open subsets. A uniform space is compact if and only if it is totally bounded and complete. If f is a uniformly continuous mapping from a uniform space X to another, then the image $f(A)$ of a totally bounded subset A of X is totally bounded.

I. Topologically Complete Spaces

A topological space (X, τ) is said to be **topologically complete** (or **Dieudonné complete**) if it admits a uniformity compatible with τ with respect to which X is complete. Each †paracompact Hausdorff space is topologically complete. Actually such a space is complete with respect to its strongest uniformity. A Hausdorff space which is homeomorphic to a †G_δ-set in a compact Hausdorff space is said to be **Čech-complete**; A metric space is homeomorphic to a complete metric space if and only if it is Čech-complete. A Hausdorff space X is paracompact and Čech-complete if and only if there is a †perfect mapping from X onto a complete metric space.

References

[1] A. Weil, Sur les espaces à structure uniforme et sur la topologie générale, Actualités Sci. Ind., Hermann, 1938.
[2] J. W. Tukey, Convergence and uniformity in topology, Ann. Math. Studies, Princeton Univ. Press, 1940.
[3] J. L. Kelley, General topology, Van Nostrand, 1955.
[4] N. Bourbaki, Eléments de mathématique, III. Topologie générale, ch. 2, Actualités Sci. Ind., 1142d, Hermann, fourth edition 1965; English translation, General topology, Addision-Wesley, 1966.
[5] J. R. Isbell, Uniform spaces, Amer. Math. Soc. Math. Surveys, 1964.
[6] H. Nakano, Uniform spaces and transformation groups, Wayne State Univ. Press, 1968.
[7] R. Emgelking, General topology, Polish Scientific Publishers, 1977.
[8] Z. Frolík, Generalization of the G_δ-property of complete metric spaces, Czech. Math. J., 10 (1960), 359–379.

437 (IV.17) Unitary Representations

A. Definitions

A homomorphism U of a †topological group G into the group of †unitary operators on a †Hilbert space $\mathfrak{H}$ ($\neq \{0\}$) is called a **unitary representation** of G if U is **strongly continuous** in the following sense: For any element $x \in \mathfrak{H}$, the mapping $g \rightarrow U_g x$ is a continuous mapping from G into $\mathfrak{H}$. The Hilbert space $\mathfrak{H}$ is called the **representation space** of U and is denoted by $\mathfrak{H}(U)$. Two unitary representations U and U' are said to be **equivalent** (**similar** or **isomorphic**), denoted by $U \cong U'$, if there exists an †isometry T from $\mathfrak{H}(U)$ onto $\mathfrak{H}(U')$ that satisfies the equality $T \circ U_g = U'_g \circ T$ for every g in G. If the representation space $\mathfrak{H}(U)$ contains no closed subspace other than $\mathfrak{H}$ and $\{0\}$ that is invariant under every U_g, the unitary representation U is said to be **irreducible**. An element x in $\mathfrak{H}(U)$ is called a **cyclic vector** if the set of all finite linear combinations of the elements $U_g x (g \in G)$ is dense in $\mathfrak{H}(U)$. A representation U having a cyclic vector is called a **cyclic representation**. Every nonzero element of the representation space of an irreducible representation is a cyclic vector.

Examples. Let G be a †topological transformation group acting on a †locally compact Hausdorff space X from the right. Suppose that there exists a †Radon measure μ that is invariant under the group G. Then a unitary representation R^μ is defined on the Hilbert space $\mathfrak{H} = L^2(X, \mu)$ by the formula $(R^\mu_g f)(x) = f(xg)$ $(f \in \mathfrak{H}, x \in X, g \in G)$. The representation R^μ is called the **regular representation** of G on (X, μ). If G acts on X from the left, then the regular representation L^μ is defined by $(L^\mu_g f)(x) = f(g^{-1}x)$. In particular, when X is the †quotient space $H \backslash G$ of a †locally compact group G by a closed subgroup H, any two invariant measures μ, μ' (if they exist) coincide up to a constant factor. Hence the regular representation R^μ on (X, μ) and the regular representation $R^{\mu'}$ on (X, μ') are equivalent. In this case, the representation R^μ is called the regular representation on X. When $H = \{e\}$, a locally compact group G has a Radon measure $\mu \neq 0$ that is invariant under every right (left) translation $h \rightarrow hg$ $(h \rightarrow gh)$ and is called a right (left) †Haar measure on G. So G has the regular representation R (L) on G. R (L) is called the right (left) regular representation of G.

B. Positive Definite Functions and Existence of Representations

A complex-valued continuous function φ on a topological group G is called **positive definite** if the matrix having $\varphi(g_i^{-1}g_j)$ as the (i,j)-component is a †positive semidefinite Hermitian matrix for any finite number of elements $g_1, \ldots, g_n$ in G. If U is a unitary representation of G, then the function $\varphi(g)=(U_g x, x)$ is positive definite for every element x in $\mathfrak{H}(U)$. Conversely, any positive definite function $\varphi(g)$ on a topological group G can be expressed as $\varphi(g)=(U_g x, x)$ for some unitary representation U and x in $\mathfrak{H}(U)$. Using this fact and the †Kreĭn-Milman theorem, it can be proved that every locally compact group G has **sufficiently many irreducible unitary representations** in the following sense: For every element g in G other than the identity element e, there exists an irreducible unitary representation U, generally depending on g, that satisfies the inequality $U_g \neq 1$. The groups having sufficiently many finite-dimensional (irreducible) unitary representations are called †maximally almost periodic. If a connected locally compact group G is maximally almost periodic, then G is the direct product of a compact group and a vector group $\mathbf{R}^m$. On the other hand, any noncompact connected †simple Lie group has no finite-dimensional irreducible unitary representation other than the unit representation $g \to 1$ (→ 18 Almost Periodic Functions).

C. Subrepresentations

Let U be a unitary representation of a topological group G. A closed subspace $\mathfrak{N}$ of $\mathfrak{H}(U)$ is called **U-invariant** if $\mathfrak{N}$ is invariant under every U_g $(g \in G)$. Let $\mathfrak{N} \neq \{0\}$ be a closed invariant subspace of $\mathfrak{H}(U)$ and V_g be the restriction of U_g on $\mathfrak{N}$. Then V is a unitary representation of G on the representation space $\mathfrak{N}$ and is called a **subrepresentation** of U. Two unitary representations L and M are called **disjoint** if no subrepresentation of L is equivalent to a subrepresentation of M; they are called **quasi-equivalent** if no subrepresentation of L is disjoint from M and no subrepresentation of M is disjoint from L.

D. Irreducible Representations

Let U be a unitary representation of G, $\mathbf{M}$ be the †von Neumann algebra generated by $\{U_g \mid g \in G\}$, and $\mathbf{M}'$ be the †commutant of $\mathbf{M}$. Then a closed subspace $\mathfrak{N}$ of $\mathfrak{H}(U)$ is invariant under U if and only if the †projection operator P corresponding to $\mathfrak{N}$ belongs to $\mathbf{M}'$. Therefore U is irreducible if and only if $\mathbf{M}'$ consists of scalar operators $\{\alpha 1 \mid \alpha \in \mathbf{C}\}$ **(Schur's lemma)**. A representation space of a cyclic or irreducible representation of a †separable topological group is †separable.

E. Factor Representations

A unitary representation U of G is called a **factor representation** if the von Neumann algebra $\mathbf{M}=\{U_g \mid g \in G\}$ is a †factor, that is, $\mathbf{M} \cap \mathbf{M}' = \{\alpha 1 \mid \alpha \in \mathbf{C}\}$. Two factor representations are quasi-equivalent if and only if they are not disjoint. U is called a **factor representation of type I, II,** or **III** if the von Neumann algebra $\mathbf{M}$ is a factor of †type I, II, or III, respectively (→ 308 Operator Algebras). A topological group G is called a **group of type I** (or **type I group**) if every factor representation of G is of type I. Compact groups, locally compact Abelian groups, connected †nilpotent Lie groups, connected †semisimple Lie groups, and real or complex †linear algebraic groups are examples of groups of type I. There exists a connected solvable Lie group that is not of type I (→ Section U), but a connected solvable Lie group is of type I if the exponential mapping is surjective (O. Takenouchi). A discrete group G with countably many elements is a type I group if and only if G has an Abelian normal subgroup with finite index (E. Thoma).

F. Representation of Direct Products

Let G_1 and G_2 be topological groups, G the †direct product of G_1 and G_2 $(G=G_1 \times G_2)$, and U_i an irreducible unitary representation of G_i $(i=1,2)$. Then the †tensor product representation $U_1 \otimes U_2 : (g_1, g_2) \to U_{g_1} \otimes U_{g_2}$ is an irreducible unitary representation of G. Conversely, if one of the groups G_1 and G_2 is of type I, then every irreducible unitary representation of G is equivalent to the tensor product $U_1 \otimes U_2$ of some irreducible representations U_i of G_i $(i=1,2)$.

G. Direct Sums

If the representation space $\mathfrak{H}$ of a unitary representation U is the †direct sum $\bigoplus_{\alpha \in I} \mathfrak{H}(\alpha)$ of mutually orthogonal closed invariant subspaces $\{\mathfrak{H}(\alpha)\}_{\alpha \in I}$, then U is called the **direct sum** of the subrepresentations $U(\alpha)$ induced on $\mathfrak{H}(\alpha)$ by U, and is denoted by $U=\bigoplus_{\alpha \in I} U(\alpha)$. Any unitary representation is the direct sum of cyclic representations. A unitary representation U is called a **representation without multiplicity** if U cannot be decomposed as a direct sum $U_1 \oplus U_2$ unless U_1 and U_2 are disjoint. If U is the direct sum of $\{U(\alpha)\}_{\alpha\in I}$ and every $U(\alpha)$ is irreducible, then U is said to be **decomposed into the direct sum of irreducible representations**. Decomposition into direct sums of irreducible representations is essentially unique if it exists; that is, if $U = \bigoplus_{\alpha\in I} U(\alpha) = \bigoplus_{\beta\in J} V(\beta)$ are two decompositions of U into direct sums of irreducible representations, then there exists a bijection φ from I onto J such that $U(\alpha)$ is equivalent to $V(\varphi(\alpha))$ for every α in I. A factor representation U of type I can be decomposed as the direct sum $U = \bigoplus_{\alpha\in I} U(\alpha)$ of equivalent irreducible representations $U(\alpha)$. In general, a unitary representation U cannot be decomposed as the direct sum of irreducible representations even if U is not irreducible. Thus it becomes necessary to use direct integrals to obtain an irreducible decomposition.

H. Direct Integrals

Let U be a unitary representation of a group G and (X,μ) be a †measure space. Assume that the following two conditions are satisfied by U: (i) There exists a unitary representation $U(x)$ of G corresponding to every element x of X, and $\mathfrak{H}(U)$ is a †direct integral (→ 308 Operator Algebras) of $\mathfrak{H}(U(x))$ $(x\in X)$ (written $\mathfrak{H}(U)=\int_X \mathfrak{H}(U(x))\,d\mu(x)$); (ii) for every g in G, the operator U_g is a decomposable operator and can be written as $U_g=\int_X U_g(x)\,d\mu(x)$. Then the unitary representation U is called the **direct integral** of the family $\{U(x)\}_{x\in X}$ of unitary representations and is denoted by $U=\int_X U(x)\,d\mu(x)$. If every point of X has measure 1, then a direct integral is reduced to a direct sum.

I. Decomposition into Factor Representations

We assume that G is a locally compact group satisfying the †second countability axiom, and also that a Hilbert space is separable. Every unitary representation U of G can be decomposed as a direct integral $U=\int_X U(x)\,d\mu(x)$ in such a way that the center $\mathbf{A}$ of the von Neumann algebra $\mathbf{M}''=\{U_g \mid g\in G\}''$ is the set of all †diagonalizable operators. In this case almost all the $U(x)$ are factor representations. Such a decomposition of U is essentially unique. There exists a †null set N in X such that for every x and x' in $X-N$ $(x\neq x')$, $U(x)$ and $U(x')$ are mutually disjoint factor representations. Hence the space X can be identified with the set G^* of all quasi-equivalence classes of factor representations of G endowed with a suitable structure of a measure space. The space G^* is called the **quasidual** of G. The measure μ is determined by U up to †equivalence of measures.

J. Duals

A topology is introduced on the set $\hat{G}$ of all equivalence classes of irreducible unitary representations of a locally compact group G in the following way. Let H_n be the n-dimensional Hilbert space $l_2(n)$ and I_n the set of all irreducible unitary representations of G realized on H_n $(1\leqslant n\leqslant\infty)$. We topologize I_n in such a way that a †net $\{U^\lambda\}_{\lambda\in L}$ in I_n converges to U if and only if $(U_g^\lambda x, y)$ converges uniformly to $(U_g x, y)$ on every compact subset of G for any x and y in H_n. Equivalence between representations in I_n is an open relation. Let $\hat{G}_n$ be the set of all equivalence classes of n-dimensional irreducible unitary representations of G with the topology of a quotient space of I_n and $\hat{G}=\bigcup_n \hat{G}_n$ be the direct sum of topological spaces $\hat{G}_n$. Then the topological space $\hat{G}$ is called the **dual** of G. $\hat{G}$ is a locally compact †Baire space with countable open base, but it does not satisfy the †Hausdorff separation axiom in general. If G is a compact Hausdorff topological group, then $\hat{G}$ is discrete. If G is a locally compact Abelian group, then $\hat{G}$ coincides with the †character group of G in the sense of Pontryagin. If G is a type I group, then there exists a dense open subset of $\hat{G}$ that is a locally compact Hausdorff space. The †σ-additive family generated by closed sets in $\hat{G}$ is denoted by $\mathfrak{B}$. In the following sections, a measure on $\hat{G}$ means a measure defined on $\mathfrak{B}$.

K. Irreducible Decompositions

In this section G is assumed to be a locally compact group of type I with countable open base. For any equivalence class x in $\hat{G}$, we choose a representative $U(x)\in x$ with the representation space $H(U(x))=l_2(n)$ if x is n-dimensional. For any measure μ on $\hat{G}$, the representation $U^\mu=\int_{\hat{G}} U(x)\,d\mu(x)$ is a unitary representation without multiplicity. Conversely, any unitary representation of G without multiplicity is equivalent to a U^μ for some measure μ on $\hat{G}$. Moreover, U^μ is equivalent to U^ν if and only if the two measures μ and ν are equivalent (that is, μ is absolutely continuous with respect to ν, and vice versa). A unitary representation U with multiplicity on a separable Hilbert space $\mathfrak{H}$ can be decomposed as follows: There exists a countable set of measures $\mu_1, \mu_2, \ldots, \mu_\infty$ whose supports are mutually disjoint such that $U \cong \int_{\hat{G}} U(x)d\mu_1(x) \oplus 2\int_{\hat{G}} U(x)d\mu_2(x) \oplus \ldots \oplus \infty\int_{\hat{G}} U(x)d\mu_\infty(x)$. The measures $\mu_1, \mu_2, \ldots, \mu_\infty$ are uniquely determined by U up to equivalence of measures. Any unitary representation U on a separable Hilbert space $\mathfrak{H}$ of an arbitrary locally compact group with countable open base (even if not of type I) can be decomposed as a direct integral of irreducible representations. In order to obtain such a decomposition, it is sufficient to decompose $\mathfrak{H}$ as a direct integral in such a way that a maximal Abelian von Neumann subalgebra $\mathbf{A}$ of $\mathbf{M}' = \{U_g \mid g \in G\}'$ is the set of all diagonalizable operators. In this case, however, a different choice of $\mathbf{A}$ induces in general an essentially different decomposition, and uniqueness of the decomposition does not hold. For a group of type I, the irreducible representations are the "atoms" of representations, as in the case of compact groups. For a group not of type I, it is more natural to take the factor representations for the irreducible representations, quasi-equivalence for the equivalence, and the quasidual for the dual of G. Therefore the theory of unitary representations for a group not of type I has different features from the one for a type I group. The theory of unitary representation for groups not of type I has not yet been successfully developed, but some important results have been obtained (e.g., L. Pukanszky, *Ann. Sci. Ecole Norm. Sup.*, 4 (1971)).

Tatsuuma [1] proved a duality theorem for general locally compact groups which is an extension of both Pontryagin's and Tannaka's duality theorems considering the direct integral decomposition of tensor product representations.

L. The Plancherel Formula

Let G be a unimodular locally compact group with countable open base, $R(L)$ be the right (left) regular representation of G, and $\mathbf{M}$, $\mathbf{N}$, and $\mathbf{P}$ be the von Neumann algebras generated by $\{R_g\}$, $\{L_g\}$, and $\{R_g, L_g\}$, respectively. Then $\mathbf{M}' = \mathbf{N}$, $\mathbf{N}' = \mathbf{M}$, and $\mathbf{P}' = \mathbf{M} \cap \mathbf{N}$. If we decompose $\mathfrak{H}$ into a direct integral in such a way that $\mathbf{P}'$ is the algebra of all diagonalizable operators, then $\mathbf{M}(x)$ and $\mathbf{N}(x)$ are factors for almost all x. This decomposition of $\mathfrak{H}$ produces a decomposition of the two-sided regular representation $\{R_g, L_g\}$ into irreducible representations and a decomposition of the regular representation $R(L)$ into factor representations. Hence the decomposition is realized as the direct integral over the quasidual G^* of G. Moreover, the factors $\mathbf{M}(x)$ and $\mathbf{N}(x)$ are of type I or II for almost all x in G^*, and there exists a †trace t in the factor $\mathbf{M}(x)$. For any f and g in $L_1(G) \cap L_2(G)$, the **Plancherel formula**

$$\int_G f(s)\overline{g(s)}\,ds = \int_{G^*} t(U_g^*(x)U_f(x))\,d\mu(x) \tag{1}$$

holds, where $U_f(x) = \int_G f(s)U_s(x)ds$ and U^* is the †adjoint of U. The **inversion formula**

$$h(s) = \int_{G^*} t(U_h(x)U_s^*(x))\,d\mu(x) \tag{2}$$

is derived from (1) for a function $h = f * g$ ($f, g \in L_1(G) \cap L_2(G)$). In (1) and (2), because of the impossibility of normalization of the trace t in a factor of type II_∞, the measure μ cannot in general be determined uniquely. However, if G is a type I group, then (1) and (2) can be rewritten as similar formulas, where the representation $U(x)$ in (1) and (2) is irreducible, the trace t is the usual trace, and the domain of integration is not the quasidual G^* but the dual $\hat{G}$ of G. The revised formula (1) is also called the Plancherel formula. In this case the measure μ on $\hat{G}$ in formulas (1) and (2) is uniquely determined by the given Haar measure on G. The measure μ is called the **Plancherel measure** of G. The support $\hat{G}_r$ of the Plancherel measure μ is called the **reduced dual** of G. The Plancherel formula gives the direct integral decomposition of the regular representation into the irreducible representations belonging to $\hat{G}_r$. Each U in $\hat{G}_r$ is contained in this decomposition, with the multiplicity equal to $\dim \mathfrak{H}(U)$.

M. Square Integrable Representations

An irreducible unitary representation U of a unimodular locally compact group G is said to be **square integrable** when for some element $x \neq 0$, in $\mathfrak{H}(U)$, the function $\varphi(g) =: (U_g x, x)$ belongs to $L^2(G, dg)$, where dg is the Haar measure of G. If U is square integrable, then $\varphi_{x,y}(g) = (U_g x, y)$ belongs to $L^2(G, dg)$ for any x and y in $\mathfrak{H}(U)$. Let U and U' be the two square integrable representations of G. Then the following **orthogonality relations** hold:

$$\int_G (U_g x, y)\overline{(U'_g u, v)}\,dg = \begin{cases} 0 & \text{if } U \text{ is not equivalent to } U', \\ d_U^{-1}(x,u)(v,y) & \text{if } U = U'. \end{cases} \tag{3}$$

When G is compact, every irreducible unitary representation U is square integrable and finite-dimensional. Moreover, the scalar d_U in (3) is the degree of U if the total measure of G is normalized to 1. In the general case, the scalar d_U in (3) is called the **formal degree** of U and is determined uniquely by the given Haar measure dg. Let y be an element in $\mathfrak{H}(U)$ with norm 1 and V be the subspace $\overline{\{\varphi_{x,y} \mid x \in \mathfrak{H}(U)\}}$ of $L^2(G)$. Then the linear mapping $T: x \to \sqrt{d_U}\,\varphi_{x,y}$ is an isometry of $\mathfrak{H}(U)$ onto V. Hence U is equivalent to a subrepresentation of the right regular representation R of G. Conversely, every irreducible subrepresentation of R is square integrable. Thus a square integrable representation is an irreducible subrepresentation of R ($\cong L$). Therefore, in the irreducible decomposition of R, the square integrable representations appear as discrete direct summands. Hence every square integrable representation U has a positive Plancherel measure $\mu(U)$ that is equal to the formal degree d_U. There exist noncompact groups that have square integrable representations. An example of such a group is $SL(2, \mathbf{R})$ (→ Section X).

N. Representations of $L_1(G)$

Let G be a locally compact group and $L_1(G)$ be the space of all complex-valued integrable functions on G. Then $L_1(G)$ is an algebra over $\mathbf{C}$, where the convolution

$$(f*g)(s) = \int_G f(st^{-1})g(t)\,dt$$

is defined to be the product of f and g. Let Δ be the †modular function of G. Then the mapping $f(s) \to f^*(s) = \Delta(s^{-1})\overline{f(s^{-1})}$ is an †involution of the algebra $L_1(G)$. Let U be a unitary representation of G, and put $U'_f = \int_G U_s f(s)\,ds$. Then the mapping $f \to U'_f$ gives a **nondegenerate representation** of the Banach algebra $L_1(G)$ with an involution, where nondegenerate means that $\{U'_f x \mid f \in L_1(G), x \in \mathfrak{H}(U)\}^{\perp}$ reduces to $\{0\}$. The mapping $U \to U'$ gives a bijection between the set of equivalence classes of unitary representations of G and the set of equivalence classes of nondegenerate representations of the Banach algebra $L_1(G)$ with an involution on Hilbert spaces. U is an irreducible (factor) representation if and only if U' is an irreducible (factor) representation. Therefore the study of unitary representations of G reduces to that of representations of $L_1(G)$. If U'_f is a †compact operator for every f in $L_1(G)$, then U is the discrete direct sum of irreducible representations, and the multiplicity of every irreducible component is finite. (See [2] for Sections A–N.)

O. Induced Representations

Induced representation is the method of constructing a representation of a group G in a canonical way from a representation of a subgroup H of G. It is a fundamental method of obtaining a unitary representation of G. Let G be a locally compact group satisfying the second countability axiom, L be a unitary representation on a separable Hilbert space $\mathfrak{H}(L)$ of a closed subgroup H of G, and m, n, Δ, and δ be the right Haar measures and the modular functions of the groups G and H, respectively. Then there exists a continuous positive function ρ on G satisfying $\rho(hg) = \delta(h)\Delta(h)^{-1}\rho(g)$ for every h in H and g in G. The †quotient measure $\mu = (\rho m)/n$ is a quasi-invariant measure on the coset space $H\backslash G$ (→ 225 Invariant Measures). Let $\mathfrak{H}$ be the vector space of weakly measurable functions f on G with values in $\mathfrak{H}(L)$ satisfying the following two conditions: (i) $f(hg) = L_h f(g)$ for every h in H and g in G; and (ii) $\|f\|^2 = \int_{H\backslash G} \|f(g)\|^2\,d\mu(\dot{g}) < +\infty$, where $\dot{g}$ represents the coset Hg. By condition (i), the norm $\|f(g)\|$ is constant on a coset $Hg = \dot{g}$ and is a function on $H\backslash G$, so the integral in condition (ii) is well defined. Then $\mathfrak{H}$ is a Hilbert space with the norm defined in (ii). A unitary representation U of G on the Hilbert space $\mathfrak{H}$ is defined by the formula

$$(U_s f)(g) = \sqrt{\rho(gs)/\rho(g)}\; f(gs).$$

U is called the **unitary representation induced by the representation L of a subgroup** H and is denoted by $U = U^L$ or $\mathrm{Ind}_H^G L$. **Induced representations** have the following properties.

(1) $U^{L_1 \oplus L_2} \cong U^{L_1} \oplus U^{L_2}$ or more generally, $U^{\int U(x)d\mu(x)} \cong \int U^{L(x)} d\mu(x)$. Therefore if U^L is irreducible, L is also irreducible (the converse does not hold in general).

(2) Let H, K be two subgroups of G such that $H \subset K$, L be a unitary representation of H, and M be the representation of K induced by L. Then two unitary representations U^M and U^L of G are equivalent.

An induced representation U^L is the representation on the space of square integrable sections of the †vector bundle with fiber $H(L)$ †associated with the principal bundle $(G, H\backslash G, H)$ (→ G. W. Mackey [3], F. Bruhat [4]).

P. Unitary Representations of Special Groups

In the following sections we describe the fundamental results on the unitary representations of certain special groups.

Q. Compact Groups

Irreducible unitary representations of a compact group are always finite-dimensional. Every unitary representation of a compact group is decomposed into the direct sum of irreducible representations. Irreducible unitary representations of a compact connected Lie group are completely classified. The characters of irreducible representations are calculated in an explicit form (→ 69 Compact Groups; 249 Lie Groups). Every irreducible unitary representation U of a connected compact Lie group G can be extended uniquely to an irreducible holomorphic representation U^C of the complexification G^C of G. U^C is holomorphically induced from a 1-dimensional representation of a Borel subgroup B of G^C (**Borel-Weil theorem**; → R. Bott [5]).

R. Abelian Groups

Every irreducible unitary representation of an Abelian group G is 1-dimensional. †Stone's theorem concerning one-parameter groups of unitary operators, $U_t=\int_{-\infty}^{\infty}e^{i\lambda t}dE_\lambda$, gives irreducible decompositions of unitary representations of the additive group **R** of real numbers. †Bochner's theorem on †positive definite functions on **R** is a restatement of Stone's theorem in terms of positive definite functions. The theory of the †Fourier transform on **R**, in particular †Plancherel's theorem, gives the irreducible decomposition of the regular representation of **R**. The theorems of Stone, Bochner, and Plancherel have been extended to an arbitrary locally compact Abelian group (→ 192 Harmonic Analysis).

S. Representations of Lie Groups and Lie Algebras

Let U be a unitary representation of a Lie group G with the Lie algebra $\mathfrak{g}$. An element x in $\mathfrak{H}(U)$ is called an **analytic vector** with respect to U if the mapping $g\rightarrow U_g x$ is a real analytic function on G with values in $\mathfrak{H}(U)$. The set of all analytic vectors with respect to U forms a dense subspace $\mathfrak{A}=\mathfrak{A}(U)$ of $\mathfrak{H}(U)$. For any elements X in $\mathfrak{g}$ and x in $\mathfrak{A}(U)$, the derivative at $t=0$ of a real analytic function $U_{\exp tX}x$ is denoted by $V(X)x$. Then $V(X)$ is a linear transformation on $\mathfrak{A}$, and the mapping $V:X\rightarrow V(X)$ is a representation of $\mathfrak{g}$ on $\mathfrak{A}$. We call V the **differential representation** of U. The representation V of $\mathfrak{g}$ can be extended uniquely to a representation of the †universal enveloping algebra $\mathfrak{B}$ of $\mathfrak{g}$. Two unitary representations $U^{(1)}$ and $U^{(2)}$ of a connected Lie group G are equivalent if and only if there exists a bijective bounded linear mapping T from $\mathfrak{H}(U^{(1)})$ onto $\mathfrak{H}(U^{(2)})$ such that T maps $\mathfrak{A}(U^{(1)})$ onto $\mathfrak{A}(U^{(2)})$ and satisfies the equality

$$(T\circ V^{(1)}(X))x=(V^{(2)}(X)\circ T)x$$

for all X in $\mathfrak{g}$ and x in $\mathfrak{A}(U^{(1)})$. Let $X_1,\dots,X_n$ be a basis of $\mathfrak{g}$ and U be a unitary representation of G. Then the element $\Delta=X_1^2+\dots+X_n^2$ in the universal enveloping algebra $\mathfrak{B}$ of $\mathfrak{g}$ is represented in the differential representation V of U by an †essentially self-adjoint operator $V(\Delta)$. Conversely, if to each element X in $\mathfrak{g}$ there corresponds a (not necessarily bounded) †skew-Hermitian operator $\rho(x)$ that satisfies the following three conditions, then there exists a unique unitary representation U of the simply connected Lie group G with the Lie algebra $\mathfrak{g}$ such that the †closure of $V(X)$ coincides with the closure of $\rho(X)$ for every X in $\mathfrak{g}$: (i) There exists a dense subspace $\mathfrak{D}$ contained in the domain of $\rho(X)\rho(Y)$ for every X and Y in $\mathfrak{g}$; (ii) for each X and Y in $\mathfrak{g}$, a and b in **R**, and x in $\mathfrak{D}$, $\rho(aX+bY)x=a\rho(X)x+b\rho(Y)x$, $\rho([X,Y])x=(\rho(X)\rho(Y)-\rho(Y)\rho(X))x$; (iii) the restriction of $\rho(X_1)^2+\dots+\rho(X_n)^2$ to $\mathfrak{D}$ is an essentially self-adjoint operator if $X_1,\dots,X_n$ is a basis of $\mathfrak{g}$ (E. Nelson [6]).

T. Nilpotent Lie Groups

For every irreducible unitary representation of a connected nilpotent Lie group G, there is some 1-dimensional unitary representation of some subgroup of G that induces it. Let G be a simply connected nilpotent Lie group, $\mathfrak{g}$ be the Lie algebra of G, and ρ be the contragredient representation of the adjoint representation of G. The representation space of ρ is the dual space $\mathfrak{g}^*$ of $\mathfrak{g}$. A subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is called **subordinate** to an element f in $\mathfrak{g}^*$ if f annihilates each bracket $[X,Y]$ for every X and Y in $\mathfrak{h}$: $(f,[X,Y])=0$. When $\mathfrak{h}$ is subordinate to f, a 1-dimensional unitary representation L of the analytic subgroup H of G with the Lie algebra $\mathfrak{h}$ is defined by the formula $\lambda_f(\exp X)=e^{2\pi i(f,X)}$ $(X\in\mathfrak{h})$. Every 1-dimensional unitary representation λ_f of H is defined as in this formula by an element f in $\mathfrak{g}^*$ to which $\mathfrak{h}$ is subordinate. The unitary representation of G induced by such a λ_f is denoted by $U(f,\mathfrak{h})$. The representation $U(f,\mathfrak{h})$ is irreducible if and only if $\mathfrak{h}$ has maximal dimension among the subalgebras subordinate to f. Two irreducible representations $U(f,\mathfrak{h})$ and $U(f,\mathfrak{h}')$ are equivalent if and only if f and f' are conjugate under the group $\rho(G)$. Therefore there exists a bijection between the set of equivalence classes of the irreducible unitary representations of a simply connected nilpotent Lie group G and the set of orbits of $\rho(G)$ on $\mathfrak{g}^*$ (A. A. Kirillov [7]).

U. Solvable Lie Groups

Let G be a simply connected solvable Lie group. If the exponential mapping is bijective, G is called an **exponential group**. All results stated above for nilpotent Lie groups hold for exponential groups except the irreducibility criterion. In this case the representation $U(f, \mathfrak{h})$ is irreducible if and only if $\mathfrak{h}$ is of maximal dimension among subordinate subalgebras and the orbit $O=\rho(G)f$ contains the affine subspace $f+\mathfrak{h}^{\perp}=f+\{g \mid g(\mathfrak{h})=0\}$ (Pukanszky condition).

The situation is more complicated for general solvable Lie groups. The isotropy subgroup $G_f=\{g\in G \mid \rho(g)f=f\}$ at $f\in\mathfrak{g}^*$ is, in general, not connected. A linear form f is called integral if there exists a unitary character η_f of G_f whose differential is the restriction of $2\pi i f$ to $\mathfrak{g}_f$ (the Lie algebra of G_f). Using the notion of "polarization," an irreducible unitary representation of G is constructed from a pair (f, η_f) of an integral form $f\in\mathfrak{g}^*$ and a character η_f. If G is of type I, then every irreducible unitary representation of G is obtained in this way. A simply connected solvable Lie group G is of type I if and only if (i) every $f\in\mathfrak{g}^*$ is integral and (ii) every G-orbit $\rho(G)f$ in $\mathfrak{g}^*$ is locally closed (Auslander and Kostant [8]).

As an example, let α be an irrational real number. Then the following Lie group G is not of type I: $G=\left\{\begin{pmatrix} e^{it} & 0 & z \\ 0 & e^{i\alpha t} & w \\ 0 & 0 & 1 \end{pmatrix} \,\middle|\, t\in\mathbf{R}, z, w\in\mathbf{C}\right\}$.

V. Semisimple Lie Groups

A connected semisimple Lie group is of type I. The character $\chi=\chi_U$ of an irreducible unitary representation U of G is defined as follows: Let $C_0^{\infty}(G)$ be the set of all complex-valued C^{∞}-functions with compact support on G. Then for any function f in $C_0^{\infty}(G)$, the operator $U_f=\int_G U_g f(g)dg$ belongs to the †trace class, and the linear form $\chi: f\rightarrow T_r U_f$ is a †distribution in the sense of Schwartz. The distribution χ is called the **character** of an irreducible unitary representation U. A character χ is invariant under any inner automorphism of G and is a simultaneous eigendistribution of the algebra of all two-sided invariant linear differential operators on G. Two irreducible unitary representations of G are equivalent if and only if their characters coincide. The distribution χ is a †locally summable function on G and coincides with a real analytic function on each connected component of the dense open submanifold G' consisting of regular elements in G. In general, χ is not real analytic on all of G (Harish-Chandra [9, III; 10].

W. Complex Semisimple Lie Groups

There are four series of irreducible representations of a complex semisimple Lie group G.

(1) A **principal series** consists of unitary representations of G induced from 1-dimensional unitary representations L of a †Borel subgroup B of G. L is uniquely determined by a unitary character $\nu\in\mathrm{Hom}(A, U(1))=A^*$ of the †Cartan subgroup A of G contained in B. Hence the representations in the principal series are parametrized by the elements in the character group A^* of the Cartan subgroup A. If we denote U^L by U^{ν}, two representations U^{ν} and $U^{\nu'}(\nu, \nu'\in A^*)$ are equivalent if and only if ν and ν' are conjugate under the †Weyl group W of G with respect to A.

(2) A **degenerate series** consists of unitary representations induced by 1-dimensional unitary representations of a †parabolic subgroup P of G other than B. (A parabolic subgroup P is any subgroup of G containing a Borel subgroup B.)

(3) A **complementary series** consists of irreducible unitary representations U^L induced by nonunitary 1-dimensional representations L of a Borel subgroup B. In this case, condition (ii) in the definition of U^L (→ Section O) must be changed. When L is a nonunitary representation, then the operator U_g^L is not a unitary operator with respect to the usual L_2-inner product (ii). However, if L satisfies a certain condition, then U_g^L leaves invariant some positive definite Hermitian form on the space of sufficiently nice functions. Completing this space, we get a unitary representation U^L. The representations thus obtained form the complementary series.

(4) A **complementary degenerate series** consists of irreducible unitary representations induced by nonunitary 1-dimensional representations of a parabolic subgroup $P\neq B$.

Representations belonging to different series are never equivalent. It seems certain that any irreducible unitary representation of a connected complex semisimple Lie group is equivalent to a representation belonging to one of the above four series, but this conjecture has not yet been proved. Moreover, E. M. Stein [11] constructed irreducible unitary representations different from any in the list obtained by I. M. Gel'fand and M. A. Naĭmark (Neumark) [12]. These representations belong to the complementary degenerate series. The characters of the representations in these four series are computed in explicit form. For example, the character χ_{ν} of the representation U^{ν} in the principal series can be calculated as follows: Let λ be a linear form on a Cartan subalgebra $\mathfrak{a}$ such that $\nu(\exp H)=e^{\lambda(H)}$ for every H in $\mathfrak{a}$, let D be the function on A defined by $D(\exp H)=\prod_{\alpha}|e^{\alpha(H)/2}-e^{-\alpha(H)/2}|^2$, where α runs over all positive roots. Then the character χ_{ν} of a representation U^{ν} in the principal series is given by the formula

$$\chi_{\nu}(\exp H)=D(\exp H)^{-1}\sum_{s\in W} e^{s\lambda(H)}.$$

In the irreducible decomposition of the regular representation of G, only irreducible representations belonging to the principal series arise. Hence the right-hand side in the Plancherel formula is an integral over the character group A^* of a Cartan subgroup A. Under a suitable normalization of the Haar measures in G and A^*, the Plancherel measure μ of G can be expressed by using the Haar measure dv of A^*:

$$d\mu(v)=w^{-1}\prod_{\alpha}|(\lambda,\alpha)/(\rho,\alpha)|^2 dv,$$

where w is the order of the Weyl group, ρ is the half-sum of all †positive roots, and α runs over all positive roots (Gel'fand and Naĭmark [12]).

X. Real Semisimple Lie Groups

As in the case of a complex semisimple Lie group, a connected real semisimple Lie group G has four series of irreducible unitary representations. However, if G has no parabolic subgroup other than a minimal parabolic subgroup B and G itself, then G has no representation in the degenerate or complementary degenerate series. Examples of such groups are $SL(2,\mathbf{R})$ and higher-dimensional †Lorentz groups. In general, the classification of irreducible unitary representations in the real semisimple case is more complicated than in the complex semisimple case. Irreducible unitary representations arising from the irreducible decomposition of the regular representation are called representations in the **principal series**. The principal series of G are divided into a finite number of subseries corresponding bijectively to the conjugate classes of the †Cartan subgroups of G.

A connected semisimple Lie group G has a square integrable representation if and only if G has a compact Cartan subgroup H. The set of all square integrable representations of G is called the **discrete series** of irreducible unitary representations. The discrete series is the subseries in the principal series corresponding to a compact Cartan subgroup H. The representations in the discrete series were classified by Harish-Chandra. Let $\mathfrak{h}$ be the Lie algebra of H, P the set of all positive roots in $\mathfrak{h}$ for a fixed linear order, π the polynomial $\prod_{\alpha\in P}H_\alpha$, and $\mathscr{F}$ the set of all real-valued linear forms on $\sqrt{-1}\mathfrak{h}$. Moreover, let L be the set of all linear forms λ in $\mathscr{F}$ such that a single-valued character ξ_λ of the group H is defined by the formula $\xi_\lambda(\exp X)=e^{\lambda(X)}$, and let L' be the set of all λ in L such that $\pi(\lambda)\neq 0$. Then for each λ in L', there exists a representation $\omega(\lambda)$ of G in the discrete series, and conversely, every representation in the discrete series is equivalent to $\omega(\lambda)$ for some λ in L'. Two representations $\omega(\lambda_1)$ and $\omega(\lambda_2)$ $(\lambda_1,\lambda_2\in L')$ are equivalent if and only if there exists an element s in $W_G=N(H)/H$ such that $\lambda_2=s\lambda_1$, where $N(H)$ is the normalizer of H in G (W_G can act on $\mathscr{F}$ as a linear transformation group in the natural way). The value of the character χ_λ on the subgroup H of the representation $\omega(\lambda)$ $(\lambda\in L')$ is given as follows: Let $\varepsilon(\lambda)$ be the signature of $\pi(\lambda)=\prod_{\alpha\in P}\lambda(H_\alpha)$, and define q and Δ by $q=(\dim G/K)/2$ and $\Delta(\exp H)=\prod_{\alpha\in P}(e^{\alpha(H)/2}-e^{-\alpha(H)/2})$. Then the character χ_λ of the representation $\omega(\lambda)$ has the value $(-1)^q\varepsilon(\lambda)\chi_\lambda(h)=\Delta(h)^{-1}\sum_{s\in W_G}(\det s)\xi_{s\lambda}(h)$ on a regular element h in H. The formal degree $d(\omega(\lambda))$ of the representation $\omega(\lambda)$ is given by the formula $d(\omega(\lambda))=C^{-1}[W_G]|\pi(\lambda)|$, where C is a positive constant (not depending on λ) and $[W_G]$ is the order of the finite group W_G (Harish-Chandra [13]). A formula expressing the character χ_λ on the whole set of regular elements in G has been given by T. Hirai [14]. The representations in discrete series are realized on L^2-cohomology spaces of homogeneous holomorphic line bundles over G/H (W. Schmid [15]). They are also realized on the spaces of harmonic spinors on the †Riemannian symmetric space G/K (M. Atiyah and Schmid [16]). They are also realized on the eigenspaces of a Casimir operator acting on the sections of vector bundles on G/K (R. Hotta, *J. Math. Soc. Japan*, 23; N. Wallach [17]). An irreducible unitary representation is called **integrable** if at least one of its matrix coefficients belongs to $L^1(G)$. Integrable representations belong to the discrete series. They have been characterized by H. Hecht and Schmid (*Math. Ann.*, 220 (1976)). The theory of the discrete series is easily extended to reductive Lie groups.

The general principal series representations of a connected semisimple Lie group G with finite center are constructed as follows. Let K be a maximal compact subgroup of G. Then there exists a unique involutive automorphism θ of G whose fixed point set coincides with K. θ is called a **Cartan involution** of G. Let H be a θ-stable Cartan subgroup of G. Then H is the direct product of a compact group $T=H\cap K$ and a vector group A. The centralizer $Z(A)$ of A in G is the direct product of a reductive Lie group $M=\theta(M)$ and A. M has a compact Cartan subgroup T. Hence the set $\hat{M}_d$ of the discrete series representations of M is not empty. Let α be an element of the dual space $\mathfrak{a}^*$ of the Lie algebra $\mathfrak{a}$ of A and put $\mathfrak{g}_\alpha=\{X\in\mathfrak{g}\,|\,[H,X]=\alpha(H)X(\forall H\in\mathfrak{a})\}$ and $\Delta=\{\alpha\in\mathfrak{a}^*\,|\,\mathfrak{g}_\alpha\neq\{0\}\}$. Let Δ^+ be the set of positive elements of Δ in a certain order of $\mathfrak{a}^*$ and put $\mathfrak{n}=\sum_{\alpha\in\Delta^+}\mathfrak{g}_\alpha$ and $N=\exp\mathfrak{n}$. Then $P=MAN$ is a closed subgroup of G. P is called a **cuspidal parabolic subgroup** of G. Let $D\in\hat{M}_d$ and $\nu\in\mathfrak{a}^*$. Then a unitary representation $D\otimes e^{i\nu}$ of P

is defined by $(D \otimes e^{i\nu})(man) = D(m)e^{i\nu(\log a)}$ ($m \in M, a \in A, n \in N$). The unitary representation $\pi_{D,\nu}$ of G induced by $D \otimes e^{i\nu}$ is independent of the choice of Δ^+ up to equivalence. Thus $\pi_{D,\nu}$ depends only on (H, D, ν). The set of representations $\{\pi_{D,\nu} | D \in \hat{M}_d, \nu \in \mathfrak{a}^*\}$ is called the **principal** H-**series**. If ν is regular in $\mathfrak{a}^*$ (i.e., $(\nu, \alpha) \neq 0$ for all $\alpha \in \Delta$), then $\pi_{D,\nu}$ is irreducible. Every $\pi_{D,\nu}$ is a finite sum of irreducible representations. The character $\theta_{D,\nu}$ of $\pi_{D,\nu}$ is a locally summable function which is supported in the closure of $\bigcup_{g \in G} g(MA)g^{-1}$. If two Cartan subgroups H_1 and H_2 are not conjugate in G, then every H_1-series representation is disjoint from every H_2-series representation. Choose a complete system $\{H_1, \ldots, H_r\}$ of conjugacy classes of Cartan subgroups of G. Then every H_i can be chosen as θ-stable. The union of the principal H_i-series $(1 \leqslant i \leqslant r)$ is the principal series of G. The right (or left) regular representation of G is decomposed as the direct integral of the principal series representations. Every complex-valued C^∞-function on G with compact support has an expansion in terms of the matrix coefficients of the principal series representations. Harish-Chandra [18] proved these theorems and determined explicitly the Plancherel measure by studying the asymptotic behavior of the Eisenstein integral [19, 20].

Y. Spherical Functions

Let G be a locally compact †unimodular group and K a compact subgroup of G. The set of all complex-valued continuous functions on G that are invariant under every left translation L_k by elements k in K is denoted by $C(K \backslash G)$. The subset of $C(K \backslash G)$ that consists of all two-sided K-invariant functions is denoted by $C(G, K)$. The subset of $C(G, K)$ consisting of all functions with compact support is denoted by $\mathbf{L} = L(G, K)$. $\mathbf{L}$ is an algebra over $\mathbf{C}$ if the product of two elements f and g in $\mathbf{L}$ is defined by the convolution.

Let λ be an algebra homomorphism from $\mathbf{L}$ into $\mathbf{C}$. Then an element of the eigenspace $F(\lambda) = \{\psi \in C(K, G) | f * \psi = \lambda(f)\psi \ (\forall f \in \mathbf{L})\}$ is called a **spherical function** on $K \backslash G$. If $F(\lambda)$ contains a nonzero element, then $F(\lambda)$ contains a unique two-sided K-invariant element ω normalized by $\omega(e) = 1$, where e is the identity element in G. This function ω is called the **zonal spherical function** associated with λ. In this case, the homomorphism λ is defined by $\lambda(f) = \int_G f(g)\omega(g^{-1})dg$. Hence the eigenspace $F(\lambda)$ is uniquely determined by the zonal spherical function ω. A function $\omega \neq 0$ in $C(G, K)$ is a zonal spherical function on $K \backslash G$ if and only if ω satisfies either of the following two conditions: (i) The mapping $f \mapsto \int f(g)\omega(g^{-1})dg$ is an algebra homomorphism of $\mathbf{L}$ into $\mathbf{C}$; (ii) ω satisfies the functional equation

$$\int_K \omega(gkh)\,dk = \omega(g)\omega(h).$$

When G is a Lie group, every spherical function is a real analytic function on $K \backslash G$.

Z. Expansion by Spherical Functions

In this section, we assume that the algebra $\mathbf{L}$ of two-sided K-invariant functions is commutative. In this case there are sufficiently many spherical functions of $K \backslash G$, and two-sided K-invariant functions are expanded by spherical functions. An irreducible unitary representation U of G is called a **spherical representation** with respect to K if the representation space $\mathfrak{H}(U)$ contains a nonzero vector invariant under every U_k, where k runs over K. By the commutativity of $\mathbf{L}$, the K-invariant vectors in $\mathfrak{H}(U)$ form a 1-dimensional subspace. Let x be a K-invariant vector in $\mathfrak{H}(U)$ with the norm $\|x\| = 1$. Then $\omega(g) = (U_g x, x)$ is a zonal spherical function on $K \backslash G$, and for every y in $\mathfrak{H}(U)$, the function $\varphi_y(g) = (U_g x, y)$ is a spherical function associated with ω. Moreover, in this case the zonal spherical function ω is a positive definite function on G. Conversely, every positive definite zonal spherical function ω can be expressed as $\omega(g) = (U_g x, x)$ for some spherical representation U and some K-invariant vector x in $\mathfrak{H}(U)$.

The set of all positive definite zonal spherical functions becomes a locally compact space Ω by the topology of compact convergence. The **spherical Fourier transform** $\hat{f}$ of a function f in $L_1(K \backslash G)$ is defined by

$$\hat{f}(\omega) = \int_G f(g)\omega(g^{-1})\,dg.$$

There exists a unique †Radon measure μ on Ω such that for every f in $\mathbf{L}$, $\hat{f}$ belongs to $L_2(\Omega, \mu)$. Also, the Plancherel formula

$$\int_G f(s)\overline{g(s)}\,ds = \int_\Omega \hat{f}(\omega)\overline{\hat{g}(\omega)}\,d\mu(\omega) \tag{4}$$

holds for every f and g in $\mathbf{L}$, and an inversion formula $f(s) = \int_\Omega \hat{f}(\omega)\omega(s)\,d\mu(\omega)$ holds for a sufficiently nice two-sided K-invariant function f [21]. Identifying a positive definite zonal spherical function with the corresponding spherical representation, we can regard Ω as a subset of the dual $\hat{G}$ of G. The Plancherel formula for two-sided K-invariant functions is obtained from the general Plancherel formula on G by restricting the domain of the integral from $\hat{G}$ to Ω. When G is a Lie group and $\mathbf{L}$ is commutative, a spherical function on $K \backslash G$ can be characterized as a simultaneous eigenfunction of G-invariant linear differential operators on $K\backslash G$.

AA. Spherical Function on Symmetric Spaces

The most important case where the algebra $\mathbf{L}=L(G,K)$ is commutative is when $K\backslash G$ is a †weakly symmetric Riemannian space or, in particular, a †symmetric Riemannian space When $K\backslash G$ is a compact symmetric Riemannian space, a spherical representation with respect to K is the irreducible component of the regular representation T on $K\backslash G$, and a spherical function on $K\backslash G$ is a function that belongs to the irreducible subspaces in $L_2(K\backslash G)$. In particular, if G is a compact connected semisimple Lie group, the highest weights of spherical representations of G with respect to K are explicitly given by using the **Satake diagram** of $K\backslash G$. The Satake diagram of $K\backslash G$ is the †Satake diagram of the noncompact symmetric Riemannian space $K\backslash G_0$ dual to $K\backslash G$ or the Satake diagram of the Lie algebra of G_0. If a symmetric space is the underlying manifold of a compact Lie group G, then G can be expressed as $G=K\backslash(G\times G)$, where K is the diagonal subgroup of $G\times G$. In this case, a zonal spherical function ω on $G=K\backslash(G\times G)$ is the normalized character of an irreducible unitary representation U of $G:\omega(g)=(\deg U)^{-1}T_rU_g$. The explicit form of ω is given by †Weyl's character formula (→ 249 Lie Groups).

The zonal spherical functions on a symmetric Riemannian space $K\backslash G$ of noncompact type are obtained in the following way: Let G be a connected semisimple Lie group with finite center, K be a maximal compact subgroup of G, and $G=NA_+K$ be an †Iwasawa decomposition. Then for any g in G there exists a unique element $H(g)$ in the Lie algebra $\mathfrak{a}_+$ of A_+ such that g belongs to $N\exp H(g)K$. Let $\mathfrak{a}$ be a Cartan subalgebra containing $\mathfrak{a}_+$, P be the set of all positive roots in $\mathfrak{a}$, and $\rho=(\sum_{\alpha\in P}\alpha)/2$. Then for any complex-valued linear form v on $\mathfrak{a}_+$, the function

$$\omega_v(g)=\int_K e^{(iv-\rho)(H(kg))}dk$$

is a zonal spherical function on the symmetric Riemannian space $K\backslash G$. Conversely, every zonal spherical function ω on $K\backslash G$ is equal to ω_v for some v. Two zonal spherical functions ω_v and $\omega_{v'}$ coincide if and only if v and v' are conjugate under the operation of the Weyl group $W_0=N_K(A)/Z_K(A)$ of $K\backslash G$ (Harish-Chandra [22], S. Helgason [23]). If v is real-valued, then ω_v is positive definite. Such a zonal spherical function ω_v is obtained from a spherical representation belonging to the principal A-series. Let Ω_0 be the set of all zonal spherical functions ω_v associated with the real-valued linear form v. Then the support of the Plancherel measure μ on $K\backslash G$ is contained in Ω_0. We can choose v as a parameter on the space Ω_0. Then the right-hand side of the Plancherel formula can be expressed as an integral over the dual space L of $\mathfrak{a}_+$. Moreover, the Plancherel measure μ is absolutely continuous with respect to the Lebesgue measure dv on the Euclidean space L and can be expressed as

$$d\mu(\omega_v)=\omega_0^{-1}|c(v)|^{-2}dv$$

under suitable normalization of μ and dv. The problem of calculating the function $c(v)$ can be reduced to the case of symmetric spaces of rank 1 and can be solved explicitly. Let p_α be the multiplicity of a restricted root α and $I(v)$ be the product

$$I(v)=\prod_\alpha B(\tfrac{1}{2}p_\alpha,\tfrac{1}{4}p_\alpha+(v,\alpha)(\alpha,\alpha)^{-1}),$$

where α runs over all positive restricted roots and B is the †beta function. Then $c(v)=I(iv)/I(\rho)$ [20, 24]. Every spherical function f on $K\backslash G$ is expressed as the Poisson integral of its "boundary values" on the Martin boundary $P\backslash G$ of $K\backslash G$, where $P=MA_+N$ is a minimal parabolic subgroup of G. The boundary values of f form a hyperfunction with values in a line bundle over $P\backslash G$ (K. Okamoto et al. [25]).

BB. Spherical Functions and Special Functions

Some important special functions are obtained as the zonal spherical functions on a certain symmetric Riemannian space $M=K\backslash G$ (G is the motion group of M). In particular when M is of rank 1, then the zonal spherical functions are essentially the functions of a single variable. For example, the zonal spherical functions on an n-dimensional Euclidean space can be expressed as

$$\omega_v(r)=2^m\Gamma(m+1)(vr)^{-m}J_m(vr),$$

where $2m=n-2$ and J_m is the †Bessel function of the mth order. The zonal spherical function on an $(n-1)$-dimensional sphere $S^{n-1}=SO(n-1)\backslash SO(n)$ is given by

$$\omega_v(\theta)=\Gamma(v+1)\Gamma(n-2)\Gamma(v+n-2)^{-1}C_v^m(\cos\theta)\qquad(v=0,1,2,\ldots),$$

where $C_v^m(z)$ is the †Gegenbauer polynomial. The zonal spherical functions on an $(n-1)$-dimensional Lobachevskiĭ space can be expressed as

$$\omega_v(t)=2^{m-1/2}\Gamma(m+1/2)\sinh^{-m+1/2}t\times\mathfrak{P}^{1/2-m}_{-1/2-m+v}(\cosh t)$$

using a generalized †associated Legendre function $\mathfrak{P}_\nu^\mu$. Many properties of special functions can be proved from a group-theoretic point of view. For example, the addition theorem is merely the homomorphism property $U_{gh}=U_g U_h$ expressed in terms of the matrix components of U. The differential equation satisfied by these special functions is derived from the fact that a zonal spherical function ω is an eigenfunction of an invariant differential operator. The integral expression of such a special function can be obtained by constructing a spherical representation U in a certain function space and calculating explicitly the inner product in the expression $\omega(g)=(U_g x, x)$ (N. Ya. Vilenkin [26]).

CC. Generalization of the Theory of Spherical Functions

The theory of spherical functions described in Sections Y–BB can be generalized in several ways. First, spherical functions are related to the trivial representation of K. A generalization is obtained if the trivial representation of K is replaced by an irreducible representation of K. The theory of such zonal spherical functions is useful for representation theory [20]. For example, the Plancherel formula for $SL(2, \mathbf{R})$ can be obtained using such spherical functions (R. Takahashi, *Japan. J. Math.*, 31 (1961)). Harish-Chandra's Eisenstein integral is such a spherical function on a general semisimple Lie group G. He used it successfully to obtain the Plancherel measure of G. Another generalization can be obtained by removing the condition that K is compact. In particular, when $K\backslash G$ is a symmetric homogeneous space of a Lie group G, the algebra $\mathscr{D}$ of all G-invariant linear differential operators is commutative if the space $K\backslash G$ has an invariant volume element. In this case, a spherical function on $K\backslash G$ can be defined as a simultaneous eigenfunction of $\mathscr{D}$. The character of a semisimple Lie group is a zonal spherical function (distribution) in this sense. The spherical functions and harmonic analysis on symmetric homogeneous space have been studied by T. Oshima and others. T. Oshima and J. Sekiguchi [27] proved the Poisson integral theorem (→ Section AA) for a certain kind of symmetric homogeneous spaces.

The spherical functions and unitary representations of topological groups that are not locally compact are studied in connection with probability theory and physics. For example, the zonal spherical functions of the rotation group of a real Hilbert space are expressed by Hermite polynomials.

DD. Discontinuous Subgroups and Representations

Let G be a connected semisimple Lie group and Γ be a discrete subgroup of G. Then the regular representation T of G on $\Gamma\backslash G$ is defined by $(T_g f)(x)=f(xg)$ $(f\in L^2(\Gamma\backslash G))$. The problem of decomposing the representation T into irreducible components is important in connection with the theory of †automorphic forms and number theory. First assume that the quotient space $\Gamma\backslash G$ is compact. Then for every function f in $L_1(G)$, the operator $T(f)$ is a compact operator. Hence the regular representation T on $\Gamma\backslash G$ can be decomposed into the discrete sum $T=\sum_{k=1}^{\infty} T^{(k)}$ of irreducible unitary representations $T^{(k)}$, and the multiplicity of every irreducible component is finite. The irreducible unitary representation U of G is related to the automorphic forms of Γ in the following way: Let x be a nonzero element in the representation space $\mathfrak{H}=\mathfrak{H}(U)$ of U. $\mathfrak{H}$ is topologized into a †locally convex topological vector space $\mathfrak{H}_x$ by the set N_x of †seminorms: $N_x=\{P_C(y)=\max_{g\in C}|(U_g x, y)|\}$, where C runs over all compact subsets in G. The topology $\mathscr{T}_x$ of $\mathfrak{H}_x$ is independent of the choice of x provided that $\dim\{T_k x\,|\,k\in K\}<\infty$, where K is a maximal compact subgroup of G. Let $\mathfrak{H}^*$ be the completion of $\mathfrak{H}_x$ with respect to the topology $\mathscr{T}_x$ (the completion is independent of the choice of x). $\mathfrak{H}^*$ contains the original Hilbert space $\mathfrak{H}$ as a subspace. Then the representation U of G on $\mathfrak{H}$ can be extended to a representation U^* of G on the space $\mathfrak{H}^*$. An element f in $\mathfrak{H}^*$ invariant under U_γ^* for every γ in Γ is called an **automorphic form** of Γ of type U. Then the multiplicity of an irreducible representation U in the regular representation T on $\Gamma\backslash G$ is equal to the dimension of the vector space consisting of all automorphic forms of type U. This theorem is called the **Gel'fand–Pyatetskiĭ-Shapiro reciprocity law** [28]. Let $T=\sum_{k=1}^{\infty} T^{(k)}$ be the irreducible decomposition of T and χ_k be the character of the irreducible unitary representation $T^{(k)}$. Then for a suitable function f on G, the integral operator K_f on $\mathfrak{H}(T)=L^2(T\backslash G)$ with kernel $k_f(x,y)=\sum_{\gamma\in\Gamma} f(x^{-1}\gamma y)$ belongs to the trace class. By calculating the trace of K_f in two ways, the following **trace formula** is obtained:

$$\sum_{k=1}^{\infty}\int_G f(g)\chi_k(g)\,dg=\sum_{\{\gamma\}}\int_{\mathfrak{D}_\gamma} f(x^{-1}\gamma x)\,dx,$$

where $\{\gamma\}$ is the conjugate class of γ in Γ and $\mathfrak{D}_\gamma$ is the quotient space of the centralizer G_γ of γ in G by the centralizer Γ_γ of γ in Γ.

When the groups G and Γ are given explicitly, the right-hand side of the trace formula can be expressed in a more explicit form, and the trace formula leads to useful consequences. A similar trace formula holds for the unitary representation U^L induced by a finite-dimensional unitary representation L of Γ instead of the regular representation T on $\Gamma\backslash G$. When the quotient space $\Gamma\backslash G$ is not compact, the irreducible decomposition of the regular representation T on $\Gamma\backslash G$ contains not only the discrete direct sum but also the direct integral (continuous spectrum). A. Selberg showed that even in this case, there are explicit examples for which the trace formula holds for the part with discrete spectrum. Also, the part with continuous spectrum can be described by the †generalized Eisenstein series. Analytic properties and the functional equation of the generalized Eisenstein series have been studied by R. Langlands [30]. Recent developments are surveyed in [31].

EE. History

Finite-dimensional unitary representations of a finite group were studied by Frobenius and Schur (1896–1905). In 1925, †Weyl studied the finite-dimensional unitary representation of compact Lie groups. The theory of infinite-dimensional unitary representation was initiated in 1939 by E. P. Wigner in his work on the inhomogeneous Lorentz group, motivated by problems of quantum mechanics.

In 1943, Gel'fand and D. A. Raikov proved the existence of sufficiently many irreducible unitary representations for an arbitrary locally compact group. The first systematic studies of unitary representations appeared in 1947 in the work of V. Bargmann on $SL(2, \mathbf{R})$ [31] and the work of Gel'fand and Neumark on $SL(2, \mathbf{C})$. Gel'fand and Naĭmark established the theory of unitary representation for complex semisimple Lie groups [12].

Harish-Chandra proved theorems concerning the unitary representations of a general semisimple Lie group; for instance, he proved that a semisimple Lie group G is of type I [7] and defined the character of a unitary representation of G and proved its basic properties [9, III; 10]. Harish-Chandra also determined the discrete series of G and their characters. Harish-Chandra [18] proved the Plancherel formula for an arbitrary connected semisimple Lie group G with finite center. Hence harmonic analysis of square integrable functions on G is established.

Further studies on harmonic analysis on semisimple Lie groups have been carried out. In particular, Paley-Wiener–type theorems, which determine the Fourier transform image of the space $C_C^\infty(G)$ of C^∞-functions with compact support, have been proved for the group $PSL(2, \mathbf{R})$ (L. Ehrenpreis and F. Mautner [33]), complex semisimple Lie groups (Zhelobenko [34]), and two-sided K-invariant functions on general semisimple Lie groups (R. Gangolli [35]). A. W. Knapp and E. M. Stein [36] studied the intertwining operators.

Concerning the construction of irreducible representations, G. W. Mackey [3] and Bruhat [4] developed the theory of induced representations of locally compact groups and Lie groups, respectively. B. Kostant [37] (see Blattner's article in [38]) noticed a relation between homogeneous †symplectic manifolds and unitary representations and proposed a method of constructing irreducible unitary representations of a Lie group. Selberg's research [29] revealed a connection between unitary representations (or spherical functions) and the theory of automorphic forms and number theory. A number of papers along these lines have since appeared [31]. In connection with number-theoretic investigations of an †algebraic group defined over an algebraic number field, unitary representations of the †adele group of G or an algebraic group over a †p-adic number field have been studied (→ [31, 38], Gel'fand, M. I. Grayev, and I. I. Pyatetskiĭ-Shapiro [39], and H. M. Jacquet and R. P. Langlands [40]).

For the algebraic approach to the infinite-dimensional representations of semisimple Lie groups and Lie algebras → [41].

For surveys of the theory of unitary representations → [2, 19, 20, 31, 38].

References

[1] N. Tatsuuma, A duality theorem for locally compact groups, J. Math. Kyoto Univ., 6 (1967), 187–293.
[2] J. Dixmier, Les C^*-algèbres et leurs représentations, Gauthier-Villars, 1964.
[3] G. W. Mackey, Induced representations of locally compact groups I, Ann. Math., 55 (1952), 101–139.
[4] F. Bruhat, Sur les representations induites des groupes de Lie, Bull. Soc. Math. France, 84 (1956), 97–205.
[5] R. Bott, Homogeneous vector bundles, Ann. Math., 66 (1957), 203–248.
[6] E. Nelson, Analytic vectors, Ann. Math., 70 (1959), 572–615.
[7] A. A. Kirillov, Unitary representations of nilpotent Lie groups, Russian Math. Surveys, 17, no. 4 (1962), 53–104. (Original in Russian, 1962.)
[8] L. Auslander and B. Kostant, Polarization

and unitary representations of solvable Lie groups, Inventiones Math., 14 (1971), 255–354.
[9] Harish-Chandra, Representations of a semisimple Lie group on a Banach space I–VI, Trans. Amer. Math. Soc., 75 (1953), 185–243; 76 (1954), 26–65; 76 (1954), 234–253; Amer. J. Math., 77 (1955), 743–777; 78 (1956), 1–41; 78 (1956), 564–628.
[10] Harish-Chandra, Invariant eigendistributions on a semisimple Lie group, Trans. Amer. Math. Soc., 119 (1965), 457–508.
[11] E. M. Stein, Analysis in matrix spaces and some new representations of $SL(N, \mathbf{C})$, Ann. Math., (2) 86 (1967), 461–490.
[12] I. M. Gel'fand and M. A. Naĭmark (Neumark), Unitäre Darstellungen der klassischen Gruppen, Akademische Verlag., 1957. (Original in Russian, 1950.)
[13] Harish-Chandra, Discrete series for semisimple Lie groups I, II, Acta Math., 113 (1965), 241–318; 116 (1966), 1–111.
[14] T. Hirai, The characters of the discrete series for semisimple Lie groups, J. Math. Kyoto Univ., 21 (1981), 417–500.
[15] W. Schmid, L^2-cohomology and the discrete series, Ann. Math., 102 (1976), 375–394.
[16] M. F. Atiyah and W. Schmid, A geometric construction of the discrete series for semisimple Lie groups, Inventiones Math., 42 (1977), 1–62.
[17] N. Wallach, On the Enright-Varadarajan modules: A construction of the discrete series, Ann. Sci. Ecole Norm. Sup., 9 (1976), 81–102.
[18] Harish-Chandra, Harmonic analysis on real reductive groups I–III, J. Functional Anal., 19 (1975), 104–204; Inventiones Math., 36 (1976), 1–55; Ann. Math., 104 (1976), 117–201.
[19] J. A. Wolf, M. Cahen, and M. De Wilde (eds.), Harmonic analysis and representations of semisimple Lie groups, Reidel, 1980.
[20] G. Warner, Harmonic analysis on semisimple Lie groups I, II, Springer, 1972.
[21] H. Yoshizawa, A proof of the Plancherel theorem, Proc. Japan Acad., 30 (1954), 276–281.
[22] Harish-Chandra, Spherical functions on a semisimple Lie group I, II, Amer. J. Math., 80 (1958), 241–310, 553–613.
[23] S. Helgason, Differential geometry and symmetric spaces, Academic Press, 1962.
[24] J. Rosenberg, A quick proof of Harish-Chandra's Plancherel theorem for spherical functions on a semisimple Lie groups, Proc. Amer. Math. Soc., 63 (1977), 143–149.
[25] M. Kashiwara, A. Kowata, K. Minemura, K. Okamoto, T. Oshima, and M. Tanaka, Eigenfunctions of invariant differential operators on a symmetric space, Ann. Math., 107 (1978), 1–39.
[26] N. Ya Vilenkin, Special functions and theory of group representations, Amer. Math. Soc. Transl. of Math. Monographs, 22, 1968. (Original in Russian, 1965.)
[27] T. Oshima and J. Sekiguchi, Eigenspaces of invariant differential operators on an affine symmetric space, Inventiones Math., 57 (1980), 1–81.
[28] I. M. Gel'fand and I. I. Pyatetskiĭ-Shapiro, Theory of representations and theory of automorphic functions, Amer. Math. Soc. Transl., (2) 26 (1963), 173–200. (Original in Russian, 1959.)
[29] A. Selberg, Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series, J. Indian Math. Soc., 20 (1956), 47–87.
[30] R. P. Langlands, On the functional equations satisfied by Eisenstein series, Lecture notes in math. 544, Springer, 1976.
[31] Automorphic forms, representations and L-functions, Amer. Math. Soc. Proc. Symposia in Pure Math., 33, 1979.
[32] V. Bargmann, Irreducible unitary representations of the Lorentz group, Ann. Math., (2) 48 (1947), 568–640.
[33] L. Ehrenpreis and F. I. Mautner, Some properties of the Fourier transform on semisimple Lie groups I–III, Ann. Math., (2) 61 (1955), 406–439; Trans. Amer. Math. Soc., 84 (1957), 1–55; 90 (1959), 431–484.
[34] D. P. Zhelobenko, Harmonic analysis of functions on semisimple Lie groups I, II, Amer. Math. Soc. Transl., (2) 54 (1966), 177–230. (Original in Russian, 1963, 1969.)
[35] R. Gangolli, On the Plancherel formula and the Paley-Wiener theorem for spherical functions on a semisimple Lie group, Ann. Math., 93 (1971), 150–165.
[36] A. W. Knapp and E. M. Stein, Intertwining operators for semisimple groups, Ann. Math., 93 (1971), 489–578.
[37] B. Kostant, Quantization and unitary representations, Lecture notes in math. 170, Springer, 1970, 87–207.
[38] Harmonic analysis on homogeneous spaces, Amer. Math. Soc. Proc. Symposia in Pure Math., 26 (1973).
[39] I. M. Gel'fand, M. I. Grayev, and I. I. Pyatetskiĭ-Shapiro, Representation theory and automorphic functions, Saunders, 1969. (Original in Russian, 1966.)
[40] H. Jacquet and R. P. Langlands, Automorphic forms on $GL(2)$, Lecture notes in math. 114, Springer, 1970.
[41] D. A. Vogan, Representations of real reductive Lie groups, Birkhäuser, 1981.

438 (XI.5) Univalent and Multivalent Functions

A. General Remarks

A single-valued †analytic function $f(z)$ defined in a domain D of the complex plane is said to be **univalent** (or **simple** or **schlicht**) if it is injective, i.e., if $f(z_1) \neq f(z_2)$ for all distinct points z_1, z_2 in D. A multiple-valued function $f(z)$ is also said to be univalent if its distinct function elements always attain distinct values at their centers. The derivative of a univalent function is never zero. The limit function of a †uniformly convergent sequence of univalent functions is univalent unless it reduces to a constant. When $f(z)$ is single-valued, the univalent function $w=f(z)$ gives rise to a one-to-one †conformal mapping between D and its image $f(D)$.

B. Univalent Functions in the Unit Disk

A systematic theory of the family of functions †holomorphic and univalent in the unit disk originates from a **distortion theorem** obtained by P. Koebe (1909) in connection with the uniformization of analytic functions. In general, distortion theorems are theorems for determining bounds of functionals, such as $|f(z)|$, $|f'(z)|$, $\arg f'(z)$, within the family under consideration. In particular, distortion theorems concerning the bounds of the arguments of $f(z)$ and $f'(z)$ are also called **rotation theorems**. Though results were at first qualitative, they were made quantitative subsequently by L. Bieberbach (1916), G. Faber (1916), and others. Any univalent function $f(z)$ holomorphic in the unit disk and normalized by $f(0)=0$ and $f'(0)=1$ satisfies the **distortion inequalities**

$$\frac{|z|}{(1+|z|)^2} \leqslant |f(z)| \leqslant \frac{|z|}{(1-|z|)^2},$$

$$\frac{1-|z|}{(1+|z|)^3} \leqslant |f'(z)| \leqslant \frac{1+|z|}{(1-|z|)^3}.$$

Here the equality holds only if $f(z)$ is of the form $z/(1-\varepsilon z)^2$ $(|\varepsilon|=1)$. In deriving these inequalities, Bieberbach centered his attention on the family of †meromorphic functions $g(\zeta) = \zeta + \sum_{\nu=0}^{\infty} b_\nu \zeta^{-\nu}$ univalent in $|\zeta|>1$. He established the **area theorem** $\sum_{\nu=1}^{\infty} \nu |b_\nu|^2 \leqslant 1$, which illustrates the fact that the area of the complementary set of the image domain is nonnegative. Bieberbach, R. Nevanlinna (1919–1920), and others constructed a systematic theory of univalent functions in the unit disk based on this theorem.

After the area theorem, the chief tools in the theory of univalent functions have been Löwner's method, the method of contour integration, the variational method, and the method of the extremal metric. In contrast to the theory of univalent functions based on Bieberbach's area theorem, K. Löwner (1923) introduced a new method. In view of a theorem on the domain kernel (C. Carathéodory, 1912), it suffices to consider an everywhere dense subfamily in order to estimate a continuous functional within the family of univalent functions holomorphic in the unit disk. Löwner used the subfamily of functions mapping the unit disk onto the so-called bounded slit domains. Namely, the range of a member of this subfamily consists of the unit disk slit along a Jordan arc that starts at a periphery point and does not pass through the origin. A mapping function of this nature is determined as the integral $f(z,t_0)$ of **Löwner's differential equation**

$$\frac{\partial f(z,t)}{\partial t} = -f(z,t)\frac{1+\kappa(t)f(z,t)}{1-\kappa(t)f(z,t)}, \qquad 0 \leqslant t \leqslant t_0,$$

with the initial condition $f(z,0)=z$, where $\kappa(t)$ is a continuous function with absolute value equal to 1. Any univalent function $f(z)$ holomorphic in the unit disk and satisfying $f(0)=0$, $f'(0)=1$ has an arbitrarily close approximation by functions of the form $e^{t_0}f(z,t_0)$. By means of this differential equation Löwner proved that $|a_3| \leqslant 3$ for any univalent function $f(z)=z+\sum_{n=2}^{\infty} a_n z^n$ $(|z|<1)$ and also derived a decisive estimate concerning a coefficient problem for the inverse function [2].

G. M. Golusin (1935) and I. E. Bazilevich (1936) first noticed that Löwner's method is also a powerful tool for deriving several distortion theorems. They showed that classical distortion theorems can be derived in more detailed form (Golusin, *Mat. Sb.*, 2 (1937), 685); in particular, Golusin (1938) obtained a precise estimate concerning the rotation theorem, i.e.,

$$|\arg f'(z)| \leqslant \begin{cases} 4\arcsin|z|, & |z|<1/\sqrt{2}, \\ \pi+\log(|z|^2/(1-|z|^2)), & 1/\sqrt{2} \leqslant |z| < 1. \end{cases}$$

Löwner's method was also investigated by A. C. Schaeffer and D. C. Spencer (1945) [8].

The method of contour integration was introduced by H. Grunsky. It starts with some 2-dimensional integral which can be shown to be positive. Transforming it into a boundary integral and using the †residue theorem, we obtain an appropriate inequality by means of this integral. By this method Grunsky established the following useful inequality (*Math.*

Z., 45 (1939)). For $g(\zeta)=\zeta+\sum_{\nu=0}^{\infty}b_\nu\zeta^{-\nu}$, which is univalent in $|\zeta|>1$, let

$$\log\frac{g(z)-g(\zeta)}{z-\zeta}=\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}c_{mn}z^{-m}\zeta^{-n}\qquad(|z|>1,|\zeta|>1).$$

The coefficients c_{mn} are polynomials in the coefficients b_ν of g. Then **Grunsky's inequality** is: For each integer N and for all complex numbers $\lambda_1,\ldots,\lambda_N$,

$$\left|\sum_{m=1}^{N}\sum_{n=1}^{N}c_{mn}\lambda_m\lambda_n\right|\leqslant\sum_{n=1}^{N}\frac{1}{n}|\lambda_n|^2.$$

It is known that if this inequality holds for an arbitrary integer N and for all complex numbers $\lambda_1,\ldots,\lambda_N$, then $g(\zeta)$ is univalent in $|\zeta|>1$. There are several generalizations of Grunsky's inequality [13].

The variational method was first developed by M. Schiffer for application to the theory of univalent functions. He first used boundary variations (*Proc. London Math. Soc.*, 44 (1938)) and later interior variations (*Amer. J. Math.*, 65 (1943)). The problem of maximizing a given real-valued functional on a family of univalent functions is called an extremal problem, and a function for which the functional attains its maximum is called an extremal function. The **variational method** is used to uncover characteristic properties of an extremal function by comparing it with nearby functions. Typical results are the qualitative information that the extremal function maps the disk $|z|<1$ onto the complement of a system of analytic arcs satisfying a differential equation and that the extremal function satisfies a differential equation. Following Schiffer, Schaeffer and Spencer [8] and Golusin (*Math. Sb.*, 19 (1946)) gave variants of the method of interior variations.

H. Grötzsch (1928–1934) treated the theory of univalent functions in a unified manner by the method of the †extremal metric. The idea of this method is to estimate the length of curves and the area of some region swept out by them together with an application of †Schwarz's inequality (→ 143 Extremal Length). After Grötzsch, the method of the extremal metric has been used by many authors. In particular, O. Teichmüller, in connection with this method, formulated the principle that the solution of a certain type of extremal problem is in general associated with a †quadratic differential, although he did not prove any general result realizing this principle in concrete form. J. A. Jenkins gave a concrete expression of the Teichmüller principle; namely, he established the general coefficient theorem and showed that this theorem contains as special cases a great many of the known results on univalent functions [11].

Univalence criteria have been given by various authors. In particular, Z. Nehari (*Bull. Amer. Math. Soc.*, 55 (1949)) proved that if $|\{f(z),z\}|\leqslant 2(1-|z|^2)^{-2}$ in $|z|<1$, then $f(z)$ is univalent in $|z|<1$, and E. Hille (*Bull. Amer. Math. Soc.*, 55 (1949)) proved that 2 is the best possible constant in the above inequality. Here, $\{f(z),z\}$ denotes the †Schwarzian derivative of $f(z)$ with respect to z:

$$\{f(z),z\}=\left(\frac{f''(z)}{f'(z)}\right)'-\frac{1}{2}\left(\frac{f''(z)}{f'(z)}\right)^2.$$

C. Coefficient Problems

In several distortion theorems **Koebe's extremal function** $z/(1-\varepsilon z)^2=\sum_{n=1}^{\infty}n\varepsilon^{n-1}z^n(|\varepsilon|=1)$ is extensively utilized. Concerning this, Bieberbach stated the following conjecture. If $f(z)=z+\sum_{n=2}^{\infty}a_nz^n$ is holomorphic and univalent in $|z|<1$, then $|a_n|\leqslant n$ $(n=2,3,\ldots)$, with equality holding only for Koebe's extremal function $z/(1-\varepsilon z)^2$ $(|\varepsilon|=1)$. This conjecture was solved affirmatively by L. de Branges in 1985 after enormous effort by many mathematicians, as described below.

Bieberbach (1916, [1]) proved $|a_2|\leqslant 2$ as a corollary to the area theorem. This result can be proved easily by most of the methods. In 1923 Löwner [2] proved $|a_3|\leqslant 3$, introducing his own method. Schaeffer and Spencer gave a proof of $|a_3|\leqslant 3$ by the variational method (*Duke Math. J.*, 10 (1943)). Furthermore, Jenkins used the method of the extremal metric to prove a coefficient inequality that implies $|a_3|\leqslant 3$ (*Analytic Functions*, Princeton Univ. Press, 1960). The problem of the fourth coefficient remained open until 1955, when P. R. Garabedian and Schiffer [3] proved $|a_4|\leqslant 4$ by the variational method. Their proof was extremely complicated. Subsequently, Z. Charzynski and Schiffer gave an alternative brief proof of $|a_4|\leqslant 4$ by using the Grunsky inequality (*Arch. Rational Mech. Anal.*, 5 (1960)). M. Ozawa (1969, [4]) and R. N. Pederson (1968, [5]) also used the Grunsky inequality to prove $|a_6|\leqslant 6$. In 1972, Pederson and Schiffer [6] proved $|a_5|\leqslant 5$. They applied the Garabedian-Schiffer inequality, a generalization of the Grunsky inequality which Garabedian and Schiffer had derived by the variational method.

On the other hand, W. K. Hayman [7] showed that for each fixed $f(z)=z+\sum_{n=2}^{\infty}a_nz^n$,

$$\lim_{n\to\infty}\frac{|a_n|}{n}=\alpha\leqslant 1,$$

with the equality holding only for Koebe's extremal function $z/(1-\varepsilon z)^2$ $(|\varepsilon|=1)$. Further, it was shown that Koebe's extremal function $z/(1-z)^2$ gives a local maximum for the nth coefficient in the sense that $\mathrm{Re}\{a_n\} \leqslant n$ whenever $|a_2-2|<\delta_n$ for some $\delta_n>0$ (Garabedian, G. G. Ross, and Schiffer, *J. Math. Mech.*, 14 (1964); E. Bombieri, *Inventiones Math.*, 4 (1967); Garabedian and Schiffer, *Arch. Rational Math. Anal.*, 26 (1967)).

In the most general form, the coefficient problem is to determine the region occupied by the points $(a_2, \ldots, a_n)$ for all functions $f(z)=z+\sum_{n=2}^{\infty} a_n z^n$ univalent in $|z|<1$. Schaeffer and Spencer [8] found explicitly the region for (a_2, a_3).

For the coefficients of functions $g(\zeta)=\zeta+\sum_{v=0}^{\infty} b_v \zeta^{-v}$ univalent in $|\zeta|>1$, the following results are known: $|b_1| \leqslant 1$ (Bieberbach [1]), $|b_2| \leqslant 2/3$ (Schiffer, *Bull. Soc. Math. France*, 66 (1938); Golusin, *Mat. Sb.*, 3 (1938)), $|b_3| \leqslant 1/2+e^{-6}$ (Garabedian and Schiffer, *Ann. Math.*, (2) 61 (1955)).

D. Other Classes of Univalent Functions

We have discussed the general family of functions univalent in the unit disk. There are also several results on distortion theorems and coefficient problems for subfamilies determined by conditions such as that the images are bounded, †starlike with respect to the origin, or †convex. For instance, if $f(z)=z+\sum_{n=2}^{\infty} a_n z^n$ is holomorphic and univalent in $|z|<1$ and its image is starlike with respect to the origin, then $|a_n| \leqslant n$ $(n=2,3,\ldots)$. If the image of $f(z)$ is convex, then $f(z)$ satisfies $|a_n| \leqslant 1$ $(n=2,3,\ldots)$ and the distortion inequalities

$$\frac{|z|}{1+|z|} \leqslant |f(z)| \leqslant \frac{|z|}{1-|z|},$$

$$\frac{1}{(1+|z|)^2} \leqslant |f'(z)| \leqslant \frac{1}{(1-|z|)^2}.$$

Here the equality sign appears at z_0 $(0<|z_0|<1)$ if and only if $f(z)$ is of the form $z/(1+\varepsilon z)$ with $\varepsilon = \pm |z_0|/z_0$.

On the other hand, problems on conformal mappings of multiply connected domains involve essential difficulties in comparison with the simply connected case. Although Bieberbach's method is unsuitable for multiply connected domains, Löwner's method, the method of contour integration, the variational method, and the method of the extremal metric remain useful (→ 77 Conformal Mappings).

E. Multivalent Functions

Multivalent functions are a natural generalization of univalent functions. There are several results that generalize classical results on univalent functions.

A function $f(z)$ that attains every value at most p times and some values exactly p times in a domain D is said to be **p-valent** in D and is called a **multivalent function** provided that $p>1$. In order for $f(z)=\sum_{n=0}^{\infty} a_n z^n$, holomorphic in $|z| \leqslant 1$, to be p-valent there, it is sufficient that it satisfies

$$p-1<\mathrm{Re}(zf'(z)/f(z))<p+1$$

on $|z|=1$. Hence it suffices to have

$$|a_p| - \sum_{n=2}^{p} n|a_{p+1-n}| > \sum_{n=2}^{\infty} n|a_{p-1+n}|.$$

If $f(z)=(1+a_1 z+a_2 z^2+\ldots)/z^p$ is holomorphic and p-valent in $0<|z| \leqslant 1$, then

$$\frac{d}{dr}\int_0^{2\pi} F(|f(re^{i\theta})|)\,d\theta \leqslant 0$$

for any increasing function $F(\rho)$ in $\rho \geqslant 0$. In particular, if $F(\rho)=\rho^2$, this becomes an area theorem from which follow coefficient estimates, etc., for p-valent functions.

Various subfamilies and generalized families of multivalent functions have been considered. Let $f(z)$ be p-valent in D, and $c_0+c_1 z+\ldots+c_{p-1} z^{p-1}+c_p f(z)$ be at most p-valent in D for any constants $c_0, c_1, \ldots, c_p$. Then $f(z)$ is said to be **absolutely p-valent** in D. If a function $f(z)$ holomorphic in a convex domain K satisfies $\mathrm{Re}(e^{i\alpha} f^{(p)}(z))>0$ for a real constant α, then $f(z)$ is absolutely p-valent in K. If $f(z)$ is absolutely p-valent in D, then

$$\left(\sum_{k=0}^{p-1} b_k z^k + b_p f(z)\right) \Bigg/ \left(\sum_{k=0}^{p-1} c_k z^k + c_p f(z)\right)$$

is at most p-value in D for any constants b_k and c_k.

If $f(z)$ is p-valent in the common part of a domain D and the disk centered at each point of D with a fixed radius ρ, then $f(z)$ is said to be **locally p-valent** in D, and ρ is called its **modulus**. A necessary and sufficient condition for $f(z)$, holomorphic in D, to be at most locally p-valent is that $f'(z), \ldots, f^{(p)}(z)$ not vanish simultaneously. In order for $f(z)$, holomorphic in D, to be **locally absolutely p-valent** it is necessary and sufficient that $f^{(p)}(z) \neq 0$. Let the number of $Re^{i\varphi}$-points of $f(z)$ in D be $n(D, Re^{i\varphi})$. If $f(z)$ satisfies

$$\frac{1}{2\pi}\int_0^{2\pi} n(D, Re^{i\varphi})\,d\varphi \leqslant p,$$

for any $R>0$, it is said to be **circumferentially mean p-valent** in D. If $f(z)$ satisfies

$$\int_0^R \int_0^{2\pi} n(D, Re^{i\varphi}) R\,dR\,d\varphi \leqslant p\pi R^2,$$

it is said to be **areally mean ρ-valent** in D. If $f(z)^q$ with $q>1$ is areally mean p-valent in D, then $f(z)$ is areally mean p/q-valent in D. For

$f(z)=(1+a_1z+a_2z^2+\dots)/z^\lambda$ holomorphic and areally mean λ-valent in $0<|z|\leqslant 1$, the following area theorem holds:

$$\sum_{n=1}^{\infty}(n-1)|a_n|^2\leqslant\lambda.$$

Let E be a set containing at least three points. If $f(z)$ in D attains every value of E at most p times and a certain value of E exactly p times (it may attain values outside E more than p times), then $f(z)$ is said to be **quasi-p-valent** in D. If $w=f(z)$ is p-valent in D and $g(w)$ is quasi-q-valent in $f(D)$, then $g(f(z))$ is at most quasi-pq-valent in D.

The first success in obtaining sharp inequalities for multivalent functions was attained by Hayman. In his work, an essential role was played by the method of †symmetrization. For instance, he obtained the following result. If $f(z)=z^p+a_{p+1}z^{p+1}+\dots$ is holomorphic and circumferentially mean p-valent in $|z|<1$, then $|a_{p+1}|\leqslant 2p$, and for $|z|=r$, $0<r<1$,

$$\frac{r^p}{(1+r)^{2p}}\leqslant|f(z)|\leqslant\frac{r^p}{(1-r)^{2p}},$$

$$|f'(z)|\leqslant\frac{p(1+r)}{r(1-r)}|f(z)|\leqslant\frac{pr^{p-1}(1+r)}{(1-r)^{2p+1}}.$$

References

[1] L. Bieberbach, Über die Koeffizienten derjenigen Potenzreihen, welche eine schlichte Abbildung des Einheitskreises vermitteln, S.-B. Preuss., Akad. Wiss., (1916), 940–955.
[2] K. Löwner, Untersuchungen über schlichte konforme Abbildungen des Einheitskreises I, Math. Ann., 89 (1923), 103–121.
[3] P. R. Garabedian and M. Schiffer, A proof of the Bieberbach conjecture for the fourth coefficient, J. Rational Mech. Anal., 4 (1955), 427–465.
[4] M. Ozawa, On the Bieberbach conjecture for the sixth coefficient, Kôdai Math. Sem. Rep., 21 (1969), 97–128.
[5] R. N. Pederson, A proof of the Bieberbach conjecture for the sixth coefficient, Arch. Rational Mech. Anal., 31 (1968–69), 331–351.
[6] R. N. Pederson and M. Schiffer, A proof of the Bieberbach conjecture for the fifth coefficient, Arch. Rational Mech. Anal., 45 (1972), 161–193.
[7] W. K. Hayman, The asymptotic behavior of p-valent functions, Proc. London Math. Soc., (3) 5 (1955), 257–284.
[8] A. C. Schaeffer and D. C. Spencer, Coefficient regions for schlicht functions, Amer. Math. Soc. Colloq. Publ., 35 (1950).
[9] G. M. Golusin, Geometric theory of functions of a complex variable, Amer. Math. Soc. Transl. of Math. Monographs, 26 (1969). (Original in Russian, 1952.)
[10] W. K. Hayman, Multivalent functions, Cambridge Univ. Press, 1958.
[11] J. A. Jenkins, Univalent functions and conformal mapping, Springer, second edition, 1965.
[12] L. V. Ahlfors, Conformal invariants: Topics in geometric function theory, McGraw-Hill, 1973.
[13] C. Pommerenke, Univalent functions (with a chapter on quadratic differentials by G. Jensen), Vandenhoeck & Ruprecht, 1975.
[14] G. Schober, Univalent functions—Selected topics, Lecture notes in math. 478, Springer, 1975.
[15] P. L. Duren, Coefficients of univalent functions, Bull. Amer. Math. Soc., 83 (1977), 891–911.

V

439 (III.19) Valuations

A. Introduction

There are two related kinds of valuations, additive (→ Section B) and multiplicative (→ Section C). The notion of valuations, originally defined on (commutative) †fields, has been extended to more general cases (→ Section K); however, we first consider the case of fields.

B. Additive Valuations

In this article, we mean by an **ordered additive group** a **totally ordered additive group**, namely, a commutative group whose operation is addition, which is a †totally ordered set satisfying the condition that $a \geqslant b$ and $c \geqslant d$ imply $a + c \geqslant b + d$ and $-a \leqslant -b$. Suppose that we are given a field K, an ordered additive group G, and an element ∞ defined to be greater than any element of G. Then a mapping $v: K \rightarrow G \cup \{\infty\}$ is called an **additive valuation** (or simply a **valuation**) of the field K if v satisfies the following three conditions: (i) $v(a) = \infty$ if and only if $a = 0$; (ii) $v(ab) = v(a) + v(b)$ for all a, $b \neq 0$; and (iii) $v(a+b) \geqslant \min\{v(a), v(b)\}$.

The set $\{v(a) \mid a \in K - \{0\}\}$ is a submodule of G and is called the **value group** of v, while the set $R_v = \{a \in K \mid v(a) \geqslant 0\}$ is a subring of K and is called the **valuation ring** of v. The ring R_v has only one †maximal ideal $\{a \mid v(a) > 0\}$, called the **valuation ideal** of v (or of R_v), and the †residue class field of R_v modulo the maximal ideal is called the **residue class field** of the valuation v. We have $v(a) \leqslant v(b)$ if and only if $aR_v \supset bR_v$. Two valuations v and v' of the field K are said to be **equivalent** when $v(a) \leqslant v(b)$ if and only if $v'(a) \leqslant v'(b)$; hence v and v' are equivalent if and only if $R_v = R_{v'}$. The **rank** of v is defined to be the †Krull dimension of the valuation ring R_v, and the **rational rank** of v to be the maximum (or supremum) of the numbers of linearly independent elements in the value group. An **extension** (or **prolongation**) of v in a field K' containing K is a valuation v' of K' whose restriction on K is v; such an extension exists for any given v and K'. Sometimes a valuation of rank 1 is called a **special valuation** (or **exponential valuation**), and a valuation of a general rank is called a **generalized valuation**. On the other hand, if k is a subfield of K such that $v(a) = 0$ for every nonzero element a of k, then v is called a **valuation over the subfield** k.

C. Multiplicative Valuations

A **multiplicative valuation** (or **valuation**) of a field K is a mapping $w: K \rightarrow \Gamma \cup \{0\}$ that satisfies the following three conditions, where Γ is the multiplicative group of positive real numbers: (i) $w(a) = 0$ if and only if $a = 0$; (ii) $w(ab) = w(a)w(b)$; and (iii) $w(a+b) \leqslant C(w(a) + w(b))$, where C is a constant (independent of the choice of a and b, but dependent on the choice of w).

The **value group** of w is defined to be $\{w(a) \mid a \in K - \{0\}\}$. Extensions of a valuation and equivalence of valuations are defined as in the case of additive valuations. Thus w' is equivalent to w if and only if there is a positive r such that for all $a \in K$, $w(a) = w'(a)^r$. In each equivalence class of valuations of a field, there exists a valuation for which the constant C in condition (iii) can be taken to be 1. A valuation w is said to be a **valuation over a subfield** k if $w(a) = 1$ for any nonzero element a of k.

We call w an **Archimedean valuation** if for any elements $a, b \in K$, $a \neq 0$, there exists a natural number n such that $w(na) > w(b)$; otherwise, w is said to be a **non-Archimedean valuation**. If w is an Archimedean valuation of a field K, then there is an injection σ from K into the complex number field **C** such that w is equivalent to the valuation w' defined by $w'(a) = |\sigma(a)|$. If w is a non-Archimedean valuation of a field K, then $w(a+b) \leqslant \max\{w(a), w(b)\}$. Hence in this case we get an additive valuation v of K when we define $v(a) = -\log w(a)$ $(a \in K)$, and either v is of rank 1 or $v(K) = \{1, 0\}$ (in the latter case, v is called **trivial**). Conversely, every additive valuation of rank 1 of K is equivalent to an additive valuation obtained in this way from a non-Archimedean valuation. (This is why an additive valuation of rank 1 is called an exponential valuation.) Therefore a non-Archimedean valuation determines a valuation ring and valuation ideal in a natural manner. Thus we can identify a non-Archimedean valuation with an additive valuation of rank 1.

D. Topology Defined by a Valuation

Let w be a multiplicative valuation of a field K. When the †distance between two elements a, b of K is defined by $w(a-b)$, K becomes a †topological field. (Although this distance may not make K into a †metric space, there exists a valuation w' equivalent to the valuation w such that K becomes a metric space with respect to the distance $w'(a-b)$ between a and b $(a, b \in K)$.) If K is †complete under the topology, then we say that K is **complete** with respect to w and w is **complete** on K. On the other hand, suppose that w' is an extension of w in a field K' containing K. If w' is complete and K is †dense in K' under the topology defined by w', then we say that the valuation w' is a **completion** of w and that the field K' is a

completion of K with respect to w. For any w, a completion exists and is unique up to isomorphism. When w is a non-Archimedean valuation, the valuation ring of the completion of w is called the **completion** of the valuation ring of w.

When v is an additive valuation of a field K, we can introduce a topology on K by taking the set of all nonzero ideals of the valuation ring R_v of v as a †base for the neighborhood system of zero. Important cases are given by valuations of rank 1, which are the same as those given by non-Archimedean valuations.

If w is a complete non-Archimedean valuation of a field K, then the valuation ring R_w of w is a †Hensel ring, which implies that if K' is a finite algebraic extension of K such that $[K':K]=n$, then w is uniquely extendable to a valuation w' of K' and $w'(a)^n = w(N(a))$, where N is the †norm $N_{K'/K}$.

E. Discrete Valuations

For a non-Archimedean valuation (or an additive valuation of rank 1) w, if the valuation ideal of w is a nonzero †principal ideal generated by an element p, then we say that p is a **prime element** for w, w is a **discrete valuation**, and the valuation ring for w is a **discrete valuation ring**. The condition on the valuation ideal of w holds if and only if the value group of w is a discrete subgroup of the (multiplicative) group Γ of positive real numbers: In the terminology of additive valuations, a valuation w is discrete if and only if it is equivalent to a valuation w' whose value group is the additive group of integers. Such a valuation w' is called a **normalized valuation** (or **normal valuation**). However, we usually mean normalization of a discrete non-Archimedean valuation as in Section H. Sometimes an additive valuation whose value group is isomorphic to the direct sum of a finite number of copies of $\mathbf{Z}$ (the additive group of integers) with a natural †lexicographic order is called a discrete valuation. Concerning a complete discrete valuation w, it is known that if the valuation ring of w contains a field, then it is isomorphic to the ring of †formal power series in one variable over a field (for other cases → 449 Witt Vectors A).

F. Examples

(1) **Trivial valuations** of a field K are the additive valuation v of K such that $v(a)=0$ for all $a \in K-\{0\}$ and the multiplicative valuation w of K such that $w(a)=1$ for all $a \in K-\{0\}$.

(2) If K is isomorphic to a subfield of the complex number field, then we get an Archimedean valuation using the absolute value, and as stated in Section C, every Archimedean valuation of K is equivalent to a valuation obtained in this way.

(3) Let $\mathfrak{p}$ be a †prime ideal of a †Dedekind domain R, $\pi \in \mathfrak{p}$ be such that $\pi \notin \mathfrak{p}^2$, and K be the field of quotients of R. Then each nonzero element α of K can be expressed in the form $\pi^r ab^{-1}$ ($r \in \mathbf{Z}; a, b \in R; a, b \notin \mathfrak{p}$), where r, the **degree** of α with respect to $\mathfrak{p}$, is uniquely determined by α. Hence, letting c be a constant greater than 1, we obtain a non-Archimedean valuation w defined by $w(\alpha)=c^{-r}$. This valuation w is called a $\mathfrak{p}$**-adic valuation**. We also get an additive valuation v defined by $v(\alpha)=r$, called a $\mathfrak{p}$**-adic exponential valuation**. The completion $K_{\mathfrak{p}}$ of K with respect to v is called the $\mathfrak{p}$**-adic extension** of K. If K is a finite †algebraic number field, the $K_{\mathfrak{p}}$ is called a $\mathfrak{p}$**-adic number field**. If $\mathfrak{p}$ is generated by an element p, then "$\mathfrak{p}$-adic" is replaced by "p-adic." For instance, given a rational prime number p, we have a p**-adic valuation** of the rational number field $\mathbf{Q}$, and we obtain the p-adic extension $\mathbf{Q}_p$ of $\mathbf{Q}$, which is called the p**-adic number field**. Every nonzero element α of $\mathbf{Q}_p$ can be written as a uniquely determined expansion $\sum_{n=r}^{\infty} a_n p^n$ ($a_r \neq 0, r \in \mathbf{Z}, a_n \in Z, 0 \leqslant a_n < p$). Then we obtain a valuation v of $\mathbf{Q}_p$ defined by $v(\alpha)=r$. This valuation v is a discrete additive valuation, and $\mathbf{Q}_p$ is complete with respect to v. The valuation ring of v is usually denoted by $\mathbf{Z}_p$, which is called the **ring of** p**-adic integers**. Each element of $\mathbf{Q}_p$ ($\mathbf{Z}_p$) is called a p**-adic number** (p**-adic integer**).

(4) Consider the field of †power series $k((t))$ in one variable t over a field k. For $0 \neq \alpha \in k((t))$, we define $v(\alpha)=r$ if $\alpha = \sum_{n=r}^{\infty} a_n t^n$ ($a_n \in k$, $a_r \neq 0$). Then v is a discrete valuation of $k((t))$, and $k((t))$ is complete with respect to this valuation.

(5) Let v be an additive valuation of a field K with the valuation ring R_v and the valuation ideal $\mathfrak{m}_v$. Let v' be an additive valuation of the field $R_v/\mathfrak{m}_v$ with the valuation ring $R_{v'}$. Then $R'' = \{a \in R_v \mid (a \bmod \mathfrak{m}_v) \in R_{v'}\}$ is a valuation ring of K. A valuation v'' whose valuation ring coincides with R'' is called the **composite** of v and v'.

G. The Approximation Theorem and the Independence Theorem

The **approximation theorem** states: Let $w_1, \ldots, w_n$ be mutually nonequivalent and nontrivial multiplicative valuations of a field K. Then for any given n elements $a_1, \ldots, a_n$ of K and a positive number ε, there exists an element a of K such that $w_i(a-a_i)<\varepsilon$ ($i=1,2,\ldots,n$).

From this follows the **independence theorem**: Let $e_1, \dots, e_n$ be real numbers, and let w_i and K be as in the approximation theorem. If $\prod_i w_i(a)^{e_i} = 1$ for all $a \in K - \{0\}$, then $e_1 = \dots = e_n = 0$.

Similar theorems hold for additive valuations. The following independence theorem is basic: Let $v_1, \dots, v_n$ be additive valuations of a field K, $R_1, \dots, R_n$ their valuation rings, and $\mathfrak{m}_1, \dots, \mathfrak{m}_n$ their maximal ideals. Let $D = \bigcap_i R_i$, $\mathfrak{p}_i = \mathfrak{m}_i \cap D$, and consider the rings of quotients $D_{\mathfrak{p}_i}$. Then $D_{\mathfrak{p}_i} = R_i$. If $R_i \not\subset R_j$ (for $i \neq j$), then D has exactly n maximal ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_n$.

H. Prime Divisors

Let K be an †algebraic number field (algebraic function field of one variable over a field k). An equivalence class of nontrivial multiplicative valuations (over k) is called a **prime divisor (prime spot)** of K.

If K is an algebraic number field of degree n, there are exactly n mutually distinct injections $\sigma_1, \dots, \sigma_n$ of K into the complex number field $\mathbf{C}$. We may assume that $\sigma_i(K)$ is contained in the real number field if and only if $i \leqslant r_1$ and $\sigma_{n-i+1}(a)$ and $\sigma_{r_1+i}(a)$ are conjugate complex numbers $(n - r_1 \geqslant i > 0,\ a \in K)$. For $i \leqslant r_1$, let $v_i(a) = |\sigma_i(a)|$, and for $1 \leqslant i \leqslant (n - r_1)/2$, let $v_{r_1+i}(a) = |\sigma_{r_1+i}(a)|^2$. Then $v_1, \dots, v_{r_1+r_2}$ $(r_2 = (n - r_1)/2)$ is a maximal set of mutually nonequivalent Archimedean valuations of K. Equivalence classes of $v_1, \dots, v_{r_1}$ are called **real (infinite) prime divisors**, and those of $v_{r_1+1}, \dots, v_{r_1+r_2}$ are called **imaginary (infinite) prime divisors**; all of them are called **infinite prime divisors**. An equivalence class of non-Archimedean valuations of K is called a **finite prime divisor**.

An Archimedean valuation of K is said to be **normal** if it is one of the valuations v_i. If v is non-Archimedean, then v is a $\mathfrak{p}$-adic valuation, where $\mathfrak{p}$ is a prime ideal of the principal order $\mathfrak{o}$ of K (→ Section F, example (3)). Hence if a is an element of K, there exists a constant c $(c > 1)$ such that $v(a) = c^{-r}$, where r is the degree of a with respect to $\mathfrak{p}$. In particular, if c is the norm of $\mathfrak{p}$ (i.e., c is the cardinality of the set $\mathfrak{o}/\mathfrak{p}$), then the valuation v is called **normal**. Any finite prime divisor is represented by a normal valuation. Then we have the **product formula** $\prod_w w(a) = 1$ for all $a \in K - \{0\}$, where w ranges over all normal valuations of K.

For a function field, a **normal valuation** is defined similarly, using e^f instead of the norm of $\mathfrak{p}$, where e is a fixed real number greater than 1 and f is the degree of the residue class field of the valuation over k. In this case we also have the product formula.

I. Extending Valuations to an Algebraic Extension of Finite Degree

Assume that a field K' is a finite algebraic extension of a field K. Let v be an additive valuation of K and v' be an extension of v to K'. We denote the valuation rings, valuation ideals, and value groups of v and v' by R_v, $R_{v'}$, $\mathfrak{m}_v$, $\mathfrak{m}_{v'}$, and G, G', respectively. Then the degree of the extension $f_{v'} = [R_{v'}/\mathfrak{m}_{v'} : R_v/\mathfrak{m}_v]$ is called the **degree** of v' over v. The group index $e_{v'} = [G' : G]$ is called the **ramification index** of v' over v. If v' ranges over all extensions of v in K', then the sum $\sum f_{v'} e_{v'}$ is not greater than $[K' : K]$ and the equality holds when v is a discrete valuation and either K' is †separable over K or v is complete.

J. Places

Let k, K, and L be fields, and suppose that $k \subset K$. Let f be a mapping of K onto $L \cup \{\infty\}$ such that $f(ab) = f(a)f(b)$ and $f(a+b) = f(a) + f(b)$, whenever the right member is meaningful, and such that the restriction of f to k is an injection. Here ∞ is an element adjoined to L and satisfying $\infty + a = a + \infty = \infty$, $\infty a = a \infty = \infty$ (for any nonzero element a of K), $1/\infty = 0$, and $1/0 = \infty$. Then f is called a **place** of K over k. In this case $R = \{x \in K \mid f(x) \neq \infty\}$ is a valuation ring of K containing k. Let $\mathfrak{m}$ be the maximal ideal of R. Then f can be identified with the mapping $g: K \to R/\mathfrak{m} \cup \{\infty\}$ defined as follows: If $a \in R$, then $g(a) = (a \bmod \mathfrak{m})$; otherwise, $g(a) = \infty$. Places of K over k can be classified in a natural way, and there exists a one-to-one correspondence between the set of classes of places of K over k and the set of equivalence classes of additive valuations over k. When K is an †algebraic function field, we usually consider the case where k is the †ground field. Then if $a_1, \dots, a_n \in R$, $(a_1, \dots, a_n) \to (g(a_1), \dots, g(a_n))$ gives a †specialization of points over k. Conversely, if $a_i, b_j \in K$ are such that $(a_1, \dots, a_n) \to (b_1, \dots, b_n)$ is a specialization over k, then there is a place f of K over k such that $(b_1, \dots, b_n)$ is isomorphic to $(f(a_1), \dots, f(a_n))$ (usually there are infinitely many such f's).

K. Pseudovaluations

A **pseudovaluation** φ of a ring A (not necessarily commutative) is a mapping of A into the set of nonnegative real numbers satisfying the following four conditions: (i) $\varphi(a) = 0$ if and only if $a = 0$; (ii) $\varphi(ab) \leqslant \varphi(a)\varphi(b)$; (iii) $\varphi(a+b) \leqslant \varphi(a) + \varphi(b)$; and (iv) $\varphi(-a) = \varphi(a)$. These conditions are weaker than those for multiplicative valuations, but with them a topology can be introduced into A as in Section D, with respect to which A becomes a topological ring.

L. History

The theory of valuations was originated by K. Hensel when he introduced p-adic numbers and applied them to number theory [1]. J. Kürschák (*J. Reine Angew. Math.*, 142 (1913)) first treated the theory of multiplicative valuations axiomatically; it was then developed remarkably by A. Ostrowski (*Acta Math.*, 41 (1918)). However, in their theory condition (iii) (→ Section C) was given only in the case $C=1$, thus excluding the normal valuation of an imaginary prime divisor in an algebraic number field. A valuation with general C was introduced by E. Artin [3]. The theory of additive valuations was originated by W. Krull (*J. Reine Angew. Math.*, 167 (1932)), although the concept of exponential valuations existed before. The theory of valuations is used to simplify †class field theory and the theory of algebraic function fields in one variable. For these purposes, the notion of multiplicative valuations is sufficient (→ 9 Algebraic Curves; 59 Class Field Theory). The idea is also used in the theory of normal rings and in algebraic geometry, for both of which the concept of additive valuations is also necessary. Pseudovaluations were used by M. Deuring (*Erg. Math.*, Springer, 1935) in the arithmetic of algebras.

References

[1] K. Hensel, Theorie der algebraischen Zahlen, Teubner, 1908.
[2] O. F. G. Schilling, The theory of valuations, Amer. Math. Soc. Math. Surveys, 1950.
[3] E. Artin, Algebraic numbers and algebraic functions, Gordon & Breach, 1967.
[4] O. Zariski and P. Samuel, Commutative algebra II, Van Nostrand, 1960.
Also → references to 67 Commutative Rings.

440 (X.35) Variational Inequalities

A. Introduction

Variational inequalities arise when we consider extremal problems of functionals under **unilateral constraints**. Some problems in physics and engineering are studied by formulating them as elliptic, parabolic, and hyperbolic variational inequalities [1–8].

B. Stationary Variational Inequality

Let D be a bounded domain in m-dimensional Euclidean space and $f \in L_2(D)$ be a given real-valued function. Consider the variational problem of minimizing the following functional J with the argument function v:

$$J[v]=\int_D |\operatorname{grad} v|^2\,dx-2\int_D fv\,dx.$$

Here, we suppose the set of admissible functions to be the closed convex subset

$$K=\{v\in H_0^1(D) \mid v\leqslant 0 \text{ a.e. in } D\}$$

of the Hilbert space $H_0^1(D)$ (→ 168 Function Spaces). It can be shown by choosing a minimizing sequence that there exists a minimum value of J which is realized by a unique $u\in K$. Since the stationary function u belongs to $H_0^1(D)$, it can be shown that the boundary condition $u|_{\partial D}=0$ is satisfied in the sense that the †trace $\gamma_0 u\in H^{1/2}(\partial D)$ (→ 224 Interpolation of Operators) of u on ∂D vanishes a.e. on ∂D. In view of the fact that $J[u]\leqslant J[v]$ is valid for any $v\in K$, it can be verified that the **stationary variational inequality**

$$\left.\begin{aligned}&-\Delta u-f\leqslant 0\\ &u\leqslant 0\\ &(-\Delta u-f)\cdot u=0\end{aligned}\right\}\qquad (1)$$

is satisfied in D in the sense of differentiation of distributions (→ 125 Distributions and Hyperfunctions). The problem (1) is a **Dirichlet problem with obstacle**. Moreover, we can prove the regularity of $u\in H^2(D)$ under an assumption of suitable smoothness for ∂D by establishing the boundedness of the solutions u_ε in $H^2(D)$ of the **penalized problems** associated with (1):

$$-\Delta u_\varepsilon+\frac{1}{\varepsilon}u_\varepsilon^+=f\qquad(\varepsilon>0),$$

$$u_\varepsilon|_{\partial D}=0.$$

Here we note that the u_ε are the stationary functions of the ordinary variational problems of minimization in $H_0^1(D)$ of the functionals

$$J_\varepsilon[v]=\int_D |\operatorname{grad} v|^2\,dx-2\int_D fv\,dx+\frac{1}{\varepsilon}\int |v^+|^2\,dx$$

with the **penalty term** (the third term of the right-hand side of the equality above). We have thus found that the stationary variational inequality (1) is the Euler equation of a conditional problem of variation (→ 46 Calculus of Variations).

C. Variational Inequality of Evolution

Let $\psi\in H^1(D)$ be a given function on D such that $\psi|_{\partial D}\geqslant 0$ and $\Delta\psi\in L_2(D)$. The **variational** inequality of evolution

$$\frac{\partial u}{\partial t}-\Delta u\leqslant 0,$$

$$u\leqslant\psi,$$

$$\left(\frac{\partial u}{\partial t}-\Delta u\right)\cdot(u-\psi)=0 \quad (t>0, x\in D),$$

$$u(0,x)=a(x) \quad (x\in D),$$

$$u(t,x)|_{\partial D}=0 \quad (t>0)$$

can be formulated as an abstract Cauchy problem (→ 286 Nonlinear Functional Analysis X)

$$\frac{du}{dt}\in Au \quad (t>0),$$

$$u(+0)=a$$

in a Hilbert space with a multivalued operator $A=-\partial\varphi$, where $\partial\varphi$ is the subdifferential of the following lower semicontinuous proper convex function on the Hilbert space $L_2(D)$:

$$\varphi(v) = \begin{cases} \frac{1}{2}\int_D |\operatorname{grad} v|^2 dx & \text{if } v\in H_0^1(D) \text{ and } v\leqslant\psi, \\ +\infty & \text{otherwise.} \end{cases}$$

Thus the solution u is given by the vector-valued function

$$u(t)=e^{tA}a.$$

Here e^{tA} is the †nonlinear semigroup generated by A (→ 88 Convex Analysis, 378 Semigroups of Operators and Evolution Equations).

D. Optimal Stopping Time Problem and Variational Inequalities

Let $\{X_t\}_{t\geqslant 0}$ be an m-dimensional Brownian motion (→ 45 Brownian Motion) and consider the problem of finding a †stopping time σ that minimizes

$$J_x[\sigma]=\mathrm{E}_x\left(\int_0^\sigma f(X_t)\,dt\right) \quad (x\in\mathbf{R}^m)$$

under the restriction that $0\leqslant\sigma\leqslant\sigma_{\partial D}$, where $\sigma_{\partial D}$ is the †hitting time for the boundary ∂D. Let us define

$$u(x)=\min_\sigma J_x[\sigma].$$

Then the †principle of optimality in dynamic programming gives the stationary variational inequality (1) with Δ replaced by $\frac{1}{2}\Delta$, and we can show by the †Dynkin formula that an optimal stopping time $\hat{\sigma}$ is the hitting time for the set $\{x\in\bar{\Omega}\mid u(x)=0\}$ (→ 127 Dynamic Programming). We can systematically discuss problems in mathematical programming and operations research by introducing quasivariational inequalities, which are slight generalizations of variational inequalities (→ 227 Inventory Control, 408 Stochastic Programming). The above-mentioned facts are applicable to general †diffusion processes described by †stochastic differential equations (→ 115 Diffusion Processes, 406 Stochastic Differential Equations). We have thus found the relation

free boundary problem ↔ variational inequality
↘ optimal stopping time problem ↙

(→ 405 Stochastic Control and Stochastic Filtering).

E. Numerical Solution of Variational Inequalities

Since the solution u of the variational inequality (1) is the stationary function for the variational problem, we can apply to the evaluation of the function u numerical methods based on the direct method of the calculus of variations (→ 300 Numerical Methods). The †finite element method, which can be regarded as a type of Ritz-Galerkin method, is extensively employed to calculate numerical solutions. In view of the unilateral constraint $u\leqslant 0$, iteration methods, such as the Gauss-Seidel iteration method, are used with modifications. An algorithm of relaxation with projection is proposed in [3] (→ 304 Numerical Solution of Partial Differential Equations).

References

[1] H. Brézis, Problèmes unilatéraux, J. Math. Pures Appl., 51 (1972), 1–168.
[2] V. Barbu, Nonlinear semigroups and differential equations in Banach spaces, Noordhoff, 1976.
[3] J.-L. Lions, Sur quelques questions d'analyse, de mécanique et de contrôle optimal, Les Presses de l'Université de Montréal, 1976.
[4] J.-L. Lions and G. Stampacchia, Variational inequalities, Comm. Pure Appl. Math., 20 (1967), 493–519.
[5] C. Baiocchi, Su un problema di frontiera libera connesso a questioni di idraulica, Ann. Mat. Pura Appl. 92 (1972), 107–127.
[6] G. Duvaut and J.-L. Lions, Les inéquations en mécanique et en physique, Dunod, 1972; English translation, Springer, 1975.
[7] A. Friedman and D. Kinderlehrer, A one phase Stefan problem, Indiana Univ. Math. J., 24 (1975), 1005–1035.
[8] R. Glowinski, J.-L. Lions, and R. Trémolières, Analyse numérique des inéquations variationelles I, II, Dunod, 1976; English translation, North-Holland, 1981.

441 (XX.3) Variational Principles

A. General Remarks

Among the principles that appear in physics are those expressed not in terms of differential forms but in terms of variational forms. These principles, describing the conditions under which certain quantities attain extremal values, are generally called **variational principles**. Besides Hamilton's principle in classical mechanics (→ Section B) and Fermat's principle in geometric optics (→ Section C), examples are found in †electromagnetism, †relativity theory, †quantum mechanics, †field theory, etc. Independence of the choice of coordinate system is an important characteristic of variational principles. Originally these principles had theological and metaphysical connotations, but a variational principle is now regarded simply as a postulate that precedes a theory and furnishes its foundation. Thus a variational principle is considered to be the supreme form of a law of physics.

B. Mechanics

In 1744 P. L. Maupertuis published an almost theological thesis, dealing with the **principle of least action**. This was the beginning of the search for a single, universal principle of mechanics, contributions to which were made successively by L. Euler, C. F. Gauss, W. R. Hamilton, H. R. Hertz, and others.

Let $\{q_r\}$ be the †generalized coordinates of a system of particles, and consider the integral of a function $L(q_r, \dot{q}_r, t)$ taken from time t_0 to t_1. If we compare the values of the integral taken along any arbitrary path starting from a fixed point P_0 in the coordinate space at time t_0 and arriving at another fixed point P_1 at time t_1, then the actual motion $q_r(t)$ (which obeys the laws of mechanics) is given by the condition that the integral is an †extremum (†stationary value), that is, $\delta\int_{t_0}^{t_1} L\,dt=0$, provided that the function L is properly chosen. This is **Hamilton's principle**, and L is the †Lagrangian function. In †Newtonian mechanics, the †kinetic energy T of a system of particles is expressed as a †quadratic form in $\dot{q}_r$. Furthermore, if the forces acting on the particles can be given by $-\operatorname{grad} V$, where the potential V does not depend explicitly on $\dot{q}_r$, we can choose $L=T-V$. Also, for a charged particle in †special relativity, we can take $L=-m_0c^2(1-v^2/c^2)^{1/2}-e\varphi+e(\mathbf{v}\cdot\mathbf{A})$, where m_0 is the rest mass of the particle, e is the charge, $\mathbf{v}$ is the velocity (with v its magnitude), c is the speed of light in vacuum, and φ and $\mathbf{A}$ are the scalar and vector potentials of the electromagnetic field, respectively.

In general relativity theory, the motion of a particle can be derived from the variational principle $\delta\int ds=0$ (ds is the Riemannian line element). Hence, geometrically, the particle moves along a †geodesic curve in 4-dimensional space-time.

C. Geometric Optics

The path of a light ray between two points P_0 and P_1 (subject to reflection and refraction) is such that the time of transit along the path among all neighboring virtual paths is an extremum (stationary value). This is called **Fermat's principle**. If the index of refraction is n, Fermat's principle can be expressed as $\delta\int_{P_0}^{P_1} n\,ds=0$ (ds is the Euclidean line element). The laws of reflection and refraction of light, as well as the law of rectilinear propagation of light in homogeneous media, can be derived from this principle.

D. Field Theory

Not only the equations of motion of a system of particles, but also various field equations (†Maxwell's equations of the electromagnetic field, †Dirac's equation of the electron field, the meson field equation, the gravitational field equation, etc.) can be derived from variational principles in terms of appropriate Lagrangian functions. In †field theory the essential virtue of the variational principle appears in the fact that the properties of various possible fields as well as conservation laws can be systematically discussed by assuming relativistic invariance and gauge invariance of the Lagrangian functions adopted. In particular, for an electromagnetic field in vacuum, the Lagrangian function density is $L=(\mathbf{H}^2-\mathbf{E}^2)/2$, and the integration is carried out over a certain 4-dimensional domain.

E. Quantum Mechanics

If H is the †Hamiltonian operator for any quantum-mechanical system, the eigenfunction ψ can be determined by the variational principle

$$\delta\int\bar{\psi}H\psi\,d\tau=0, \quad \text{with} \quad \int\bar{\psi}\psi\,d\tau=1,$$

where $\bar{\psi}$ is the complex conjugate of ψ and $d\tau$ is the volume element. Based on this variational principle, the †direct method of the calculus of variations is often employed for an approximate numerical calculation of the energy eigenvalues and eigenfunctions. In particular, by restricting the functional form of ψ to the product of one-body wave functions, we can obtain Hartree's equation. A further suitable symmetrization of ψ leads to Fock's equation.

F. Statistical Mechanics

Let φ be a statistical-mechanical state of a system, and let $S(\varphi)$ and $E(\varphi)$ be the state's entropy and energy (mean entropy and mean energy for an infinitely extended system); T is the thermodynamical temperature, and $f(\varphi)=E(\varphi)-TS(\varphi)$ is the free energy. Then the equilibrium state for $T\geqslant 0$ is determined as the state φ that gives the minimum value of $f(\varphi)$ (maximum for $T<0$).

References

[1] R. Courant and D. Hilbert, Methods of mathematical physics, Interscience, I, 1953; II, 1962.
[2] B. L. Moiseiwitsch, Variational principles, Interscience, 1966.
[3] R. Weinstock, Calculus of variations, with applications to physics and engineering, McGraw-Hill, 1952.
Also → references to 46 Calculus of Variations.

442 (VI.12) Vectors

A. Definitions

The **vector** concept originated in physics from such well-known notions as velocity, acceleration, and force. These physical quantities are supplied with length and direction; they can be added or multiplied by scalars. In the Euclidean space E^n (or, in general, an †affine space), a vector **a** is represented by an **oriented segment** $\overrightarrow{pq}$. Two oriented segments $\overrightarrow{p_1q_1}$ and $\overrightarrow{p_2q_2}$ are considered to represent the same vector **a** if and only if the following two conditions are satisfied: (1) The four points p_1, q_1, p_2, q_2 lie in the same plane π. (2) $p_1q_1//p_2q_2$ and $p_1q_2//q_1q_2$. Hence a vector in E^n is an equivalence class of oriented segments $\overrightarrow{pq}$, where the equivalence relation $\overrightarrow{p_1q_1}\sim\overrightarrow{p_2q_2}$ is defined by the two conditions just given. Hereafter, we denote the vector by $[\overrightarrow{pq}]$, or simply $\overrightarrow{pq}$. The points p and q are called the **initial point** and **terminal point** of the vector $\overrightarrow{pq}$.

Given a vector $\mathbf{a}=\overrightarrow{pq}$ and a real number λ, we define the **scalar multiple** $\lambda\mathbf{a}$ as the vector $\overrightarrow{pr}$, where r is the point on the straight line containing both p and q such that the ratio $[\overrightarrow{pr}:\overrightarrow{pq}]$ is equal to λ (if $p=q$, then we put $r=p$). The operation $(\lambda,\mathbf{a})\to\lambda\mathbf{a}$ is called **scalar multiplication**. Given two vectors $\mathbf{a}=\overrightarrow{pq}$ and $\mathbf{b}=\overrightarrow{qs}$, the vector $\mathbf{c}=\overrightarrow{ps}$ is called the **sum** of **a** and **b** and is denoted by $\mathbf{c}=\mathbf{a}+\mathbf{b}$. The vector $\overrightarrow{pp}=\mathbf{0}$ is called the **zero vector**. If $\mathbf{a}=\overrightarrow{pq}$, we put $-\mathbf{a}=\overrightarrow{qp}$.

Scalar multiplication and addition of vectors satisfy the following seven conditions: (1) $\mathbf{a}+\mathbf{b}=\mathbf{b}+\mathbf{a}$ (commutative law); (2) $\mathbf{a}+(\mathbf{b}+\mathbf{c})=(\mathbf{a}+\mathbf{b})+\mathbf{c}$ (associative law); (3) $\mathbf{a}+\mathbf{0}=\mathbf{a}$; (4) for each **a** there is $-\mathbf{a}$ such that $\mathbf{a}+(-\mathbf{a})=\mathbf{0}$; (5) $\lambda(\mathbf{a}+\mathbf{b})=\lambda\mathbf{a}+\lambda\mathbf{b}$, $(\lambda+\mu)\mathbf{a}=\lambda\mathbf{a}+\mu\mathbf{a}$ (distributive laws); (6) $\lambda(\mu\mathbf{a})=(\lambda\mu)\mathbf{a}$ (associative law for scalar multiplication); and (7) $1\mathbf{a}=\mathbf{a}$. Hence the set V of all vectors in E^n forms a †real linear space. Sometimes, a set satisfying (1)–(7), that is, by definition, a linear space, is called a vector space, and its elements are called vectors.

The pair consisting of a vector $\overrightarrow{pq}$ and a specific initial point p of $\overrightarrow{pq}$ is sometimes called a **fixed vector**. An illustration of this is given by the force vector with its initial point being where the force is applied. By contrast, a vector $\overrightarrow{pq}$ is sometimes called a **free vector**. If we fix the origin o in E^n, then for any point p in E^n, the vector $\overrightarrow{op}$ is called the **position vector** of p.

If two vectors $\mathbf{a}=\overrightarrow{op}$ and $\mathbf{b}=\overrightarrow{oq}$ are †linearly dependent, they are sometimes said to be **collinear**. If there vectors $\mathbf{a}=\overrightarrow{op}$, $\mathbf{b}=\overrightarrow{oq}$, and $\mathbf{c}=\overrightarrow{or}$ are linearly dependent, they are sometimes said to be **coplanar**.

If a set of vectors $\mathbf{e}_1,\ldots,\mathbf{e}_n$ forms a †basis of a vector space V, then the vectors $\mathbf{e}_i$ are called **fundamental vectors** in V. Each vector $\mathbf{a}\in V$ is uniquely expressed as $\mathbf{a}=\sum\alpha_i\mathbf{e}_i$ $(\alpha_i\in\mathbf{R})$. We call $(\alpha_1,\ldots,\alpha_n)$ the **components** of the vector **a** with respect to the fundamental vectors $\mathbf{e}_1,\ldots,\mathbf{e}_n$.

B. Inner Product

In the Euclidean space E^n, the length of the line segment $\overline{pq}$ is called the **absolute value** (or **magnitude**) of the vector $\mathbf{a}=\overrightarrow{pq}$ and is denoted by $|\mathbf{a}|$. A vector of length one is called a **unit vector**. For two vectors $\mathbf{a}=\overrightarrow{op}$ and $\mathbf{b}=\overrightarrow{oq}$, the value $(\mathbf{a},\mathbf{b})=|\mathbf{a}||\mathbf{b}|\cos\theta$ is called the **inner product** (or **scalar product**) of **a** and **b**, where θ is the angle $\angle poq$. Instead of $(\mathbf{a},\mathbf{b})$, the notations $\mathbf{a}\cdot\mathbf{b}$, or $\mathbf{ab}$ are also used. If neither vector **a** nor vector **b** is equal to **0**, then $(\mathbf{a},\mathbf{b})=0$ implies $\angle poq=\pi/2$, that is, the orthogonality of the two vectors $\overrightarrow{op}$ and $\overrightarrow{oq}$. If we take an

†orthonormal basis $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ in E^n (i.e., a set of fundamental vectors with $|\mathbf{e}_i| = 1$, $(\mathbf{e}_i, \mathbf{e}_j) = 0$ $(i \neq j)$), the inner product of vectors $\mathbf{a} = \sum \alpha_i \mathbf{e}_i$, $\mathbf{b} = \sum \beta_i \mathbf{e}_i$ is equal to $\sum_{i=1}^n \alpha_i \beta_i$. The inner product has the following three properties (i) $(\mathbf{x}, \mathbf{x}) \geqslant 0$ and is zero if and only if $\mathbf{x} = \mathbf{0}$; (ii) $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$; (iii) $(\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}) = (\mathbf{x}_1, \mathbf{y}) + (\mathbf{x}_2, \mathbf{y})$, $(\alpha \mathbf{x}, \mathbf{y}) = \alpha(\mathbf{x}, \mathbf{y})$ $(\alpha \in \mathbf{R})$. Similar linearity holds for $\mathbf{y}$.

Generally, an $\mathbf{R}$-valued †bilinear form $(\mathbf{x}, \mathbf{y})$ on a linear space V satisfying the previous three conditions is also called an inner product. If a linear space V is equipped with an inner product, the space is called an **inner product space** (→ 256 Linear Spaces H; 197 Hilbert Spaces). If V is an inner product space, the absolute value $|\mathbf{x}|$ of $\mathbf{x} \in V$ is defined to be $\sqrt{(\mathbf{x}, \mathbf{x})}$.

C. Vector Product

In the 3-dimensional Euclidean space E^3, we take an orthonormal basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors in E^3 whose components with respect to $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$. The vector

$$\begin{vmatrix} \alpha_2 & \alpha_3 \\ \beta_2 & \beta_3 \end{vmatrix} \mathbf{e}_1 + \begin{vmatrix} \alpha_3 & \alpha_1 \\ \beta_3 & \beta_1 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{vmatrix} \mathbf{e}_3,$$

which is symbolically written as

$$\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{vmatrix},$$

is called the **exterior product** or **vector product** of $\mathbf{a}$ and $\mathbf{b}$ and is denoted by $[\mathbf{a}, \mathbf{b}]$ or $\mathbf{a} \times \mathbf{b}$. The vector $[\mathbf{a}, \mathbf{b}]$ is determined uniquely up to its sign by $\mathbf{a}$ and $\mathbf{b}$ and is independent of the choice of the orthonormal basis.

Suppose that we have $\mathbf{a} = \overrightarrow{op}$, $\mathbf{b} = \overrightarrow{oq}$. Then $|[\mathbf{a}, \mathbf{b}]| = |\mathbf{a}| \cdot |\mathbf{b}| \sin\theta$, where $\theta = \angle poq$. Also $|[\mathbf{a}, \mathbf{b}]|$ is equal to the area of the parallelogram determined by $\mathbf{a}$ and $\mathbf{b}$. To illustrate the orientation of $[\mathbf{a}, \mathbf{b}]$, we sometimes use the idea of a turning screw. That is, the direction of a right-handed screw advancing while turning at o from p to q (within the angle less than 180°) coincides with the direction of $[\mathbf{a}, \mathbf{b}]$ (Fig. 1). The exterior product has the following three properties: (1) $[\mathbf{a}, \mathbf{b}] = -[\mathbf{b}, \mathbf{a}]$ (antisymmetric law); (2) $[\lambda\mathbf{a}, \mathbf{b}] = \lambda[\mathbf{a}, \mathbf{b}]$ (associative law for scalar multiplication); (3) $[\mathbf{a}, \mathbf{b} + \mathbf{c}] = [\mathbf{a}, \mathbf{b}] + [\mathbf{a}, \mathbf{c}]$ (distributive law). The vector product does not satisfy the associative law, but it does satisfy the †Jacobi identity $[\mathbf{a}, [\mathbf{b}, \mathbf{c}]] + [\mathbf{b}, [\mathbf{c}, \mathbf{a}]] + [\mathbf{c}, [\mathbf{a}, \mathbf{b}]] = \mathbf{0}$. The vector $[\mathbf{a}, [\mathbf{b}, \mathbf{c}]]$ is sometimes called the **vector triple product**, and for this we have **Lagrange's formula** $[\mathbf{a}, [\mathbf{b}, \mathbf{c}]] = (\mathbf{a}, \mathbf{c})\mathbf{b} - (\mathbf{a}, \mathbf{b})\mathbf{c}$.

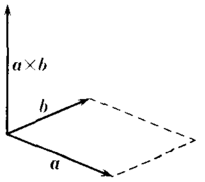


Fig. 1

Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be vectors in E^3 whose components with respect to an orthonormal fundamental basis are $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$, and $(\gamma_1, \gamma_2, \gamma_3)$. Then $(\mathbf{a}, [\mathbf{b}, \mathbf{c}]) = (\mathbf{b}, [\mathbf{c}, \mathbf{a}]) = (\mathbf{c}, [\mathbf{a}, \mathbf{b}]) = [\mathbf{a}, \mathbf{b}, \mathbf{c}]$, and the common value is equal to the determinant of the 3×3 matrix

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \\ \gamma_1 & \gamma_2 & \gamma_3 \end{pmatrix}.$$

The value denoted by $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$ is called the **scalar triple product** of $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and is equal to the volume of the parallelotope whose three edges are $\mathbf{a} = \overrightarrow{op}$, $\mathbf{b} = \overrightarrow{oq}$, and $\mathbf{c} = \overrightarrow{or}$ with common initial point o. The triple $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is called a right-hand system or a left-hand system according as $[\mathbf{a}, \mathbf{b}, \mathbf{c}]$ is positive or negative. We have $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = 0$ if and only if $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are coplanar. (For the †exterior product of vectors in E^n and the concept of †p-vectors → 256 Linear Spaces O.)

D. Vector Fields

In this section we consider the case of a 3-dimensional Euclidean space E^3 (for the general case → 105 Differentiable Manifolds). A scalar-valued or a vector-valued function defined on a set D in E^3 is called a **scalar field** or a **vector field**, respectively. The continuity or the differentiability of a vector field is defined by the continuity or the differentiability of its components.

For a differentiable scalar field $f(x, y, z)$, the vector field with the components $(\partial f/\partial x, \partial f/\partial y, \partial f/\partial z)$ is called the **gradient** of f and is denoted by $\mathbf{grad} f$. For a differentiable vector field $\mathbf{V}(x, y, z)$ whose components are $(u(x, y, z), v(x, y, z), w(x, y, z))$, the vector field with components

$$\left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}, \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}, \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)$$

is called the **rotation** (or **curl**) of $\mathbf{V}$ and is denoted by $\mathbf{rot}\,\mathbf{V}$ (or $\mathbf{curl}\,\mathbf{V}$). Also, for a differentiable vector field $\mathbf{V}$, the scalar field defined by $\partial u/\partial x + \partial v/\partial y + \partial w/\partial z$ is called the **divergence** of $\mathbf{V}$ and is denoted by $\mathbf{div}\,\mathbf{V}$. Utilizing the vector operator ∇ having differential operators $(\partial/\partial x, \partial/\partial y, \partial/\partial z)$ as its components, we may write simply $\operatorname{grad} f = \nabla f$, $\operatorname{div} \mathbf{V} = (\nabla, \mathbf{V})$, $\operatorname{rot} \mathbf{V} =$

[∇, V]. The symbol ∇ is called **nabla**, **atled** (inverse of delta), or **Hamiltonian**.

A vector field **V** with rot **V** = 0 is said to be **irrotational**, (**lamellar**, or **without vortex**). A vector field **V** with div **V** = 0 is said to be **solenoidal** (or **without source**). Thus grad f is irrotational and rot **V** is solenoidal. In a small neighborhood or in a †simply-connected domain, an irrotational field is a gradient, a solenoidal field is a rotation, and an arbitrary vector field **V** is the sum of these two kinds of vector fields: $V = \operatorname{grad}\varphi + \operatorname{rot}\mathbf{u}$ (**Helmholtz theorem**); the function φ is called the **scalar potential** of **V**, and the vector field **u** is called the **vector potential** of **V**. Furthermore, the operator $\nabla^2 = \nabla\nabla = \operatorname{div}\operatorname{grad} = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ is called the **Laplace operator** (or **Laplacian**) and is denoted by Δ. A function that satisfies $\Delta\varphi = 0$ is called a †harmonic function. Locally, an irrotational and solenoidal vector field is the gradient of a harmonic function. If **A** is a vector field whose components are $(\varphi_1, \varphi_2, \varphi_3)$ (i.e., $\mathbf{A}(\mathbf{v}) = (\varphi_1(\mathbf{v}), \varphi_2(\mathbf{v}), \varphi_3(\mathbf{v}))$), we can let Δ operate on **A** by setting $\Delta\mathbf{A} = (\Delta\varphi_1, \Delta\varphi_2, \Delta\varphi_3)$. We then have $\Delta\mathbf{A} = \nabla^2\mathbf{A} = \operatorname{grad}\operatorname{div}\mathbf{A} - \operatorname{rot}\operatorname{rot}\mathbf{A}$.

Suppose that we are given a vector field **V** and a curve C such that the vector $\mathbf{V}(p)$ is tangent to the curve at each point $p \in C$. The curve C is the †integral curve of the vector field **V** and is called the **vector line** of the vector field **V**. The set of all vector lines intersecting with a given closed curve C is called a **vector tube**. Given a closed curve C and a vector field **V**, the †curvilinear integral $\int(\mathbf{V}, d\mathbf{s})$ (where $d\mathbf{s}$ is the line element of C) is called the **circulation** (of **V**) along the closed curve C. A vector field is irrotational if its circulation along every closed curve vanishes; the converse is true in a simply connected domain. Further, let v_n be the †normal component of a vector field **V** with respect to a surace S, and let dS be the volume element of the surface. We put $\mathbf{n}dS = d\mathbf{S}$, where **n** is the unit normal vector in the positive direction of the surface S. Then the †surface integral $\int v_n dS = \int(\mathbf{V}, d\mathbf{S})$ is called the **vector flux** through the surface S. A vector field whose vector flux vanishes for every closed surface is solenoidal. (For the corresponding formulas → 94 Curvilinear Integrals and Surface Integrals. For generalizations to higher-dimensional manifolds → 105 Differentiable Manifolds; 194 Harmonic Integrals; Appendix A, Table 3.)

References

[1] H. Weyl, Raum, Zeit, Materie, Springer, fifth edition, 1923; English translation, Space, time, matter, Dover, 1952.
[2] S. Banach, Mechanics, Warsaw, 1951.
[3] H. Flanders, Differential forms, with applications to the physical sciences, Academic Press, 1963.
[4] H. K. Nickerson, D. C. Spencer, and N. E. Steenrod, Advanced calculus, Van Nostrand, 1959.

443 (XII.8) Vector-Valued Integrals

A. General Remarks

Integrals whose values are elements (or subsets) of †topological linear spaces are generally called **vector-valued integrals** or **vector integrals**. As in the scalar case, there are vector-valued integrals of Riemann type (→ 37 Banach Spaces K) and of Lebesgue type. In this article we consider only the latter. There are cases where integrands are vector-valued, where measures are vector-valued, and where both are vector-valued. The methods of integration are also divided into the strong type, in which the integrals are defined by means of the original topology of the topological linear space X, and the weak type, in which they are reduced to numerical integrals by applying continuous linear functionals on X. Combining these we can define many kinds of integrals.

Historically, D. Hilbert's †spectral resolution is the first example of vector-valued integrals, but the general theory of vector-valued integrals started only after S. Bochner [1] defined in 1933 an integral of strong type for functions with values in a Banach space with respect to numerical measures. Then G. Birkhoff [2] defined a more general integral by replacing absolutely convergent sums with unconditionally convergent sums. At approximately the same time, N. Dunford introduced integrals equivalent to these. Later, R. S. Phillips (*Trans. Amer. Math. Soc.*, 47 (1940)) generalized the definition to the case where values of functions are in a †locally convex topological linear space, and C. E. Rickart (*Trans. Amer. Math. Soc.*, 52 (1942)) to the case where functions take subsets of a locally convex topological linear space as their values. The theory of integrals of weak type for functions with values in a Banach space and numerical measures was constructed by I. M. Gel'fand [3], Dunford [4], B. J. Pettis [5], and others (1936–1938). N. Bourbaki [6] dealt with the case where integrands take values in a locally convex topological linear space. As for integrals of numerical functions by vector-valued measures, a representative of strong type integrals is the integral of R. G. Bartle, Dunford, and J. T. Schwartz [7] (1955). Weak type integrals have been discussed by Bourbaki [6], D. R. Lewis (*Pacific J. Math.*, 33 (1970)), and I. Kluvánek (*Studia Math.*, 37 (1970)). The bilinear integral of Bartle (*Studia Math.*, 15 (1956)) is typical of integrals in the case where both integrands and measures are vector-valued. For interrelations of these integrals → the papers by Pettis and Bartle cited above and T. H. Hildebrandt's report in the *Bulletin of the American Mathematical Society*, 59 (1953).

Since the earliest investigations [1–3] the main aim of the theory of vector-valued integrals has been to obtain integral representations of vector-valued (set) functions and various linear operators [8]. However, there is the fundamental difficulty of the nonvalidity of the Radon-Nikodým theorem. Whatever definition of integrals we take, the theorem does not hold for vector-valued set functions unconditionally. Many works sought conditions for functions, operators, or spaces such that the conclusion of the theorem would be restored; the works of Dunford and Pettis [9] and Phillips (*Amer. J. Math.*, 65 (1943)) marked a summit of these attempts. Later, after A. Grothendieck's investigations (1953–1956), this problem began to be studied again, beginning in the late 1960s, by many mathematicians (→ J. Diestel and J. J. Uhl, Jr., *Rocky Mountain J. Math.*, 6 (1976); [10]).

Recently, integrals of multivalued vector-valued functions have also been employed in mathematical statistics, economics, control theory, and many other fields. Some contributions are, besides Rickart cited above, G. B. Price (*Trans. Amer. Math. Soc.*, 47 (1940), H. Kudo (*Sci. Rep. Ochanomizu Univ.*, 4 (1953)), H. Richter (*Math. Ann.*, 150 (1963)), R. J. Aumann [11], G. Debreu [12], and M. Hukuhara (*Funkcial. Ekvac.*, 10 (1967)). Furthermore, C. Castaing and M. Varadier [13] have defined weak type integrals of multivalued functions and introduced many results concerning them. In the following we shall give explanations of typical vector-valued integrals with values in a Banach space only.

B. Measurable Vector-Valued Functions

Let $x(s)$ be a function defined on a †σ-finite measure space $(S, \mathfrak{S}, \mu)$ with values in a Banach space X. This is called a **simple function** or **finite-valued function** if there exists a partition of S into a finite number of mutually disjoint measurable sets $A_1, A_2, \ldots, A_n$ in each of which $x(s)$ takes a contant value c_j. Then $x(s)$ can be written as $\sum_{j=1}^{n} c_j \chi_{A_j}(s)$, where $\chi_{A_j}(s)$ is the †characteristic function of A_j. A function $x(s)$ is said to be **measurable** or **strongly measurable** if it is the strong limit of a sequence of simple functions almost everywhere, that is, $\lim_{n\to\infty} \|x_n(s)-x(s)\|=0$ a.e. Then the numerical function $\|x(s)\|$ is measurable. If μ is a †Radon measure on a compact Hausdorff space S, then the measurable functions can be characterized by †Luzin's property (→ 270 Measure Theory I).

A function $x(s)$ is said to be **scalarly measurable** or **weakly measurable** if the numerical function $\langle x(s), x'\rangle$ is measurable for any †continuous linear functional $x' \in X'$. A function $x(s)$ is measurable if and only if it is scalarly measurable and there are a †null set $E_0 \subset S$ and a †separable closed subspace $Y \subset X$ such that $x(s) \in Y$ whenever $s \notin E_0$ (**Pettis measurability theorem**).

C. Bochner Integrals

A measurable vector-valued function $x(s)$ is said to be **Bochner integrable** if the norm $\|x(s)\|$ is †integrable. If $x(s)$ is a Bochner integrable simple function $\sum c_j \chi_{A_j}(s)$, then its Bochner integral is defined by

$$\int_S x(s)\,d\mu = \sum \mu(A_j) c_j.$$

For a general Bochner integrable function $x(s)$ there exists a sequence of simple functions satisfying the following conditions: (i) $\lim_{n\to\infty} \|x_n(s)-x(s)\|=0$ a.e. (ii) $\lim_{n\to\infty} \int_S \|x_n(s)-x(s)\|\,d\mu=0$. Then $\int_S x_n(s)\,d\mu$ converges strongly and its limit does not depend on the choice of the sequence $\{x_n(s)\}$. We call the limit the **Bochner integral** of $x(s)$ and denote it by $\int_S x(s)\,d\mu$ or by $(\mathrm{Bn})\int_S x(s)\,d\mu$ to distinguish it from other kinds of integrals. A Bochner integrable function on S is Bochner integrable on every measurable subset of S. The Bochner integral has the basic properties of Lebesgue integrals, such as linearity, †complete additivity, and †absolute continuity, with absolute values replaced by norms. †Lebesgue's convergence theorem and †Fubini's theorem also hold. However, the Radon-Nikodým theorem does not hold in general (→ Section H). Let T be a †closed linear operator from X to another Banach space Y. If both $x(s)$ and $Tx(s)$ are Bochner integrable, then the integral of $x(s)$ belongs to the domain of T and

$$T\left(\int_S x(s)\,d\mu\right) = \int_S Tx(s)\,d\mu.$$

If, in particular, T is bounded, then the assumption is always satisfied. If μ is the †Lebesgue measure on the Euclidean space $\mathbf{R}^n$, then Lebesgue's differentiability theorem holds for the Bochner integrals regarded as a set function on the regular closed sets (→ 380 Set Functions D).

D. Unconditionally Convergent Series

Let $\sum_{j=1}^{\infty} x_j$ be a series of elements x_j of a Banach space X. It is said to be **absolutely convergent** if $\sum \|x_j\| < \infty$. It is called **unconditionally convergent** if for any rearrangement α the resulting series $\sum x_{\alpha(j)}$ converges strongly. Then the sum does not depend on α. Clearly, an absolutely convergent series is unconditionally convergent. If X is the number space or is finite-dimensional, then the converse holds. However, if X is infinite-dimensional, there is always an unconditionally convergent series which is not absolutely convergent (**Dvoretzky-Rogers theorem**).

A series $\sum x_j$ is unconditionally convergent if and only if each subseries converges weakly (**Orlicz-Pettis theorem**). If $\sum x_j$ is an unconditionally convergent series, then $\sum \langle x_j, x' \rangle$ converges absolutely for any continuous linear functional $x' \in X'$. If X is a Banach space containing no closed linear subspace isomorphic to the †sequence space c_0, then conversely a series $\sum x_j$ converges unconditionally whenever $\sum |\langle x_j, x' \rangle| < \infty$ for any $x' \in X'$ (**Bessaga-Pełczyński theorem**). A Banach space that is †sequentially complete relative to the weak topology, such as a †reflexive Banach space, and a separable Banach space that is the dual of another Banach space, such as l_1 and the †Hardy space $H_1(\mathbf{R}^n)$, satisfy the assumption, while c_0, l_∞, and $L_\infty(\Omega)$ for an infinitely divisible Ω do not. The totality of absolutely convergent series (resp. unconditionally convergent series) in X is identified with the †topological tensor product $l_1 \hat{\otimes} X$ (resp. $l_1 \hat{\hat{\otimes}} X$) (Grothendieck).

E. Birkhoff Integrals

We say that a series $\sum B_j$ of subsets of X converges unconditionally if for any $x_j \in B_j$ the series $\sum x_j$ converges unconditionally. Then $\sum B_j$ denotes the set of such sums. A vector-valued function $x(s)$ is said to be **Birkhoff integrable** if there is a countable partition Δ: $S = \bigcup_{j=1}^{\infty} A_j$ ($A_j \in \mathfrak{S}$, $A_j \cap A_k = \varnothing$ $(j \neq k)$, $\mu(A_j) < \infty$) such that the set $x(A_j)$ of values on A_j are bounded and $\sum \mu(A_j) x(A_j)$ converges unconditionally and if the sum converges to an element of X as the partition is subdivided. The limit is called the **Birkhoff integral** of $x(s)$ and is denoted by $(\mathrm{Bk})\int_S x(s)\,d\mu$ or simply by $\int_S x(s)\,d\mu$. A Birkhoff integrable function is Birkhoff integrable on any measurable set. The Birkhoff integral has, as a set function, complete additivity and absolute continuity in μ. It is linear in the integrand but Fubini's theorem and the Radon-Nikodým theorem do not hold. A Bochner integrable function is Birkhoff integrable, and the integrals coincide. The converse does not hold.

F. Gel'fand-Pettis Integrals

A scalarly measurable function $x(s)$ is said to be **scalarly integrable** or **weakly integrable** if for each $x' \in X'$, $\langle x(s), x' \rangle$ is integrable. Then the linear functional x^* on X' defined by

$$\int_S \langle x(s), x' \rangle\, d\mu = \langle x', x^* \rangle$$

is called the **scalar integral** of $x(s)$. Gel'fand [3] and Dunford [4] proved that x^* belongs to the bidual X''. Hence scalarly integrable functions are often called **Dunford integrable** and the integrals x^* the **Dunford integrals**. More generally, Gel'fand [3] showed that if $x'(s)$ is a function with values in the dual X' of a Banach space X such that $\langle x, x'(s) \rangle$ is integrable for any $x \in X$, then there is an $x' \in X'$ satisfying

$$\int_S \langle x, x'(s) \rangle\, d\mu = \langle x, x' \rangle.$$

This element is sometimes called the **Gel'fand integral** of $x'(s)$. A scalarly integrable function $x(s)$ is scalarly integrable on any measurable subset A. If the scalar integral is always in X, i.e., for each A there is an $x_A \in X$ such that

$$\int_A \langle x, x'(s) \rangle\, d\mu = \langle x_A, x' \rangle, \quad x' \in X',$$

then $x(s)$ is said to be **Pettis integrable** or **Gel'fand-Pettis integrable** and x_A is called the **Pettis integral** or **Gel'fand-Pettis integral** on A and is denoted by $(\mathrm{P})\int_A x(s)\,d\mu$ or simply by $\int_A x(s)\,d\mu$. The Pettis integral has complete additivity and absolute continuity as a set function, similarly to the Birkhoff integral. Again, Fubini's theorem and the Radon-Nikodým theorem do not hold. The scalar integral on measurable sets of a scalarly integrable function $x(s)$ is completely additive and absolutely continuous with respect to the †weak* topology of X'' as the dual to X'. It is completely additive or absolutely continuous in the norm topology if and only if $x(s)$ is Pettis integrable (Pettis [5]; [10]). If $x(s)$ is Pettis integrable and $f(s)$ is a numerical function in $L_\infty(S)$, then the product $f(s)x(s)$ is Pettis integrable. Birkhoff integrable functions are Pettis integrable, and the integrals coincide. Conversely, if a measurable function is Pettis integrable, then it is Birkhoff integrable. When X satisfies the Bessaga-Pełczyński condition (→ Section D), a measurable scalarly integrable function is Pettis integrable.

G. Vector Measures

Let Φ be a set function defined on a completely additive class $\mathfrak{S}$ of subsets of the space S and with values in a Banach space X. It is called a **finitely additive vector measure** (resp. a **completely additive vector measure** or simply a **vector measure**) if $\Phi(A_1 \cup A_2) = \Phi(A_1) + \Phi(A_2)$ whenever A_1 and $A_2 \in \mathfrak{S}$ are disjoint (resp. $\Phi(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \Phi(A_j)$ in the norm topology for all $A_j \in \mathfrak{S}$ such that $A_j \cap A_k = \varnothing$ $(j \neq k)$). We remark that the latter sum always converges unconditionally. A set function Φ is completely additive if and only if $\langle \Phi(A), x' \rangle$ is completely additive for all $x' \in X'$ (**Pettis complete additivity theorem**).

Let Φ be a finitely additive vector measure and E be a measurable set. The **total variation** of Φ on E and the **semivariation** of Φ on E are defined by

$$V(\Phi)(E) = \sup \sum_{j=1}^{n} \|\Phi(A_j)\| \tag{1}$$

and

$$\|\Phi\|(E) = \sup \left\| \sum_{j=1}^{n} \alpha_j \Phi(A_j) \right\|, \tag{2}$$

respectively, where the suprema are taken over all finite partitions of $E: E = \bigcup A_j$ $(A_j \in \mathfrak{S}, A_j \cap A_k = \varnothing (j \neq k))$ and all numbers α_j with $|\alpha_j| \leqslant 1$. If $V(\Phi)(S) < \infty$, then Φ is called a measure **of bounded variation**. $\|\Phi\|(S) < \infty$ if and only if $\sup\{\|\Phi(A)\| \mid A \in \mathfrak{S}\} < \infty$. Then Φ is said to be **bounded**. The function $V(\Phi)(E)$ of E is finitely additive but $\|\Phi\|(E)$ is only subadditive: $\|\Phi\|(A \cup B) \leqslant \|\Phi\|(A) + \|\Phi\|(B)$. If Φ is a vector measure of bounded variation, then $V(\Phi)$ is a positive measure. Every vector measure is bounded. A completely additive vector measure on a †finitely additive class $\mathfrak{L}$ can uniquely be extended to a vector measure on the completely additive class $\mathfrak{S}$ generated by $\mathfrak{L}$ (Kluvánek).

Let μ be a positive measure and Φ be a vector measure. Then we have $\Phi(A) \to 0$ as $\mu(A) \to 0$ if and only if Φ vanishes on every A with $\mu(A) = 0$. Then Φ is said to be **absolutely continuous** with respect to μ. For every vector measure Φ there is a measure μ such that $\|\Phi\|(A) \to 0$ as $\mu(A) \to 0$ and that $0 \leqslant \mu(A) \leqslant \|\Phi\|(A)$ (Bartle, Dunford, and Schwartz). As a set function, the Bochner integral is a vector measure of bounded variation and the Pettis integral is a bounded vector measure. Both are absolutely continuous with respect to the integrating measure. Let X be $L_p(0,1)$ for $1 \leqslant p \leqslant \infty$, and define $\Phi(E)$ for a Lebesgue measurable set E to be the characteristic function of E. If $p = 1$, Φ is a vector measure of bounded variation. If $1 < p < \infty$, Φ is a bounded vector measure, but it is not of bounded variation on any set E with $\mu(E) > 0$. If $p = \infty$, then Φ is no longer completely additive. These vector measures are absolutely continuous with respect to the Lebesgue measure, but they cannot be represented as the Bochner integral or the Pettis integral.

Let Φ be a vector measure on $\mathfrak{S}$. An $\mathfrak{S}$-measurable numerical function $f(s)$ is said to be Φ-integrable if there exists a sequence of simple functions $f_n(s)$ such that $f_n(s) \to f(s)$ a.e. and that for each $E \in \mathfrak{S}$, $\int_E f_n(s)\, d\Phi$ converges in the norm of X. Then the limit is independent of the choice of f_n. It is called the **Bartle-Dunford-Schwartz integral** and is denoted by $\int_E f(s)\, d\Phi$. Lebesgue's convergence theorem holds for this integral. If Φ is absolutely continuous with respect to the measure μ, then every $f \in L_\infty(\mu)$ is Φ-integrable, and the operator that maps f to $\int_S f\, d\Phi$ is continuous with respect to the weak* topology in $L_\infty(\mu)$ and the weak topology of X. Hence it is a †weakly compact operator. In particular, the range of a vector measure is relatively compact in the weak topology [7]. If Φ is the vector measure of the Pettis integral of a vector-valued function $x(s)$, then the above integral is equal to the Pettis integral of $f(s)x(s)$.

A vector measure Φ is said to be **nonatomic** if for each set A with $\Phi(A) \neq 0$ there is a subset B of A such that $\Phi(B) \neq 0$ and $\Phi(A \setminus B) \neq 0$. If X is finite-dimensional, then the range of a nonatomic vector measure is a compact convex set (**Lyapunov convexity theorem**). This has been generalized to the infinite-dimensional case in many ways, but the conclusion does not hold in the original form (→ Kluvánek and G. Knowles [15]; [10]).

H. The Radon-Nikodým Theorem

As the above examples show, the †Radon-Nikodým theorem does not hold for vector measures in the original form. From 1967 to 1971, M. Metivier, M. A. Rieffel, and S. Moedomo and Uhl improved the classical result of Phillips (1943) and proved the following theorem.

Radon-Nikodým theorem for vector measures. The following conditions are equivalent for μ-absolutely continuous vector measures Φ defined on a finite measure space $(S, \mathfrak{S}, \mu)$: (i) There is a Pettis integrable measurable function $x(s)$ such that

$$\Phi(A)=(\mathrm{P})\int_A x(s)\,d\mu.$$

(ii) For each $\varepsilon>0$ there is an $E\in\mathfrak{S}$ such that $\mu(S\setminus E)<\varepsilon$ and such that $\{\Phi(A)/\mu(A)\,|\,A\in\mathfrak{S}, A\subset E\}$ is relatively compact. (iii) For each $E\in\mathfrak{S}$ with $\mu(E)>0$ there is a subset F of E with $\mu(F)>0$ such that $\{\Phi(A)/\mu(A)\,|\,A\in\mathfrak{S}, A\subset F\}$ is relatively weakly compact. Then Φ is of bounded variation if and only if $x(s)$ is Bochner integrable.

On the other hand, since Birkhoff and Gel'fand it has been known that for special Banach spaces X every μ-absolutely continuous vector measure of bounded variation with values in X can be represented as a Bochner integral with respect to μ. Such spaces are said to have the **Radon-Nikodým property**. Separable dual spaces (Gel'fand, Pettis; Dunford and Pettis), reflexive spaces (Gel'fand, Pettis, Phillips), and $l_1(\Omega)$, Ω arbitrary, etc., have the Radon-Nikodým property, while $L_\infty(0,1)$ (Bochner), c_0 (J. A. Clarkson), $L_1(\Omega)$ on a nonatomic Ω (Clarkson, Gel'fand), and $C(\Omega)$ on an infinite compact Hausdorff space Ω, etc., do not. Gel'fand proved that $L_1(0,1)$ (and c_0) is not a dual by means of this fact. From 1967 to 1974, Riefell, H. B. Maynard, R. E. Huff, and W. J. Davis and R. P. Phelps succeeded in characterizing geometrically the Banach spaces with the Radon-Nikodým property. We know today that the following conditions for Banach spaces X are equivalent [10]: (i) X has the Radon-Nikodým property. (ii) Every separable closed linear subspace of X has the Radon-Nikodým property. (iii) Every function $f:[0,1]\to X$ of bounded variation is (strongly or weakly) differentiable a.e. (iv) For any finite measure space $(S,\mathfrak{S},\mu)$ and bounded linear operator $T: L_1(S)\to X$, there is an essentially bounded measurable function $x(s)$ with values in X such that

$$Tf=\int_S f(s)x(s)\,d\mu, \qquad f\in L_1(S).$$

(v) Each nonvoid bounded closed convex set K in X is the †closed convex hull of the set of its strongly exposed points, where a point $x_0\in K$ is called a **strongly exposed point** of K if there is an $x'\in X'$ such that $\langle x_0,x'\rangle>\langle x,x'\rangle$ for all $x\in K\setminus\{x_0\}$ and that any sequence $x_n\in K$ with $\lim\langle x_n,x'\rangle=\langle x_0,x'\rangle$ converges to x_0 strongly.

A Banach space X is said to have the **Kreĭn-Mil'man property** if each bounded closed convex set in X is the closed convex hull of its †extreme points. A Banach space X with the Radon-Nikodým property has the Kreĭn-Mil'man property (J. Lindenstrauss). If X is a dual space, then the converse holds (Huff and P. D. Morris). A Banach space with the Kreĭn-Mil'man property clearly has no closed linear space isomorphic to c_0, but there are Banach spaces that do not contain c_0 and do not have the Kreĭn-Mil'man property. The dual X of a Banach space Y has the Radon-Nikodým property if and only if the dual of every separable closed linear subspace of Y is separable (Uhl, C. Stegall).

I. Integrals of Multivalued Vector Functions

Let $\Gamma(s)$ be a multivalued function defined on a σ-finite complete measure space $(S,\mathfrak{S},\mu)$ with values that are nonempty closed subsets of a separable Banach space X. The inverse image of a subset E of X under $\Gamma(s)$ is, by definition, the set of all s such that $\Gamma(s)\cap E\neq\varnothing$. $\Gamma(s)$ is said to be **measurable** or **strongly measurable** if the inverse image of each open set in X under $\Gamma(s)$ belongs to $\mathfrak{S}$. Let $\mathfrak{B}(X)$ be the †Borel field of X, and $\mathfrak{S}\times\mathfrak{B}(X)$ be the product completely additive class, that is, the smallest completely additive class containing all direct products $A\times B$ of $A\in\mathfrak{S}$ and $B\in\mathfrak{B}(X)$. Then the measurability of $\Gamma(s)$ is equivalent to each of the following: (i) The graph $\{(s,x)\,|\,x\in\Gamma(s), s\in\mathfrak{S}\}$ of $\Gamma(s)$ belongs to $\mathfrak{S}\times\mathfrak{B}(X)$. (ii) The inverse image of every Borel set in X under $\Gamma(s)$ belongs to $\mathfrak{S}$. (iii) For each $x\in X$, the distance $d(x,\Gamma(s))=\inf\{\|x-y\|\,|\,y\in\Gamma(s)\}$ between x and $\Gamma(s)$ is measurable as a function on S.

A measurable function $x(s)$ on S with values in X is called a **measurable selection** of $\Gamma(s)$ if $x(s)$ is in $\Gamma(s)$ for all s. (X being separable, we need not discriminate between strong and weak measurability.) The measurability of $\Gamma(s)$ is also equivalent to the following important statement on the existence of measurable selections of $\Gamma(s)$: (iv) There are a countable number of measurable selections $\{x_n(s)\}$ of $\Gamma(s)$ such that the closure of the set $\{x_n(s)\,|\,n=1,2,\dots\}$ coincides with $\Gamma(s)$ for all $s\in S$. $\Gamma(s)$ is said to be **scalarly measurable** or **weakly measurable** if the support function $\delta'(x',\Gamma(s))=\sup\{\langle x,x'\rangle\,|\,x\in\Gamma(s)\}$ is measurable on S for all $x'\in X'$. The strong measurability of $\Gamma(s)$ clearly implies the weak one. If the values of $\Gamma(s)$ are nonempty weakly compact convex sets, then the measurabilities are equivalent. Hereafter we shall assume that $\Gamma(s)$ takes the values in the weakly compact convex sets. If the support function $\delta'(x',\Gamma(s))$ is integrable on S for all $x'\in X'$, then $\Gamma(s)$ is said to be **scalarly integrable**. Then the **scalar integral** of $\Gamma(s)$ is defined to be the set in X'' of all scalar integrals of its measurable selections, i.e.,

$$\int_S \Gamma(s)\,d\mu=\left\{\int_S x(s)\,d\mu \,\middle|\, x(s) \text{ is a measurable selection of } \Gamma(s)\right\}.$$

If $\|\Gamma(s)\| = \sup\{\|x\| \mid x \in \Gamma(s)\}$ is integrable, then every measurable selection is Bochner integrable and the integral $\int_S \Gamma(s)\,d\mu$ becomes a nonempty weakly compact convex set in X. When the values of $\Gamma(s)$ are nonempty compact convex sets, there is another method, by G. Debreu, of defining the integral. Let $\mathfrak{K}$ be the class of all nonempty compact convex sets in X and δ be the Hausdorff metric, i.e., for K_1 and $K_2 \in \mathfrak{K}$ define $\delta(K_1, K_2) = \max[\sup\{d(x, K_2) \mid x \in K_1\}, \sup\{d(x, K_1) \mid x \in K_2\}]$. Further, for K_1, $K_2 \in \mathfrak{K}$ and $\alpha \geqslant 0$ define the sum and the nonnegative scalar multiple by $K_1 + K_2 = \{x_1 + x_2 \mid x_1 \in K_1, x_2 \in K_2\}$ and $\alpha \cdot K_1 = \{\alpha x \mid x \in K_1\}$, respectively. Then $\mathfrak{K}$ endowed with the Hausdorff metric and the above addition and scalar multiplication is isometrically embedded in a closed convex cone in a separable Banach space Y by the Rådsröm embedding theorem (*Proc. Amer. Math. Soc.*, 3 (1952)). Let φ be this isometry. Then the **(strong) measurability** and the **(strong) integrability** of $\Gamma(s)$ are defined by the measurability and the Bochner integrability of the Y-valued function $\varphi(\Gamma(s))$, respectively, and its (strong) integral as the inverse image of the Bochner integral of $\varphi(\Gamma(s))$ under φ:

$$\int_S \Gamma(s)\,d\mu = \varphi^{-1}\left(\int_S \varphi(\Gamma(s))\,d\mu\right).$$

This definition of integral for strongly measurable $\Gamma(s)$ is shown to be compatible with that mentioned before. It is clear by the definition that the integral value in this case is a nonempty compact convex set and that most properties of Bochner integrals also hold for this integral.

References

[1] S. Bochner, Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind, Fund. Math., 20 (1933), 262–276.
[2] G. Birkhoff, Integration of functions with values in a Banach space, Trans. Amer. Math. Soc., 38 (1935), 357–378.
[3] I. Gel'fand, Abstrakte Funktionen und lineare Operatoren, Mat. Sb., 4 (46) (1938), 235–286.
[4] N. Dunford, Uniformity in linear spaces, Trans. Amer. Math. Soc., 44 (1938), 305–356.
[5] B. J. Pettis, On integration in vector spaces, Trans. Amer. Math. Soc., 44 (1938), 277–304.
[6] N. Bourbaki, Eléments de mathématique, Intégration, Hermann, ch. 6, 1959.
[7] R. G. Bartle, N. Dunford, and J. Schwartz, Weak compactness and vector measures, Canad. J. Math., 7 (1955), 289–305.
[8] N. Dunford and J. T. Schwartz, Linear operators I, Interscience, 1958.
[9] N. Dunford and B. J. Pettis, Linear operations on summable functions, Trans. Amer. Math. Soc., 47 (1940), 323–392.
[10] J. Diestel and J. J. Uhl, Jr., Vector measures, Amer. Math. Soc. Math. Surveys 13 (1977).
[11] R. J. Aumann, Integrals of set-valued functions, J. Math. Anal. Appl., 12 (1965), 1–22.
[12] G. Debreu, Integration of correspondences, Proc. Fifth Berkeley Symp. Math. Statist. Probab., II, pt. I (1967), 351–372.
[13] C. Castaing and M. Valadier, Convex analysis and measurable multifunctions, Springer, 1977.
[14] N. Dinculeanu, Vector measures, Pergamon, 1967.
[15] I. Kluvánek and G. Knowles, Vector measures and control systems, North-Holland, 1975.

444 (XXI.42) Viète, François

François Viète (1540–December 13, 1603) was born in Fontenay-le-Comte, Poitou, in western France. He served under Henri IV, first as a lawyer and later as a political advisor. His mathematics was done in his leisure time. He used symbols for known variables for the first time and established the methodology and principles of symbolic algebra. He also systematized the algebra of the time and used it as a method of discovery. He is often called the father of algebra. He improved the methods of solving equations of the third and fourth degrees obtained by G. Cardano and L. Ferrari. Realizing that solving the algebraic equation of the 45th degree proposed by the Belgian mathematician A. van Roomen can be reduced to searching for $\sin(\alpha/45)$ knowing $\sin\alpha$, he was able to solve it almost immediately. However, he would not acknowledge negative roots and refused to add terms of different degrees because of his belief in the Greek principle of homogeneity of magnitudes. He also contributed to trigonometry and represented the number π as an infinite product.

References

[1] Francisci Vietae, Opera mathematica, F. van Schooten (ed.), Leyden, 1646 (Georg Olms, 1970).
[2] Jacob Klain, Die griechische Logistik und die Entstehung der Algebra I, II, Quellen und

Studien zur Gesch. Math., (B) 3 (1934), 18–105; (B) 3 (1936), 122–235.

445 (XXI.43) Von Neumann, John

John von Neumann (December 28, 1903–February 8, 1957) was born in Budapest, Hungary, the son of a banker. By the time he graduated from the university there in 1921, he had already published a paper with M. Fekete. He was later influenced by H. Weyl and E. Schmidt at the universities of Zürich and Berlin, respectively, and he became a lecturer at the universities of Berlin and Hamburg. He moved to the United States in 1930 and in 1933 became professor at the Institute for Advanced Study at Princeton. In 1954 he was appointed a member of the US Atomic Energy Commission. The fields in which he was first interested were †set theory, theory of †functions of real variables, and †foundations of mathematics. He made important contributions to the axiomatization of set theory. At the same time, however, he was deeply interested in theoretical physics, especially in the mathematical foundations of quantum mechanics. From this field, he was led into research on the theory of †Hilbert spaces, and he obtained basic results in the theory of †operator rings of Hilbert spaces. To extend the theory of operator rings, he introduced †continuous geometry. Among his many famous works are the theory of †almost periodic functions on a group and the solving of †Hilbert's fifth problem for compact groups. In his later years, he contributed to †game theory and to the design of computers, thus playing a major role in all fields of applied mathematics.

References

[1] J. von Neumann, Collected works I–VI, Pergamon, 1961–1963.
[2] J. von Neumann, 1903–1957, J. C. Oxtoby, B. J. Pettis, and G. B. Price (eds.), Bull. Amer. Math. Soc., 64 (1958), 1–129.
[3] J. von Neumann, Mathematische Grundlagen der Quantenmechanik, Springer, 1932.
[4] J. von Neumann, Functional operators I, II, Ann. Math. Studies, Princeton Univ. Press, 1950.
[5] J. von Neumann, Continuous geometry, Princeton Univ. Press, 1960.
[6] J. von Neumann and O. Morgenstern, Theory of games and economic behavior, Princeton Univ. Press, third edition, 1953.

W

446 (XX.13) Wave Propagation

A disturbance originating at a point in a medium and propagating at a finite speed in the medium is called a **wave**. For example, a sound wave propagates a change of density or stress in a gas, liquid, or solid. A wave in an elastic solid body is called an elastic wave. **Surface waves** appear near the surface of a medium, such as water or the earth. When electromagnetic disturbances are propagated in a gas, liquid, or solid or in a vacuum, they are called **electromagnetic waves**. Light is a kind of electromagnetic wave. According to †general relativity theory, gravitational action can also be propagated as a wave.

It many cases waves can be described by the **wave equation**:

$$\frac{\partial^2\psi}{\partial t^2}=c^2\left(\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}+\frac{\partial^2\psi}{\partial z^2}\right).$$

Here t is time, x, y, z are the Cartesian coordinates of points in the space, c is the propagation velocity, and ψ represents the state of the medium.

If we take a closed surface surrounding the origin of the coordinate system, the state $\psi(0,t)$ at the origin at time t can be determined by the state at the points on the closed surface at time $t-r/c$, with r the distance of the point from the origin. More precisely, we have

$$\psi(0,t)=\frac{1}{4\pi}\int\left(\psi\frac{\partial}{\partial n}\left(\frac{1}{r}\right)-\frac{1}{r}\frac{\partial\psi}{\partial n}-\frac{1}{cr}\frac{\partial\psi}{\partial t}\frac{\partial r}{\partial n}\right)_{t-r/c}df.$$

Here n is the inward normal at any point of the closed surface, and the integral is taken over the surface, while the value of the integrand is taken at time $t-r/c$. This relation is a mathematical representation of **Huygens's principle**, which is valid for the 3-dimensional case but does not hold for the 2-dimensional case (→ 325 Partial Differential Equations of Hyperbolic Type).

A **plane wave** propagating in the direction of a unit vector $\mathbf{n}$ can be represented by $\psi=F(t-\mathbf{n}\cdot\mathbf{r}/c)$, where F is an arbitrary function and $\mathbf{r}(x,y,z)$ is the position vector. The simplest case is given by a **sine wave (sinusoidal wave)**: $\psi=A\sin(\omega t-\mathbf{k}\cdot\mathbf{r}+\delta)$. Here A **(amplitude)** and δ **(phase constant)** are arbitrary constants, $\mathbf{k}$ is in the direction of wave propagation and satisfies the relation $|\mathbf{k}|c=\omega$. ω is the **angular frequency**, $\omega/2\pi$ the **frequency**, $\mathbf{k}$ the **wave number vector**, $|\mathbf{k}|$ the **wave number**, $2\pi/\omega$ the **period**, and $2\pi/|\mathbf{k}|$ the **wavelength**. The velocity with which the crest of the wave advances is equal to $\omega/|\mathbf{k}|=c$ and is called the **phase velocity**.

A **spherical wave** radiating from the origin can generally be represented by

$$\psi=\sum_n\varphi_n\left(\frac{d}{r\,dr}\right)^n\frac{1}{r}F\left(t-\frac{r}{c}\right),$$

where φ_n is the †solid harmonic of order n.

Waves are not restricted to those governed by the wave equation. In general, ψ is not a scalar, but has several components (e.g., ψ may be a vector), which satisfy a set of simultaneous differential equations of various kinds. Usually they have solutions in the form of sinusoidal waves, but the phase velocity $c=\omega/|\mathbf{k}|$ is generally a function of the wavelength λ. Such a wave, called a **dispersive wave**, has a propagation velocity (velocity of propagation of the disturbance through the medium) that is not equal to the phase velocity. A disturbance of finite extent that can be approximately represented by a plane wave is propagated with a velocity $c-\lambda dc/d\lambda$, called the **group velocity**. Often there exists a definite relationship between the amplitude vector $\mathbf{A}$ (and the corresponding phase constant δ) and wave number vector $\mathbf{k}$, in which case the wave is said to be **polarized**. In particular, when $\mathbf{A}$ and $\mathbf{k}$ are parallel (perpendicular), the wave is called a **longitudinal (transverse) wave**. Usually equations governing the wave are linear, and therefore superposition of two solutions gives a new solution (†principle of superposition). Superposition of two sinusoidal waves traveling in opposite directions gives rise to a wave whose crests do not move (e.g., $\psi=A\sin\omega t\sin\mathbf{k}\cdot\mathbf{r}$). Such a wave is called a **stationary wave**. Since the energy of a wave is proportional to the square of ψ, the energy of the resultant wave formed by superposition of two waves is not equal to the sum of the energies of the component waves. This phenomenon is called **interference**. When a wave reaches an obstacle it propagates into the shadow region of the obstacle, where there is formed a special distribution of energy dependent on the shape and size of the obtacle. This phenomenon is called **diffraction**.

For aerial sound waves and water waves, if the amplitude is so large that the wave equation is no longer valid, we are faced with †nonlinear problems. For instance, **shock waves** appear in the air when surfaces of discontinuity of density and pressure exist. They appear in explosions and for bodies traveling at high speeds. Concerning wave mechanics dealing with atomic phenomena → 351 Quantum Mechanics.

References

[1] H. Lamb, Hydrodynamics, Cambridge Univ. Press, sixth edition, 1932.
[2] Lord Rayleigh, The theory of sound, Macmillan, second revised edition, I, 1937; II, 1929.
[3] M. Born and E. Wolf, Principles of optics, Pergamon, fourth edition, 1970.
[4] F. S. Crawford, Jr., Waves, Berkeley phys. course III, McGraw-Hill, 1968.
[5] C. A. Coulson, Waves; A mathematical theory of the common type of wave motion, Oliver & Boyd, seventh edition, 1955.
[6] L. Brillouin, Wave propagation and group velocity, Academic Press, 1960.
[7] I. Tolstoy, Wave propagation, McGraw-Hill, 1973.
[8] J. D. Achenbach, Wave propagation in elastic solids, North-Holland, 1973.
[9] K. F. Graff, Wave motion in elastic solids, Ohio State Univ. Press, 1975.
[10] J. Lighthill, Waves in fluids, Cambridge Univ. Press, 1978.
[11] R. Courant and D. Hilbert, Methods of mathematical physics II, Interscience, 1962.

447 (XXI.44) Weierstrass, Karl

Karl Weierstrass (October 31, 1815–February 19, 1897) was born into a Catholic family in Ostenfelde, in Westfalen, Germany. From 1834 to 1838 he studied law at the University of Bonn. In 1839 he moved to Münster, where he came under the influence of C. Gudermann, who was then studying the theory of elliptic functions. From this time until 1855, he taught in a parochial junior high school; during this period he published an important paper on the theory of analytic functions. Invited to the University of Berlin in 1856, he worked there with L. Kronecker and E. E. Kummer. In 1864, he was appointed to a full professorship, which he held until his death.

His foundation of the theory of analytic functions of a complex variable at about the same time as Riemann is his most fundamental work. In contrast to Riemann, who utilized geometric and physical intuition, Weierstrass stressed the importance of rigorous analytic formulation. Aside from the theory of analytic functions, he contributed to the theory of functions of real variables by giving examples of continuous functions that were nowhere differentiable. With his theory of †minimal surfaces, he also contributed to geometry. His lectures at the University of Berlin drew many listeners, and in his later years he was a respected authority in the mathematical world.

References

[1] K. Weierstrass, Mathematische Werke I–VII, Mayer & Miller, 1894–1927.
[2] F. Klein, Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert I, Springer, 1926 (Chelsea, 1956).

448 (XXI.45) Weyl, Hermann

Hermann Weyl (November 9, 1885–December 8, 1955) was born in Elmshorn in the state of Schleswig-Holstein in Germany. Entering the University of Göttingen in 1904, he also audited courses for a time at the University of Münich. In 1908, he obtained his doctorate from the University of Göttingen with a paper on the theory of integral equations, and by 1910 he was a lecturer at the same university. In 1913, he became a professor at the Federal Technological Institute at Zürich; in 1928–1929, a visiting professor at Princeton University; in 1930, a professor at the University of Göttingen; and in 1933, a professor at the Institute for Advanced Study at Princeton. He retired from his professorship there in 1951, when he became professor emeritus. He died in Zürich in 1955.

Weyl contributed fresh and fundamental works covering all aspects of mathematics and theoretical physics. Among the most notable are results on problems in †integral equations, †Riemann surfaces, the theory of †Diophantine approximation, the representation of groups, in particular compact groups and †semisimple Lie groups (whose structure he elucidated), the space-time problem, the introduction of †affine connections in differential geometry, †quantum mechanics, and the foundations of mathematics. In his later years, with his son Joachim he studied meromorphic functions. In addition to his many mathematical works he left works in philosophy, history, and criticism.

References

[1] H. Weyl, Gesammelte Abhandlungen I–IV, Springer, 1968.
[2] H. Weyl, Die Idee der Riemannschen Fläche, Teubner, 1913, revised edition, 1955; English translation, The concept of a Riemann surface, Addison-Wesley, 1964.

[3] H. Weyl, Raum, Zeit, Materie, Springer, 1918, fifth edition, 1923; English translation, Space, time, matter, Dover, 1952.
[4] H. Weyl, Das Kontinuum, Veit, 1918.
[5] H. Weyl, Gruppentheorie und Quantenmechanik, Hirzel, 1928.
[6] H. Weyl, Classical groups, Princeton Univ. Press, 1939, revised edition, 1946.
[7] H. Weyl and F. J. Weyl, Meromorphic functions and analytic curves, Princeton Univ. Press, 1943.
[8] H. Weyl, Philosophie der Mathematik und Naturwissenschaften, Oldenbourg, 1926; English translation, Philosophy of mathematics and natural science, Princeton Univ. Press, 1949.
[9] H. Weyl, Symmetry, Princeton Univ. Press, 1952.

449 (III.18) Witt Vectors

A. General Remarks

Let Γ be an †integral domain of characteristic 0, and p a fixed prime number. For each infinite-dimensional vector $x=(x_0, x_1, \ldots)$ with components in Γ, we define its **ghost components** $x^{(0)}, x^{(1)}, \ldots$ by $x^{(0)}=x_0, x^{(n)}=x_0^{p^n}+px_1^{p^{n-1}}+\ldots+p^n x_n$. We define the sum of the vectors x and $y=(y_0, y_1, \ldots)$ to be the vector with ghost components $x^{(0)}+y^{(0)}, x^{(1)}+y^{(1)}, \ldots$, and their product to be the vector with ghost components $x^{(0)}y^{(0)}, x^{(1)}y^{(1)}, \ldots$. The sum and product are uniquely determined vectors with components in Γ. Writing their first two terms explicitly, we have

$$x+y$$
$$=\left[x_0+y_0, x_1+y_1-\sum_{\nu=1}^{p-1}\frac{1}{p}\binom{p}{\nu}x_0^\nu y_0^{p-\nu}, \ldots\right],$$
$$xy=(x_0y_0, x_1y_0^p+y_1x_0^p+px_1y_1, \ldots).$$

In general, it can be proved that the nth components $\sigma_n(x,y)$ and $\pi_n(x,y)$ of the sum and product are polynomials in $x_0, y_0, x_1, y_1, \ldots, x_n, y_n$ whose coefficients are rational integers. With these operations of addition and multiplication, the set of these vectors forms a †commutative ring, of which the zero element is $(0,0,\ldots)$ and the unity element is $(1,0,\ldots)$. Let k be a field of characteristic p. For vectors $(\xi_0, \xi_1, \ldots)$, and $(\eta_0, \eta_1, \ldots)$ with components in k, we define their sum and product by $(\xi_0, \xi_1, \ldots)+(\eta_0, \eta_1, \ldots)=(\ldots, \sigma_n(\xi, \eta), \ldots)$ and $(\xi_0, \xi_1, \ldots)(\eta_0, \eta_1, \ldots)=(\ldots, \pi_n(\xi, \eta), \ldots)$. Since the coefficients of σ_n and π_n are rational integers, these operations are well defined. With these operations, the set of such vectors becomes an integral domain $W(k)$ of characteristic 0. Elements of $W(k)$ are called **Witt vectors** over k.

If we put $V(\xi_0, \xi_1, \ldots)=(0, \xi_0, \xi_1, \ldots)$ and $(\xi_0, \xi_1, \ldots)^p=(\xi_0^p, \xi_1^p, \ldots)$, we get the formula $p\xi=V\xi^p$. (Note that this ξ^p is not the pth power of ξ in $W(k)$ in the usual sense.) Therefore, if we put $|\xi|=p^{-n}$ for a vector ξ whose first nonzero component is ξ_n, then this absolute value $|\ |$ gives a †valuation of $W(k)$. In particular, when k is a †perfect field, denoting the vector $(\xi_0, 0, \ldots)$ by $\{\xi_0\}$ we get $(\xi_0, \xi_1, \ldots)=\sum p^i\{\xi_i^{p^{-i}}\}$, and $W(k)$ is a †complete valuation ring with respect to this valuation. Therefore the †field of quotients of $W(k)$ is a complete valuation field of which p is a prime element and k is the †residue class field. Conversely, let K be a field of characteristic 0 that is complete under a †discrete valuation v, $\mathfrak{o}$ be the valuation ring of v, and k be the residue class field of v. Assume that k is a perfect field of characteristic p. If p is a prime element of v, then $\mathfrak{o}=W(k)$. If $v(p)=v(\pi^e)$ $(e>1)$ with a prime element π of $\mathfrak{o}$, we have $\mathfrak{o}=W(k)[\pi]$, and π is a root of an †Eisenstein polynomial $X^e+a_1X^{e-1}+\ldots+a_e$ $(a_i\in pW(k), a_e\notin p^2W(k))$. In this way we can determine explicitly the structure of a †p-adic number field (→ 257 Local Fields).

B. Applications to Abelian p-Extensions and Cyclic Algebras of Characteristic p

Next we consider $W_n(k)=W(k)/V^nW(k)$. The elements of $W_n(k)$ can be viewed as the n-dimensional vectors $(\xi_0, \ldots, \xi_{n-1})$, but their laws of composition are defined as in the previous section. They are called **Witt vectors of length** n. We define an operator $\wp$ by $\wp\xi=\xi^p-\xi$. Using it, we can generalize the theory of †Artin-Schreier extensions (→ 172 Galois Theory) to the case of Abelian extensions of exponent p^n over a field of characteristic p. Indeed, let k be a field of characteristic p and $\xi=(\xi_0, \ldots, \xi_{n-1})$ an element of $W(k)$. If $\eta=(\eta_0, \ldots, \eta_{n-1})$ is a root of the vector equation $\wp X-\xi=0$, then the other roots are of the form $\eta+\alpha(\alpha=(\alpha_0, \ldots, \alpha_{n-1}), \alpha_i\in\mathbf{F}_p)$. In particular, if $\xi_0\notin\wp k=\{\alpha^p-\alpha \mid \alpha\in k\}$, the field $K=k(\eta_0, \ldots, \eta_{n-1})$ is a cyclic extension of degree p^n over k, and conversely, every cyclic extension of k of degree p^n is obtained in this way. Let $(1/\wp)\xi$ denote the set of all roots of $\wp X-\xi=0$. Then more generally, any finite Abelian extension of exponent p^n of k can be obtained as $K=k((1/\wp)\xi \mid \xi\in H)$ with a suitable finite subgroup $H/\wp W_n(k)$ of $W_n(k)/\wp W_n(k)$, and the Galois group of K/k is isomorphic to $H/\wp W_n(k)$.

Moreover, for a †cyclic extension $K = k((1/\wp)\beta)$ of exponent p^n over k and for $\alpha \in k(\alpha \neq 0)$, we can define a †cyclic algebra $(\alpha, \beta]$ generated by an element u over K by the fundamental relations $u^{p^n} = \alpha$, $\wp\theta = \beta$, $u\theta u^{-1} = \theta + (1, 0, \ldots, 0)$ (where $\theta = (\theta_0, \ldots, \theta_{n-1})$, $u\theta u^{-1} = (u\theta_0 u^{-1}, \ldots, u\theta_{n-1}u^{-1})$), and $(\alpha, \beta]$ is a central simple algebra over k.

Using these results, we can develop the structure theory of the †Brauer group of exponent p^n of a †field of power series in one variable with coefficients in a finite field $\mathbf{F}_q$ (of a †field of algebraic functions in one variable over $\mathbf{F}_q$) exactly as in the case of a p-adic field (of an algebraic number field) (E. Witt [1]; → 29 Associative Algebras G).

On the other hand, $W_n(k)$ is a commutative †algebraic group over k and is important in the theories of algebraic groups and †formal groups (→ 13 Algebraic Groups).

References

[1] E. Witt, Zyklische Körper und Algebren der Charakteristik p vom Grad p^n, J. Reine Angew. Math., 176 (1937), 126–140.

[2] H. Hasse, Zahlentheorie, Akademie-Verlag., 1949.

[3] N. Jacobson, Lectures in abstract algebra III, Van Nostrand, 1964.

N

450 (V.19) Zeta Functions

A. Introduction

Since the 19th century, many special functions called ζ-functions (zeta functions) have been defined and investigated. The four main problems concerning ζ-functions are: (1) Methods of defining ζ-functions. (2) Investigation of the properties of ζ-functions. Generally, ζ-functions have the following four properties in common: (i) They are meromorphic on the whole complex plane; (ii) they have †Dirichlet series expansions; (iii) they have Euler product expansions; and (iv) they satisfy certain functional equations. Also, it is an important problem to find the poles, residues, and zeros of ζ-functions. (3) Application to number theory, in particular to the theory of decomposition of prime ideals in finite extensions of algebraic number fields (→ 59 Class Field Theory). (4) Study of the relations between different ζ-functions.

Most of the functions called ζ-functions or L-functions have the four properties of problem (2). The following is a classification of the important types of ζ-functions that are already known, which will be discussed later in this article:

(1) The ζ- and L-functions of algebraic number fields: the Riemann ζ-function, Dirichlet L-functions (study of these functions gave impetus to the theory of ζ-functions), Dedekind ζ-functions, Hecke L-functions, Hecke L-functions with †Grössencharakters, Artin L-functions, and Weil L-functions. (2) The p-adic L-functions related to the works of H. W. Leopoldt, T. Kubota, K. Iwasawa, etc. (3) The ζ-functions of quadratic forms: Epstein ζ-functions, ζ-functions of indefinite quadratic forms (C. L. Siegel), etc. (4) The ζ- and L-functions of algebras: Hey ζ-functions and the ζ-functions given by R. Godement, T. Tamagawa, etc. (5) The ζ-functions associated with Hecke operators, related to the work of E. Hecke, M. Eichler, G. Shimura, H. Jacquet, R. P. Langlands, etc. (6) The congruence ζ- and L-functions attached to algebraic varieties defined over finite fields (E. Artin, A. Weil, A. Grothendieck, P. Deligne), ζ- and L-functions of schemes. (7) Hasse ζ-functions attached to the algebraic varieties defined over algebraic number fields. (8) The ζ-functions attached to discontinuous groups: Selberg ζ-functions, the Eisenstein series defined by A. Selberg, Godement, and I. M. Gel'fand, etc. (9) Y. Ihara's ζ-function related to non-Abelian class field theory over a function field over a finite field. (10) ζ-functions associated with prehomogeneous vector spaces (M. Sato, T. Shintani).

B. The Riemann ζ-Function

Consider the series

$$\zeta(s)=1+\frac{1}{2^s}+\frac{1}{3^s}+\ldots+\frac{1}{n^s}+\ldots,$$

which converges for all real numbers $s>1$. It was already recognized by L. Euler that $\zeta(s)$ can also be expressed by a convergent infinite product $\prod_p(1-p^{-s})^{-1}$, where p runs over all prime numbers (*Werke*, ser. I, vol. VII, ch. XV, § 274). This expansion is called **Euler's infinite product expansion** or simply the **Euler product**. However, Riemann was the first to treat $\zeta(s)$ successfully as a function of a complex variable s (1859) [R1]; for this reason, it is called the **Riemann ζ-function**. As can be seen from its Euler product expansion, $\zeta(s)$ is holomorphic and has no zeros in the domain $\operatorname{Re}s>1$. Riemann proved, moreover, that it has an analytic continuation to the whole complex plane, is meromorphic everywhere, and has a unique pole $s=1$. The functions $(s-1)\zeta(s)$ and $\zeta(s)-1/(s-1)$ are †integral functions of s. This can be seen by considering the integral expression

$$\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)=\int_0^\infty x^{s/2-1}\left(\sum_{n=1}^\infty e^{-n^2\pi x}\right)dx$$

$$=-\frac{1}{s}-\frac{1}{1-s}+\int_1^\infty (x^{(1-s)/2-1}+x^{(s/2)-1})\times\left(\sum_{n=1}^\infty e^{-n^2\pi x}\right)dx.$$

From this last formula, we also obtain an equality

$$\xi(s)=\xi(1-s),$$

where

$$\xi(s)=\pi^{-s/2}\Gamma(s/2)\zeta(s).$$

This equality is called the **functional equation** for the ζ-function. The residue of $\zeta(s)$ at $s=1$ is 1, and around $s=1$,

$$\zeta(s)=\frac{1}{s-1}+C+O(|s-1|),$$

where C is †Euler's constant. This is called the **Kronecker limit formula** for $\zeta(s)$.

The function $\zeta(s)$ has no zeros in $\operatorname{Re}s\geqslant 1$, and its only zeros in $\operatorname{Re}s\leqslant 0$ are simple zeros at $s=-2, -4, \ldots, -2n, \ldots$. But $\zeta(s)$ has infinitely many zeros in $0<\operatorname{Re}s<1$, which are called the nontrivial zeros. B. Riemann conjectured that all nontrivial zeros lie on the line $\mathrm{Re}\, s = 1/2$ (1859). This is called the **Riemann hypothesis**, which has been neither proved nor disproved (→ Section I).

If $N(T)$ denotes the number of zeros of $\zeta(s)$ in the rectangle $0 < \mathrm{Re}\, s < 1$, $0 < \mathrm{Im}\, s < T$, we have an asymptotic formula

$$N(T) = \frac{1}{2\pi} T \log T - \frac{1 + \log 2\pi}{2\pi} T + O(\log T)$$

(H. von Mangoldt, 1905). Also, $\zeta(s)$ has the following infinite product expansion:

$$(s-1)\zeta(s) = \frac{1}{2} e^{bs} \frac{1}{\Gamma\left(\frac{s}{2}+1\right)} \prod_{\rho} \left(1 - \frac{s}{\rho}\right) e^{s/\rho},$$

where b is a constant and ρ runs over all nontrivial zeros of $\zeta(s)$ (J. Hadamard, 1893). Hadamard and C. de La Vallée-Poussin proved the †prime number theorem, almost simultaneously, by using some properties of $\zeta(s)$ (→ 123 Distribution of Prime Numbers B).

The following **approximate functional equation** is important in investigating the values of $\zeta(s)$:

$$\zeta(s) = \sum_{n<x} \frac{1}{n^s} + \varphi(s) \sum_{n<y} \frac{1}{n^{1-s}} + O(x^{-\sigma}) + O(y^{\sigma-1} |t|^{(1/2)-\sigma}),$$

where φ is †Euler's function, and $\zeta(s) = \varphi(s)\zeta(1-s)$, $s = \sigma + it$, $2\pi xy = |t|$, and the approximation is uniform for $-h \leqslant \sigma \leqslant h$, $x > k$, $y > k$ with h and k positive constants (G. H. Hardy and J. E. Littlewood, 1921).

Euler obtained the values of $\zeta(s)$ for positive even integers s:

$$\zeta(2m) = \frac{2^{2m-1} \pi^{2m} B_{2m}}{(2m)!}$$

($m = 1, 2, 3, \ldots$, and the B_{2m} are †Bernoulli numbers). The values of $\zeta(s)$ for positive odd integers s, however, have not been expressed in such a simple form. The values of $\zeta(s)$ for negative integers s are given by $\zeta(0) = B_1(0) = -\frac{1}{2}$, $\zeta(1-n) = -\frac{B_n(0)}{n}$, $n = 2, 3, \ldots$, where the $B_n(x)$ are †Bernoulli polynomials.

As a slight generalization of $\zeta(s)$, A. Hurwitz (1862) considered

$$\zeta(s, a) = \sum_{n=0}^{\infty} \frac{1}{(n+a)^s}, \quad 0 < a \leqslant 1.$$

This is called the **Hurwitz ζ-function**. Thus $\zeta(s, 1) = \zeta(s)$, and $\zeta(s, 1/2) = (2^s - 1)\zeta(s)$. This function $\zeta(s, a)$ can also be continued analytically to the whole complex plane and satisfies a certain functional equation. But in general it has no Euler product expansion.

C. Dirichlet L-Functions

Let m be a positive integer, and classify all rational integers modulo m. The set of all classes coprime to m forms a multiplicative Abelian group of order $h = \varphi(m)$. Let χ be a †character of this group. Call (n) the residue class of n mod m, and put $\chi(n) = \chi((n))$ when $(n, m) = 1$ and $\chi(n) = 0$ when $(n, m) \neq 1$. Now, the function of a complex variable s defined by

$$L(s) = L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

is called a **Dirichlet L-function**. This function converges absolutely for $\mathrm{Re}\, s > 1$ and has an Euler product expansion

$$L(s, \chi) = \prod_{p} \frac{1}{1 - \chi(p) p^{-s}}.$$

If there exist a divisor f of m ($f \neq m$) and a character χ^0 modulo f such that $\chi(n) = \chi^0(n)$ for all n with $(n, m) = 1$, we call χ a **nonprimitive character**. Otherwise, χ is called a **primitive character**. If χ is nonprimitive, there exists such a unique primitive χ^0. In this situation, the divisor f of m associated with χ^0 is called the **conductor** of χ (and of χ^0). We have

$$L(s, \chi) = L(s, \chi^0) \prod_{p|m} (1 - \chi^0(p) p^{-s}).$$

Let χ be primitive. If the conductor $f = 1$, then χ is a trivial character ($\chi = 1$), and $L(s)$ is equal to the Riemann ζ-function $\zeta(s)$. On the other hand, if $f > 1$, then $L(s)$ is an entire function of s. In particular, if χ is a nontrivial primitive character, $L(1) = L(1, \chi)$ is finite and nonzero. P. G. L. Dirichlet proved the theorem of existence of prime numbers in arithmetic progressions using this fact (→ 123 Distribution of Prime Numbers D).

$L(s, \chi)$ has a functional equation similar to that of $\zeta(s)$; namely, if χ is a primitive character with conductor f and we put

$$\xi(s, \chi) = (f/\pi)^{s/2} \Gamma((s+a)/2) L(s, \chi),$$

where $a = 0$ for $\chi(-1) = 1$ and $a = 1$ for $\chi(-1) = -1$, then we have

$$\xi(s, \chi) = W(\chi) \xi(1-s, \bar{\chi}),$$

where

$$W(\chi) = (-i)^a f^{-1/2} \tau(\chi), \quad \tau(\chi) = \sum_{n \bmod f} \chi(n) \zeta_f^n$$

($\zeta_f = \exp(2\pi i/f)$). The latter sum is called the **Gaussian sum**. Note that $|W(\chi)| = 1$.

The values of $L(s)$ for negative integers s are given by $L(1-m, \chi) = -B_{\chi,m}/m$ ($m = 1, 2, 3, \ldots$), where the $B_{\chi,m}$ are defined by

$$\sum_{\mu=1}^{f} \frac{\chi(\mu) t e^{\mu t}}{e^{ft} - 1} = \sum_{m=0}^{\infty} B_{\chi,m} t^m.$$

Moreover, if $\chi(-1)=-1$, we have

$$L(1,\chi)=\frac{\pi}{i}\frac{\tau(\chi)}{f}\sum_{x=1}^{f}(-\overline{\chi(x)}\cdot x)$$

$$=\pi i\tau(\chi)B_{\bar{\chi},1},$$

and if $\chi(-1)=1$, $\chi\neq 1$, we have

$$L(1,\chi)=2\frac{\tau(\chi)}{f}\sum_{x=1}^{[f/2]}(-\chi(x)\log|1-\zeta_f^x|).$$

In certain cases, the functional equation can be utilized to obtain the values of $L(m,\chi)$ from those of $L(1-m,\chi)$. Actually, if $\chi(-1)=1$, $m=2n=2,4,6,\ldots$, we have

$$L(2n,\chi)=\frac{(-1)^n}{(2n)!}\left(\frac{2\pi}{f}\right)^{2n}\tau(\chi)(-B_{\chi,2n}),$$

and if $\chi(-1)=-1$, $m=2n+1=3,5,7,\ldots$, we have

$$L(2n+1,\chi)$$

$$=(-i)\frac{(-1)^n}{(2n+1)!}\left(\frac{2\pi}{f}\right)^{2n+1}\tau(\chi)(-B_{\chi,2n+1}).$$

Dirichlet L-functions are important not only in the arithmetic of rational number fields but also in the arithmetic of quadratic or cyclotomic fields.

D. ζ-Functions of Algebraic Number Fields (Dedekind ζ-Functions)

The Riemann ζ-function can be generalized to ζ-functions of algebraic number fields (→ 14 Algebraic Number Fields). Let k be an algebraic number field of degree n, and let $\mathfrak{a}$ run over all integral ideals of k. Consider the sequence $\zeta_k(s)=\sum_{\mathfrak{a}}N(\mathfrak{a})^{-s}$. This sequence converges for $\mathrm{Re}\,s>1$ and has an Euler product expansion $\zeta_k(s)=\prod_{\mathfrak{p}}(1-N(\mathfrak{p})^{-s})^{-1}$, where $\mathfrak{p}$ runs over all prime ideals of k. This function, which is continued analytically to the whole complex plane as a meromorphic function, is called a **Dedekind ζ-function**. Its only pole is a simple pole at $s=1$, with the residue $h_k\kappa_k$. Here h_k is the †class number of k, and $\kappa_k=2^{r_1+r_2}\pi^{r_2}R/(w|d|^{1/2})$, where r_1 $(2r_2)$ is the number of isomorphisms of k into the real (complex) number field, w is the number of roots of unity in k, d is the †discriminant of k, and R is the †regulator of k (R. Dedekind, 1877) [D1].

The function $\zeta_k(s)$ has no zeros in $\mathrm{Re}\,s\geqslant 1$, while in $\mathrm{Re}\,s\leqslant 0$ it has zeros of order r_2 at -1, -3, $-5,\ldots$, zeros of order r_1+r_2 at -2, -4, $-6,\ldots$, and a zero of order r_1+r_2-1 at $s=0$. All other zeros lie in the open strip $0<\mathrm{Re}\,s<1$, which actually contains infinitely many zeros. It is conjectured that all these zeros lie on the line $\mathrm{Re}\,s=1/2$ (the Riemann hypothesis for Dedekind ζ-functions). To obtain a generalization of the functional equation for the Riemann $\zeta(s)$ to the case of $\zeta_k(s)$, we put

$$\Xi_k(s)=\left(\frac{\sqrt{|d|}}{2^{r_2}\pi^{n/2}}\right)^s\Gamma\left(\frac{s}{2}\right)^{r_1}\Gamma(s)^{r_2}\zeta_k(s).$$

Then $\Xi_k(s)=\Xi_k(1-s)$ (Hecke, 1917). If K is a Galois extension of k, then $\zeta_K(s)/\zeta_k(s)$ is an integral function (H. Aramata, 1933; R. Brauer, 1947).

E. Hecke L-Functions

As a generalization of Dirichlet L-functions to algebraic number fields, Hecke (1917) defined the following L-function $L_k(s,\chi)$: Let k be an algebraic number field of finite degree, and let $\tilde{\mathfrak{m}}=\mathfrak{m}\prod\mathfrak{p}_\infty$ be an †integral divisor ($\mathfrak{m}$ the finite part, $\prod\mathfrak{p}_\infty$ the †infinite part). Consider the †ideal class group of k modulo $\tilde{\mathfrak{m}}$ and its character χ (here we put $\chi(\mathfrak{a})=0$ for $(\mathfrak{a},\mathfrak{m})\neq 1$). Then the L-functions are defined by

$$L_k(s,\chi)=\sum_{\mathfrak{a}}\chi(\mathfrak{a})/N(\mathfrak{a})^s$$

[H2], where $\mathfrak{a}$ runs over all integral ideals of k. $L_k(s,\chi)$ is called a **Hecke L-function**. It converges for $\mathrm{Re}\,s>1$ and has an Euler product expansion

$$L_k(s,\chi)=\prod_{\mathfrak{p}}\frac{1}{1-\chi(\mathfrak{p})N(\mathfrak{p})^{-s}}.$$

Here $\mathfrak{p}$ runs over all prime ideals of k. If there is a divisor $\tilde{\mathfrak{f}}\,|\,\tilde{\mathfrak{m}}$ $(\tilde{\mathfrak{f}}\neq\tilde{\mathfrak{m}})$ and a character χ^0 modulo $\tilde{\mathfrak{f}}$ such that $\chi^0(\mathfrak{a})=\chi(\mathfrak{a})$ for all $\mathfrak{a}$ with $(\mathfrak{a},\mathfrak{m})=1$, then χ is called **nonprimitive**: otherwise, χ is called a **primitive character**. In general, there exist unique such $\tilde{\mathfrak{f}}$ and χ^0. In this situation, $\tilde{\mathfrak{f}}$ is called the **conductor** of χ. If χ is primitive and the conductor $\tilde{\mathfrak{f}}$ is (1), then χ is a trivial character and $L_k(s,\chi)$ coincides with $\zeta_k(s)$. If χ is primitive and $\chi\neq 1$, then $L_k(s,\chi)$ is an integral function of s, and $L_k(1,\chi)\neq 0$. Utilizing this fact, it can be proved that there exist infinitely many prime ideals in each class of the ideal class group modulo an integral divisor $\tilde{\mathfrak{m}}$ of k.

Let χ be a primitive character with the conductor $\tilde{\mathfrak{f}}$, d be the discriminant of k, $\sigma_1,\ldots,\sigma_{r_1}$ be all distinct isomorphisms of k into the real number field $\mathbf{R}$, and $\mathfrak{f}$ be the finite part of $\tilde{\mathfrak{f}}$. Then if ξ is an integer of k such that $\xi\equiv 1$ $(\mathrm{mod}\,\mathfrak{f})$, we have

$$\chi((\xi))=(\mathrm{sgn}\,\xi^{\sigma_1})^{a_1}\cdot\ldots\cdot(\mathrm{sgn}\,\xi^{\sigma_{r_1}})^{a_{r_1}},$$

where a_m $(m=1,\ldots,r_1)$ is either 0 or 1, depending on χ. By putting

$$\xi_k(s,\chi)=$$

$$\left(\frac{\sqrt{|d|N(\mathfrak{f})}}{2^{r_2}\pi^{n/2}}\right)^s\cdot\prod_{m=1}^{r_1}\Gamma\left(\frac{s+a_m}{2}\right)\Gamma(s)^{r_2}\cdot L_k(s,\chi),$$

we have the following functional equation for the Hecke L-function:

$$\xi_k(s,\chi)=W(\chi)\xi_k(1-s,\bar{\chi}),$$

where $W(\chi)$ is a complex number with absolute value 1 and the exact value of $W(\chi)$ is given as a Gaussian sum. Just as some properties concerning the distribution of prime numbers can be proved using the Riemann ζ-function and Dirichlet L-functions, some properties concerning the distribution of prime ideals can be proved using the Hecke L-functions (→ 123 Distribution of Prime Numbers F).

T. Takagi used Hecke L-functions in founding his †class field theory. In the other direction, this theory implies $L(1,\chi)\neq 0$ $(\chi\neq 1)$.

Let K be a †class field over k that corresponds to an ideal class group H of k with index h. By using class field theory, we obtain $\zeta_K(s)=\prod_\chi L_k(s,\chi)$, where the product is over all characters χ of ideal class groups of k, such that $\chi(H)=1$. This formula can be regarded as an alternative formulation of the decomposition theorem of class field theory (→ 59 Class Field Theory). By taking the residues of both sides of the formula at $s=1$, we obtain $h_K\kappa_K=h_k\kappa_k\prod_{\chi\neq 1}L_k(1,\chi)$.

In particular, if $k=\mathbf{Q}$ (the rational number field) and K is a quadratic number field $\mathbf{Q}(\sqrt{d})$ (d is the discriminant of K), then we have

$$\zeta_K(s)=\zeta(s)\cdot L(s),\quad L(s)=\sum_{n=1}^{\infty}\left(\frac{d}{n}\right)n^{-s},$$

where (d/n) is the †Kronecker symbol, and we put $(d/n)=0$ when $(n,d)\neq 1$. From this, we obtain the class number formula for quadratic number fields (→ 347 Quadratic Fields). A similar method is used for computation of class numbers of cyclotomic fields K (→ 14 Algebraic Number Fields L).

In general, the computation of the relative class number h_K/h_k when K/k is an Abelian extension is reduced to the evaluation of $L(1,\chi)$. This computation has been made successfully for the following cases (besides for the examples in the previous paragraph): k is imaginary quadratic and K is the absolute class field of k or the class field corresponding to †ray $S(\mathfrak{m})$; k is totally real and K is a totally imaginary quadratic extension of k. H. M. Stark and T. Shintani made conjectures about the values of $L(1,\chi)$ [S25, S19].

Let $L(s,\chi)$ be a Hecke L-function for the character χ. Then it follows from the functional equation that the values of $L(s,\chi)$ at $s=0,-1,-2,-3,\ldots$ are zero if k is not totally real. Furthermore, if k is a totally real finite algebraic number field, then these values of $L(s,\chi)$ are algebraic numbers (C. L. Siegel, H. Klingen, T. Shintani).

F. Hecke L-Functions with Grössencharakters

E. Hecke (1918, 1920) extended the notion of characters by introducing the †Grössencharakter χ and defined L-functions with such characters:

$$L_k(s,\chi)=\sum_{\mathfrak{a}}\frac{\chi(\mathfrak{a})}{N(\mathfrak{a})^s}.$$

He also proved the existence of their Euler product expansions and showed that they satisfy certain functional equations [H3]. Moreover, by estimating the sum $\sum_{N(\mathfrak{p})<x}\chi(\mathfrak{p})$, he obtained some results on the distribution of prime ideals.

Later, Iwasawa and J. Tate independently gave clearer definitions of the Grössencharakter χ and $L_k(s,\chi)$ by using harmonic analysis on the adele and idele groups of k (→ 6 Adeles and Ideles) [L3].

Let $\mathbf{J}_k$ be the idele group of k, $\mathbf{P}_k$ be the group of †principal ideles, and $\mathbf{C}_k=\mathbf{J}_k/\mathbf{P}_k$ be the idele class group. Then a Grössencharakter is a continuous character χ of $\mathbf{C}_k$, and χ induces a character of $\mathbf{J}_k$, which is also denoted by χ. Let $\mathbf{J}_k=\mathbf{J}_\infty\times\mathbf{J}_0$ be the decomposition of $\mathbf{J}_k$ into the infinite part $\mathbf{J}_\infty$ and the finite part $\mathbf{J}_0$. Let $\mathbf{U}_0$ be the unit group of $\mathbf{J}_0$, and for each integral ideal $\mathfrak{m}$ of k, put $\mathbf{U}_{\mathfrak{m},0}=\{\mathfrak{u}\in\mathbf{U}_0\mid\mathfrak{u}\equiv 1\ (\mathrm{mod}\ \mathfrak{m})\}$, so that $\{\mathbf{U}_{\mathfrak{m},0}\}$ forms a base for the neighborhood system of 1 in $\mathbf{J}_0$. Put $\mathbf{J}_{\mathfrak{m},0}=\{\mathfrak{a}\in\mathbf{J}_0\mid\mathfrak{a}_\mathfrak{p}=1 \text{ for all } \mathfrak{p}\mid\mathfrak{m}\}$, and with each $\mathfrak{a}\in\mathbf{J}_{\mathfrak{m},0}$, associate an ideal $\tilde{\mathfrak{a}}=\prod_\mathfrak{p}\mathfrak{p}^{v_\mathfrak{p}(\mathfrak{a})}$, where $\mathfrak{a}=(\mathfrak{a}_\mathfrak{p})$ and the ideal in $k_\mathfrak{p}$ generated by $\mathfrak{a}_\mathfrak{p}$ is equal to $\mathfrak{p}^{v_\mathfrak{p}(\mathfrak{a})}$. Then the mapping $\mathfrak{a}\rightarrow\tilde{\mathfrak{a}}$ gives a homomorphism of $\mathbf{J}_{\mathfrak{m},0}$ into the group $G(\mathfrak{m})=\{\tilde{\mathfrak{a}}\mid(\mathfrak{a},\mathfrak{m})=1\}$, and its kernel is contained in $\mathbf{U}_{\mathfrak{m},0}$. Since χ is continuous, $\chi(\mathbf{U}_{\mathfrak{m},0})=1$ for some $\mathfrak{m}$. The greatest common divisor $\mathfrak{f}$ of all such ideals $\mathfrak{m}$ is called the **conductor** of χ. For each $\mathfrak{a}\in\mathbf{J}_{\mathfrak{f},0}$, $\chi(\mathfrak{a})$ depends only on the ideal $\tilde{\mathfrak{a}}$ $(\in G(\mathfrak{f}))$; hence by putting $\chi(\mathfrak{a})=\tilde{\chi}(\tilde{\mathfrak{a}})$, we obtain a character $\tilde{\chi}$ of $G(\mathfrak{f})$. Now put $L_k(s,\chi)=\sum\tilde{\chi}(\tilde{\mathfrak{a}})/N(\tilde{\mathfrak{a}})^s$, where the sum is over all integral ideals $\tilde{\mathfrak{a}}\in G(\mathfrak{f})$. This is called a **Hecke L-function with Grössencharakter** χ. For $\chi\neq 1$, it is an entire function. On the other hand, if we restrict χ to $\mathbf{J}_\infty=\mathbf{R}^{*r_1}\times\mathbf{C}^{*r_2}$, then for $u=(a_1,\ldots,a_{r_1},a_{r_1+1},\ldots,a_{r_1+r_2})\in\mathbf{J}_\infty$, we have

$$\chi(u)=\prod_{j=1}^{r_1+r_2}|a_j|^{\lambda_j\sqrt{-1}}\cdot\prod_{j=1}^{r_1}(\operatorname{sgn}a_j)^{e_j}\cdot\prod_{l=r_1+1}^{r_1+r_2}\left(\frac{a_l}{|a_l|}\right)^{e_l},$$

where $e_j=0$ or 1, $e_l\in\mathbf{Z}$, $\lambda_j\in\mathbf{R}$. The numbers e_j, e_l, λ_j are determined uniquely by χ. Putting

$$\xi_k(s,\chi)=\left(\frac{\sqrt{|d|N(\mathfrak{f})}}{2^{r_2}\pi^{n/2}}\right)^s\cdot\prod_{j=1}^{r_1}\Gamma\left(\frac{s+e_j+\lambda_j\sqrt{-1}}{2}\right)\times\prod_{l=r_1+1}^{r_1+r_2}\Gamma\left(s+\frac{|e_l|+\lambda_l\sqrt{-1}}{2}\right)\cdot L(s,\chi),$$

we have a functional equation

$$\xi_k(s,\chi)=W(\chi)\xi_k(1-s,\bar{\chi}),$$

where $W(\chi)$ is a complex number with absolute value 1.

We can express $\xi_k(s,\chi)$ by an integral form on $\mathbf{J}_k$ as

$$\xi_k(s,\chi)=c\int_{\mathbf{J}_k}\varphi(\mathfrak{r})\chi(\mathfrak{r})V(\mathfrak{r})^s d^*\mathfrak{r},$$

where $V(\mathfrak{r})$ is the †total volume of the idele $\mathfrak{r}$, c is a constant that depends on the †Haar measure $d^*\mathfrak{r}$ of $\mathbf{J}_k$, and $\varphi(\mathfrak{r})$ is defined by

$$\varphi(\mathfrak{r})=\prod_{\mathfrak{p}}\varphi_{\mathfrak{p}}(x_{\mathfrak{p}}),\qquad \mathfrak{r}=(\ldots x_{\mathfrak{p}}\ldots),$$

$$\begin{aligned}&\varphi_{\mathfrak{p}_{\infty,i}}(x)\\&=x^{e_i}e^{-\pi x^2},\quad i\leqslant r_1,\quad k_{\mathfrak{p}_{\infty,i}}=\mathbf{R},\\&\left.\begin{aligned}&=\frac{1}{2\pi}\bar{x}^{e_i}e^{-2\pi|x|^2},\quad e_i\geqslant 0,\\&=\frac{1}{2\pi}x^{-e_i}e^{-2\pi|x|^2},\quad e_i<0,\end{aligned}\right\}\ i>r_1,\quad k_{\mathfrak{p}_{\infty,i}}=\mathbf{C},\end{aligned}$$

$$\left.\begin{aligned}\varphi_{\mathfrak{p}}(x)&=e^{2\pi i\lambda(x)},\quad x\in(\mathfrak{b}\mathfrak{f})_{\mathfrak{p}}^{-1},\\&=0,\qquad\qquad x\notin(\mathfrak{b}\mathfrak{f})_{\mathfrak{p}}^{-1}\end{aligned}\right\}\ \mathfrak{p}\text{ finite.}$$

Hence $(\mathfrak{b}\mathfrak{f})_{\mathfrak{p}}^{-1}$ is the $\mathfrak{p}$-component (→ 6 Adeles and Ideles B) of the ideal $(\mathfrak{b}\mathfrak{f})^{-1}$ ($\mathfrak{b}$ is the †different of $k/\mathbf{Q}$) and $\lambda(x)$ is an additive character of $k_{\mathfrak{p}}$ defined as follows. $\mathbf{Q}_p$ is the †p-adic field, $\mathbf{Z}_p$ is the ring of p-adic integers, λ_0 is the mapping $\mathbf{Q}_p\to\mathbf{Q}_p/\mathbf{Z}_p\subset\mathbf{Q}/\mathbf{Z}\subset\mathbf{R}/\mathbf{Z}$, and $\lambda=\lambda_0\circ Tr_{k_{\mathfrak{p}}/\mathbf{Q}_p}$. By putting $\chi(\mathfrak{r})=\prod_{\mathfrak{p}}\chi_{\mathfrak{p}}(x_{\mathfrak{p}})$, $\mathfrak{r}=(\ldots x_{\mathfrak{p}}\ldots)$, we have

$$\xi_k(s,\chi)=c\prod_{\mathfrak{p}}\int_{k_{\mathfrak{p}}}\varphi_{\mathfrak{p}}(x)\chi_{\mathfrak{p}}(x)V_{\mathfrak{p}}(x)^{-s}d^*x,$$

where $\mathfrak{p}$ runs over all prime divisors of k, finite or infinite. Moreover, with a constant $C_{\mathfrak{p}}$, we have

$$\begin{aligned}&\int_{k_{\mathfrak{p}_{\infty,i}}}\varphi_{\mathfrak{p}_{\infty,i}}(x)\chi_{\mathfrak{p}_{\infty,i}}(x)V_{\mathfrak{p}_{\infty,i}}(x)^{-s}d^*x\\&=C_{\mathfrak{p}_{\infty,i}}\cdot\pi^{-(s+\sqrt{-1}\lambda_i+e_i)/2}\\&\qquad\times\Gamma((s+\sqrt{-1}\,\lambda_i+e_i)/2),\quad k_{\mathfrak{p}_{\infty,i}}=\mathbf{R},\\&=C_{\mathfrak{p}_{\infty,i}}\cdot(2\pi)^{(s+(\sqrt{-1}\lambda_i+|e_i|)/2)}\\&\qquad\times\Gamma(s+(\sqrt{-1}\,\lambda_i+|e_i|)/2),\quad k_{\mathfrak{p}_{\infty,i}}=\mathbf{C},\end{aligned}$$

$$\begin{aligned}&\int_{k_{\mathfrak{p}}}\varphi(x)\chi_{\mathfrak{p}}(x)V_{\mathfrak{p}}(x)^{-s}d^*x\\&=C_{\mathfrak{p}}N(\mathfrak{b}_{\mathfrak{p}})^{s-1/2}\tilde{\chi}(\mathfrak{b}_{\mathfrak{p}}^{-1})\frac{1}{1-\tilde{\chi}(\mathfrak{p})/N(\mathfrak{p})^s},\\&=C_{\mathfrak{p}}N((\mathfrak{b}\mathfrak{f})_{\mathfrak{p}})^s\tau_{\mathfrak{p}}(\chi_{\mathfrak{p}})\cdot\mu(U_{\mathfrak{f},\mathfrak{p}}),\quad \mathfrak{p}\,|\,\mathfrak{f}.\end{aligned}$$

Here $\tau_{\mathfrak{p}}(\chi_{\mathfrak{p}})$ is a constant called the **local Gaussian sum**, and $\mu(U_{\mathfrak{f},\mathfrak{p}})$ is the volume of $\{u\in k_{\mathfrak{p}}\,|\,u\equiv 1\ (\mathrm{mod}\ \mathfrak{f})\}$. These integrals over $k_{\mathfrak{p}}$ are the Γ-factors and Euler factors of $\xi_k(s,\chi)$, according as $\mathfrak{p}$ is infinite or finite. The functional equation is obtained by applying the †Poisson summation formula for $\varphi(x)$ and its †Fourier transform on the adele group $\mathbf{A}_k$ (→ 6 Adeles and Ideles).

Let $\mathbf{D}_k$ be the connected component of 1 in $\mathbf{C}_k$. If $\chi(\mathbf{D}_k)=1$, the corresponding $\tilde{\chi}$ is a character of an ideal class group of k with a conductor $\mathfrak{f}$. Conversely, all such characters can be obtained in this manner.

As stated in Section E, the Hecke L-functions with characters (of ideal class groups) can be used to describe the decomposition law of prime divisors in class field theory. However, for L-functions with Grössencharakter, such arithmetic implications have not been found yet, except that in the case of Grössencharakters of A_0 type, Y. Taniyama discovered, following the suggestion of A. Weil, that the L-function has a deep connection with the arithmetic of a certain infinite Abelian extension of k [T2, W7]. In particular, when $L(s,\chi)$ is a factor of the †Hasse ζ-function of an Abelian variety A with †complex multiplication, it describes the arithmetic of the field generated by the coordinates of the division points of A.

G. Artin L-Functions

Let K be a finite Galois extension of an algebraic number field k (of degree n), $G=G(K/k)$ be its Galois group, $\sigma\to A(\sigma)$ be a matrix representation (characteristic 0) of G, and χ be its character. Let $\mathfrak{p}$ be a prime ideal of k, and define $L_{\mathfrak{p}}(s,\chi)$ by

$$\log L_{\mathfrak{p}}(s,\chi)=\sum_{m=1}^{\infty}\frac{\chi(\mathfrak{p}^m)}{mN(\mathfrak{p}^m)^s},\quad \mathrm{Re}\,s>1,$$

with $\chi(\mathfrak{p}^m)=(1/e)\sum_{\tau\in T}\chi(\sigma^m\tau)$, where T is the †inertia group of $\mathfrak{p}$, $|T|=e$, and σ is a †Frobenius automorphism of $\mathfrak{p}$. Then we have

$$L_{\mathfrak{p}}(s,\chi)=\det(E-A_{\mathfrak{p}}\cdot N(\mathfrak{p})^{-s})^{-1},$$

$$A_{\mathfrak{p}}=\frac{1}{e}\sum_{\tau\in T}A(\sigma\tau).$$

In particular, if $T=\{1\}$ (i.e., $\mathfrak{p}$ is †unramified in K/k), then

$$L_{\mathfrak{p}}(s,\chi)=\det(E-A(\sigma)\cdot N(\mathfrak{p})^{-s})^{-1}.$$

Now put

$$L(s,\chi,K/k)=\prod_{\mathfrak{p}}L_{\mathfrak{p}}(s,\chi),\quad \mathrm{Re}\,s>1,$$

and call $L(s,\chi,K/k)$ an **Artin L-function** [A2].

(1) The most important property of $L(s,\chi,K/k)$ is that if K/k is an Abelian extension and χ is a linear character, it follows from class field theory that $\chi(\mathfrak{p})$ is the character of the ideal class group of k (modulo the †conductor of K/k) and that the Artin L-function equals

a Hecke L-function. This equality is equivalent to Artin's †reciprocity law, and in fact Artin obtained his reciprocity law after he conjectured the equality.

(2) If $K' \supset K \supset k$ and K'/k is a Galois extension, then $L(s, \chi, K/k) = L(s, \chi, K'/k)$.

(3) If $K \supset \Omega \supset k$ and ψ is a character of $G(K/\Omega)$, then $L(s, \psi, K/\Omega) = L(s, \chi_\psi, K/k)$, where χ_ψ is the character of $G(K/k)$ †induced from ψ.

(4) If $\chi_1 = 1$, then $L(s, \chi_1, K/k) = \zeta_k(s)$.

(5) $L(s, \chi_1 + \chi_2, K/k) = L(s, \chi_1, K/k) \cdot L(s, \chi_2, K/k)$.

Conversely, the Artin L-function $L(s, \chi, K/k)$ is characterized by properties (1)–(5).

(6) If χ_R is the †regular representation of G, then $L(s, \chi_R, K/k) = \zeta_K(s)$; hence

$$\zeta_K(s) = \zeta_k(s) \prod_{\chi \neq 1} L(s, \chi, K/k)^{\chi(1)},$$

where χ runs over all irreducible characters $\neq 1$ of G.

(7) Every character of a finite group G can be expressed as $\chi = \sum m_i \chi_{\psi_i}$ $(m_i \in \mathbf{Z})$, where each χ_{ψ_i} is an †induced character from a certain linear character ψ_i of an elementary subgroup of G (**Brauer's theorem**). (Here an elementary subgroup is a subgroup that is the direct product of a cyclic group and a p-group for some prime p.) Hence (3) and (5) imply that an Artin L-function is the product of integral powers (positive or negative) of Hecke L-functions $L_{\Omega_i}(s, \psi_i)$:

$$L(s, \chi, K/k) = \prod_i L_{\Omega_i}(s, \psi_i)^{m_i}.$$

Hence an Artin L-function is a univalent meromorphic function defined over the whole complex plane. Artin made the still open conjecture that if χ is irreducible and $\chi \neq 1$, then $L(s, \chi, K/k)$ is an entire function (**Artin's conjecture**).

This conjecture holds obviously if all m_i are nonnegative. Except for such a case, Artin's conjecture had no affirmative examples until 1974, when Deligne and Serre [D9] proved that each "new cusp form" of weight 1 gives rise to an entire Artin L-function $L(s, \chi, K/k)$ with $\chi(1) = 2$ and $\chi(\rho) = 0$ (ρ is the complex conjugation); by this method, some nontrivial examples were computed by J. Tate and J. Buhler (*Lecture notes in math.* 654 (1978)). Then R. P. Langlands [L5] constructed nontrivial examples of Artin's conjecture for certain 2-dimensional representations

$$\mathrm{Gal}(K/k) \ni \sigma \mapsto A(\sigma) \in GL(2, \mathbf{C})$$

by using ideas of H. Saito and T. Shintani [S1, S20]. This method works for all representations for which the image of the $A(\sigma)$ in $PGL(2, \mathbf{C})$ is the †tetrahedral group. It also works for some †octahedral cases, but a new idea is needed in the †icosahedral case.

(8) Let $\mathfrak{p}_{\infty,i}$ $(i = 1, \ldots, r_1 + r_2)$ be the infinite primes of k. Put

$$\gamma(s, \chi, \mathfrak{p}_{\infty,i}, K/k)$$
$$= (\Gamma(s/2)\Gamma((s+1)/2))^{\chi(1)}$$
for complex $\mathfrak{p}_{\infty,i}$,
$$= \Gamma(s/2)^{(\chi(1)+\chi(\sigma))/2}\Gamma((s+1)/2)^{(\chi(1)-\chi(\sigma))/2}$$
for real $\mathfrak{p}_{\infty,i}$,

where $\sigma \in G$ is the complex conjugation determined by a prime factor of $\mathfrak{p}_{\infty,i}$ in K. Next we introduce the notion of the **conductor** $\mathfrak{f}_\chi$ **with the group character** χ defined by Artin (*J. Reine Angew. Math.*, 164 (1931)). First, for any subset $\mathfrak{m} \subset G$, we put $\chi(\mathfrak{m}) = \sum_{m \in \mathfrak{m}} \chi(m)$; then $\mathfrak{f}_\chi$ is given by

$$\mathfrak{f}_\chi = \mathfrak{f}(\chi, K/k) = \prod_{\mathfrak{p}} \mathfrak{p}^{f(\mathfrak{p})},$$

where

$$f(\mathfrak{p}) = \frac{1}{e}[\{e\chi(1) - \chi(T)\} + \{p^{e_1}\chi(1) - \chi(V_1)\} + \{p^{e_2}\chi(1) - \chi(V_2)\} + \ldots],$$

and where $V_1, V_2, \ldots$, are the higher †ramification groups of prime factors of $\mathfrak{p}$ in K (in lower numbering) and $p^{e_i} = |V_i|$ (→ 14 Algebraic Number Fields I).

Now put

$$\xi(s, \chi, K/k) = \left(\frac{|d|^{\chi(1)} N_k(\mathfrak{f}_\chi)}{\pi^{n\chi(1)}}\right)^{s/2} \times \prod_{\mathfrak{p}_{\infty,i}} \gamma(s, \chi, \mathfrak{p}_{\infty,i}, K/k) \cdot L(s, \chi, K/k).$$

Then the functional equation is written

$$\xi(1-s, \bar{\chi}, K/k) = W(\chi)\xi(s, \chi, K/k), \quad |W(\chi)| = 1.$$

The known proof of this functional equation depends on (7) and the functional equations of Hecke L-functions discussed in Section E. As for the constants $W(\chi)$, there are significant results by B. Dwork, Langlands, and Deligne [D6].

(9) There are some applications to the theory of the distribution of prime ideals.

H. Weil L-Functions

Weil defined a new L-function that is a generalization of both Artin L-functions and Hecke L-functions with Grössencharakter [W5]. Let K be a finite Galois extension of an algebraic number field k, let C_K be the idele class group $K_A^\times/K^\times$ of K, and let $\alpha_{K/k} \in H^2(\mathrm{Gal}(K/k), C_K)$ be the †canonical cohomology class of †class field theory. Then this $\alpha_{K/k}$ determines an extension $W_{K/k}$ of $\mathrm{Gal}(K/k)$ by C_K: $1 \to C_K \to W_{K/k} \to \mathrm{Gal}(K/k) \to 1$ (exact), and the transfer induces an isomorphism $W_{K/k}^{ab} \cong C_k$, where ab denotes the topological commutator quotient. If L is a Galois extension of k containing K, then there is a canonical homomorphism $W_{L/k} \rightarrow W_{K/k}$. Hence we define the **Weil group** W_k for $\bar{k}/k$ as the †projective limit group $\mathrm{proj}_K \lim W_{K/k}$ of the $W_{K/k}$. It is obvious that we have a surjective homomorphism $\varphi: W_k \rightarrow \mathrm{Gal}(\bar{k}/k)$ and an isomorphism $r_k: C_k \rightarrow W_k^{ab}$, where W_k^{ab} is the maximal Abelian Hausdorff quotient of W_k. For $w \in W_k$, let $\|w\|$ be the adelic norm of $r_k^{-1}(w)$.

If k_v is a †local field, then we define the Weil group W_{k_v} for $\bar{k}_v/k_v$ by replacing the idele class group C_K with the multiplicative group $K_w^\times$ in the above definition, where K_w denotes a Galois extension of k_v. If k_v is the completion of a finite algebraic number field k at a place v, then we have natural homomorphisms $k_v^\times \rightarrow C_k$ and $\mathrm{Gal}(\bar{k}_v/k_v) \rightarrow \mathrm{Gal}(\bar{k}/k)$. Accordingly, we have a homomorphism $W_{k_v} \rightarrow W_k$ that commutes with these homomorphisms.

Let W_k be the Weil group of an algebraic number field k, and let $\rho: W_k \rightarrow GL(V)$ be a continuous representation of W_k on a complex vector space V. Let $v = \mathfrak{p}$ be a finite prime of k, and let ρ_v be the representation of W_{k_v} induced from ρ. Let Φ be an element of W_{k_v} such that $\varphi(\Phi)$ is the inverse Frobenius element of $\mathfrak{p}$ in $\mathrm{Gal}(\bar{k}_v/k_v)$, and let I be the subgroup of W_{k_v} consisting of elements w such that $\varphi(w)$ belongs to the †inertia group of $\mathfrak{p}$ in $\mathrm{Gal}(\bar{k}_v/k_v)$. Let V^I be the subspace of elements in V fixed by $\rho_v(I)$, let $N\mathfrak{p}$ be the norm of $\mathfrak{p}$, and let

$$L_{\mathfrak{p}}(V,s) = \det(1 - (N\mathfrak{p})^{-s}\rho_v(\Phi) \mid V^I)^{-1}.$$

We can define $L_v(V,s)$ for each Archimedean prime v also, and let

$$L(V,s) = \prod_v L_v(V,s).$$

Then this product converges for s in some right half-plane and defines a function $L(V,s)$. We call $L(V,s)$ the **Weil L-function** for the representation $\rho: W_k \rightarrow GL(V)$. This function $L(V,s)$ can be extended to a meromorphic function on the complex plane and satisfies the functional equation

$$L(V,s) = \varepsilon(V,s) L(V^*, 1-s)$$

(T. Tamagawa), where V^* is the dual of V, and $\varepsilon(V,s)$ is an exponential function of s of the form ab^s [T6].

P. Deligne generalized these results in the following manner: Let W_k' be a †group scheme over $\mathbf{Q}$ which is the †semidirect product of W_k by the additive group $\mathbf{G}_a$, on which W_k acts by the rule $wxw^{-1} = \|w\| x$. We can define the notion of representations of W_k' and the L-functions of them in the natural manner [T6].

I. The Riemann Hypothesis

As mentioned in Section B, the Riemann hypothesis asserts that all zeros of the Riemann ζ-function in $0 < \mathrm{Re}\, s < 1$ lie on the line $\mathrm{Re}\, s = 1/2$. In his celebrated paper [R1], Riemann gave six conjectures (including this), and assuming these conjectures, proved the †prime number theorem:

$$\pi(x) \sim \frac{x}{\log x} \sim \mathrm{Li}(x) = \int_2^x \frac{dx}{\log x}, \quad x \rightarrow \infty.$$

Here $\pi(x)$ denotes the number of prime numbers smaller than x. Among his six conjectures, all except the Riemann hypothesis have been proved (a detailed discussion is given in [L1]). The prime number theorem was proved independently by Hadamard and de La Vallée-Poussin without using the Riemann hypothesis (→ Section B; 123 Distribution of Prime Numbers B).

R. S. Lehman showed that there are exactly 2,500,000 zeros of $\zeta(\sigma + it)$ for which $0 < t < 170{,}571.35$, all of which lie on the critical line $\sigma = 1/2$ and are simple (*Math. Comp.*, 20 (1966)). Later R. P. Brent extended this computation up to 75,000,000 first zeros (1979).

Hardy proved that there are infinitely many zeros of $\zeta(s)$ on the line $\mathrm{Re}\, s = 1/2$ (1914). Furthermore, A. Selberg [S6] proved that if $N_0(T)$ is the number of zeros of $\zeta(s)$ on the line with $0 < \mathrm{Im}\, s < T$, then $N_0(T) > AT \log T$ (A is a positive constant) (1942). Thus if $N(T)$ is the number of zeros of $\zeta(s)$ in the rectangle $0 < \mathrm{Re}\, s < 1$, $0 < \mathrm{Im}\, s < T$, then $\liminf_{T \rightarrow \infty} N_0(T)/N(T) > 0$. N. Levinson proved $\liminf_{T \rightarrow \infty} N_0(T)/N(T) > 1/3$ (*Advances in Math.*, 13 (1974)). If $N_\varepsilon(T)$ is the number of zeros of $\zeta(s)$ in $1/2 - \varepsilon < \mathrm{Re}\, s < 1/2 + \varepsilon$, $0 < \mathrm{Im}\, s < T$, then $\lim_{T \rightarrow \infty} N_\varepsilon(T)/N(T) = 1$ for any positive number ε (H. Bohr and E. Landau, 1914). Bohr studied the distribution of the values of $\zeta(s)$ in detail and initiated the theory of †almost periodic functions (1925).

D. Hilbert remarked in his lecture at the Paris Congress that the Riemann hypothesis is equivalent to

$$\pi(x) = \mathrm{Li}(x) + O(\sqrt{x} \log x), \quad x \rightarrow \infty$$

(H. von Koch, 1901). It is also equivalent to

$$\sum_{n=1}^{N} \mu(n) = O(N^{1/2+\varepsilon}), \quad N \rightarrow \infty,$$

for any $\varepsilon > 0$, where $\mu(n)$ is the Möbius function. Assuming the Riemann hypothesis, we get

$$N(T) = \frac{1}{2\pi} T \log T - \frac{1 + \log 2\pi}{2\pi} T + o(\log T)$$

(Littlewood, 1924).

The computation of the zeros of the ζ-functions and the L-functions of general algebraic number fields is more difficult, but conjectures similar to the Riemann hypothesis have been proposed.

Weil showed that a necessary and sufficient condition for the validity of the Riemann hypothesis for all Hecke L-functions $L(s,\chi)$ is that a certain †distribution on the idele group $\mathbf{J}_k$ be positive definite [W1 (1952b)].

It is not known whether the general ζ- and L-functions of algebraic number fields have any zeros in the interval (0, 1) on the real axis (see the works of A. Selberg and S. Chowla). Similar problems are considered for the various ζ-functions given in Sections P, Q, and T.

J. p-Adic L-Functions

Let χ be a †primitive Dirichlet character with conductor f, and let $L(s,\chi)$ be the †Dirichlet L-function for χ. Then the values $L(1-n,\chi)$ of $L(s,\chi)$ at nonpositive integers $1-n$ $(n=1,2,\dots)$ are algebraic numbers (→ Section E). Let p be a prime number, let $\mathbf{Q}_p$ be the †p-adic number field, and let $\mathbf{C}_p$ be the completion of the algebraic closure $\bar{\mathbf{Q}}_p$ of $\mathbf{Q}_p$. It is known that $\mathbf{C}_p$ is also algebraically closed. Since $\mathbf{Q}\subset\mathbf{Q}_p$, we fix an embedding $\bar{\mathbf{Q}}\subset\bar{\mathbf{Q}}_p$ and consider $\{L(1-n,\chi)\}_{n=1}^{\infty}$ as a sequence in $\mathbf{C}_p$.

Let $|\ |_p$ be the extension to $\mathbf{C}_p$ of the standard p-adic valuation of $\mathbf{Q}_p$. Let q be p or 4 according as $p\neq 2$ or $p=2$, and let ω be the primitive Dirichlet character with conductor q satisfying $\omega(n)\equiv n \pmod q$ for any integer n prime to p. Then T. Kubota and H. W. Leopoldt proved that there exists a unique function $L_p(s,\chi)$ satisfying the conditions [K5]:

(1) $L_p(s,\chi)=\dfrac{a_{-1}}{s-1}+\sum_{n=0}^{\infty}a_n(s-1)^n \quad (a_n\in\mathbf{C}_p)$;

(2) $a_{-1}=0$ if $\chi\neq 1$ and the series $\sum_{n=0}^{\infty}a_n(s-1)^n$ converges for $|s-1|_p<|q^{-1}p^{1/(p-1)}|_p$;

(3) $L_p(1-n,\chi)=(1-\chi\omega^{-n}p^{1-n})L(1-n,\chi\omega^{-n})$ holds for $n=1,2,3,\dots$.

The function $L_p(s,\chi)$ satisfying these three conditions is called the **p-adic L-function** for the character χ. It is easy to see that $L_p(s,\chi)$ is identically zero if $\chi(-1)=-1$, but $L_p(s,\chi)$ is nontrivial if $\chi(-1)=1$.

Let B_n be the Bernoulli number. Then B_n satisfies the conditions: (1) B_n/n is p-integral if $(p-1)\nmid n$ (von Staudt) and (2) $(1/n)B_n\equiv(1/(n+p-1))B_{n+p-1} \pmod p$ holds in this case (Kummer). The generalization of these results for the generalized Bermoulli number $B_{\chi,n}$ was obtained by Leopoldt. Since $L(1-n,\chi)=-(1/n)\mathrm{B}_{\chi,n}$, such p-integrabilities and congruences can be naturally interpreted and generalized in terms of the p-adic L-functions $L(s,\chi)$.

We assume $\chi(-1)=1$. Then $L_p(0,\chi)=(1-\chi\omega^{-1}(p))L(0,\chi\omega^{-1})$ and $\chi\omega^{-1}(-1)=-1$. Hence we can express the first factor h_N^- of the class number of a cyclotomic field $\mathbf{Q}(\exp(2\pi i/N))$ as a product of some $L_p(0,\chi)$'s. By using this fact, K. Iwasawa proved [I7] that the p-part $p^{e_n^-}$ of the †first factor $h_{Np^n}^-$ $(N\in\mathbf{N})$ satisfies

$$e_n^-=\lambda n+\mu p^n+\nu \qquad (\lambda,\mu,\nu\in\mathbf{Z};\lambda,\mu\geqslant 0)$$

for any sufficiently large n. Here Iwasawa conjectured $\mu=0$, which was proved by B. Ferrero and L. Washington [F1]. Also, we can obtain some congruences involving the first factor h_N^- of $\mathbf{Q}(\exp(2\pi i/N))$ from this formula.

Let χ be a nontrivial primitive Dirichlet character with conductor f, let

$$\tau(\chi)=\sum_{a=1}^{f}\chi(a)e^{2\pi i a/f}$$

be the †Gaussian sum for χ, and let $\log_p$ be the p-adic logarithmic function. Then Leopoldt [L6] calculated the value $L(1,\chi)$ and obtained

$$L_p(1,\chi)$$
$$=-\left(1-\frac{\chi(p)}{p}\right)\frac{\tau(\chi)}{f}\sum_{a=1}^{f}\chi(a)\log_p(1-e^{-2\pi i a/f}).$$

As an application of this formula, Leopoldt obtained a p-adic †class number formula for the maximal real subfield $F=\mathbf{Q}(\cos(2\pi/N))$ of $\mathbf{Q}(\exp(2\pi i/N))$: Let $\zeta_p(s,F)$ be the product of the $L_p(s,\chi)$ for all primitive Dirichlet characters χ such that (1) $\chi(-1)=1$ and (2) the conductor of χ is a divisor of N. We define the **p-adic regulator** R_p by replacing the usual log by the p-adic logarithmic function $\log_p$. Let h be the class number of F, $m=[F:\mathbf{Q}]$, and let d be the discriminant of F. Then the residue of $\zeta_p(s,F)$ at $s=1$ is

$$\prod_{\chi}\left(1-\frac{\chi(p)}{p}\right)\frac{2^{m-1}hR_p}{\sqrt{d}}.$$

Hence $\zeta_p(s,F)$ has a simple pole at $s=1$ if and only if $R_p\neq 0$. In general, for any totally real finite algebraic number field F, Leopoldt conjectured that the p-adic regulator R_p of F is not zero (**Leopoldt's conjecture**). This conjecture was proved by J. Ax and A. Brumer for the case when F is an Abelian extension of $\mathbf{Q}$ [A4, B7].

By making use of the Stickelberger element, Iwasawa gave another proof of the existence of the p-adic L-function [I7]. In particular, he obtained the following result: Let χ be a primitive Dirichlet character with conductor f. Then there exists a primitive Dirichlet character θ such that the p-part of the conductor of θ is either 1 or q and such that the conductor and the order of $\chi\theta^{-1}$ are both powers of p. Let $\mathfrak{o}_\theta$ be the ring generated over the ring $\mathbf{Z}_p$ of p-adic integers by the values of θ. Then there exists a unique element $f(x,\theta)$ of the quotient field of $\mathfrak{o}_\theta[[x]]$ depending only on θ and satisfying

$$L_p(s,\chi)=2f(\zeta(1+q_0)^s-1,\theta),$$

where q_0 is the least common multiple of q and the conductor of θ, and $\zeta=\chi(1+q_0)^{-1}$. Furthermore, Iwasawa proved that $f(x,\theta)$ belongs to $\mathfrak{o}_\theta[[x]]$ if θ is not trivial.

Let $P=\mathbf{Q}(\exp(2\pi i/q))$ and, for any $n\geqslant 1$, let $P_n=\mathbf{Q}(\exp(2\pi i/qp^n))$. Let $P_\infty=\bigcup_{n\geqslant 1}P_n$. Then P_∞ is a Galois extension of $\mathbf{Q}$ satisfying $\mathrm{Gal}(P_\infty/\mathbf{Q})\cong\mathbf{Z}_p^\times$ (the multiplicative group of p-adic units), and P is the subfield of $P_\infty/\mathbf{Q}$ corresponding to the subgroup $1+q\mathbf{Z}_p$ of $\mathbf{Z}_p^\times$.

Let ψ be a $\mathbf{C}_p$-valued primitive Dirichlet character such that (1) $\psi(-1)=-1$ and (2) the p-part of the conductor f_ψ of ψ is either 1 or q. Let K_ψ be the cyclic extension of $\mathbf{Q}$ corresponding to ψ by class field theory. Let $K=K_\psi\cdot P$, $K_n=K\cdot P_n$ and $K_\infty=K\cdot P_\infty$. Let A_n be the p-primary part of the ideal class group of K_n, let $A_n\to A_m$ $(n\geqslant m)$ be the mapping induced by the †relative norm N_{k_n/K_m}, and let $X_K=\varprojlim A_n$. Since each A_n is a finite p-group, X_K is a $\mathbf{Z}_p$-module. Let $V_K=X_k\otimes_{\mathbf{Z}_p}C_p$, and let

$$V_\psi=\{v\in V_K\,|\,\delta(v)=\psi(\delta)v\text{ for all }\delta\in\mathrm{Gal}(K/\mathbf{Q})\}.$$

Let q_0 be the least common multiple of f_ψ and q, and let γ_0 be the element of $\mathrm{Gal}(K_\infty/K)$ that corresponds to

$$1+q_0\in 1+q\mathbf{Z}_p=\mathrm{Gal}(P_\infty/P)$$

by the restriction mapping $\mathrm{Gal}(K_\infty/K)\hookrightarrow\mathrm{Gal}(P_\infty/P)$. Let $f_\psi(x)$ be the characteristic polynomial of γ_0-1 acting on V_ψ. Hence $f_\psi(x)$ is an element of $\mathfrak{o}_\psi[x]$.

We assume that $\omega\psi^{-1}$ is not trivial. Let $f(x,\omega\psi^{-1})$ be as before. Then $f(x,\omega\psi^{-1})$ is an element of $\mathfrak{o}_\psi[[x]]$. Iwasawa conjectured that $f_\psi(x)$ and $f(x,\omega\psi^{-1})$ coincide up to a unit of $\mathfrak{o}_\psi[[x]]$ (**Iwasawa's main conjecture**). This conjecture was proved recently by B. Mazur and A. Wiles in the case where ψ is a power of ω.

Let F be a totally real finite algebraic number field, let K be a totally real Abelian extension of F, and let χ be a character of $\mathrm{Gal}(K/F)$. Let $L(s,\chi)$ be the †Artin L-function for χ. Then we can construct the p-adic analog $L_p(s,\chi)$ of $L(s,\chi)$ (J.-P. Serre, J. Coates, W. Sinnott, P. Deligne, K. Ribet, P. Cassou-Noguès). But, at present, we have no formula for $L_p(1,\chi)$. Coates generalized Iwasawa's main conjecture to this case, but it has not yet been proven.

P-adic L-functions have been defined in some other cases (e.g. → [K3, M1, M3]).

K. ζ-Functions of Quadratic Forms

Dirichlet defined a Dirichlet series associated with a binary quadratic form and also considered a sum of such Dirichlet series extended over all classes of binary quadratic forms with a given discriminant D, which is actually equivalent to the Dedekind ζ-function of a quadratic field. Dirichlet obtained a formula for the class numbers of binary quadratic forms. The formula is interpreted nowadays as a formula for the class numbers of quadratic fields in the narrow sense.

According as the binary quadratic form is definite or indefinite, we apply different methods to obtain its class number.

Epstein ζ-functions: P. Epstein generalized the definition of the ζ-function of a positive definite binary quadratic form to the case of n variables (*Math. Ann.*, 56 (1903), 63 (1907)). Let V be a real vector space of dimension m with a positive definite quadratic form Q. Let M be a †lattice in V, and put

$$\zeta_Q(s,M)=\sum_{\substack{x\in M\\x\neq 0}}\frac{1}{Q(x)^s},\quad \mathrm{Re}\,s>\frac{m}{2}.$$

This series is absolutely convergent in $\mathrm{Re}\,s>m/2$, and

$$\lim_{s\to m/2}\left(s-\frac{m}{2}\right)\zeta_Q(s,M)=D(M)^{-1/2}\pi^{m/2}\Gamma\left(\frac{m}{2}\right)^{-1},$$

$$D(M)=\det|Q(x_i,x_j)|,$$

where $x_1,\dots,x_m$ is a basis of M and $Q(x,y)=(Q(x+y)-Q(x)-Q(y))/2$. If the $Q(x)$ $(x\in M, x\neq 0)$ are all positive integers, we can write

$$\zeta_Q(s,M)=\sum_{n=1}^{\infty}\frac{a(n)}{n^s},$$

where $a(n)$ is the number of distinct $x\in M$ with $Q(x)=n$. In general, let $x_1,\dots,x_m$ be a basis of the lattice M and $x_1^*,\dots,x_m^*$ be its dual basis $(Q(x_i,x_j^*)=\delta_{ij})$. Call $M^*=\sum_i x_i^*\mathbf{Z}$ the **dual lattice** of M. If we consider the ϑ-series (†theta series)

$$\vartheta_Q(u,M)=\sum_{x\in M}\exp(-\pi uQ(x))\quad(\mathrm{Re}\,u>0),$$

then

$$\vartheta_Q(u,M)=(u^{-m/2}D(M)^{-1/2})\vartheta_Q(u^{-1},M^*).$$

With $\xi_Q(s,M)=\pi^{-s}\Gamma(s)\zeta_Q(s,M)$, the displayed equality leads to the functional equation

$$\xi_Q(s,M)=D(M)^{-1/2}\cdot\xi_Q\left(\frac{m}{2}-s,M^*\right).$$

In general, $\zeta_Q(s,M)$ has no Euler product expansion.

Consider the case where $M=\sum\mathbf{Z}x_i$ $(x_i=$

$(0, \ldots, 0, 1, 0, \ldots, 0))$, $Q(x) = \sum_{i=1}^{m} u_i^2$, for $x = (u_1, \ldots, u_m)$. If we put $\zeta_m(s) = \zeta_Q(s, M)$, $L(s) = \sum_{n=1}^{\infty} (-4/n) n^{-s}$, then we have

$$\zeta_1(s) = 2\zeta(2s),$$

$$\zeta_2(s) = 4\zeta(s) \cdot L(s) = 4 \times (\text{the Dedekind } \zeta\text{-function of } \mathbf{Q}(\sqrt{-1})),$$

$$\zeta_4(s) = 8(1 - 2^{2-2s})\zeta(s)\zeta(s-1),$$

$$\zeta_6(s) = -4(\zeta(s)L(s-2) - 4\zeta(s-2)L(s)),$$

$$\zeta_8(s) = 16(1 - 2^{1-s} + 2^{4-2s})\zeta(s)\zeta(s-3),$$

$$\zeta_{10}(s) = (4/5)(\zeta(s)L(s-4) + 4^2\zeta(s-4)L(s)) - 2 \sum_{\substack{\mu \in \mathbf{Z}[\sqrt{-1}] \\ \mu \neq 0}} \frac{\mu^4}{(\mu\bar{\mu})^s},$$

$$\zeta_{12}(s) = c_1 2^{-s}\zeta(s)\zeta(s-5)(2^6 - 2^{6-s}) + c_2\varphi\{\sqrt{\Delta(\tau)}\},$$

where $\varphi\{\sqrt{\Delta(\tau)}\}$ is the Dirichlet series corresponding to $\sqrt{\Delta(\tau)}$ by the †Mellin transform and $\Delta(\tau) = z\{\prod_{n=1}^{\infty}(1 - z^n)\}^{24}$ with $z = e^{2\pi i \tau}$. $\zeta_m(s)$ has zeros on the line $\mathrm{Re}\, s = \sigma = m/4$, given explicitly for $m = 4, 8$ as follows:

$$m = 4: \quad s = 1 + l\pi i/\log 2, \quad l = 1, 2, \ldots,$$

$$m = 8: \quad s = 2 + (i/\log 2)(2l\pi \pm \arctan\sqrt{15}), \quad l = 0, \pm 1, \ldots.$$

Regarding the Epstein ζ-function of binary quadratic forms

$$\zeta_Q(s) = \sum_{m,n}{}' Q(m, n)^{-s},$$

with

$$Q(x, y) = ax^2 + bxy + cy^2, \qquad a, b, c \in \mathbf{R}, a > 0, c > 0, \Delta = 4ac - b^2 > 0,$$

we have the Chowla-Selberg formula (1949):

$$\zeta_Q(s) = \left(2\zeta(2s)a^{-s} + \frac{2^{2s}a^{s-1}\sqrt{\pi}}{\Gamma(s)\Delta^{s-1/2}} \times \zeta(2s-1)\Gamma\left(s - \frac{1}{2}\right)\right) + \left(\frac{\pi^s 2^{s+3/2}}{a^{1/2}\Gamma(s)\Delta^{s/2-1/4}} \times \sum_{n=1}^{\infty} n^{s-1/2}\sigma_{1-2s}(n)\cos\frac{n\pi b}{a}\right) \times \int_0^{\infty} \varphi^{s-3/2}\exp\left\{-\frac{\pi n \Delta^{1/2}}{2a}(\varphi + \varphi^{-1})\right\} d\varphi,$$

where $\sigma_k(n) = \sum_{d|n} d^k$ and $\zeta(s)$ is the Riemann ζ-function. By using this formula, we can give another proof of the following result of H. Heilbronn: Let $h(-\Delta)$ be the class number of the imaginary quadratic field with discriminant $-\Delta$. Then $h(-\Delta) \to \infty$ $(\Delta \to \infty)$.

The following generalization of this result was obtained by C. L. Siegel [S22]: Let k be a fixed finite algebraic number field. Let K be a finite Galois extension of k, and let $d = d(K)$, $h = h(K)$, and $R = R(K)$ be the discriminant of K, the class number of K, and the regulator of K, respectively. We assume that K runs over extensions of k such that $[K:k]/\log d \to 0$; then we have

$$\log(hR) \sim \log\sqrt{|d|}.$$

Siegel ζ-functions of indefinite quadratic forms: Siegel defined and investigated some ζ-functions attached to nondegenerate indefinite quadratic forms, which are also meromorphic on the whole complex plane and satisfy certain functional equations [S24].

The case of quadratic forms with irrational algebraic coefficients was treated by Tamagawa and K. G. Ramanathan.

L. ζ-Functions of Algebras

K. Hey defined the ζ-function of a †simple algebra A over the rational number field $\mathbf{Q}$ (M. Deuring [D10]) (→ 27 Arithmetic of Associative Algebras). Consider an arbitrary †maximal order $\mathfrak{o}$ of A, and let

$$\zeta_A(s) = \sum_{\mathfrak{a}} \frac{1}{N(\mathfrak{a})^s}, \quad \mathrm{Re}\, s > 1,$$

with the summation taken over all left integral ideals $\mathfrak{a}$ of $\mathfrak{o}$. Then ζ_A is independent of the choice of a maximal order $\mathfrak{o}$. Let k be the †center of A, and put $[A:k] = n^2$. First, ζ_A is decomposed into Euler's infinite product expansion $\zeta_A(s) = \prod_{\mathfrak{p}} \mathbf{Z}_{\mathfrak{p}}(s)$ ($\mathfrak{p}$ runs over the prime ideals of k). For $\mathfrak{p}$ not dividing the discriminant $\mathfrak{d}$ of A, $Z_{\mathfrak{p}}(s)$ coincides with the $\mathfrak{p}$-component of $\prod_{j=0}^{n-1}\zeta_k(ns - j)$. Hence $\zeta_A(s)$ coincides with $\prod_{j=0}^{n-1}\zeta_k(ns - j)$ up to a product of $\mathfrak{p}$-factors for $\mathfrak{p} | \mathfrak{d}$ which are explicit rational functions of $N(\mathfrak{p})^{-ns}$.

Moreover, if A is the total matrix algebra of degree r over the division algebra $\mathfrak{D}$, then we have $\zeta_A(s) = \prod_{j=0}^{r-1}\zeta_{\mathfrak{D}}(rs - j)$, and $\zeta_{\mathfrak{D}}(s)$ satisfies a functional equation similar to that of $\zeta_k(s)$ (Hey). Also, $\zeta_A(s)$ is meromorphic over the whole complex plane, and at $s = 1, (n-1)/n, \ldots, 1/n$, it has poles of order 1. Using analytic methods, M. Zorn (1931) showed that the simple algebra A with center k such that $A_{\mathfrak{p}}$ is a matrix algebra over $k_{\mathfrak{p}}$ for every finite or infinite prime divisor $\mathfrak{p}$ of k is itself a matrix algebra over k (→ 27 Arithmetic of Associative Algebras D). A purely algebraic proof of this was given by Brauer, H. Hasse, and E. Noether. G. Fujisaki (1958) gave another proof using the Iwasawa-Tate method. As a direct application of the ζ-function, the computation of the residue at $s=1$ of ζ_A leads to the formula containing the class number of maximal order $\mathfrak{D}$.

Godement defined the ζ-function of fairly general algebras [G1], and Tamagawa investigated in detail the explicit ζ-functions of division algebras, deriving their functional equations [T1].

Let $\tilde{A}=\prod'_{\mathfrak{p}} A_{\mathfrak{p}}$ be the adele ring of A, and let $G=\prod'_{\mathfrak{p}} G_{\mathfrak{p}}$ be the idele group (of A). We take a maximal order $\mathfrak{D}_{\mathfrak{p}}$ of $A_{\mathfrak{p}}$ and a maximal compact subgroup $U_{\mathfrak{p}}$ of $G_{\mathfrak{p}}$. Let $\omega_{\mathfrak{p}}$ be a †zonal spherical function of $G_{\mathfrak{p}}$ with respect to $U_{\mathfrak{p}}$; that is, $\omega_{\mathfrak{p}}$ is a function in $G_{\mathfrak{p}}$ and satisfies

$$\omega_{\mathfrak{p}}(ugv)=\omega_{\mathfrak{p}}(g) \quad (u,v\in U_{\mathfrak{p}}), \quad \omega_{\mathfrak{p}}(1)=1,$$

$$\int_{U_{\mathfrak{p}}}\omega_{\mathfrak{p}}(guh)\,du=\omega_{\mathfrak{p}}(g)\omega_{\mathfrak{p}}(h).$$

In addition, we define the weight function $\varphi_{\mathfrak{p}}$ on $A_{\mathfrak{p}}$ by

$$\varphi_{\mathfrak{p}}(x)=\begin{cases}\text{the characteristic function of } \mathfrak{D}_{\mathfrak{p}} \\ \text{when } \mathfrak{p} \text{ is finite,} \\ \\ \exp(-\pi T_{\mathfrak{p}}(xx^*)) \\ \text{when } \mathfrak{p} \text{ is infinite,}\end{cases}$$

where $T_{\mathfrak{p}}$ is the †reduced trace of $A_{\mathfrak{p}}/\mathbf{R}$ and $*$ is a positive †involution. Tamagawa gave an explicit form of the local ζ-function with the character $\omega_{\mathfrak{p}}$ defined by

$$\zeta_{\mathfrak{p}}(s,\omega_{\mathfrak{p}})=\int_{G_{\mathfrak{p}}}\varphi_{\mathfrak{p}}(g)\omega_{\mathfrak{p}}(g^{-1})|N_{\mathfrak{p}}(g)|_{\mathfrak{p}}^{s}\,dg,$$

where $N_{\mathfrak{p}}$ is the †reduced norm of $A_{\mathfrak{p}}/k_{\mathfrak{p}}$, and $|\ |_{\mathfrak{p}}$ is the valuation of $k_{\mathfrak{p}}$. Then $\omega=\prod_{\mathfrak{p}}\omega_{\mathfrak{p}}$ is the zonal spherical function of G with respect to $\prod U_{\mathfrak{p}}=U$. In particular, if ω is a positive definite zonal spherical function belonging to the spectrum of the discrete subgroup $\Gamma=A^*$ $=\{$all the invertible elements of $A\}$ of G, then the **Tamagawa ζ-function** with character ω is given by

$$\zeta(s,\omega)=\prod_{\mathfrak{p}}\zeta_{\mathfrak{p}}(s,\omega_{\mathfrak{p}})=\int_G\varphi(g)\omega(g^{-1})\|g\|^s\,dg,$$

where $\varphi(g)=\prod\varphi_{\mathfrak{p}}(g_{\mathfrak{p}})$ and $\|\ \|$ is the volume of the element g of G. When A is a division algebra, $\zeta(s,\omega)$ is analytically continued to a meromorphic function over the whole complex plane and satisfies the functional equation. The Tamagawa ζ-function may also be considered as one type of ζ-function of the Hecke operator. When A is an indefinite quaternion algebra over a totally real algebraic number field Φ, the groups of units of various orders of A operate discontinuously on the product of complex upper half-planes. Thus the spaces of holomorphic forms are naturally associated with A. The investigation of ζ-functions associated with these holomorphic automorphic forms was initiated by M. Eichler and extended by G. Shimura, H. Shimizu, and others. Eichler investigated the case $\Phi=\mathbf{Q}$, and Shimura and Shimizu investigated the case for an arbitrary totally real field Φ by defining general holomorphic automorphic forms, Hecke operators, and corresponding ζ-functions. The functional equations of these ζ-functions were proved by Shimizu. Shimizu generalized Eichler's work and found relations among ζ-functions of orders of various quaternion algebras belonging to different discriminants and levels [S10]. For the related results, see, e.g., the work of K. Doi and H. Naganuma [D12].

M. ζ-Functions Defined by Hecke Operators

The ζ-functions of algebraic number fields, algebras, or quadratic forms, and the L-functions are all defined by Dirichlet series, are analytically continued to univalent functions on the complex plane, and satisfy functional equations. One problem is to characterize the functions having such properties.

(1) H. Hamburger (1921–1922) characterized the Riemann ζ-function (up to constant multiples) by the following three properties: (i) It can be expanded as $\zeta(s)=\sum_{n=1}^{\infty}a_n/n^s$ $(\mathrm{Re}\,s\gg 0)$; (ii) it is holomorphic on the complex plane except as $s=1$, and $(s-1)\zeta(s)$ is an entire function of finite †genus; (iii) $G(s)=G(1-s)$, where $G(s)=\pi^{-s/2}\Gamma(s/2)\zeta(s)$.

(2) E. Hecke's theory [H4]: Fixing $\lambda>0$, $k>0$, $\gamma=\pm 1$, and putting

$$R(s)=(2\pi/\lambda)^{-s}\Gamma(s)\varphi(s)$$

for an analytic function $\varphi(s)$, we make the following three assumptions: (i) $(s-k)\varphi(s)$ is an entire function of finite genus; (ii) $R(s)=\gamma R(k-s)$; (iii) $\varphi(s)$ can be expanded as $\varphi(s)=\sum_{n=1}^{\infty}a_n/n^s$ $(\mathrm{Re}\,s>\sigma_0)$. Then we call $\varphi(s)$ a function belonging to the sign (λ,k,γ).

The functions $\zeta(2s)$, $L(2s)$, and $L(2s-1)$ satisfy assumptions (i)–(iii), where L may be either a Dirichlet L-function, an L-function with Grössencharakter of an imaginary quadratic field, or an L-function with class character of a real quadratic form whose Γ-factors are of the form $\Gamma(s/2)\Gamma((s+1)/2)\sim\Gamma(s)$. If $\varphi(s)$ belongs to the sign (λ,k,γ), then $n^{-s}\varphi(s)$ belongs to the sign $(n\lambda,k,\gamma)$. To each Dirichlet series $\varphi(s)=\sum_{n=1}^{\infty}a_n/n^s$ with the sign (λ,k,γ), we attach the series $f(\tau)=a_0+\sum_{n=1}^{\infty}a_n e^{2\pi in\tau/\lambda}$, where

$$a_0=\gamma(2\pi/\lambda)^{-k}\Gamma(k)\mathrm{Res}_{s=k}(\varphi(s))$$
$$=\gamma\,\mathrm{Res}_{s=k}(R(s)).$$

This correspondence $\varphi(s)\to f(\tau)$ may also be realized by the †Mellin transform

$$R(s)=\int_0^{\infty}\left(\sum_{n=1}^{\infty}a_n e^{-2\pi ny/\lambda}\right)y^{s-1}\,dy$$
$$=\int_0^{\infty}(f(iy)-a_0)y^{s-1}\,dy,$$
$$f(iy)-a_0=\frac{1}{2\pi i}\int_{\mathrm{Re}\,s=\sigma_0}R(s)y^{-s}\,ds.$$

In this case, (i) $f(\tau)$ is holomorphic in the upper half-plane and $f(\tau+\lambda)=f(\tau)$, (ii) $f(-1/\tau)/(-i\tau)^k=\gamma f(\tau)$, and (iii) $f(x+iy)=O(y^{\mathrm{const}})$ $(y\to+0)$ uniformly for all x.

Conversely, the Dirichlet series $\varphi(s)=\sum_{n=1}^{\infty}a_n n^{-s}$ formed by the transformation in the previous paragraph from $f(\tau)$ satisfying (i)–(iii) belongs to the sign (λ,k,γ). We also say that the function $f(\tau)$ belongs to the sign (λ,k,γ).

If k is an even integer, then the functions $f(\tau)$ belonging to $(1,k,(-1)^{k/2})$ are the †modular forms of level 1 and weight k. A necessary and sufficient condition for a function $\varphi(s)$ belonging to $(1,k,(-1)^{k/2})$ to have an Euler product is that the corresponding modular form $f(\tau)$ be a simultaneous eigenfunction of the ring formed by the †Hecke operators T_n $(n=1,2,\dots)$. In this case, the coefficient a_n of $\varphi(s)=\sum a_n n^{-s}$ coincides with the eigenvalue of T_n. Namely, if $f\,|\,T_n=t_n f$, we have $\varphi(s)=a_1(\sum_{n=1}^{\infty}t_n n^{-s})$, and this is decomposed into the Euler product $\varphi(s)=a_1\prod_p(1-t_p p^{-s}+p^{k-1-2s})^{-1}$. We call $\varphi(s)/a_1$ a **ζ-function defined by Hecke operators** (Hecke [H5]). For example, $\zeta(s)\cdot\zeta(s-k+1)$ and the Ramanujan function

$$\sum_{n=1}^{\infty}\tau(n)n^{-s}=\prod_p(1-\tau(p)p^{-s}+p^{11-2s})^{-1}$$

are ζ-functions defined by Hecke operators. Hecke applied the theory of Hecke operators to study the group $\Gamma(N)$ [H5]; the situation is more complicated than the case of $\Gamma(1)=SL(2,\mathbf{Z})$. The space of automorphic forms of weight k belonging to the †congruence subgroup

$$\Gamma_0(N)=\left\{\begin{pmatrix}a&b\\c&d\end{pmatrix}\in SL(2,\mathbf{Z})\,\middle|\,c\equiv 0\pmod N\right\}$$

is denoted by $\mathfrak{M}_k(\Gamma_0(N))$. The essential part of $\mathfrak{M}_k(\Gamma_0(N))$ is spanned by the functions $f(\tau)=\sum a_n e^{2\pi in\tau}$ satisfying the conditions: (1) $\varphi(s)=\sum a_n n^{-s}$ has the Euler product expansion

$$\varphi(s)=\prod_{p|N}(1-a_p p^{-s})^{-1}$$
$$\times\prod_{p\nmid N}(1-a_p p^{-s}+p^{k-1-2s})^{-1}.$$

(2) The functional equation $R(s)=\gamma R(k-s)$ holds, where $R(s)=(2\pi/\sqrt{N})^{-s}\Gamma(s)\varphi(s)$. (3) When χ is an arbitrary primitive character of $\mathbf{Z}$ such that the conductor f is coprime to N, then

$$R(s,\chi)=(2\pi/\sqrt{N}\,f)^{-s}\Gamma(s)\sum a_n\chi(n)n^{-s}$$

extends to an entire function satisfying the functional equation $R(s,\chi)=\omega R(k-s,\bar{\chi})$ $(|\omega|=1)$ (Shimura). Conversely, (2) and (3) characterize the Dirichlet series $\varphi(s)$ corresponding to $f(\tau)\in\mathfrak{M}_k(\Gamma_0(N))$ (Weil [W1 (1967a)]).

Considering the correspondence $f(\tau)=\sum a_n q^n\to\varphi(s)=\sum a_n n^{-s}$ not as a Mellin transformation but rather as a correspondence effected through Hecke operators, we can derive the ζ-function defined by Hecke operators. When the Hecke operator T_n is defined with respect to a discontinuous group Γ and we have a representation space $\mathfrak{M}$ of the Hecke operator ring $\mathscr{H}$, we denote the matrix of the operation of $T_n\in\mathscr{H}$ on $\mathfrak{M}$ by $(T_n)=(T_n)_{\mathfrak{M}}$ and call the matrix-valued function $\sum_n(T_n)_{\mathfrak{M}}n^{-s}$ the ζ-function defined by Hecke operators. The equation $\varphi(s)=\sum a_n n^{-s}$ is a specific instance of the correspondence in the first sentence, where $\Gamma=\Gamma(N)$, $\mathfrak{M}\subset\mathfrak{M}_k(\Gamma_0(N))$, $\dim\mathfrak{M}=1$. One advantage of this definition is that it may be applied whenever the concept of Hecke operators can be defined with respect to the group Γ (for instance, even for the Fuchsian group without a †cusp). Thus when Γ is a Fuchsian group given by the unit group of a quaternion algebra Φ over the rational number field $\mathbf{Q}$ and $\mathfrak{M}$ is the space of automorphic forms with respect to Γ, the ζ-function $\sum(T_n)n^{-s}$ is defined (Eichler). Moreover, by using its integral expression over the idele group $\mathbf{J}_\Phi$ of Φ, we can obtain its functional equation following the Iwasawa-Tate method (Shimura). Furthermore, by algebrogeometric consideration of T_n, it can be shown that

$$\zeta(s)\zeta(s-1)\det(\textstyle\sum(T_n)_{\mathfrak{S}_2}n^{-s})$$
$$=\zeta(s)\zeta(s-1)\det\left(\prod_p(1-(T_p)_{\mathfrak{S}_2}p^{-s}+(R_p)_{\mathfrak{S}_2}p^{1-2s})^{-1}\right)$$

coincides (up to a trivial factor) with the Hasse ζ-function of some model of the Riemann surface defined by Γ when $\mathfrak{M}$ is the space $\mathfrak{S}_2$ of all †cusp forms of weight 2 (Eichler [E1], Shimura [S12]).

The algebrogeometric meaning of $\det(\sum(T_n)_{\mathfrak{S}_k}n^{-s})$, when $\mathfrak{M}$ is the space $\mathfrak{S}_k$ of all cusp forms of weight k, has been made clear for the case where Γ is obtained from $\Gamma_0(N)$, $\Gamma(N)$, and the quaternion algebra (M. Kuga, M. Sato, Shimura, Y. Ihara). From these facts, it becomes possible to express $(T_p)_{\mathfrak{S}_k}$, the decomposition of the prime number p in some type of Galois extension (Shimura [S14], Kuga), in terms of Hecke operators. These works gave the first examples of non-Abelian class field theory. Note that this type of ζ-function may be regarded as the analog (or generalization) of L-functions of algebraic number fields, as can be seen from the comparison in Table 1.

Table 1

Algebraic number field	k	Ideal group	Character χ	$\Sigma\chi(n)n^{-s}$
	$\updownarrow$	$\updownarrow$	$\updownarrow$	$\updownarrow$
Algebraic group	G	Hecke ring	Representation space $\mathfrak{M}$	$\Sigma(T_n)_{\mathfrak{M}}n^{-s}$

As for special values of ζ-functions defined by Hecke operators, the following fact is known: Let $f(\tau)=\Sigma a_n q^n \in \mathfrak{M}_k(SL(2,\mathbf{Z}))$ be a common eigenfunction of the Hecke operators, and let $\varphi(s)=\Sigma a_n n^{-s}$ be the corresponding Dirichlet series. Let K_f be the field generated over the rational number field $\mathbf{Q}$ by the coefficients a_n of f. Then, for any two integers m and m' satisfying $0<m, m'<k$ and $m\equiv m' \pmod 2$, the ratio $(R(m):R(m'))$ of the special values of

$$R(s)=\frac{\Gamma(s)}{(2\pi)^s}\varphi(s)=\int_0^\infty (f(iy)-a_0)y^{s-1}\,dy$$

at m and m' belongs to the field K_f.

G. Shimura discovered this fact for Ramanujan's function $\Delta(\tau)$ (*J. Math. Soc. Japan*, 11 (1959)), and then Yu. I. Manin generalized it to the above case and, by constructing a p-adic analog of $\varphi(s)$ from it, pointed out the importance of such results [M1]. R. M. Damerell also used such results to study special values of Hecke's L-function with Grössencharakter of an imaginary quadratic field (*Acta Arith.*, 17 (1970), 19 (1971)). Furthermore, Shimura generalized these results to congruence subgroups of $SL(2,\mathbf{Z})$ (*Comm. Pure Appl. Math.*, 29 (1976)), and to Hilbert modular groups (*Ann. Math.*, 102 (1975)). The connection between special values of ζ-functions and the periods of integrals has been studied further by Shimura, Deligne, and others.

In addition, in connection with nonholomorphic automorphic forms H. Maass considered L-functions of real quadratic fields (with class characters) having $\Gamma(s/2)^2$ or $\Gamma((s+1)/2)^2$ as Γ-factors. Furthermore, T. Kubota studied the relation of ζ-functions $\zeta_k(s)$ of an arbitrary algebraic field k or ζ-functions of simple rings to (nonanalytic) automorphic forms of several variables and considered the reciprocity law for the Gaussian sum from a new viewpoint.

N. L-Functions of Automorphic Representations (I)

R. P. Langlands reconstructed the theory of †Hecke operators from the viewpoint of representation theory and defined very general L-functions. He proposed many conjectures about them in [L4], and he and H. Jacquet proved most of them in [J1] for the case $G=GL_2$.

First Langlands defined the L-group ${}^LG$ for any connected reductive algebraic group G defined over a field k in the following manner [B6].

There is a canonical bijection between isomorphism classes of connected †reductive algebraic groups defined over a fixed algebraically closed field $\bar{k}$ and isomorphism classes of †root systems. It is defined by associating to G the root data $\Psi(G)=(X^*(T),\Phi,X_*(T),\Phi^\vee)$, where T is a †maximal torus of G, $X^*(T)$ ($X_*(T)$) the group of characters (†1-parameter subgroups) of T, Φ ($\Phi^\vee$) the set of roots (coroots) of G with respect to T.

Since the choice of a †Borel subgroup B of G containing T is equivalent to that of a basis Δ of Φ, the aforementioned bijection yields one between isomorphism classes of triples (G,B,T) and isomorphism classes of based root data $\Psi_0(G)=(X^*(T),\Delta,X_*(T),\Delta^\vee)$. There is a split exact sequence

$$1\to \operatorname{Int} G\to \operatorname{Aut} G\to \operatorname{Aut}\Psi_0(G)\to 1.$$

and this mapping induces a canonical bijection $\operatorname{Aut}\Psi_0(G)\cong\operatorname{Aut}(G,B,T,\{x_\alpha\}_{\alpha\in\Delta})$ if $x_\alpha\in G_\alpha$ $(\alpha\in\Delta)$ are fixed.

Let G be a connected reductive algebraic group defined over $\bar{k}$. Let T be a maximal torus of G, and let B be a Borel subgroup of G containing T. Let $\Psi_0(G)=(X^*(T),\Delta,X_*(T),\Delta^\vee)$ be as before. Then there is a connected reductive algebraic group ${}^LG^0$ over $\mathbf{C}$ such that $\Psi_0(G)^\vee=(X_*(T),\Delta^\vee,X^*(T),\Delta)$ corresponds to the triple $({}^LG^0,{}^LB^0,{}^LT^0)$, where ${}^LB^0$ and ${}^LT^0$ are a Borel subgroup of ${}^LG^0$ and the maximal torus of ${}^LB^0$. For example, (1) if $G=GL_n$, then ${}^LG^0=GL_n(\mathbf{C})$; (2) if $G=Sp_{2n}$, then ${}^LG^0=SO_{2n+1}(\mathbf{C})$.

We assume that $\bar{k}$ is the algebraic closure of k and G is defined over k. Then $\gamma\in\operatorname{Gal}(\bar{k}/k)$ induces an automorphism of the $\bar{k}$-group $G\times_k\bar{k}$. Hence γ defines an element of $\operatorname{Aut}({}^LG^0,{}^LB^0,{}^LT^0)$ because it is a holomorphic image of $\operatorname{Aut}\Psi_0(G\times_k\bar{k})=\operatorname{Aut}\Psi_0(G\times_k\bar{k})^\vee$. Hence we can define the †semidirect product ${}^LG={}^LG^0\rtimes\operatorname{Gal}(\bar{k}/k)$, and call it the **$L$-group** of G.

Let k be a †local field, and let G be a connected reductive algebraic group defined over k. We identify G with the group of its k-rational points. Let W_k' be the Weil-Deligne group of k (→ Section H), and let $\Phi(G)$ be the set of homomorphisms $\varphi: W_k'\to{}^LG$ over $\operatorname{Gal}(\bar{k}/k)$. Let $\Pi(G)$ be the set of infinitesimal equivalence classes of irreducible **admissible** representations of G. If k is a non-Archimedean field, then

$\Pi(G)$ is the set consisting of equivalence classes of irreducible representations $\pi: G \to \operatorname{Aut} V$ on complex vector spaces V such that the space V^K of vectors invariant by K is finite dimensional for every compact open subgroup K of G and such that $V = \bigcup V^K$, where K runs over the compact open subgroups of G. If k is an Archimedean field, then $\Pi(G)$ is the set consisting of equivalence classes of representations π of the pair $(\mathfrak{g}, K)$ of the Lie algebra $\mathfrak{g}$ of G and a maximal compact subgroup K satisfying similar conditions [B6]. Then Langlands conjectured that we can parametrize $\Pi(G)$ by $\Phi(G)$ as $\Pi(G) = \bigcup_\varphi \Pi(G)_\varphi$. Let $\pi \in \Pi(G)_\varphi$ ($\varphi \in \Phi(G)$), and let r be a representation of ${}^LG$. Then we can define the **L-function** $L(s, \pi, r)$ and the **ε-factor** $\varepsilon(s, \pi, r)$ of π by

$$L(s, \pi, r) = L(s, r \circ \varphi), \quad \varepsilon(s, \pi, r) = \varepsilon(s, r \circ \varphi, \psi),$$

where the right-hand sides are those of the Weil-Deligne group (→ Section H) and ψ is a nontrivial character of k.

Let G be a connected reductive group over a global field k (i.e., an algebraic number field of finite degree or an algebraic function field of one variable over a finite field), let π be an irreducible admissible representation of $G_{\mathbf{A}}$, where $G_{\mathbf{A}}$ is the group of rational points of G over the †adele ring $k_{\mathbf{A}}$ of k, and let r be a finite-dimensional representation of ${}^LG$. Let ψ be a nontrivial character of $k_{\mathbf{A}}$ which is trivial on k. For any place v of k, let r_v be the representation of the L-group of $G_v = G \times_k k_v$ induced by r, and let ψ_v be the additive character of k_v associated with ψ. It is known that π is decomposed into the tensor product $\bigotimes \pi_v$ of $\pi_v \in \Pi(G(k_v))$ [B6]. Hence we put

$$L(s, \pi, r) = \prod_v L(s, \pi_v, r_v),$$

$$\varepsilon(s, \pi, r) = \prod_v \varepsilon(s, \pi_v, r_v).$$

The local factor $L(s, \pi_v, r_v)$ is in fact defined if v is Archimedean, or G is a †torus, or φ is unramified (i.e., G_v is quasisplit and splits over an unramified extension of k_v, and $G(\mathfrak{o}_v)$ is a special maximal compact subgroup of $G(k_v)$, and π_v is of class one with respect to $G(\mathfrak{o}_v)$, where $\mathfrak{o}_v$ is the integer ring of k_v). It follows that the right-hand side $\prod L(s, \pi_v, r_v)$ is defined up to a finite number of non-Archimedean places v. Furthermore, Langlands proved that $\prod \varepsilon(s, \pi_v, r_v)$ is in fact a finite product, and the infinite product $\prod L(s, \pi_v, r_v)$ converges in some right half-plane if π is automorphic (i.e., if π is a subquotient of the right regular representation of $G_{\mathbf{A}}$ in $G_k \backslash G_{\mathbf{A}}$). It is conjectured that $L(s, \pi, r)$ admits a meromorphic continuation to the whole complex plane and satisfies a functional equation

$$L(s, \pi, r) = \varepsilon(s, \pi, r) L(1 - s, \tilde{\pi}, r)$$

if π is automorphic, where $\tilde{\pi}$ is the †contragredient representation of π. Furthermore, if $G = GL_n$ and r is the standard representation of GL_n, then we can construct $L(s, \pi, r)$ and $\varepsilon(s, \pi, r)$ by generalizing the Iwasawa-Tate method. We can also show in this case that $L(s, \pi, r)$ is entire if π is cuspidal. The conjectures are studied in some other cases [B6].

O. L-Functions of Automorphic Representations (II)

A. Weil generalized the theory of †Hecke operators and the corresponding L-functions to the case of †automorphic forms (for holomorphic and nonholomorphic cases together) of GL_2 over a global field [W9]. Then H. Jacquet and Langlands developed a theory from the viewpoint of †representation theory [J1, J2]. They attached L-functions not to automorphic forms but to †automorphic representations of $GL_2(k)$.

Let k be a non-Archimedean local field, and let $\mathfrak{o}_k$ be the maximal order of k. Let $\mathscr{H}_k$ be the space of functions on $G_k = GL_2(k)$ that are locally constant and compactly supported. Then $\mathscr{H}_k$ becomes an algebra with the convolution product

$$(f_1 * f_2)(h) = \int_{G_k} f_1(g) f_2(g^{-1}h)\, dg,$$

where dg is the †Haar measure of G_k that assigns 1 to the maximal compact subgroup $K_k = GL_2(\mathfrak{o}_k)$. Let π be a representation of $\mathscr{H}_k$ on a complex vector space V. Then we say that π is **admissible** if and only if π satisfies the following two conditions: (1) For every v in V, there is an f in $\mathscr{H}_k$ so that $\pi(f)v = v$; (2) Let σ_i $(i = 1, \ldots, r)$ be a family of inequivalent irreducible finite-dimensional representations of K_k, and let

$$\xi(g) = \sum_{i=1}^{r} \dim(\sigma_i)^{-1} \operatorname{tr} \sigma_i(g^{-1}).$$

Then ξ is an idempotent of $\mathscr{H}_k$. We call such a ξ an **elementary idempotent** of $\mathscr{H}_k$. Then for every elementary idempotent ξ of $\mathscr{H}_k$, the operator $\pi(\xi)$ has a finite-dimensional range. If π is an admissible representation of $GL_2(k)$ (→ Section N), then

$$\pi(f) = \int_{G_k} f(g)\pi(g)\, dg \qquad (f \in \mathscr{H}_k)$$

gives an admissible representation of $\mathscr{H}_k$ in this sense. Furthermore, any admissible representation of $\mathscr{H}_k$ can be obtained from an admissible representation of $GL_2(k)$.

Let k be the real number field. Let $\mathscr{H}_1$ be the space of infinitely differentiable compactly supported functions on $G_k(=GL_2(k))$ that are $K_k(=O(2,k))$ finite on both sides, let $\mathscr{H}_2$ be the space of functions on K_k that are finite sums of matrix elements of irreducible representations of K_k, and let $\mathscr{H}_k=\mathscr{H}_1\oplus\mathscr{H}_2$. Then $\mathscr{H}_1$, $\mathscr{H}_2$, and $\mathscr{H}_k$ become algebras with the convolution product. Let π be a representation of $\mathscr{H}_k$ on a complex vector space V. Then π is **admissible** if and only if the following three conditions are satisfied: (1) Every vector v in V is of the form $v=\sum_{i=1}^r\pi(f_i)v_i$ with $f_i\in\mathscr{H}_1$ and $v_i\in V$; (2) for every elementary idempotent $\xi(g)=\sum_{i=1}^r\dim(\sigma_i)^{-1}\operatorname{tr}\sigma_i(g^{-1})$, where the σ_i are a family of inequivalent irreducible representations of K_k, the range of $\pi(\xi)$ is finite-dimensional; (3) for every elementary idempotent ξ of $\mathscr{H}_k$ and for every vector v in $\pi(\xi)V$, the mapping $f\mapsto\pi(f)v$ of $\xi\mathscr{H}_1\xi$ into the finite-dimensional space $\pi(\xi)V$ is continuous. We can define the Hecke algebra $\mathscr{H}_k$ and the notion of admissible representations also in the case $k=\mathbf{C}$. In these cases, an admissible representation of $\mathscr{H}_k$ comes from a representation of the †universal enveloping algebra of $GL_2(k)$ but may not come from a representation of $GL_2(k)$. It is known that for any local field k, the †character of each irreducible representation is a locally integrable function.

Let k be a global field, $G_k=GL_2(k)$, and let $G_\mathbf{A}=GL_2(k_\mathbf{A})$ be the group of rational points of G_k over the adele ring $k_\mathbf{A}$ of k. For any place v of k, let k_v be the completion of k at v, let $G_v=GL_2(k_v)$, and let k_v be the standard maximal compact subgroup of G_v. Let $\mathscr{H}_v$ be the Hecke algebra $\mathscr{H}_{k_v}$ of G_v, and let ε_v be the normalized Haar measure of K_v. Then ε_v is an elementary idempotent of $\mathscr{H}_v$. Let $\mathscr{H}=\bigotimes_{\varepsilon_v}\mathscr{H}_v$ be the restricted tensor product of the local Hecke algebra $\mathscr{H}_v$ with respect to the family $\{\varepsilon_v\}$. We call $\mathscr{H}$ the **global Hecke algebra** of G_A.

Let π be a representation of $\mathscr{H}$ on a complex vector space V. We define the notion of admissibility of π as before. Then we can show that, for any irreducible admissible representation π of $\mathscr{H}$ and for any place v of k, there exists an irreducible admissible representation π_v of $\mathscr{H}_v$ on a complex vector space V_v such that (1) for almost all v, $\dim V_v^{K_v}=1$ and (2) π is equivalent to the restricted tensor product $\bigotimes\pi_v$ of the π_v with respect to a family of nonzero $x_v\in V_v^{K_v}$. Furthermore, the factors $\{\pi_v\}$ are unique up to equivalence.

Let k be a local field, let ψ be a nontrivial character of k, and let $\mathscr{H}_k$ be the Hecke algebra of $G_k=GL_2(k)$. Let π be an infinite-dimensional admissible irreducible representation of $\mathscr{H}_k$. Then there is exactly one space $W(\pi,\psi)$ of continuous functions on G_k with the following three properties: (1) If W is in $W(\pi,\psi)$, then for all g in G_k and for all x in k,

$$W(\begin{pmatrix}1 & x\\ 0 & 1\end{pmatrix}g)=\psi(x)W(g);$$

(2) $W(\pi,\psi)$ is invariant under the right translations of $\mathscr{H}_k$, and the representation on $W(\pi,\psi)$ is equivalent to π; (3) if k is Archimedean and if W is in $W(\pi,\psi)$, then there is a positive number N such that

$$W(\begin{pmatrix}t & 0\\ 0 & 1\end{pmatrix})=O(|t|^N)$$

as $|t|\to\infty$. We call $W(\pi,\psi)$ the **Whittaker model** of π. The Whittaker model exists in the global case if and only if each factor π_v of $\pi=\bigotimes\pi_v$ is infinite-dimensional.

Let k be a local field, and let π be as before. Then the L-function $L(s,\pi)$ and the ε-factor $\varepsilon(s,\pi,\psi)$ are defined in the following manner: Let ω be the quasicharacter of $k^\times$ (i.e., the continuous homomorphism $k^\times\to\mathbf{C}^\times$) defined by

$$\pi(\begin{pmatrix}a & 0\\ 0 & a\end{pmatrix})=\omega(a)id_V.$$

Then the †contragredient representation $\tilde{\pi}$ of π is equivalent to $\omega^{-1}\otimes\pi$. For any g in G_k and W in $W(\pi,\psi)$, let

$$\Psi(g,s,W)=\int_{k^\times}W(\begin{pmatrix}a & 0\\ 0 & 1\end{pmatrix}g)|a|^{s-1/2}\,d^\times a,$$

$$\tilde{\Psi}(g,s,W)=\int_{k^\times}W(\begin{pmatrix}a & 0\\ 0 & 1\end{pmatrix}g)|a|^{s-1/2}\omega^{-1}(a)\,d^\times a.$$

Then there is a real number s_0 such that these integrals converge for $\operatorname{Re}(s)>s_0$ for any $g\in G_k$ and $W\in W(\pi,\psi)$. If k is a non-Archimedean local field with $\mathbf{F}_q$ as its residue field, then there is a unique factor $L(s,\pi)$ such that $L(s,\pi)^{-1}$ is a polynomial of q^{-s} with constant term 1,

$$\Phi(g,s,W)=\Psi(g,s,W)/L(s,\pi)$$

is a holomorphic function of s for all g and W, and there is at least one W in $W(\pi,\psi)$ so that $\Phi(e,s,W)=a^s$ with a positive constant a. If k is an Archimedean local field, then we can define the gamma factor $L(s,\pi)$ in the same manner. Furthermore, for any local field k, if

$$\tilde{\Phi}(g,s,W)=\tilde{\Psi}(g,s,W)/L(s,\tilde{\pi}),$$

then there is a unique factor $\varepsilon(s,\psi,\pi)$ which, as a function of s, is an exponential such that

$$\tilde{\Phi}(\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}g,1-s,W)=\varepsilon(s,\pi,\psi)\Phi(g,s,W)$$

for all $g\in G_k$ and $W\in W(\pi,\psi)$.

Let π and π' be two infinite-dimensional irreducible admissible representations of G_k. Then π and π' are equivalent if and only if the quasicharacters ω and ω' are equal and

$$\frac{L(1-s,\chi^{-1}\otimes\tilde{\pi})\varepsilon(s,\chi\otimes\pi,\psi)}{L(s,\chi\otimes\pi)} = \frac{L(1-s,\chi^{-1}\otimes\tilde{\pi}')\varepsilon(s,\chi\otimes\pi',\psi)}{L(s,\chi\otimes\pi')}$$

holds for any quasicharacter χ. In particular, the set $\{L(s,\chi\otimes\pi)$ and $\varepsilon(s,\chi\otimes\pi,\psi)$ for all $\chi\}$ characterizes the representation π.

Let k be a global field, $G_k=GL_2(k)$, $G_{\mathbf{A}}=GL_2(k_{\mathbf{A}})$, and let $K_{\mathbf{A}}=\prod K_v$ be the standard maximal compact subgroup of $G_{\mathbf{A}}$. Then the †global Hecke algebra $\mathscr{H}$ acts on the space of continuous functions on $G_k\backslash G_{\mathbf{A}}$ by the right translations. Let φ be a continuous function on $G_k\backslash G_{\mathbf{A}}$. Then φ is an **automorphic form** if and only if (1) φ is $K_{\mathbf{A}}$-finite on the right, (2) for every †elementary idempotent ξ in $\mathscr{H}$, the space $(\xi\mathscr{H})\varphi$ is finite-dimensional, and (3) φ is slowly increasing if k is an algebraic number field. An automorphic form φ is a **cusp form** if and only if

$$\int_{k\backslash k_{\mathbf{A}}}\varphi(\begin{pmatrix}1 & x\\ 0 & 1\end{pmatrix}g)\,dx=0$$

for all g in $G_{\mathbf{A}}$. Let $\mathscr{A}$ be the space of automorphic forms on $G_k\backslash G_{\mathbf{A}}$, and let $\mathscr{A}_0$ be the space of cusp forms on $G_k\backslash G_{\mathbf{A}}$. They are $\mathscr{H}$-modules. Let $\psi=\prod\psi_v$ be a nontrivial character of $k\backslash k_{\mathbf{A}}$, and let π be an irreducible admissible representation $\pi=\bigotimes_v\pi_v$ of the global Hecke algebra $\mathscr{H}=\bigotimes_{\varepsilon_v}\mathscr{H}_v$. If π is a †constituent of the $\mathscr{H}$-module $\mathscr{A}$, then we can define the local factors $L(s,\pi_v)$ and $\varepsilon(s,\pi_v,\psi_v)$ for all v, although π_v may not be infinite-dimensional. Further, the infinite products

$$L(s,\pi)=\prod L(s,\pi_v)\text{ and }L(s,\tilde{\pi})=\prod L(s,\tilde{\pi}_v)$$

converge absolutely in a right half-plane, and the functions $L(s,\pi)$ and $L(s,\tilde{\pi})$ can be analytically continued to the whole complex plane as meromorphic functions of s. If π is a constituent of $\mathscr{A}_0$, then all π_v are infinite-dimensional, $L(s,\pi)$ and $L(s,\tilde{\pi})$ are entire functions, and π is contained in $\mathscr{A}_0$ with multiplicity one. If k is an algebraic number field, then they have only a finite number of poles and are bounded at infinity in any vertical strip of finite width. If k is an algebraic function field of one variable with field of constant $\mathbf{F}_q$, then they are rational functions of q^{-s}. In either case, $\varepsilon(s,\pi_v,\psi_v)=1$ for almost all v, and hence

$$\varepsilon(s,\pi)=\prod\varepsilon(s,\pi_v,\psi_v)$$

is well defined. Furthermore, the functional equation

$$L(s,\pi)=\varepsilon(s,\pi)\,L(1-s,\tilde{\pi})$$

is satisfied.

As for the condition for π being a constituent of $\mathscr{A}_0$, we have the following: Let $\pi=\bigotimes\pi_v$ be an irreducible admissible representation of $\mathscr{H}$. Then π is a constituent of $\mathscr{A}_0$ if and only if (1) for every v, π_v is infinite-dimensional; (2) the quasicharacter η defined by

$$\pi(\begin{pmatrix}a & 0\\ 0 & a\end{pmatrix})=\eta(a)id.$$

is trivial on $k^\times$; (3) π satisfies a certain condition so that, for any quasicharacter ω of $k^\times\backslash k_{\mathbf{A}}^\times$, $L(s,\omega\otimes\pi)=\prod L(s,\omega_v\otimes\pi_v)$ and $L(s,\omega^{-1}\otimes\tilde{\pi}_v)=\prod L(s,\omega_v^{-1}\otimes\tilde{\pi}_v)$ converge on a right half-plane; and (4) for any quasicharacter ω of $k^\times\backslash k_{\mathbf{A}}^\times$, $L(s,\omega\otimes\pi)$ and $L(s,\omega^{-1}\otimes\tilde{\pi})$ are entire functions of s which are bounded in vertical strips and satisfy the functional equation

$$L(s,\omega\otimes\pi)=\varepsilon(s,\omega\otimes\pi)\,L(1-s,\omega^{-1}\otimes\tilde{\pi}).$$

P. Congruence ζ-Functions of Algebraic Function Fields of One Variable or of Algebraic Curves

Let K be an †algebraic function field of one variable over $k=\mathbf{F}_q$ (finite field with q elements). The **ζ-function of the algebraic function field** K/k, denoted by $\zeta_K(s)$, is defined by the infinite sum $\sum_{\mathfrak{A}}N(\mathfrak{A})^{-s}$, where the summation is over all integral divisors $\mathfrak{A}$ of K/k and where the norm $N(\mathfrak{A})$ equals $q^{\deg(\mathfrak{A})}$. Equivalently, $\zeta_K(s)$ is defined by the infinite product $\prod_{\mathfrak{p}}(1-N(\mathfrak{p})^{-s})^{-1}$, where $\mathfrak{p}$ runs over all prime divisors of K/k. By the change of variable $u=q^{-s}$, $\zeta_K(s)=Z_K(u)$ becomes a formal power series in u. $\zeta_K(s)$ and $Z_K(u)$ are sometimes called the **congruence ζ-functions** of K/k.

The fundamental theorem states that (i) (Rationality) $Z_K(u)$ is a rational function of u of the form $Z_K(u)=P(u)/(1-u)(1-qu)$, where $P(u)\in\mathbf{Z}[u]$ is a polynomial of degree $2g$, g being the genus of K; (ii) (Functional equation) $Z_K(u)$ satisfies the functional equation

$$Z_K(1/u)=q^{g-1}u^{2-2g}Z_K(u/q);$$

and (iii) if $P(u)$ is decomposed into linear factors in $\mathbf{C}[u]$: $P(u)=\prod_{i=1}^{2g}(1-\alpha_i u)$, then all the reciprocal roots α_i are complex numbers of absolute value $\sqrt{q}$. Statement (iii) is the analog of the **Riemann hypothesis** because it is equivalent to saying that all the zeros of $\zeta_K(s)=Z_K(q^{-s})$ lie on the line $\mathrm{Re}\,s=1/2$.

The congruence ζ-function was introduced by E. Artin [A1 (1924)] as an analog of the Riemann or Dedekind ζ-functions. Of its fundamental properties, the rationality (i) and the functional equation (ii) were proven by

F. K. Schmidt (1931), using the †Riemann-Roch theorem for the function field K/k. The Riemann hypothesis (iii) was verified first in the elliptic case ($g=1$) by H. Hasse [H1] and then in the general case by A. Weil [W2 (1948)]. For the proof of (iii), it was essential to consider the geometry of algebraic curves that correspond to given function fields.

Let C be a nonsingular complete curve over k with function field K. Then $Z_K(u)$ coincides with the ζ-function of C/k, denoted by $Z(u, C)$, which is defined by the formal power series $\exp(\sum_{m=1}^{\infty} N_m u^m/m)$. Here N_m is the number of rational points of C over the extension k_m of k of degree m. The rationality of $Z_K(u)$ is then equivalent to the formula

$$N_m = 1 + q^m - \sum_{i=1}^{2g} \alpha_i^m \qquad (m \in \mathbf{N}),$$

and the Riemann hypothesis for $Z_K(u)$ is equivalent to the estimate

$$(*) \quad |N_m - 1 - q^m| \leqslant 2g\, q^{m/2} \qquad (m \in \mathbf{N}).$$

Now if F is the qth power morphism of C to itself (the **Frobenius morphism** of C relative to k), then an important observation is that N_m is the number of fixed points of the mth iterate F^m of F. In other words, N_m is equal to the intersection number of the graph of F^m with the diagonal on the surface $C \times C$, and is related to the "trace" of the Frobenius correspondence. Then (*) follows from †Castelnuovo's lemma in the theory of correspondences on a curve. This is Weil's proof of the Riemann hypothesis in [W2]; compare the proof by A. Mattuck and J. Tate (*Abh. Math. Sem. Hamburg* 22 (1958)) and A. Grothendieck (*J. Reine Angew. Math.*, 200 (1958)) using the Riemann-Roch theorem for an algebraic surface.

On the other hand, let J be the †Jacobian variety of C over k. For each prime number l different from the characteristic of k, let $M_l(\alpha)$ denote the †l-adic representation of an endomorphism α of J obtained from its action on points of J of order l^n ($n=1,2,\dots$). Letting π be the endomorphism of J induced from F (which is the same as the Frobenius morphism of J), we have $P(u) = \det(1 - M_l(\pi)u)$, i.e., the numerator of the ζ-function coincides with the characteristic polynomial of $M_l(\pi)$. In this setting, the Riemann hypothesis is a consequence of the positivity of the Rosati antiautomorphism [E1]. This is the second proof given by Weil [W2], and applies to arbitary Abelian varieties.

Recently E. Bombieri, inspired by Stepanov's idea, gave an elementary proof of (*) using only the Riemann-Roch theorem for a curve (*Sém. Bourbaki*, no. 430 (1973)).

Q. ζ-Functions of Algebraic Varieties over Finite Fields

Let V be an algebraic variety over the finite field with q elements $\mathbf{F}_q$, and let N_m be the number of $\mathbf{F}_q{}^m$-rational points of V. Then the **ζ-function of V over $\mathbf{F}_q$** is the formal power series in $\mathbf{Z}[[u]]$ defined by

$$Z(u, V) = \exp\left(\sum_{m=1}^{\infty} N_m u^m/m\right);$$

alternatively it can be defined by the infinite product $\prod_P (1 - u^{\deg P})^{-1}$, where P runs over the set of prime divisors of V and $\deg P$ is the degree of the residue field of P over $\mathbf{F}_q$ (in other words, P runs over prime rational 0-cycles of V over $\mathbf{F}_q$).

Weil Conjecture. In 1949, the following properties of the ζ-function were conjectured by Weil [W3]. Let V be an n-dimensional complete nonsingular (absolutely irreducible) variety over $\mathbf{F}_q$. Then (1) $Z(u, V)$ is a rational function of u. (2) $Z(u, V)$ satisfies the functional equation

$$Z((q^n u)^{-1}, V) = \pm q^{n\chi/2} u^{\chi} Z(u, V),$$

where the integer χ is the intersection number (the degree of $\Delta_V \cdot \Delta_V$) of the diagonal subvariety Δ_V with itself in the product $V \times V$, which is called the Euler-Poincaré characteristic of V. (3) Moreover, we have

$$Z(u, V) = \frac{P_1(u) \cdot P_3(u) \cdot \ldots \cdot P_{2n-1}(u)}{P_0(u) \cdot P_2(u) \cdot \ldots \cdot P_{2n}(u)},$$

where $P_h(u) = \prod_{j=1}^{B_h} (1 - \alpha_j^{(h)} u)$ is a polynomial with $\mathbf{Z}$-coefficients such that $\alpha_j^{(h)}$ are algebraic integers of absolute value $q^{h/2}$ ($0 \leqslant h \leqslant 2n$); the latter statement is the **Riemann hypothesis for** $V/\mathbf{F}_q$. (4) When V is the reduction mod p of a complete nonsingular variety V^* of characteristic 0, then the degree B_h of $P_h(u)$ is the hth Betti number of V^* considered as a complex manifold.

This conjecture, called the **Weil conjecture**, has been completely proven. To give a brief history, first B. Dwork [D13] proved the rationality of the ζ-function for any (not necessarily complete or nonsingular) variety over $\mathbf{F}_q$. Then A. Grothendieck [A3, G2, G3] developed the l-adic étale cohomology theory with M. Artin and others, and proved the above statements (1)–(4) (except for the Riemann hypothesis) with $P_h(u)$ replaced by some $P_{h,l}(u) \in \mathbf{Q}_l[u]$; and S. Lubkin [L7] obtained similar results for liftable varieties. Finally Deligne [D4] proved the Riemann hypothesis and the independence of l of $P_{h,l}(u)$. More details will be given below. Before the final solution for the general case was obtained, the conjecture had been verified for some special types of varieties. For curves and Abelian varieties, its truth was previously shown by Weil (→ Section P). In the paper [W3] in which he proposed the above conjecture, Weil verified it for Fermat hypersurfaces, i.e., those defined by the equation $a_0x_0^m + \ldots + a_{n+1}x_{n+1}^m = 0$ ($a_i \in \mathbf{F}_q^\times$); in this case, the ζ-function is of the form $P(u)^{(-1)^{n+1}}/\prod_{j=0}^{n}(1-q^j u)$ with a polynomial $P(u)$ that can be explicitly described in terms of Jacobi sums. Dwork [D14] studied by p-adic analysis the case of hypersurfaces in a projective space, verifying the conjecture for them except for the Riemann hypothesis. Further nontrivial examples were provided by †$K3$ surfaces (Deligne [D2], Pyatetskiĭ-Shapiro, Shafarevich [P1]) and cubic 3-folds (E. Bombieri, H. Swinnerton-Dyer [B5]); in these cases the proof of the Riemann hypothesis was reduced to that of certain Abelian varieties naturally attached to these varieties. It can be said that the Weil conjecture has greatly influenced the development of algebraic geometry, as regards both the foundations and the methods of proof of the conjecture itself; see the expositions by N. Katz [K2] or B. Mazur [M2].

Weil Cohomology, l-Adic Cohomology. The Weil conjecture suggested the possibility of a good cohomology theory for algebraic varieties over a field of arbitrary characteristic. We first formulate the desired properties of a good cohomology (S. Kleiman [K4]). Let $\bar{k}$ be an algebraically closed field and K a field of characteristic 0, which is called the coefficient field. A contravariant functor $V \to H^*(V)$ from the category of complete connected smooth varieties over $\bar{k}$ to the category of augmented $\mathbf{Z}^+$-graded finite-dimensional anticommutative K-algebras (cup product as multiplication) is called a **Weil cohomology** with coefficients in K if it has the following three properties. (1) **Poincaré duality**: If $n = \dim V$, then a canonical isomorphism $H^{2n}(V) \cong K$ exists and the cup product $H^j(V) \times H^{2n-j}(V) \to H^{2n}(V) \cong K$ induces a perfect pairing. (2) **Künneth formula**: For any V_1 and V_2 the mapping $H^*(V_1) \otimes H^*(V_2) \to H^*(V_1 \times V_2)$ defined by $a \otimes b \to \mathrm{Proj}_1^*(a) \cdot \mathrm{Proj}_2^*(b)$ is an isomorphism. (3) Good relation with **algebraic cycles**: Let $C^j(V)$ be the group of algebraic cycles of codimension j on V. There exists a fundamental-class homomorphism FUND: $C^j(V) \to H^{2j}(V)$ for all j, which is functorial in V, compatible with products via Künneth's formula, has compatibility of the intersection with the cup product, and maps 0-cycle$\in C^n(V)$ to its degree as an element of $K \cong H^{2n}(V)$. If a Weil cohomology theory H exists for the V's over $\bar{k}$, we can prove the **Lefschetz fixed-point formula**:

$$((\text{graph of } F) \cdot (\text{diagonal}))_{V \times V} = \sum_{j=0}^{2n} (-1)^j \mathrm{tr}(F^* \mid H^j(V))$$

for a morphism $F: V \to V$.

If $\bar{k} = \mathbf{C}$ (the field of complex numbers), the classical cohomology $V \to H^*(V^{an}, \mathbf{Q})$, where V^{an} denotes the complex manifold associated with V, gives a Weil cohomology. If $\bar{k}$ is an arbitrary algebraically closed field and if l is a prime number different from the characteristic of $\bar{k}$, then the principal results in the theory of the étale cohomology state that the l-adic cohomology $V \to H^*_{\text{ét}}(V, \mathbf{Q}_l)$ is a Weil cohomology with coefficient field $\mathbf{Q}_l$ (the field of l-adic numbers) [A3, D5, G3, M4]. In defining this, Grothendieck introduced a new concept of topology, which is now called Grothendieck topology. In the étale topology of a variety V, for example, any étale covering of a Zariski open subset is regarded as an "open set." With respect to the étale topology, the cohomology group $H^*(V, \mathbf{Z}/n)$ of V with coefficients in $\mathbf{Z}/n$ is defined in the usual manner and is a finite $\mathbf{Z}/n$-module. If l is a prime number as above, $\varprojlim_\nu H^*(V, \mathbf{Z}/l^\nu)$ is a module over $\mathbf{Z}_l = \varprojlim_\nu \mathbf{Z}/l^\nu$ of finite rank, and

$$H^*_{\text{ét}}(V, \mathbf{Q}_l) = (\varprojlim_\nu H^*(V, \mathbf{Z}/l^\nu)) \otimes_{\mathbf{Z}_l} \mathbf{Q}_l$$

defines the l-adic cohomology group, giving rise to a Weil cohomology.

For the characteristic p of k, p-adic étale cohomology does not give Weil cohomology; but the crystalline cohomology (Grothendieck and P. Berthelot [B2, B3]) takes the place of p-adic cohomology and is almost a Weil cohomology: in this theory the fundamental class is defined only for smooth subvarieties.

Now fix a Weil cohomology for $\bar{k} = \bar{\mathbf{F}}_q$, an algebraic closure of a finite field $\mathbf{F}_q$. Given an algebraic variety V over $\mathbf{F}_q$, let $\bar{V} = V \otimes \bar{k}$ denote the base extension of V to $\bar{k}$; then $\mathbf{F}_{q^m}$-rational points of V can be identified with the fixed points of the mth iterate of the Frobenius morphism F of V relative to $\mathbf{F}_q$. Then the Lefschetz fixed-point formula implies the rationality of $Z(u, V)$; more precisely, letting $P_j(u) = \det(1 - uF^* \mid H^j(\bar{V}))$ be the characteristic polynomial of the automorphism F^* of $H^j(\bar{V})$ induced by F, we have

$$Z(u, V) = \prod_{j=0}^{2n} P_j(u)^{(-1)^{j+1}}.$$

The functional equation of the ζ-function then follows from the Poincaré duality. This proves (1), (2), and a part of (3) in the statement of the Weil conjecture. Further, in the case of l-adic cohomology, (4) means that $\deg P_j(u) =$

$\dim_{\mathbf{Q}_l} H^j(\bar{V}, \mathbf{Q}_l)$ is equal to the jth Betti number of a lifting of V to characteristic 0; this follows from the comparison theorem of M. Artin for the l-adic cohomology and the classical cohomology, combined with the invariance of l-adic cohomology under specialization.

Proof of the Riemann Hypothesis. In 1974, Deligne [D4, I] completed the proof of the Weil conjecture for projective nonsingular varieties by proving that, given such a V over $\mathbf{F}_q$, any eigenvalue of F^* on $H^j_{\text{ét}}(\bar{V}, \mathbf{Q}_l)$ is an algebraic integer, all the conjugates of which are of absolute value $q^{j/2}$. (This implies that $P_j(u) = \det(1 - uF^* \mid H^j_{\text{ét}}(\bar{V}, \mathbf{Q}_l))$ is in $\mathbf{Z}[u]$ and is independent of l.) The proof is done by induction on $n = \dim V$; by the general results in l-adic cohomology (the weak Lefschetz theorem on a hyperplane section, the Poincaré duality, and the Künneth formula), the proof is reduced to the assertion that (*) any eigenvalue α of F^* on $H^n_{\text{ét}}(\bar{V}, \mathbf{Q}_l)$ is an algebraic integer such that $|\alpha'| \leqslant q^{(n+1)/2}$ for all conjugates α' of α. The main ingredients in proving (*) are (1) Grothendieck's theory of L-functions, based on the étale cohomology with compact support and with coefficients in a $\mathbf{Q}_l$-sheaf [G2, G3]; (2) the theory of Lefschetz pencils (Deligne and Katz [D7]), and the Kajdan-Margulis theorem on the monodromy of a Lefschetz pencil (J. L. Verdier, *Sém. Bourbaki*, no. 423 (1972)); and (3) Rankin's methods to estimate the coefficients of modular forms, as adapted to the Grothendieck's L-series. By means of these geometric and arithmetic techniques, Deligne achieved the proof of the Riemann hypothesis for projective nonsingular varieties. For the generalization to complete varieties, see Deligne [D4, II].

Applications of the (Verified) Weil Conjecture. (1) The Ramanujan conjecture (→ 32 Automorphic Functions D): The connection of this conjecture and the Weil conjecture for certain fiber varieties over a modular curve was observed by M. Sato and partially verified by Y. Ihara [I1] and then established by Deligne [D3]. The Weil conjecture as proven above implies the truth of the Ramanujan conjecture and its generalization by H. Petersson.

(2) Estimation of trigonometric sums: Let q be the power of a prime number p. Then

$$\left|\sum_{(x_1,\ldots,x_n)\in\mathbf{F}_q^n} \exp\frac{2\pi i}{p}\operatorname{tr}_{\mathbf{F}_q/\mathbf{F}_p}(F(x_1,\ldots,x_n))\right| \leqslant (d-1)^n q^{n/2},$$

where $F(X_1,\ldots,X_n)\in\mathbf{F}_q[X_1,\ldots,X_n]$ is a polynomial of degree d that is not divisible by p, and the homogeneous part of the highest degree of F defines a smooth irreducible hypersurface in $\mathbf{P}^{n-1}$. This is a generalization of the Weil estimation of the Kloosterman sum ([D4, W1 (1948c)]; → 4 Additive Number Theory D).

(3) The **hard Lefschetz theorem**: Let $L \in H^2(V)$ be the class of a hyperplane section of an n-dimensional projective nonsingular variety V over an algebraically closed field. Then the cup product by $L^i: H^{n-i}(V) \to H^{n+i}(V)$ is an isomorphism for all $i \leqslant n$. Deligne [D4, II] proved this for l-adic cohomology, from which N. Katz and W. Messing [K1] deduced its validity in any Weil cohomology or in the crystalline cohomology.

Also some geometric properties of an algebraic variety V are reflected in the properties of $Z(u, V)$. The ζ-function $Z(u, A)$ of an Abelian variety A determines the isogeny class of A [T4]. For any algebraic integer α, every conjugate of which has absolute value $q^{1/2}$, there exists an Abelian variety $A/\mathbf{F}_q$ such that α is a root of $\det(1 - uF^* \mid H^1(A)) = 0$ [H6]. J. Tate [T3] conjectured that the rank of the space cohomology classes of algebraic cycles of codimension r is equal to the order of the pole at $u = 1/q^r$ of $Z(u, V)$. This conjecture is still open but has been verified in certain nontrivial cases, e.g., (1) products of curves and Abelian varieties, $r = 1$ (Tate [T4]), (2) Fermat hypersurfaces of dimension $2r$ with some condition on the degree and the characteristic (Tate [T3], T. Shioda, *Proc. Japan Acad.* 55 (1979)), and (3) elliptic $K3$ surfaces, $r = 1$ (M. Artin and Swinnerton-Dyer, *Inventiones Math.* 20 (1973)).

R. ζ- and L-Functions of Schemes

Let X be a †scheme of finite type over $\mathbf{Z}$, and let $|X|$ denote the set of closed points of X; for each $x \in |X|$, the residue field $k(x)$ is finite, and its cardinality is called the norm $N(x)$ of x. The **ζ-function of a scheme** X is defined by the product $\zeta(s, X) = \prod_{x\in|X|}(1 - N(x)^{-s})^{-1}$. This converges absolutely for $\operatorname{Re} s > \dim X$, and it is conjectured to have an analytic continuation in the entire s-plane (Serre [S7]). It reduces to the Riemann (resp. Dedekind) ζ-function if $X = \operatorname{Spec}(\mathbf{Z})$ (resp. $\operatorname{Spec}(\mathfrak{o})$, $\mathfrak{o}$ being the ring of integers of an algebraic number field), and to the ζ-function $Z(q^{-s}, X)$ (→ Section Q) if X is a variety over a finite field $\mathbf{F}_q$. The case of varieties defined over an algebraic number field is discussed in Section S.

Let G be a finite group of automorphisms of a scheme X, and assume that the quotient $Y = X/G$ exists (e.g., X is quasiprojective). For an element x in $|X|$, let y be its image in $|Y|$, and let $D(x) = \{g \in G \mid g(x) = x\}$, the decomposition group of x over y. The natural mapping $D(x) \to \operatorname{Gal}(k(x)/k(y))$ is surjective, and its kernel $I(x)$ is called the inertia group at x. An element of $D(x)$ is called a Frobenius element at x if its image in $\mathrm{Gal}(k(x)/k(y))$ corresponds to the $N(y)$th-power automorphism of $k(x)$. Now let R be a representation of G with character χ. The **Artin L-function** $L(s, X, \chi)$ is defined by

$$L(s,X,\chi)=\exp\left(\sum_{y\in|Y|}\sum_{n=1}^{\infty}\chi(y^n)N(y)^{-ns}/n\right)$$
$$=\prod_{y\in|Y|}\det(1-R(F_y)N(y)^{-s})^{-1},$$

where $\chi(y^n)$ denotes the mean value of χ on the nth power of Frobenius elements F_x at x (x any point of $|X|$ over y), and similarly $R(F_y)$ denotes the mean value of $R(F_x)$; it converges absolutely for $\mathrm{Re}\, s > \dim X$. Again this is reduced to the usual Artin L-function (→ Section G) if X is the spectrum of the ring of integers of an algebraic number field. The Artin L-functions of a scheme have many formal properties analogous to those of Artin L-functions of a number field (Serre [S7]).

Let us consider the case where X is an algebraic variety over a finite field $\mathbf{F}_q$ and elements of G are automorphisms of X over $\mathbf{F}_2$; in this case, $L(s, X, \chi)$ is a formal power series in $u = q^{-s}$, which is called a congruence Artin L-function. For the case where X is a complete nonsingular algebraic curve and χ is an irreducible character of G different from the trivial one, Weil [W2] proved that $L(s, X, \chi)$ is a polynomial in $u = q^{-s}$; thus the analog of †Artin's conjecture holds here. More generally, for any algebraic variety X over $\mathbf{F}_q$, Grothendieck [G2, G3] proved the rationality of L-functions together with the alternating product expression by polynomials in u, as in the case of ζ-functions, by the methods of l-adic cohomology. Actually, Grothendieck treated a more general type of L-function associated with l-adic sheaves on X, which also play an important role in Deligne's proof of the Riemann hypothesis (→ Section Q).

S. Hasse ζ-Functions

For a nonsingular complete algebraic variety V defined over a finite algebraic number field K, let $V_\mathfrak{p}$ be the reduction of V modulo a prime ideal $\mathfrak{p}$ of K, $K_\mathfrak{p}$ be the residue field of $\mathfrak{p}$, and $Z(u, V_\mathfrak{p})$ be the ζ-function of $V_\mathfrak{p}$ over $K_\mathfrak{p}$. The ζ-function $\zeta(s, V)$ of the complex variable s, determined by the infinite product (excluding the finite number of $\mathfrak{p}$'s where $V_\mathfrak{p}$ is not defined),

$$\zeta(s,V)=\prod_{\mathfrak{p}}{}' Z(N(\mathfrak{p})^{-s},V_\mathfrak{p}),$$

is called the **Hasse ζ-function** of V over the algebraic number field K. For this function, we have **Hasse's conjecture** [W4]: $\zeta(s, V)$ is a meromorphic function over the whole complex plane of s and satisfies the functional equation of ordinary type. Sometimes it is more natural to consider

$$\zeta_j(s,V)=\prod_{\mathfrak{p}}{}' P_j(N(\mathfrak{p})^{-s},V_\mathfrak{p})^{-1}\quad(0\leqslant j\leqslant 2\dim V),$$

where $P_j(u, V_\mathfrak{p})$ is the jth factor of $Z(u, V_\mathfrak{p})$, and we have a similar conjecture for them. For the definition of $\zeta_j(s, V)$ taking into account the factors for bad primes and the precise form of the conjectural functional equation, see Serre [S8]. Note that $\zeta_j(s, V)$ converges absolutely for $\mathrm{Re}\, s > j/2+1$ as a consequence of the Weil conjecture.

Hasse's conjecture remains unsolved for the general case, but has been verified when V is one of the following varieties:

($\mathrm{I_a}$) Algebraic curves defined by the equation $y^e = \gamma x^f + \delta$ and Fermat hypersurfaces (Weil [W6]).

($\mathrm{I_b}$) Elliptic curves with complex multiplication (Deuring [D11]).

($\mathrm{I_c}$) Abelian varieties with complex multiplication (Taniyama [T2], Shimura and Taniyama [S11], Shimura, H. Yoshida).

($\mathrm{I_d}$) Singular $K3$ surfaces, i.e., $K3$ surfaces with 20 Picard numbers (Shafarevich and Pyatetskiĭ-Shapiro [P1], Deligne [D2], T. Shioda and H. Inose [S21]).

($\mathrm{II_a}$) Algebraic curves that are suitable models of the elliptic modular function fields (Eichler [E1], Shimura [S12]).

($\mathrm{II_b}$) Algebraic curves that are suitable models of the automorphic function fields obtained from a quaternion algebra (Shimura [S13, S15]).

($\mathrm{II_c}$) Certain fiber varieties of which the base is a curve of type ($\mathrm{II_a}$) or ($\mathrm{II_b}$) and the fibers are Abelian varieties (Kuga and Shimura [K6], Ihara [I1], Deligne [D3]).

($\mathrm{II_d}$) Certain Shimura varieties of higher dimension (Langlands and others; → [B6]).

In these cases, $\zeta(s, V)$ can be expressed by known functions, i.e., by Hecke L-functions with Grössencharakters of algebraic number fields in cases (I) or by Dirichlet series corresponding to modular forms in cases (II). This fact has an essential meaning for the arithmetic properties of these functions. For example, the extended †Ramanujan conjecture concerning the Hecke operator of the automorphic form reduces to Weil's conjecture on varieties related to those in cases II. Moreover, for ($\mathrm{II_a}$)–($\mathrm{II_c}$) the essential point is the congruence relation $\tilde{T}_p = \Pi + \Pi^*$ (Kronecker, Eichler [E1], Shimura). In particular, for ($\mathrm{II_b}$) this formula is related to the problem of constructing class fields over totally imaginary quadratic extensions of a totally real field F utilizing special values of automorphic functions and class fields over F. Actually, the formula is equivalent to the reciprocity law for class fields (Shimura).

One of the facts that makes the Hasse ζ-function important is that it describes the decomposition law of prime ideals of algebraic number fields when V is an algebraic curve or an Abelian variety (Weil, Shimura [S14], Taniyama [T2], T. Honda [H6]). In that case, its Hasse ζ-function has the following arithmetic meaning.

Let C be a complete, nonsingular algebraic curve defined over an algebraic number field K, and let J be the Jacobian variety of C defined over K. For a prime number l, fix an l-adic coordinate system Σ_l on J, and let $K(J, l^\infty)$ be the extension field of K obtained by adjoining to K all the coordinates of the l^νth division points ($\nu = 1, 2, \ldots$) of J. Then $K(J, l^\infty)/K$ is an infinite Galois extension of K. The corresponding Galois group $\mathfrak{G}(J, l^\infty)$ has the l-adic representation $\sigma \rightarrow M_l^*(\sigma)$ by the l-adic coordinates Σ_l. Almost all prime ideals $\mathfrak{p}$ of K are unramified in $K(J, l^\infty)/K$. Thus when we take an arbitrary prime factor $\mathfrak{P}$ of $\mathfrak{p}$ in $K(J, l^\infty)$, the Frobenius substitution of $\mathfrak{P}$,

$$\sigma_{\mathfrak{P}} = \left[\frac{K(J, l^\infty)/K}{\mathfrak{P}}\right],$$

is uniquely determined. Furthermore, the characteristic polynomial $\det(1 - M_l^*(\sigma_{\mathfrak{P}})u)$ is determined only by $\mathfrak{p}$ and does not depend on the choice of the prime factor $\mathfrak{P}$; we denote this polynomial by $P_{\mathfrak{p}}(u, C)$. In this case, for almost all $\mathfrak{p}$, $P_{\mathfrak{p}}(u, C)$ is a polynomial with rational integral coefficients independent of l; namely, the numerator of the ζ-function of the reduction of $C \bmod p$. Thus

$$\zeta_1(s, C) = \prod_{\mathfrak{p}}{}' P_{\mathfrak{p}}(N(\mathfrak{p})^{-s}, C)^{-1}$$

$$\sim \prod_{\mathfrak{p}}{}' \det(1 - M_l^*(\sigma_{\mathfrak{P}})N(\mathfrak{p})^{-s})^{-1}.$$

Here the product $\prod' \det(1 - M_l^*(\sigma_{\mathfrak{P}})N(\mathfrak{p})^{-s})^{-1}$ has the same expression as the Artin L-function if we ignore the fact that M_l^* is the l-adic representation and $K(J, l^\infty)$ is the infinite extension. Thus if we can describe $\zeta(s, C)$ explicitly, then the decomposition process of the prime ideal for intermediate fields between $K(J, l^\infty)$ and K can be made fairly clear. In fact, this is the case for examples ($\mathrm{I_a}$)–($\mathrm{I_c}$) and ($\mathrm{II_a}$)–($\mathrm{II_c}$), from which the relations between the arithmetic of the field of division points $K(J, l^\infty)/K$ and the eigenvalues of the Hecke operator have been obtained. Thus for curves and Abelian varieties, $\zeta(s, V)$ is related to the arithmetic of some number fields; but it is not known whether similar arithmetical relations exist for other kinds of varieties except in a few cases.

Tate's Conjecture. For a projective nonsingular variety V over a finite algebraic number field K, let $\mathfrak{A}^r(\bar{V})$ denote the group of algebraic cycles of codimension r on $\bar{V} = V \otimes_K \mathbf{C}$ modulo homological equivalence and let $\mathfrak{A}^r(V)$ be the subgroup of $\mathfrak{A}^r(\bar{V})$ generated by algebraic cycles rational over K. Then Tate [T3] conjectured that the rank of $\mathfrak{A}^r(V)$ is equal to the order of the pole of $\zeta_{2r}(s, V)$ at $s = r + 1$. This conjecture is closely connected with **Hodge's conjecture** that the space of rational cohomology classes of type (r, r) on $\bar{V}$ is spanned by $\mathfrak{A}^r(\bar{V})$; in fact, the equivalence of these conjectures is known for Abelian varieties of †CM type (H. Pohlmann, *Ann. Math.*, 88 (1968)) and for Fermat hypersurfaces of dimension $2r$ (Tate [T3], Weil [W6]). Thus, when $r = 1$, Tate's conjecture for these varieties holds by Lefschetz's theorem, and when $r > 1$, it holds in certain cases where the Hodge conjecture is verified (Shioda, *Math. Ann.*, 245 (1979); Z. Ran, *Compositio Math.*, 42 (1981)). Further examples are given by $K3$ surfaces with large Picard numbers (Shioda and Inose [S21]; T. Oda, *Proc. Japan Acad.*, 56 (1980)).

***L*-Functions of Elliptic Curves.** Let E be an elliptic curve (with a rational point) over the rational number field $\mathbf{Q}$, and let N be its conductor; a prime number p divides N if and only if E has bad reduction $\bmod p$ (Tate [T5]). The L-function of E over $\mathbf{Q}$ is defined as follows:

$$L(s, E) = \prod_{p \mid N} (1 - \varepsilon_p p^{-s})^{-1} \prod_{p \nmid N} (1 - a_p p^{-s} + p^{1-2s})^{-1},$$

where $\varepsilon_p = 0$ or ± 1 and $1 - a_p u + p u^2 = P_1(u, E \bmod p)$. There are many interesting results and conjectures concerning $L(s, E)$ [T5]:

(1) Functional equation. Let

$$\xi(s, E) = N^{s/2}(2\pi)^{-s}\Gamma(s)L(s, E).$$

Then it is conjectured that $\xi(s, E)$ is holomorphic in the entire s-plane and satisfies the functional equation $\xi(s, E) = \pm\xi(2 - s, E)$. This is true if E has complex multiplication (Deuring) or E is a certain modular curve (Eichler, Shimura).

(2) **Taniyama-Weil conjecture.** Weil [W1 (1967a)] conjectured that, if $L(s, E) = \sum_{n=1}^\infty a_n n^{-s}$, then $f(\tau) = \sum_{n=1}^\infty a_n e^{2\pi i n \tau}$ is a cusp form of weight 2 for the congruence subgroup $\Gamma_0(N)$ which is an eigenfunction for Hecke operators; moreover E is isogenous to a factor of the Jacobian variety of the modular curve for $\Gamma_0(N)$ in such a way that $f(\tau)d\tau$ corresponds to the differential of the first kind on E. If this conjecture is true, then the statements in (1) follow.

(3) **Birch–Swinnerton-Dyer conjecture**. Assuming analytic continuation of $L(s, E)$, B. Birch and H. Swinnerton-Dyer [B4] conjectured that the order of the zero of $L(s, E)$ at $s = 1$ is equal to the rank r of the group $E(\mathbf{Q})$ of rational points of E which is finitely generated by the Mordell-Weil theorem. They verified this for many examples, especially for curves of the type $y^2 = x^3 - ax$. J. Coates and A. Wiles (*Inventiones Math.*, 39 (1977)) proved that if E has complex multiplication and if $r > 0$ then $L(s, E)$ vanishes at $s = 1$. This conjecture has a refinement which extends also to Abelian varieties over a global field (Tate, *Sém. Bourbaki*, no. 306 (1966)).

(4) **Sato's conjecture**. Let

$$1 - a_p u + p u^2 = (1 - \pi_p u)(1 - \bar{\pi}_p u),$$

with $\pi_p = \sqrt{p}\, e^{i\theta_p}$ $(0 < \theta_p < \pi)$. When E has complex multiplication, the distribution of θ_p for half of p is uniform in the interval $[0, \pi]$, and θ_p is $\pi/2$ for the remaining half of p. Suppose that E does not have complex multiplication. Then Sato conjectured that

$$\lim_{x \to \infty} \frac{\text{(the number of prime numbers } p \text{ less than } x \text{ such that } \theta_p \in [\alpha, \beta])}{\text{(the number of prime numbers less than } x)}$$

$$= \frac{2}{\pi} \int_\alpha^\beta \sin^2 \theta \, d\theta \quad (0 < \alpha < \beta < \pi)$$

(Tate [T3]).

H. Yoshida [Y1] posed an analog of Sato's conjecture for elliptic curves defined over function fields with finite constant fields and proved it in certain cases.

(5) Formal groups. Letting $L(s, E) = \sum a_n n^{-s}$ as before, set $f(x) = \sum_{n=1}^{\infty} a_n x^n / n$. Honda [H6] showed that $f^{-1}(f(x) + f(y))$ is a †formal (Lie) group with coefficients in $\mathbf{Z}$ and that this group is isomorphic over $\mathbf{Z}$ to a formal group obtained by power series expansion of the group law of E with respect to suitable †local uniformizing coordinates at the origin. Such an interpretation of the ζ-function also applies to other cases in which ζ-functions of †group varieties may be characterized as Dirichlet series whose coefficients give a normal form of the group law; e.g., the case of algebraic tori (T. Ibukiyama, *J. Fac. Sci. Univ. Tokyo*, (IA) 21 (1974)).

T. Selberg ζ-Functions and ζ-Functions Associated with Discontinuous Groups

Let $\Gamma \subset SL(2, \mathbf{R})$ be a †Fuchsian group operating on the complex upper half-plane $H = \{z = x + iy \mid y > 0\}$. When the two eigenvalues of an element $\gamma \in \Gamma$ are distinct real numbers ξ_1, ξ_2 $(\xi_1 \xi_2 = 1, \xi_1 < \xi_2)$, we call γ †hyperbolic. Then the number ξ_2^2 is denoted by $N(\gamma)$ and is called the norm of γ. When γ is hyperbolic, γ^n $(n = 1, 2, 3, \dots)$ is also hyperbolic. When $\pm\gamma$ is not a positive power of other hyperbolic elements, γ is called a primitive hyperbolic element. The elements conjugate to primitive hyperbolic elements are also primitive hyperbolic elements and have the same norm as γ. Let P_1, $P_2, \dots$ be the conjugacy classes of primitive hyperbolic elements of Γ, and let $\gamma_i \in P_i$ be their representatives. Suppose that a matrix representation $\gamma \to M(\gamma)$ of Γ is given. Then the analytic function given by

$$Z_\Gamma(s, M) = \prod_i \prod_{n=0}^{\infty} \det(I - M(\gamma_i) N(\gamma_i)^{-s-n})$$

is called the **Selberg ζ-function** (Selberg [S5]). When $\Gamma \backslash H$ is compact and Γ is torsion-free, then $Z_\Gamma(s, M)$ has the following properties.

(1) It can be analytically continued to the whole complex plane of s and gives an †integral function of genus at most 2.

(2) It has zeros of order $(2n+1)(2g-2)v$ at $-n$ $(n = 0, 1, 2, 3, \dots)$. Here g is the genus of the Riemann surface $\Gamma \backslash H$ and v is the degree of the representation M. All other zeros lie on the line $\operatorname{Re} s = 1/2$, except for a finite number that lie on the interval $(0, 1)$ of the real axis.

(3) It satisfies the functional equation

$$Z_\Gamma(1 - s, M) = Z_\Gamma(s, M) \exp\left(-vA(\Gamma \backslash H) \times \int_0^{s-1/2} v \tan(\pi v)\, dv \right),$$

where

$$A(\Gamma \backslash H) = \iint_{\Gamma \backslash H} \frac{dx\, dy}{y^2} = 2\pi(2g - 2), \qquad x + iy \in H.$$

Property (2) shows that the Riemann hypothesis is almost valid for $Z_\Gamma(s, M)$. The proof is based on the following fact concerning the eigenvalue problem for the variety $\Gamma \backslash H$: The eigenvalue λ of the equation

$$y^2(\partial^2/\partial x^2 + \partial^2/\partial y^2)u + \lambda u = 0, \quad u \in L^2(\Gamma \backslash H)$$

cannot be a negative number.

Using this function, T. Yamada (1965) investigated the unit distribution of real quadratic fields.

Selberg ζ-functions are defined similarly when $\Gamma \backslash G$ has finite volume but is noncompact. In this case, however, the decomposition of $L_2(\Gamma \backslash G)$ into irreducible representation spaces has a continuous spectrum; hence the properties of the Selberg ζ-function of Γ are quite different from the case when $\Gamma \backslash G$ is compact. Selberg defined the **generalized Eisenstein series** to give the eigenfunctions of this continuous spectrum explicitly. When $\Gamma =$

$SL(2, \mathbf{Z})$, the series is given by $\sum_{(c,d)=1} \frac{y^s}{|c\tau+d|^{2s}}$.

This type of generalized Eisenstein series is also defined for the general semisimple algebraic group G and its arithmetic subgroup. It has been studied by Selberg, Godement, Gel'fand, Harish-Chandra, Langlands, D. Zagier, and others.

U. Ihara ζ-Functions

Let k_p be a p-adic field, $\mathfrak{o}_p$ the ring of integers in k_p, and $G = PSL_2(\mathbf{R}) \times PSL_2(k_p)$. Suppose that Γ is a subgroup of G such that (1) Γ is discrete, (2) $\Gamma \backslash G$ is compact, (3) Γ has no torsion element except the identity, (4) $\Gamma_{\mathbf{R}}$ (the projection of Γ in $PSL_2(\mathbf{R})$) is dense in $PSL_2(\mathbf{R})$, and (5) Γ_p (the projection of Γ in $PSL_2(k_p)$) is dense in $PSL_2(k_p)$. Then $\Gamma \cong \Gamma_{\mathbf{R}} \cong \Gamma_p$. Let $X = \{x + iy \mid y > 0\}$ be the upper half-plane, and let Γ act on X via $\Gamma_{\mathbf{R}}$. The action of Γ on X is not discontinuous, but the subgroup $\Gamma_0 = \{\gamma \in \Gamma \mid \text{projection of } \gamma \text{ to } \Gamma_p \in PSL_2(\mathfrak{o}_p)\}$ operates on X properly discontinuously. For each $z \in X$, define $\Gamma_z = \{\gamma \in \Gamma \mid \gamma(z) = z\}$. Then Γ_z is isomorphic to $\mathbf{Z}$ or $\{1\}$. Let $\tilde{\mathbf{P}}(\Gamma) = \{z \in X \mid \Gamma_z \cong \mathbf{Z}\}$. The group Γ acts on $\tilde{\mathbf{P}}(\Gamma)$, since Γ_z and $\Gamma_{\gamma z}$ are conjugate in Γ. Let $\mathbf{P}(\Gamma) = \tilde{\mathbf{P}}(\Gamma)/\Gamma$. Suppose that $P \in \mathbf{P}(\Gamma)$ is represented by $z \in X$. Choose a generator γ of Γ_z and project γ to Γ_p. Then γ is equivalent to a diagonal matrix $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$ with $\lambda \in k_p$. We denote the valuation of k_p by ord_p and consider $|\operatorname{ord}_p(\lambda)|$. This value depends only on P and we denote it by $\deg(P)$. The **Ihara ζ-function** of Γ is defined by

$$Z_\Gamma(u) = \prod_{P \in \mathbf{P}(\Gamma)} (1 - u^{\deg(P)})^{-1}.$$

Ihara proved that

$$Z_\Gamma(u) = \frac{\prod_{i=1}^{g} (1 - \pi_i u)(1 - \pi_i' u)}{(1-u)(1-q^2 u)} (1-u)^H,$$

where q is the number of elements in the residue class field of $\mathfrak{p}$, and g is the genus of the Riemann surface $\Gamma_0 \backslash X$ and $H = (g-1)q(q-1)$. Similar results hold even if Γ has torsion elements and the quotient $\Gamma \backslash G$ is only assumed to have finite volume.

Aside from the factor $(1-u)^H$, this looks like Weil's formula for the congruence ζ-function of an algebraic curve defined over $\mathbf{F}_{q^2}$. Ihara conjectured that the first factor of $Z_\Gamma(u)$ is always the congruence ζ-function of some algebraic curve over $\mathbf{F}_{q^2}$, and furthermore that Γ could be regarded as the fundamental group of a certain Galois covering of this curve which describes the decomposition law of prime divisors in this covering [12, 13, 14]. He verified the conjecture in the case $\Gamma = PGL_2(\mathbf{Z}[1/p])$ by using the †moduli of elliptic curves. Related results have been obtained by Shimura, Ihara, Y. Morita, and others.

V. ζ-Functions Associated with Prehomogeneous Vector Spaces

M. Sato posed a notion of prehomogeneous vector spaces and defined ζ-functions associated with them. Sato's program has been carried on by himself and T. Shintani [S2, S3, S17, S18]. Let G be a linear algebraic group, V a finite-dimensional linear space of dimension n, and ρ a rational representation $G \to GL(V)$, where G, V, and ρ are defined over $\mathbf{Q}$. The triple (G, ρ, V) is called a **prehomogeneous vector space** if there exists a proper algebraic subset S of $V_{\mathbf{C}}$ such that $V_{\mathbf{C}} - S$ is a single $G_{\mathbf{C}}$-orbit. The algebraic set S is called the set of singular points of V. We also assume that G is reductive and S is an irreducible hypersurface of V. Let V^* be the dual vector space of V, and ρ^* the dual (contragredient) representation of G. Then (G, ρ^*, V^*) is again a prehomogeneous vector space, and we denote its set of singular points by S^*. There are homogeneous polynomials P and Q of the same degree d on V and V^*, respectively, such that $S = \{x \in X \mid P(x) = 0\}$ and $S^* = \{x^* \in V^* \mid Q(x^*) = 0\}$. P and Q are relative invariants of G, i.e., $P(\rho(g)x) = \chi(g)P(x)$ and $Q(\rho^*(g)x^*) = \chi(g)^{-1} Q(x^*)$ (for $g \in G$, $x \in V$, and $x^* \in V^*$) hold with a rational character χ of G. Put $G^1 = \ker \chi = \{g \in G \mid \chi(g) = 1\}$. Denote by $G_{\mathbf{R}}^+$ the connected component of 1 of the Lie group $G_{\mathbf{R}}$. Let $V_{\mathbf{R}} - S = V_1 \cup \ldots \cup V_l$, $V_{\mathbf{R}}^* - S^* = V_1^* \cup \ldots \cup V_l^*$ be the decompositions of $V_{\mathbf{R}} - S$ and $V_{\mathbf{R}}^* - S^*$ into their topologically connected components. Then V_i and V_j^* are $G_{\mathbf{R}}^+$-orbits. We further assume that $V_{\mathbf{R}} \cap S$ decomposes into the union of a finite number of $G_{\mathbf{R}}^1$-orbits. Set $\Gamma = G_{\mathbf{R}}^+ \cap G_{\mathbf{Z}}^1$, and take Γ-invariant lattices L and L^* in $V_{\mathbf{Q}}$ and $V_{\mathbf{Q}}^*$, respectively. Consider the following functions in s:

$$\Phi_i(f, s) = \int_{V_i} f(x) |P(x)|^s dx,$$

$$\Phi_j^*(f, s) = \int_{V_j^*} f^*(x^*) |Q(x^*)|^s dx^*,$$

and

$$Z_i(f, L, s) = \int_{G_{\mathbf{R}}^+/\Gamma} \chi(g)^s \sum_{x \in L \cap V_i} f(\rho(g)x)\, dg,$$

$$Z_j^*(f^*, L^*, s)$$

$$= \int_{G_{\mathbf{R}}^+/\Gamma} \chi(g)^{-s} \sum_{x^* \in L^* \cap V_j^*} f^*(\rho^*(g)x^*)\, dg,$$

where f and f^* are †rapidly decreasing functions on $V_{\mathbf{R}}$ and $V_{\mathbf{R}}^*$, respectively, dx and dx^* are Haar measures of $V_{\mathbf{R}}$ and $V_{\mathbf{R}}^*$, respectively, and dg is a Haar measure of G. Then the ratios

$$\frac{Z_i(f, L, s)}{\Phi_i(f, s-n/d)}=\xi_i(s, L),$$

$$\frac{Z_j^*(f^*, L^*, s)}{\Phi_j^*(f^*, s-n/d)}=\xi_j^*(s, L^*)$$

are independent of the choice of f and f^* and are Dirichlet series in s. These Dirichlet series $\xi_i(s, L)$ and $\xi_j^*(s, L^*)$ are called **ξ-functions associated with the prehomogeneous space**. Considering Fourier transforms of $|P(x)|^s$ and $|Q(x^*)|^s$, we obtain functional equations for ξ_i and ξ_j^* under some additional (but mild) conditions on (G, ρ, V) as follows. The Dirichlet series ξ_i and ξ_j^* are analytically continuable to meromorphic functions on the whole s-plane, and they satisfy

$$v(L^*)\xi_j^*(n/d-s, L^*)$$
$$=\gamma(s-n/d)(2\pi)^{-ds}|b_0|^s\exp(\pi d\sqrt{-1}\,s/2)$$
$$\times\sum_{j=1}^{l} u_{ij}(s)\xi_i(s, L),$$

with a Γ-factor $\gamma(s)=\prod_{i=1}^{d}\Gamma(s-c_i+1)$. Here $u_{ij}(s)$ $(1\leqslant i,j\leqslant l)$ are polynomials in $\exp(-\pi\sqrt{-1}\,s)$ with degree $\leqslant d$, and b_0 and c_i are constants depending only on (G, ρ, V).

Epstein's ζ-functions and Siegel's Dirichlet series associated with indefinite quadratic forms are examples of the above-defined ζ-functions. Shintani defined such ζ-functions related to integral binary cubic forms and obtained asymptotic formulas concerning the class numbers of irreducible integral binary cubic forms with discriminant n, which are improvements on the results of Davenport [S17].

Recently M. Sato studied ζ-functions of prehomogeneous vector spaces without assuming the conditions that G is reductive and S is irreducible. In this case, ζ-functions of several complex variables are obtained. For examples and classification of prehomogeneous vector spaces → [S4].

References

[A1] E. Artin, Collected papers, S. Lang and J. T. Tate (eds.), Addison-Wesley, 1965.
[A2] E. Artin, Zur Theorie der L-Reihen mit allgemeinen Gruppencharakteren, Abh. Math. Sem. Univ. Hamburg, 8 (1930), 292–306 [A1, 165–179].
[A3] M. Artin, A. Grothendieck, and J. L. Verdier, Théorie des topos et cohomologie étale des schémas (SGA 4), Lecture notes in math. 269, 270, 305, Springer, 1972–1973.
[A4] J. Ax, On the units of an algebraic number fields, Illinois J. Math., 9 (1965), 584–589.
[B1] P. Bayer and J. Neukirch, On values of zeta functions and l-adic Euler characteristics, Inventiones Math., 50 (1978), 35–64.
[B2] P. Berthelot, Cohomogie cristalline des schéma de caracteristique $p>0$, Lecture notes in math. 407, Springer, 1974.
[B3] P. Berthelot and A. Ogus, Notes on crystalline cohomology, Math. notes 21, Princeton Univ. Press, 1978.
[B4] B. Birch and H. Swinnerton-Dyer, Notes on elliptic curves II, J. Reine Angew. Math., 218 (1965), 79–108.
[B5] E. Bombieri and H. P. F. Swinnerton-Dyer, On the local zeta function of a cubic threefold, Ann. Scuola Norm. Sup. Pisa, (3) 21 (1967), 1–29.
[B6] A. Borel and W. Casselman (eds.), Automorphic forms, representations, and L-functions, Amer. Math. Soc. Proc. Symp. Pure Math., 33 (pts. 1 and 2) (1979).
[B7] A. Brumer, On the units of algebraic number fields, Mathematika, 14 (1967), 121–124.
[C1] K. Chandrasekharan, Lectures on the Riemann zeta-function, Tata Inst., 1953.
[C2] S. Chowla and A. Selberg, On Epstein's zeta function I, Proc. Nat. Acad. Sci. US, 35 (1949), 371–374.
[C3] J. Coates, P-adic L-functions and Iwasawa's theory, Algebraic Number Fields (L-functions and Galois Properties), A. Fröhlich (ed.), Academic Press, 1977.
[D1] R. Dedekind, Über die Theorie der ganzen algebraischen Zahlen, Dirichlet's Vorlesungen über Zahlentheorie, Supplement XI, fourth edition, 1894, secs. 184, 185, 186 (Gesammelte mathematische Werke III, Braunschweig, 1932, 297–314; Chelsea, 1968).
[D2] P. Deligne, La conjecture de Weil pour les surfaces $K3$, Inventiones Math., 75 (1972), 206–226.
[D3] P. Deligne, Formes modulaires et représentations l-adiques, Sém. Bourbaki, exp. 355, Lecture notes in math. 349, Springer, 1973.
[D4] P. Deligne, La conjecture de Weil, I, II, Publ. Math. Inst. HES, 43 (1974), 273–307; 52 (1980), 137–252.
[D5] P. Deligne, Cohomologie étale (SGA $4\frac{1}{2}$), Lecture notes in math. 569, Springer, 1977.
[D6] P. Deligne, Les constantes des équations fonctionnelles des fonctions L, Lecture notes in math. 349, Springer, 1973.
[D7] P. Deligne and N. Katz, Groupes de monodromie en géométrie algébrique (SGA

7II), Lecture notes in math. 340, Springer, 1973.
[D8] P. Deligne and K. Ribet, Values of Abelian L-functions at negative integers over totally real fields, Inventiones Math., 59 (1980), 227–286.
[D9] P. Deligne and J-P. Serre, Formes modulaires de poids 1, Ann. Sci. Ecole Norm. Sup. (4), 7 (1974), 507–530.
[D10] M. Deuring, Algebren, Erg. Math., Springer, second edition, 1968.
[D11] M. Deuring, Die Zetafunktionen einer algebraischen Kurve vom Geschlechte Eins I, II, III, IV, Nachr. Akad. Wiss. Göttingen (1953), 85–94; (1955), 13–42; (1956), 37–76; (1957), 55–80.
[D12] K. Doi and H. Naganuma, On the functional equation of certain Dirichlet series, Inventiones Math., 9 (1969–1970), 1–14.
[D13] B. Dwork, On the rationality of the zeta-function of an algebraic variety, Amer. J. Math., 82 (1960), 631–648.
[D14] B. Dwork, On the zeta function of a hypersurface I, II, III, Publ. Math. Inst. HES, 1962, 5–68; Ann. Math., (2) 80 (1964), 227–299; (2) 83 (1966), 457–519.
[E1] M. Eichler, Einführung in die Theorie der algebraischen Zahlen und Funktionen, Birkhäuser, 1963; English translation, Introduction to the theory of algebraic numbers and functions, Academic Press, 1966.
[F1] B. Ferrero and L. C. Washington, The Iwasawa invariant μ_p vanishes for Abelian number fields, Ann. Math., (2) 109 (1979), 377–395.
[G1] R. Godement, Les fonctions ζ des algèbres simples I, II, Sém. Bourbaki, exp. 171, 176, 1958–1959.
[G2] A. Grothendieck, Formule de Lefschetz et rationalité des fonction L, Sém. Bourbaki, exp. 379, 1964–1965 (Benjamin, 1966).
[G3] A. Grothendieck, Cohomologie l-adique et fonctions L (SGA 5), Lecture notes in math. 589, Springer, 1977.
[H1] H. Hasse, Zur Theorie der abstrakten elliptischen Funktionenkörper I, II, III, J. Reine Angew. Math., 175 (1936), 55–62, 69–88, 193–208.
[H2] E. Hecke, Mathematische Werke, Vandenhoeck & Ruprecht, 1959.
[H3] E. Hecke, Eine neue Art von Zetafunctionen und ihre Beziehung zur Verteilung der Primzahlen I, II, Math. Z., 1 (1918), 357–376; 6 (1920), 11–51 [H2, 215–234, 249–289].
[H4] E. Hecke, Über die Bestimmung Dirichletscher Reihen durch ihre Funktionalgleichung, Math. Ann., 112 (1936), 664–699 (H2, 591–626].
[H5] E. Hecke, Über Modulfunktionen und die Dirichletschen Reihen mit Eulerscher Produktentwicklung I, II, Math. Ann., 114 (1937), 1–28; 316–351 [H2, 644–671, 672–707].
[H6] T. Honda, Formal groups and zeta functions, Osaka J. Math., 5 (1968), 199–213.
[H7] T. Honda, Isogeny classes of Abelian varieties over finite fields, J. Math. Soc. Japan, 20 (1968), 83–95.
[I1] Y. Ihara, Hecke polynomials as congruence ζ functions in elliptic modular case (To validate Sato's identity), Ann. Math., (2) 85 (1967), 267–295.
[I2] Y. Ihara, On congruence monodromy problems, I, II, Lecture notes, Univ. of Tokyo, 1968–1969.
[I3] Y. Ihara, Non-Abelian class fields over function fields in special cases, Actes Congr. Intern. Math., 1970, Nice, Gauthier-Villars, p. 381–389.
[I4] Y. Ihara, Congruence relations and Shimura curves. Amer. Math. Soc. Proc. Symp. Pure Math., 33 (1979), pt. 2, 291–311.
[I5] K. Iwasawa, On Γ-extensions of algebraic number fields, Bull. Amer. Math. Soc., 65 (1959), 183–226.
[I6] K. Iwasawa, On p-adic L-functions, Ann. Math., (2) 89 (1969), 198–205.
[I7] K. Iwasawa, Lectures on p-adic L-functions, Ann. math. studies 74, Princeton Univ. Press, 1972.
[J1] H. Jacquet and R. P. Langlands, Automorphic forms on $GL(2)$, Lecture notes in math. 114, Springer, 1970.
[J2] H. Jacquet, Automorphic forms on $GL(2)$, II, Lecture notes in math. 278, Springer, 1972.
[K1] N. Katz and W. Messing, Some consequences of the Riemann hypothesis for varieties over finite fields, Inventiones Math., 23 (1974), 73–77.
[K2] N. Katz, An overview of Deligne's proof of the Riemann hypothesis for varieties over finite fields, Amer. Math. Soc. Proc. Symp. Pure Math., 28 (1976), 275–305.
[K3] N. Katz, P-adic L-functions for CM-fields, Inventiones Math., 49 (1978), 199–297.
[K4] S. Kleiman, Algebraic cycles and the Weil conjectures, Dix exposés sur la cohomologies des schémas, North-Holland, 1968, 359–386.
[K5] T. Kubota and H. W. Leopoldt, Eine p-adischer Theorie der Zetawerte, J. Reine Angew. Math., 214/215 (1964), 328–339.
[K6] M. Kuga and G. Shimura, On the zeta function of a fiber variety whose fibers are Abelian varieties, Ann. Math., (2) 82 (1965), 478–539.
[L1] E. G. H. Landau, Handbuch der Lehre von der Verteilung der Primzahlen I, II, Teubner, 1909 (Chelsea, 1953).
[L2] S. Lang, Sur les series L d'une variété algébrique, Bull. Soc. Math. France, 84 (1956), 385–407.

[L3] S. Lang, Algebraic number theory, Addison-Wesley, 1970.
[L4] R. P. Langlands, Problems in the theory of automorphic forms, Lecture notes in math. 170, Springer, 1970, 18–86.
[L5] R. P. Langlands, Base change for $GL(2)$, Ann. math. studies 96, Princeton Univ. Press, 1980.
[L6] H. W. Leopoldt, Eine p-adische Theorie der Zetawerte II, J. Reine Angew. Math., 274/275 (1975), 224–239.
[L7] S. Lubkin, A p-adic proof of the Weil conjecture, Ann. Math., (2) 87 (1968), 105–194; (2) (1968), 195–225.
[M1] Yu. I. Manin, Periods of parabolic forms and p-adic Hecke series, Math. USSR-Sb., 21 (1973), 371–393.
[M2] B. Mazur, Eigenvalues of Frobenius acting on algebraic varieties over finite fields, Amer. Math. Soc. Proc. Symp. Pure Math., 29 (1975).
[M3] B. Mazur and H. Swinnerton-Dyer, On the p-adic L-series of an elliptic curve, Inventiones Math., 25 (1974), 1–61.
[M4] J. S. Milne, Etale cohomology, Princeton series in mathematics 33, Princeton Univ. Press, 1980.
[P1] I. I. Pyatetskiĭ-Shapiro and I. R. Shafarevich, The arithmetic of $K3$ surfaces, Proc. Steklov Inst. Math., 132 (1973), 45–57.
[P2] K. Prachar, Primzahlverteilung, Springer, 1957.
[R1] B. Riemann, Über die Anzahl der Primzahlen unter einer gegebenen Grösse, Gesammelte mathematische Werke, 1859, 145–153 (Dover, 1953).
[S1] H. Saito, Automorphic forms and algebraic extensions of number fields, Lectures in math. 8, Kinokuniya, Tokyo, 1975.
[S2] M. Sato, Theory of prehomogeneous vector spaces (notes by T. Shintani) (in Japanese), Sûgaku no Ayumi, 15 (1970), 85–157.
[S3] M. Sato and T. Shintani, On zeta functions associated with prehomogeneous vector spaces, Ann. Math., (2) 100 (1974), 131–170.
[S4] M. Sato and T. Kimura, A classification of irreducible prehomogeneous vector spaces and their invariants, Nagoya Math. J., 65 (1977), 1–155.
[S5] A. Selberg, Harmonic analysis and discontinuous groups on weakly symmetric Riemannian spaces with applications to Dirichlet series, J. Indian Math. Soc., 20 (1956), 47–87.
[S6] A. Selberg, On the zeros of Riemann's zeta-function, Skr. Norske Vid. Akad. Oslo, no. 10 (1942), 1–59.
[S7] J.-P. Serre, Zeta and L functions, Arithmetical Algebraic Geometry, Harper & Row, 1965, 82–92.
[S8] J.-P. Serre, Facteurs locaux des fonctions zêta des variétés algébriques (Définitions et conjectures), Sém. Delange-Pisot-Poitou, 11 (1969/70), no. 19.
[S9] J.-P. Serre, Formes modulaires et fonctions zêta p-adiques, Lecture notes in math. 317, Springer, 1973, 319–338.
[S10] H. Shimizu, On zeta functions of quaternion algebras, Ann. Math., (2) 81 (1965), 166–193.
[S11] G. Shimura and Y. Taniyama, Complex multiplication of Abelian varieties and its applications to number theory, Publ. Math. Soc. Japan, no. 6, 1961.
[S12] G. Shimura, Introduction to the arithmetic theory of automorphic functions, Publ. Math. Soc. Japan, no. 11, 1971.
[S13] G. Shimura, On the zeta-functions of the algebraic curves uniformized by certain automorphic functions, J. Math. Soc. Japan, 13 (1961), 275–331.
[S14] G. Shimura, A reciprocity law in nonsolvable extensions, J. Reine Angew. Math., 221 (1966), 209–220.
[S15] G. Shimura, Construction of class fields and zeta functions of algebraic curves, Ann. Math., (2) 85 (1967), 58–159.
[S16] G. Shimura, On canonical models of arithmetic quotients of bounded symmetric domains, I, II, Ann. Math., (2) 91 (1970) 144–222; 92 (1970), 528–549.
[S17] T. Shintani, On Dirichlet series whose coefficients are class numbers of integral binary cubic forms, J. Math. Soc. Japan, 24 (1972), 132–188.
[S18] T. Shintani, On zeta functions associated with the vector space of quadratic forms, J. Fac. Sci. Univ. Tokyo, (IA) 22 (1975), 25–65.
[S19] T. Shintani, On a Kronecker limit formula for real quadratic fields, J. Fac. Sci. Univ. Tokyo, (IA) 24 (1977), 167–199.
[S20] T. Shintani, On liftings of holomorphic automorphic forms, Amer. Math. Soc. Proc. Symp. Pure Math., 33 (1979), pt. 2, 97–110.
[S21] T. Shioda and H. Inose, On singular $K3$ surfaces, Complex Analysis and Algebraic Geometry, Iwanami and Cambridge Univ. Press, 1977, 119–136.
[S22] C. L. Siegel, Gesammelte Abhandlungen I, II, III, IV, Springer, 1966, 1979.
[S23] C. L. Siegel, Über die Classenzahl quadratischer Zahlkörper, Acta Arith., 1 (1935), 83–86 [S22, I. 406–409].
[S24] C. L. Siegel, Über die Zetafunktionen indefiniter quadratischer Formen I, II, Math. Z., 43 (1938), 682–708; 44 (1939), 398–426 [S22, II, 41–67, 68–96].
[S25] H. M. Stark, L-functions at $s=1$, II, III, Advances in Math., 17 (1975), 60–92; 22 (1976), 64–84.
[T1] T. Tamagawa, On the ζ-functions of a

division algebra, Ann. Math., (2) 77 (1963), 387–405.
[T2] Y. Taniyama, *L*-functions of number fields and zeta functions of Abelian varieties, J. Math. Soc. Japan, 9 (1957), 330–366.
[T3] J. Tate, Algebraic cycles and poles of zeta functions, Arithmetical Algebraic Geometry, Harper & Row, 1965, 93–110.
[T4] J. Tate, Endomorphisms of Abelian varieties over finite fields, Inventiones Math., 2 (1966), 134–144.
[T5] J. Tate, The arithmetic of elliptic curves, Inventiones Math., 23 (1974), 179–206.
[T6] J. Tate, Number theoretic background, Amer. Math. Soc. Proc. Symp. Pure Math., 33 (1979), pt. 2, 3–26.
[T7] T. Tatuzawa, On the Hecke-Landau *L*-series, Nagoya Math. J., 16 (1960), 11–20.
[T8] E. C. Titchmarch, The theory of the Riemann zeta-function, Clarendon Press, 1951.
[W1] A. Weil, Collected papers, I–III, Springer, 1980.
[W2] A. Weil, Courbes algébriques et variétés abéliennes, Hermann, 1971 (orig., 1948).
[W3] A. Weil, Numbers of solutions of equations in finite fields, Bull. Amer. Math. Soc., 55 (1949), 497–508 [W1, I, 399–400].
[W4] A. Weil, Number theory and algebraic geometry, Proc. Intern. Congr. Math., 1950, Cambridge, vol. 2, 90–100 [W1, I, 442–449].
[W5] A. Weil, Sur la théorie du corps de classes, J. Math. Soc. Japan, 3 (1951), 1–35 [W1, I, 483–517].
[W6] A. Weil, Jacobi sums as "Grössencharaktere," Trans. Amer. Math. Soc., 73 (1952), 487–495 [W1, II, 63–71].
[W7] A. Weil, On a certain type of characters of the idele-class group of an algebraic number field, Proc. Intern. Symposium on Algebraic Number Theory, Tokyo-Nikko, 1955, 1–7 [W1, II, 255–261].
[W8] A. Weil, Basic number theory, Springer, 1967.
[W9] A. Weil, Dirichlet series and automorphic forms, Lecture notes in math. 189, Springer, 1971.
[Y1] H. Yoshida, On an analogue of the Sato conjecture, Inventiones Math., 19 (1973), 261–277.

Appendix A
Tables of Formulas

1. Algebraic Equations (→ 10 Algebraic Equations)

(I) Quadratic Equation $ax^2+bx+c=0 \quad (a\neq 0)$

The roots are

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}=\frac{-b'\pm\sqrt{b'^2-ac}}{a} \qquad (b\equiv 2b').$$

The discriminant is b^2-4ac.

(II) Cubic Equation $ax^3+bx^2+cx+d=0 \quad (a\neq 0)$

By the translation $\xi=x+b/3a$, the equation is transformed into $\xi^3+3p\xi+q=0$, where

$p\equiv(3ac-b^2)/9a^2, \qquad q\equiv(2b^3-9abc+27a^2d)/27a^3.$

Its discriminant is $-27(q^2+4p^3)$. The roots of the latter equation are

$$\xi=\sqrt[3]{\alpha}+\sqrt[3]{\beta}, \quad \omega\sqrt[3]{\alpha}+\omega^2\sqrt[3]{\beta}, \quad \omega^2\sqrt[3]{\alpha}+\omega\sqrt[3]{\beta},$$

where

$$\omega=e^{2\pi i/3}=\frac{-1+\sqrt{3}\,i}{2}, \quad \left.\begin{matrix}\alpha\\ \beta\end{matrix}\right\}=\frac{-q\pm\sqrt{q^2+4p^3}}{2} \qquad \text{(Cardano's formula).}$$

Casus irreducibilis (the case when $q^2+4p^3<0$). Putting $\alpha\equiv re^{i\theta}$ ($\beta=\bar{\alpha}$), the roots are

$$\xi=2\sqrt[3]{r}\cos(\theta/3), \quad 2\sqrt[3]{r}\cos[(\theta+2\pi)/3], \quad 2\sqrt[3]{r}\cos[(\theta+4\pi)/3].$$

(III) Quartic Equation (Biquadratic Equation) $ax^4+bx^3+cx^2+dx+e=0 \quad (a\neq 0)$

By the translation $\xi=x+b/4a$, the equation is transformed into

$\xi^4+p\xi^2+q\xi+r=0.$

The cubic resolvent of the latter is $t^3-pt^2-4rt+(4pr-q^2)=0$. If t_0 is one of the roots of the cubic resolvent, the roots ξ of the above equation are the solutions of two quadratic equations

$$\xi^2\pm\sqrt{t_0-p}\,[\xi-q/2(t_0-p)]+t_0/2=0 \qquad \text{(Ferrari's formula).}$$

2. Trigonometry

(I) Trigonometric Functions (→ 432 Trigonometry)

(1) In Fig. 1, $OA=OB=OP=1$, and

$MP=\sin\theta, \quad OM=\cos\theta, \quad AT=\tan\theta,$

$BL=\cot\theta, \quad OT=\sec\theta, \quad OL=\operatorname{cosec}\theta.$

(2) $\sin^2\theta+\cos^2\theta=1,$

$\tan\theta=\sin\theta/\cos\theta, \quad \cot\theta=1/\tan\theta, \quad \sec\theta=1/\cos\theta,$

$\operatorname{cosec}\theta=1/\sin\theta, \quad 1+\tan^2\theta=\sec^2\theta, \quad 1+\cot^2\theta=\operatorname{cosec}^2\theta.$

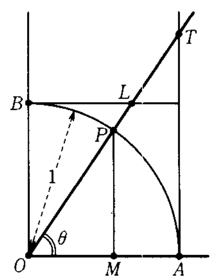


Fig. 1

(3)

	θ	$-\theta$	$\pi/2\pm\theta$	$\pi\pm\theta$	$n\pi\pm\theta$
sin	s	$-s$	c	$\mp s$	$\pm(-1)^n s$
cos	c	c	$\mp s$	$-c$	$(-1)^n c$
tan	t	$-t$	$\mp 1/t$	$\pm t$	$\pm t$

(4)

α	$0°$	$15°$	$18°$	$22.5°$	$30°$	$36°$	$45°$	
↗	0	$\pi/12$	$\pi/10$	$\pi/8$	$\pi/6$	$\pi/5$	$\pi/4$	
$\sin\alpha$	0	$\frac{\sqrt{3}-1}{2\sqrt{2}}$	$\frac{\sqrt{5}-1}{4}$	$\frac{\sqrt{2-\sqrt{2}}}{2}$	$\frac{1}{2}$	$\frac{\sqrt{10-2\sqrt{5}}}{4}$	$\frac{1}{\sqrt{2}}$	$\cos\alpha$
$\cos\alpha$	1	$\frac{\sqrt{3}+1}{2\sqrt{2}}$	$\frac{\sqrt{10+2\sqrt{5}}}{4}$	$\frac{\sqrt{2+\sqrt{2}}}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{5}+1}{4}$	$\frac{1}{\sqrt{2}}$	$\sin\alpha$
	$\pi/2$	$5\pi/12$	$2\pi/5$	$3\pi/8$	$\pi/3$	$3\pi/10$	$\pi/4$	↙
	$90°$	$75°$	$72°$	$67.5°$	$60°$	$54°$	$45°$	α

(5) Addition Formulas

$$\sin(\alpha\pm\beta)=\sin\alpha\cos\beta\pm\cos\alpha\sin\beta,\qquad \cos(\alpha\pm\beta)=\cos\alpha\cos\beta\mp\sin\alpha\sin\beta,$$

$$\tan(\alpha\pm\beta)=(\tan\alpha\pm\tan\beta)/(1\mp\tan\alpha\tan\beta).$$

(6) $\sin 2\alpha=2\sin\alpha\cos\alpha,\qquad \cos 2\alpha=\cos^2\alpha-\sin^2\alpha=2\cos^2\alpha-1=1-2\sin^2\alpha,$

$$\tan 2\alpha=2\tan\alpha/(1-\tan^2\alpha).$$

$$\sin 3\alpha=3\sin\alpha-4\sin^3\alpha,\qquad \cos 3\alpha=4\cos^3\alpha-3\cos\alpha,$$

$$\tan 3\alpha=(3\tan\alpha-\tan^3\alpha)/(1-3\tan^2\alpha).$$

$$\sin n\alpha=\sum_{i=0}^{[(n-1)/2]}\binom{n}{2i+1}(-1)^i\sin^{2i+1}\alpha\cos^{n-(2i+1)}\alpha,$$

$$\cos n\alpha=\sum_{i=0}^{[n/2]}\binom{n}{2i}(-1)^i\sin^{2i}\alpha\cos^{n-2i}\alpha.$$

(7) $\sin^2(\alpha/2)=(1-\cos\alpha)/2,\qquad \cos^2(\alpha/2)=(1+\cos\alpha)/2,$

$$\tan^2(\alpha/2)=(1-\cos\alpha)/(1+\cos\alpha).$$

(8) $2\sin\alpha\cos\beta=\sin(\alpha+\beta)+\sin(\alpha-\beta),\qquad 2\cos\alpha\sin\beta=\sin(\alpha+\beta)-\sin(\alpha-\beta),$

$$2\cos\alpha\cos\beta=\cos(\alpha+\beta)+\cos(\alpha-\beta),\qquad -2\sin\alpha\sin\beta=\cos(\alpha+\beta)-\cos(\alpha-\beta).$$

$$\sin\alpha+\sin\beta=2\sin[(\alpha+\beta)/2]\cos[(\alpha-\beta)/2],$$

$$\sin\alpha-\sin\beta=2\cos[(\alpha+\beta)/2]\sin[(\alpha-\beta)/2],$$

$$\cos\alpha+\cos\beta=2\cos[(\alpha+\beta)/2]\cos[(\alpha-\beta)/2],$$

$$\cos\alpha-\cos\beta=-2\sin[(\alpha+\beta)/2]\sin[(\alpha-\beta)/2].$$

(II) Plane Triangles

As shown in Fig. 2, we denote the interior angles of a triangle ABC by α, β, γ; the corresponding side lengths by a, b, c; the area by S; the radii of inscribed, circumscribed, and escribed circles by r, R, r_A, respectively; the perpendicular line from the vertex A to the side BC by AH; the midpoint of the side BC by M; bisector of the angle A by AD; and the lengths of AH, AM, AD by h_A, m_A, f_A, respectively. Similar notations are used for B and C. Put $s\equiv(a+b+c)/2$. The symbol ... means similar formulas by the cyclic permutation of the letters A, B, C, and corresponding quantities.

$$\frac{a}{\sin\alpha}=\frac{b}{\sin\beta}=\frac{c}{\sin\gamma}=2R\qquad\text{(law of sines).}$$

$$a=b\cos\gamma+c\cos\beta,\quad\ldots\qquad\text{(the first law of cosines).}$$

$$a^2=b^2+c^2-2bc\cos\alpha,\quad\ldots\qquad\text{(the second law of cosines).}$$

$$\sin^2(\alpha/2)=(s-b)(s-c)/bc,\quad\ldots;\qquad \cos^2(\alpha/2)=s(s-a)/bc,\quad\ldots.$$

$$(b+c)\sin(\alpha/2)=a\cos[(\beta-\gamma)/2],\quad\ldots;\qquad (b-c)\cos(\alpha/2)=a\sin[(\beta-\gamma)/2],\quad\ldots.$$

$$\frac{a+b}{a-b}=\frac{\tan[(\alpha+\beta)/2]}{\tan[(\alpha-\beta)/2]},\quad\ldots\qquad\text{(Napier's rule).}$$

$$S = ah_A/2 = (1/2)bc\sin\alpha = (1/2)a^2\sin\beta\sin\gamma/\sin\alpha = abc/4R = 2R^2\sin\alpha\sin\beta\sin\gamma$$
$$= rs = r_A(s-a) = \sqrt{rr_Ar_Br_C}$$
$$= \sqrt{s(s-a)(s-b)(s-c)} \qquad \text{(Heron's formula).}$$
$$r = (s-a)\tan(\alpha/2) = 4R\sin(\alpha/2)\sin(\beta/2)\sin(\gamma/2).$$
$$r_A = s\tan(\alpha/2) = (s-b)\cot(\gamma/2) = 4R\sin(\alpha/2)\cos(\beta/2)\cos(\gamma/2).$$
$$1/r = (1/h_A) + (1/h_B) + (1/h_C).$$
$$m_A^2 = (2b^2 + 2c^2 - a^2)/4 = (b^2 + c^2 + 2bc\cos\alpha)/4.$$
$$f_A = 2bc\cos(\alpha/2)/(b+c) = 2\sqrt{bcs(s-a)}\ /(b+c).$$
$$f_Af_Bf_C = 8abcrs^2/(b+c)(c+a)(a+b).$$

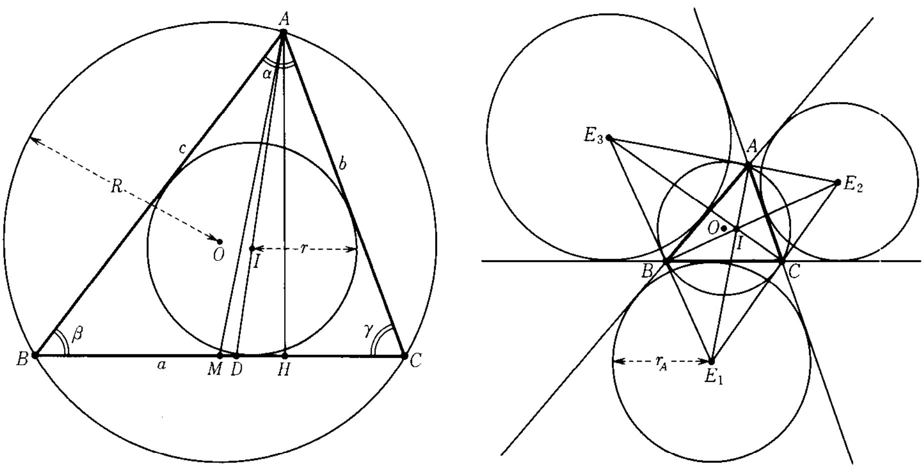


Fig. 2

(III) Spherical Triangles

We denote the interior angles of a spherical triangle by α, β, γ; the corresponding sides by a, b, c; the area by S; and the radius of the supporting sphere by ρ. We have

$$\sin a : \sin b : \sin c = \sin\alpha : \sin\beta : \sin\gamma \qquad \text{(law of sines).}$$

$$\cos a = \cos b\cos c + \sin b\sin c\cos\alpha, \quad \ldots; \qquad \cos\alpha = -\cos\beta\cos\gamma + \sin\beta\sin\gamma\cos a, \quad \ldots$$

(law of cosines).

$$\sin a\cos\beta = \cos b\sin c - \sin b\cos c\cos\alpha, \quad \ldots \qquad \text{(law of sines and cosines).}$$

$$\cot a\sin b = \cos b\cos\gamma + \cot\alpha\sin\gamma, \quad \ldots \qquad \text{(law of cotangents).}$$

$$\tan[(a+b)/2]/\tan[(a-b)/2] = \tan[(\alpha+\beta)/2]/\tan[(\alpha-\beta)/2], \quad \ldots \qquad \text{(law of tangents).}$$

$$\tan[(\alpha+\beta)/2]\tan(\gamma/2) = \cos[(a-b)/2]/\cos[(a+b)/2], \quad \ldots;$$

$$\tan[(\alpha-\beta)/2]\tan(\gamma/2) = \sin[(a-b)/2]/\sin[(a+b)/2], \quad \ldots;$$

$$\tan[(a+b)/2]\cot(c/2) = \cos[(\alpha-\beta)/2]/\cos[(\alpha+\beta)/2], \quad \ldots;$$

$$\tan[(a-b)/2]\cot(c/2) = \sin[(\alpha-\beta)/2]/\sin[(\alpha+\beta)/2], \quad \ldots \qquad \text{(Napier's analogies).}$$

$$S = (\alpha+\beta+\gamma-\pi)\rho^2 = 2\rho^2\arccos\frac{\cos^2(a/2R) + \cos^2(b/2R) + \cos^2(c/2R)}{2\cos(a/2R)\cos(b/2R)\cos(c/2R)} \qquad \text{(Heron's formula).}$$

For a right triangle ($\gamma = \pi/2$), we have Napier's rule of circular parts: taking the subscripts modulo 5 in Fig. 3,

$$\sin\theta_i = \tan\theta_{i+1}\tan\theta_{i-1} = \cos\theta_{i+2}\cos\theta_{i-2}.$$

For example, we have

$$\cos c = \cos a \cos b = \cot\alpha \cot\beta,$$

$$\cos\beta = \tan a \cot c = \cos b \sin\alpha,$$

$$\sin a = \tan b \cot\beta = \sin c \sin\alpha.$$

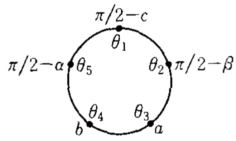


Fig. 3

3. Vector Analysis and Coordinate Systems

We denote a 3-dimensional vector by $\mathbf{A} \equiv (A_x, A_y, A_z) = A_x \boldsymbol{i} + A_y \boldsymbol{j} + A_z \boldsymbol{k}$, $|\mathbf{A}| = \sqrt{A_x^2 + A_y^2 + A_z^2}$.

(I) Vector Algebra (→ 442 Vectors)

Scalar product $\mathbf{A} \cdot \mathbf{B} \equiv \mathbf{AB} \equiv (\mathbf{A}, \mathbf{B}) = A_x B_x + A_y B_y + A_z B_z = |\mathbf{A}|\,|\mathbf{B}|\cos\theta$ (where θ is the angle between $\mathbf{A}$ and $\mathbf{B}$).

Vector product

$$\mathbf{A} \times \mathbf{B} \equiv [\mathbf{A}, \mathbf{B}] = (A_y B_z - A_z B_y)\mathbf{i} + (A_z B_x - A_x B_z)\mathbf{j} + (A_x B_y - A_y B_x)\mathbf{k} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}.$$

$|\mathbf{A} \times \mathbf{B}| = |\mathbf{A}|\,|\mathbf{B}|\sin\theta.$

$\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}.$ $\mathbf{A} \cdot \mathbf{A} \equiv \mathbf{A}^2 = |\mathbf{A}|^2.$ $\mathbf{A} \times \mathbf{A} = 0.$ $\mathbf{A} \cdot (\mathbf{A} \times \mathbf{B}) = 0.$ $(\mathbf{A} \times \mathbf{B})^2 = |\mathbf{A}|^2|\mathbf{B}|^2 - (\mathbf{A} \cdot \mathbf{B})^2.$

$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}.$ $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) + \mathbf{B} \times (\mathbf{C} \times \mathbf{A}) + \mathbf{C} \times (\mathbf{A} \times \mathbf{B}) = 0.$

$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = \mathbf{A} \cdot \{\mathbf{B} \times (\mathbf{C} \times \mathbf{D})\} = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{B} \cdot \mathbf{C})(\mathbf{A} \cdot \mathbf{D}).$

Scalar triple product $[\mathbf{ABC}] \equiv \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix}.$

$$[\mathbf{BCD}]\mathbf{A} + [\mathbf{ACD}]\mathbf{B} + [\mathbf{ABD}]\mathbf{C} = [\mathbf{ABC}]\mathbf{D}. \quad [\mathbf{ABC}][\mathbf{EFG}] = \begin{vmatrix} \mathbf{A}\cdot\mathbf{E} & \mathbf{A}\cdot\mathbf{F} & \mathbf{A}\cdot\mathbf{G} \\ \mathbf{B}\cdot\mathbf{E} & \mathbf{B}\cdot\mathbf{F} & \mathbf{B}\cdot\mathbf{G} \\ \mathbf{C}\cdot\mathbf{E} & \mathbf{C}\cdot\mathbf{F} & \mathbf{C}\cdot\mathbf{G} \end{vmatrix}.$$

(II) Differentiation of a Vector Field (→ 442 Vectors)

$$\nabla \equiv \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \qquad \text{(Nabla)},$$

$$\operatorname{grad}\varphi \equiv \nabla\varphi = \frac{\partial\varphi}{\partial x}\mathbf{i} + \frac{\partial\varphi}{\partial y}\mathbf{j} + \frac{\partial\varphi}{\partial z}\mathbf{k} \qquad \text{(gradient of } \varphi\text{)},$$

$$\operatorname{rot}\mathbf{A} \equiv \nabla \times \mathbf{A} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\mathbf{i} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\mathbf{j} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\mathbf{k} \qquad \text{(rotation of } \mathbf{A}\text{)},$$

$$\operatorname{div}\mathbf{A} \equiv \nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \qquad \text{(divergence of } \mathbf{A}\text{)},$$

$$\Delta\varphi \equiv \nabla^2\varphi \equiv \operatorname{div}\operatorname{grad}\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2} \qquad \text{(Laplacian of } \varphi\text{)}.$$

$\operatorname{grad}(\varphi\psi) = \varphi\operatorname{grad}\psi + \psi\operatorname{grad}\varphi,$

$\operatorname{grad}(\mathbf{A} \cdot \mathbf{B}) = (\mathbf{B} \cdot \operatorname{grad})\mathbf{A} + (\mathbf{A} \cdot \operatorname{grad})\mathbf{B} + \mathbf{A} \times \operatorname{rot}\mathbf{B} + \mathbf{B} \times \operatorname{rot}\mathbf{A},$

$\operatorname{rot}(\varphi\mathbf{A}) = \varphi\operatorname{rot}\mathbf{A} - \mathbf{A} \times \operatorname{grad}\varphi,$ $\operatorname{rot}(\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \operatorname{grad})\mathbf{A} - (\mathbf{A} \cdot \operatorname{grad})\mathbf{B} + \mathbf{A}\operatorname{div}\mathbf{B} - \mathbf{B}\operatorname{div}\mathbf{A},$

$\operatorname{div}(\varphi\mathbf{A}) = \varphi\operatorname{div}\mathbf{A} + \mathbf{A} \cdot \operatorname{grad}\varphi,$ $\operatorname{div}(\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \operatorname{rot}\mathbf{A} - \mathbf{A} \cdot \operatorname{rot}\mathbf{B}.$

$\operatorname{rot}\operatorname{grad}\varphi = 0,$ $\operatorname{div}\operatorname{rot}\mathbf{A} = 0.$ $\Delta\mathbf{A} = \operatorname{grad}\operatorname{div}\mathbf{A} - \operatorname{rot}\operatorname{rot}\mathbf{A}.$

$\Delta(f \circ \varphi) = (df/d\varphi)\Delta\varphi + (d^2f/d\varphi^2)(\operatorname{grad}\varphi)^2,$ $\Delta(\varphi\psi) = \varphi\Delta\psi + \psi\Delta\varphi + 2(\operatorname{grad}\varphi \cdot \operatorname{grad}\psi).$

(III) Integration of a Vector Field (→ 94 Curvilinear Integrals and Surface Integrals, 442 Vectors)

Let D be a 3-dimensional domain, B its boundary, dV the volume element of D, dS the surface element of B, and $d\mathbf{S}=\mathbf{n}\,dS$, where $\mathbf{n}$ is the outer normal vector of the surface B. We have

Gauss's formula $$\iiint_D \operatorname{div}\mathbf{A}\,dV=\iint_B d\mathbf{S}\cdot\mathbf{A}=\iint_B(\mathbf{n}\cdot\mathbf{A})\,dS,$$

$$\iiint_D \operatorname{rot}\mathbf{A}\,dV=\iint_B d\mathbf{S}\times\mathbf{A}=\iint_B(\mathbf{n}\times\mathbf{A})\,dS,$$

$$\iiint_D \operatorname{grad}\varphi\,dV=\iint_B \varphi\,d\mathbf{S};$$

Green's formula $$\iint_B \varphi\frac{\partial\psi}{\partial n}\,dS=\iiint_D(\varphi\Delta\psi+\operatorname{grad}\varphi\cdot\operatorname{grad}\psi)\,dV,$$

$$\iint_B\left(\varphi\frac{\partial\psi}{\partial n}-\psi\frac{\partial\varphi}{\partial n}\right)dS=\iiint_D(\varphi\Delta\psi-\psi\Delta\varphi)\,dV,$$

$$4\pi\varphi(x_0)=-\iiint_D\frac{\Delta\varphi}{r}\,dV+\iint_B\left\{\frac{1}{r}\frac{\partial\varphi}{\partial n}-\varphi\frac{\partial}{\partial n}\left(\frac{1}{r}\right)\right\}dS,$$

where r is the distance from the point x_0.

Let B be a bordered surface with a boundary curve Γ, ds the line element of Γ, dS the surface element of B, and $d\mathbf{s}=\mathbf{t}\,ds$, $d\mathbf{S}=\mathbf{n}\,dS$, for $\mathbf{t}$ the unit tangent vector of Γ and under the proper choice of the positive direction for the surface normal $\mathbf{n}$. We have

Stokes's formula $$\iint_B d\mathbf{S}\cdot\operatorname{rot}\mathbf{A}=\oint_\Gamma\mathbf{A}\cdot d\mathbf{s}=\oint_\Gamma(\mathbf{t}\cdot\mathbf{A})\,ds,\qquad \iint_B d\mathbf{S}\times\operatorname{grad}\varphi=\oint_\Gamma\varphi\,d\mathbf{s}.$$

If the domain D is simply connected, and the vector field $\mathbf{V}$ tends sufficiently rapidly to 0 near the boundary of D and at infinity, we have

Helmholtz's theorem $$\mathbf{V}=\operatorname{grad}\varphi+\operatorname{rot}\mathbf{A},\quad \varphi=-\iiint_D\frac{\operatorname{div}\mathbf{V}}{4\pi r}\,dV,\quad \mathbf{A}=\iiint_D\frac{\operatorname{rot}\mathbf{V}}{4\pi r}\,dV.$$

(IV) Moving Coordinate System

Denote differentiation with respect to the rest and the moving systems by d/dt, d^*/dt, respectively. Let the relative velocity of the systems be $\mathbf{v}$. Then we have

$$\frac{d\varphi}{dt}=\frac{d^*\varphi}{dt}-\mathbf{v}\cdot\operatorname{grad}\varphi,\quad \frac{d\mathbf{A}}{dt}=\frac{d^*\mathbf{A}}{dt}-[\mathbf{v}\cdot\operatorname{grad}\mathbf{A}-(\mathbf{A}\cdot\operatorname{grad})\mathbf{v}].$$

With respect to rotating coordinates we have

$$\mathbf{v}=\mathbf{w}\times\mathbf{r}.$$

$$\frac{d\mathbf{A}}{dt}=\frac{d^*\mathbf{A}}{dt}+[\mathbf{w}+\mathbf{A}-((\mathbf{w}\times\mathbf{r})\cdot\operatorname{grad})\mathbf{A}]$$

When the domain of integration is also a function of t,

$$\frac{d}{dt}\int\mathbf{A}\cdot d\mathbf{s}=\int\left\{\frac{\partial\mathbf{A}}{\partial t}+\operatorname{grad}(\mathbf{v}\cdot\mathbf{A})-\mathbf{v}\times\operatorname{rot}\mathbf{A}\right\}\cdot d\mathbf{s},$$

$$\frac{d}{dt}\iint\mathbf{A}\cdot d\mathbf{S}=\iint\left\{\frac{\partial\mathbf{A}}{\partial t}+\operatorname{rot}(\mathbf{A}\times\mathbf{v})+\mathbf{v}\operatorname{div}\mathbf{A}\right\}\cdot d\mathbf{S},$$

$$\frac{d}{dt}\iiint\varphi\,dV=\iiint\left\{\frac{\partial\varphi}{\partial t}+(\mathbf{v}\cdot\operatorname{grad}\varphi)+\varphi\operatorname{div}\mathbf{v}\right\}dV=\iiint\frac{\partial\varphi}{\partial t}\,dV+\iint\varphi\mathbf{v}\cdot d\mathbf{S}.$$

(V) Curvilinear Coordinates (→ 90 Coordinates)

Let $(x_1,\ldots,x_n)$ be rectangular coordinates in an n-dimensional Euclidean space. If

$$x_j=\varphi_j(u_1,\ldots,u_n)\quad(j=1,\ldots,n),\qquad J\equiv\det(\partial\varphi_j/\partial u_k)\neq0,$$

the system $(u_1,\ldots,u_n)$ may be taken as a coordinate system of an n-dimensional space, and the original space is a Riemannian manifold with the first fundamental form

$$g_{jk}=\sum_{i=1}^{n}\frac{\partial\varphi_i}{\partial u_j}\frac{\partial\varphi_i}{\partial u_k}\qquad (j,k=1,\dots,n),$$

$$g\equiv\det(g_{jk})=J^2.$$

When the metric is of the diagonal form $g_{jk}=g_j^2\delta_{jk}$, the coordinate system $(u_1,\dots,u_n)$ is called an orthogonal curvilinear coordinate system or an isothermal curvilinear coordinate system. In such a case we have $J=g_1\dots g_n$, and the line element is given by $ds^2=\sum_{j=1}^n g_j^2\,du_j^2$.

For a scalar f and a vector $\xi=(\xi_1,\dots,\xi_n)$, we have

$$(\operatorname{grad} f)_j=\frac{1}{g_j}\frac{\partial f}{\partial u_j}\quad (j=1,\dots,n),\qquad \Delta f=\frac{1}{J}\sum_{j=1}^{n}\frac{\partial}{\partial u_j}\left(\frac{J}{g_j^2}\frac{\partial f}{\partial u_j}\right),$$

$$\operatorname{div}\xi=\frac{1}{J}\sum_{j=1}^{n}\frac{\partial}{\partial u_j}\left(\frac{J}{g_j}\xi_j\right),\quad (\operatorname{rot}\xi)_{jk}=\frac{1}{g_jg_k}\left[\frac{\partial(g_k\xi_k)}{\partial u_j}-\frac{\partial(g_j\xi_j)}{\partial u_k}\right]\qquad (j,k=1,\dots,n).$$

When $n=2$, the rot may be considered a scalar, $\operatorname{rot}\xi=(\operatorname{rot}\xi)_{12}$, and when $n=3$, the rot may be considered a vector, with components

$$\operatorname{rot}\xi=((\operatorname{rot}\xi)_{23},(\operatorname{rot}\xi)_{31},(\operatorname{rot}\xi)_{12}).$$

The following are examples of orthogonal coordinates.

(1) Planar Curvilinear Coordinates. In the present Section (1), we put

$$x_1=x,\quad x_2=y,\quad u_1=u,\quad u_2=v,\quad g_1=p,\quad g_2=q.$$

$$ds^2=p^2dx^2+q^2dy^2,\qquad J\equiv\partial(x,y)/\partial(u,v)=\sqrt{pq}\ .$$

Planar orthogonal curvilinear coordinates may be represented in the form $x+iy=F(U+iV)$, F being a complex analytic function, by suitable choice of the functions $U=U(u)$, $V=V(v)$.

(i) Polar Coordinates (r,θ) (Fig. 4).

$$x=r\cos\theta,\quad y=r\sin\theta;\quad x+iy=\exp(\log r+i\theta).$$

$$r=\sqrt{x^2+y^2}\ ,\quad \theta=\arctan(y/x).$$

$$p=1,\quad q=r,\quad J=r,\quad ds^2=dr^2+r^2d\theta^2.$$

$$\Delta f=\frac{\partial^2 f}{\partial r^2}+\frac{1}{r}\frac{\partial f}{\partial r}+\frac{1}{r^2}\frac{\partial^2 f}{\partial\theta^2}.$$

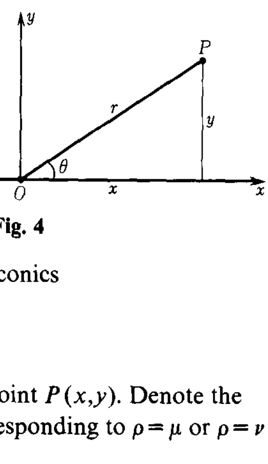


Fig. 4

(ii) Elliptic Coordinates (μ,ν) (Fig. 5). Among the family of confocal conics

$$\frac{x^2}{a^2+\rho}+\frac{y^2}{b^2+\rho}=1\quad (a>b),$$

there are two values of ρ for which the curve passes through a given point $P(x,y)$. Denote the two values of ρ by μ and ν, where $\mu>-b^2>\nu>-a^2$. The curve corresponding to $\rho=\mu$ or $\rho=\nu$ is an ellipse or a hyperbola, respectively. Then we have the relations

$$x^2=(\mu+a^2)(\nu+a^2)/(a^2-b^2),\qquad y^2=(\mu+b^2)(\nu+b^2)/(b^2-a^2).$$

Let the common foci be $(\pm c,0)$ $(c^2=a^2-b^2)$. Then we have

$$r_1=\sqrt{(x-c)^2+y^2}\ ,\qquad r_2=\sqrt{(x+c)^2+y^2}$$

where r_1, r_2 are the distances from the two foci as in Fig. 5, and

$$4(a^2+\mu)=(r_1+r_2)^2,\qquad 4(a^2+\nu)=(r_1-r_2)^2.$$

$$p=\frac{1}{2}\sqrt{\frac{\mu-\nu}{(\mu+a^2)(\mu+b^2)}}\ ,\qquad q=\frac{1}{2}\sqrt{\frac{\nu-\mu}{(\nu+a^2)(\nu+b^2)}}$$

(iii) Parabolic Coordinates (α,β) (Fig. 6). Among the family of parabolas $y^2=4\rho(x+\rho)$ with the focus at the origin and having the x-axis as the principal axis, there are two values of ρ for which the curve passes through a given point $P(x,y)$. Denote the two values of ρ by α, β $(\alpha>0>\beta)$. We have $x=-(\alpha+\beta)$, $y=\sqrt{-4\alpha\beta}$.

(iv) Equilateral (or Rectangular) Hyperbolic Coordinates (u,v) (Fig. 7). This is a system that replaces $x/2$, $y/2$ in (iii) by $-y$ and x, respectively, with $\sqrt{\alpha}=u$, $\sqrt{-\beta}=v$. The relations are

$$x=uv,\quad y=(u^2-v^2)/2;\quad x+iy=i(u-iv)^2/2,\quad u^2,v^2=\sqrt{x^2+y^2}\pm y,\quad p=q=\sqrt{u^2+v^2}.$$

The curves $x=$constant or $y=$constant are equilateral hyperbolas.

(v) Bipolar Coordinates (ξ,η) (Fig. 8). These coordinates represent a point $P(x,y)$ on a plane as the intersection of the family of circles passing through two fixed points $(\pm a,0)$ and the family of loci on which the ratio of distances from the same two fixed points $(\pm a,0)$ is constant. The latter is the set of Apollonius' circles. The relations are

$$x=\frac{a\sinh\xi}{\cosh\xi+\cos\eta},\qquad y=\frac{a\sin\eta}{\cosh\xi+\cos\eta}\qquad(-\infty<\xi<\infty,\ 0\leqslant\eta\leqslant 2\pi).$$

$$p=q=\frac{a}{\cosh\xi+\cos\eta}.$$

(2) Curvilinear Coordinates in 3-Dimensional Space. In the present Section (2), we put $x_1=x$, $x_2=y$, $x_3=z$.

(i) Circular Cylindrical Coordinates (Cylindrical Coordinates) (ρ,φ,z) (Fig. 9).

$$x=\rho\cos\varphi,\quad y=\rho\sin\varphi,\quad z=z.$$

$$ds^2=d\rho^2+\rho^2d\varphi^2+dz^2,\qquad J=\rho.\qquad \Delta f=\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial f}{\partial\rho}\right)+\frac{1}{\rho^2}\frac{\partial^2 f}{\partial\varphi^2}+\frac{\partial^2 f}{\partial z^2}.$$

(ii) Polar Coordinates (Spherical Coordinates) (Fig. 9).

$$x=r\sin\theta\cos\varphi,\quad y=r\sin\theta\sin\varphi,\quad z=r\cos\theta.$$

$$r=\sqrt{x^2+y^2+z^2},\quad \varphi=\arctan(y/x),\quad \theta=\arctan\left(\sqrt{x^2+y^2}\,/z\right).$$

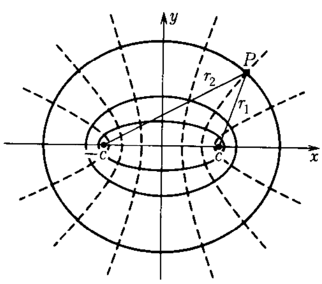


Fig. 5

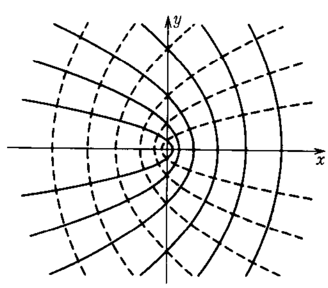


Fig. 6

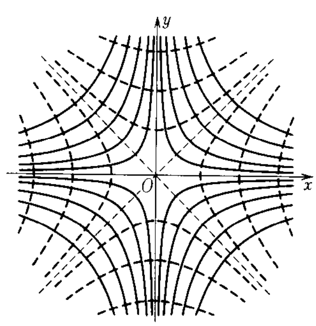


Fig. 7

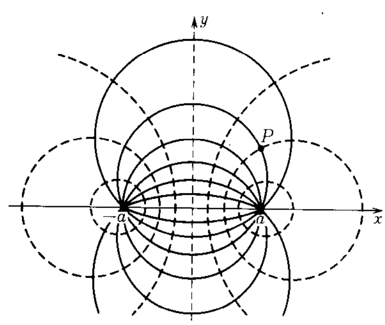


Fig. 8

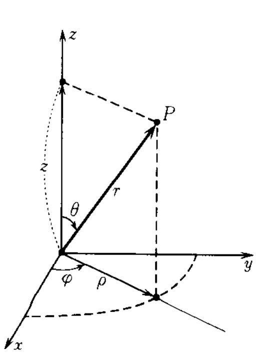


Fig. 9

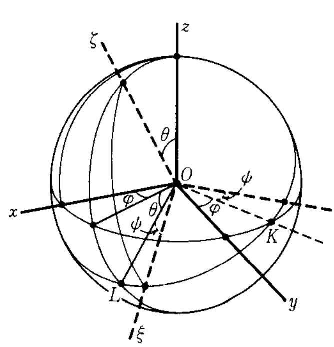


Fig. 10

The angles φ and θ are called azimuth and zenith angle, respectively. We further have

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\varphi^2, \qquad J = r^2 \sin\theta.$$

$$\Delta f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2}.$$

(iii) Euler's Angles (Fig. 10). Let (x,y,z) and (ξ,η,ζ) be two linear orthogonal coordinate systems with common origin O. Denote the angle between the z-axis and the ζ-axis by θ; the angle between the zx-plane and the $z\zeta$-plane by φ; and the angle between the η-axis and the intersection OK of the xy-plane and the $\xi\eta$-plane (or the angle between the ξ-axis and the intersection OL of the $z\zeta$-plane and the $\xi\eta$-plane) by ψ. The angles θ, φ, and ψ are called Euler's angles. The direction cosines of one coordinate axis with respect to the other coordinate system are as follows:

	x	y	z
ξ	$\cos\varphi\cos\theta\cos\psi - \sin\varphi\sin\psi$	$\sin\varphi\cos\theta\cos\psi + \cos\varphi\sin\psi$	$-\sin\theta\cos\psi$
η	$-\cos\varphi\cos\theta\sin\psi - \sin\varphi\cos\psi$	$-\sin\varphi\cos\theta\sin\psi + \cos\varphi\cos\psi$	$\sin\theta\sin\psi$
ζ	$\cos\varphi\sin\theta$	$\sin\varphi\sin\theta$	$\cos\theta$

(iv) Rotational (or Revolutional) Coordinates (u,v,ρ). Let (u,v) be curvilinear coordinates (Section (1)) on the $z\rho$-plane. The rotational coordinates (u,v,ρ) are given by the combination of $x=\rho\cos\varphi$, $y=\rho\sin\varphi$ with the coordinates on the $z\rho$-plane. We have

$$ds^2 = p^2 du^2 + q^2 dv^2 + \rho^2 d\varphi^2,$$

where p, q are the corresponding values for the coordinates (u,v).

(v) Generalized Cylindrical Coordinates (u,v,z). These are a combination of curvilinear coordinates (u,v) on the xy-plane with z. We have

$$ds^2 = p^2 du^2 + q^2 dv^2 + dz^2.$$

For various selections of (u,v) we have coordinates as follows:

(u,v)	Rotational Coordinate System	Generalized Cylindrical Coordinate System
Linear rectangular coordinates	Circular cylindrical coordinates	Linear rectangular coordinates
Polar coordinates ((1)(i))	Spherical coordinates	Circular cylindrical coordinates
Elliptic coordinates ((1)(ii))	Spheroidal coordinates(1)	Elliptic cylindrical coordinates
Parabolic coordinates ((1)(iii))	Rotational parabolic coordinates(2)	Parabolic cylindrical coordinates
Equilateral hyperbolic coordinates ((1)(iv))	Rotational hyperbolic coordinates	Hyperbolic cylindrical coordinates
Bipolar coordinates ((1)(v))	Toroidal coordinates(3) Bipolar coordinates(4)	Bipolar cylindrical coordinates

Notes

(1) When the ρ-axis is a minor or major axis, we have prolate or oblate spheroidal coordinates, respectively.

(2) We take the z-axis as the common principal axis of the parabolas.

(3) Where the line passing through two fixed points is the ρ-axis.

(4) Where the line passing through two fixed points is the z-axis.

(vi) Ellipsoidal Coordinates (λ,μ,ν) (Fig. 11). Among the family of confocal quadrics

$$\frac{x^2}{a^2+\rho} + \frac{y^2}{b^2+\rho} + \frac{z^2}{c^2+\rho} = 1 \quad (a>b>c>0),$$

there are three values of ρ for which the surface passes through a given point $P(x,y,z)$. Denote the three values of ρ by λ, μ, ν, where $\lambda > -c^2 > \mu > -b^2 > \nu > -a^2$. The surfaces corresponding to $\rho=\lambda$, $\rho=\mu$, and $\rho=\nu$ are an ellipsoid, a hyperboloid of one sheet, and a hyperboloid of two sheets, respectively. We have

$$x^2=\frac{h(a)}{(a^2-b^2)(a^2-c^2)}, \qquad y^2=\frac{h(b)}{(b^2-c^2)(b^2-a^2)},$$

$$z^2=\frac{h(c)}{(c^2-a^2)(c^2-b^2)}; \qquad h(\alpha)\equiv(\lambda+\alpha^2)(\mu+\alpha^2)(\nu+\alpha^2).$$

$$g_1=\frac{\sqrt{(\lambda-\mu)(\lambda-\nu)}}{2\rho(\lambda)}, \qquad g_2=\frac{\sqrt{(\mu-\nu)(\mu-\lambda)}}{2\rho(\mu)},$$

$$g_3=\frac{\sqrt{(\nu-\lambda)(\nu-\mu)}}{2\rho(\nu)}; \qquad \rho(t)\equiv\sqrt{(t+a^2)(t+b^2)(t+c^2)}\ .$$

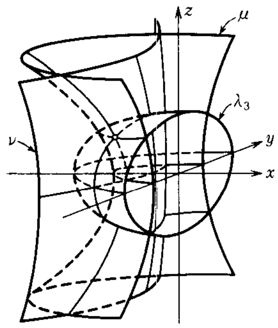


Fig. 11

4. Differential Geometry

(I) Classical Differential Geometry (→ 111 Differential Geometry of Curves and Surfaces)

(1) Plane Curves (Fig. 12). At a point $P(x_0,y_0)$ on a curve $y=f(x)$, the equation of the tangent line is $y-y_0=f'(x_0)(x-x_0)$,

$$PT=\left|y_0\sqrt{1+y_0'^2}\,/y_0'\right|,$$

and the tangential shadow $TM=y_0/y_0'$. The equation of the normal line is $f'(x_0)(y-y_0)+(x-x_0)=0$,

$$PN=\left|y_0\sqrt{1+y_0'^2}\,\right|,$$

and the normal shadow $MN=y_0y_0'$. The slope of the tangent is $\tan\alpha=f'(x_0)=y_0'$. The curvature at P is

$$\kappa=1/PQ=f''(x_0)/\left[1+f'(x_0)^2\right]^{3/2}$$

The coordinates of the center of curvature Q are

$$\left(x_0-f'(x_0)\left[1+f'(x_0)^2\right]/f''(x_0),\ \ f(x_0)+[1+f'(x_0)^2]/f''(x_0)\right).$$

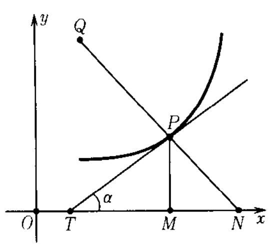


Fig. 12

(2) Space Curves $x_i=x_i(t)$ $(i=1,2,3)$, or $\mathbf{x}=\mathbf{x}(t)$. The line element of a curve $\mathbf{x}=\mathbf{x}(t)$ is

$$ds=\sqrt{(dx_1)^2+(dx_2)^2+(dx_3)^2}=\sqrt{\sum_{\alpha=1}^{3}\dot{x}_\alpha^2}\ dt \qquad \left(\cdot=\frac{d}{dt}\right).$$

The curvature is

$$\kappa=\sqrt{\sum \ddot{x}_\alpha^2-\ddot{s}^2/\dot{s}^2}$$

For $t=s$ (arc length), the curvature is $\kappa=\sqrt{(\sum x_\alpha''^2)}$, and the torsion is $\tau=[\det(x_\alpha', x_\alpha'', x_\alpha''')_{\alpha=1,2,3}]/\kappa^2$, where $'=d/ds$. When we denote Frenet's frame by (ξ,η,ζ), we have $\xi=\mathbf{x}'$, $\eta=\xi'/\kappa$, $\zeta=\xi\times\eta$ (vector product).

The Frenet-Serret formulas are

$$\xi' = \kappa\eta, \quad \eta' = -\kappa\xi + \tau\zeta, \quad \zeta' = -\tau\eta.$$

(3) Surface in 3-Dimensional Space $x_\alpha = x_\alpha(u_1, u_2)$ $(\alpha = 1,2,3)$. The first fundamental form of the surface is

$$g_{jk} = \sum_{\alpha=1}^{3} \frac{\partial x_\alpha}{\partial u_j}\frac{\partial x_\alpha}{\partial u_k} \quad (j,k=1,2). \quad g = \det(g_{jk}) > 0.$$

Let (g^{jk}) be the inverse matrix of (g_{jk}). The tangent plane at the point $x_\alpha^{(0)}$ is given by

$$\det\left(x_\alpha - x_\alpha^{(0)}, \quad (\partial x_\alpha/\partial u_1)^{(0)}, \quad (\partial x_\alpha/\partial u_2)^{(0)}\right) = 0.$$

The normal line at the point $x_\alpha^{(0)}$ is given by $x_\alpha - x_\alpha^{(0)} = t\nu_\alpha^{(0)}$, where t is a parameter, and ν_α is the unit normal vector, given by

$$\nu_\alpha = \frac{1}{\sqrt{g}}\frac{\partial(x_\beta, x_\gamma)}{\partial(u_1,u_2)} \quad \left(\delta_{\alpha\beta\gamma}^{123} = +1\right).$$

The second fundamental form is

$$h_{jk} \equiv \sum_{\alpha=1}^{3} \nu_\alpha \frac{\partial^2 x_\alpha}{\partial u_j \partial u_k} = -\sum_{\alpha=1}^{3} \frac{\partial \nu_\alpha}{\partial u_j}\frac{\partial x_\alpha}{\partial u_k}. \quad h \equiv \det(h_{jk}).$$

The principal radii of curvature R_1, R_2 are the roots of the quadratic equation

$$\frac{1}{R^2} - \sum_{j,k} g^{jk} h_{jk} \frac{1}{R} + \frac{h}{g} = 0.$$

The mean curvature (or Germain's curvature) is

$$H \equiv \frac{1}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) = \frac{1}{2}\sum_{j,k} g^{jk} h_{jk},$$

and $H=0$ is the condition for the given surface to be a minimal surface. The Gaussian curvature (or total curvature) is

$$K \equiv \frac{1}{R_1 R_2} = \frac{h}{g},$$

and $K=0$ is the condition for the surface to be developable.

We use the notations of Riemannian geometry, with g_{jk} the fundamental tensor:

$$\frac{\partial^2 x_\alpha}{\partial u_j \partial u_k} = \sum_{a=1}^{3} \left\{ {a \atop jk} \right\} \frac{\partial x_\alpha}{\partial u_a} + h_{jk}\nu_\alpha \qquad \text{(Gauss's formula)}.$$

$$R_{ijkl} = h_{jl}h_{ik} - h_{jk}h_{il} \quad \text{(Gauss's equation)}. \qquad h_{jk;l} = h_{jl;k} \quad \text{(Codazzi-Mainardi equation)}.$$

$$K = \frac{R_{1212}}{g} = \frac{R}{2} = \frac{1}{\sqrt{g}}\left[\frac{\partial}{\partial u_2}\left(\frac{\sqrt{g}}{g_{11}}\left\{ {2 \atop 11} \right\}\right) - \frac{\partial}{\partial u_1}\left(\frac{\sqrt{g}}{g_{11}}\left\{ {2 \atop 12} \right\}\right)\right]$$

$$= \frac{1}{\sqrt{g}}\left[\frac{\partial}{\partial u_1}\left(\frac{\sqrt{g}}{g_{22}}\left\{ {1 \atop 22} \right\}\right) - \frac{\partial}{\partial u_2}\left(\frac{\sqrt{g}}{g_{22}}\left\{ {1 \atop 12} \right\}\right)\right].$$

$$\frac{\partial \nu_\alpha}{\partial u_j} = -\sum_{k,l} h_{jk} g^{kl} \frac{\partial x_\alpha}{\partial u_l} \qquad \text{(Weingarten's formula)}.$$

The third fundamental form is given by

$$l_{jk} \equiv \sum_{\alpha=1}^{3} \frac{\partial \nu_\alpha}{\partial u_j}\frac{\partial \nu_\alpha}{\partial u_k} = \sum_{s,t} g^{st} h_{js} h_{kt} = 2Hh_{jk} - Kg_{jk}. \quad \det(l_{jk}) = K^2 g = Kh.$$

(4) Geodesic Curvature. Let $C: u_i = u_i(s)$ be a curve on a surface S and ρ be the curvature of C at a point P. Let θ be the angle between the osculating plane of C and the plane tangent to S. The geodesic curvature ρ_g of C at P is given by

$$\rho_g = \rho\cos\theta = \sqrt{g}\ \det\left(\frac{du_i}{ds},\ \frac{d^2u_i}{ds^2} + \sum_{j,k}\left\{ {i \atop jk} \right\}\frac{du_j}{ds}\frac{du_k}{ds}\right)_{i=1,2}$$

$\rho_g = 0$ is the condition for C to be a geodesic. Let D be a simply connected domain on the surface S, whose boundary Γ consists of n smooth curves. Let θ_α be the outer angle at the intersection of two consecutive curves ($\alpha = 1, \ldots, n$). Then we have the Gauss-Bonnet formula:

$$\int_\Gamma \rho_g ds + \iint_D K dS = 2\pi - \sum_{\alpha=1}^{n} \theta_\alpha.$$

(II) Riemannian Geometry, Tensor Calculus (→ 417 Tensor Calculus)

In the present section, we use Einstein's convention (omission of the summation symbol).

(1) Numerical Tensor.

Kronecker's δ $\quad \delta_{jk},\ \delta^{jk},\ \delta^j_k = \begin{cases} 1 & (j=k) \\ 0 & (j \neq k). \end{cases}$

$$\delta^{j_1 \ldots j_p}_{k_1 \ldots k_p} = \det\left(\delta^{j_\mu}_{k_\nu}\right)_{\mu,\nu=1,\ldots,p} = \begin{cases} 0 & (\{j_\mu\} \neq \{k_\nu\}), \\ +1 & (\{j_\mu\} = \{k_\nu\} \text{ and } (j_\mu) \text{ is an even permutation of } (k_\nu)), \\ -1 & (\{j_\mu\} = \{k_\nu\}) \text{ and } (j_\mu) \text{ is an odd permutation of } (k_\nu). \end{cases}$$

Eddington's ε $\quad \varepsilon_{j_1 \ldots j_n} = \delta^{1 \ldots n}_{j_1 \ldots j_n}, \quad \varepsilon^{j_1 \ldots j_n} = \delta^{j_1 \ldots j_n}_{1 \ldots n}.$

$\delta^{j_1 \ldots j_p j_{p+1} \ldots j_n}_{k_1 \ldots k_p j_{p+1} \ldots j_n} = (n-p)!\, \delta^{j_1 \ldots j_p}_{k_1 \ldots k_p}. \quad \det(a^u_\nu)_{\mu,\nu=1,\ldots,n} = \varepsilon^{j_1 \ldots j_n} a^1_{j_1} a^2_{j_2} \ldots a^n_{j_n} = \varepsilon_{j_1 \ldots j_n} a^{j_1}_1 a^{j_2}_2 \ldots a^{j_n}_n.$

(2) Fundamental Objects in Riemannian Geometry. Let g_{jk} be the fundamental tensor, and (g^{jk}) be the inverse matrix of (g_{jk}). We put $g \equiv \det(g_{jk})$.

The Christoffel symbol is

$$\left\{ {i \atop jk} \right\} = \frac{1}{2} g^{ia} \left[\frac{\partial g_{aj}}{\partial x^k} + \frac{\partial g_{ak}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^a} \right] = \left\{ {i \atop kj} \right\}, \quad \left\{ {a \atop ak} \right\} = \frac{\partial \log \sqrt{g}}{\partial x^k},$$

which has the transformation rule

$$\left\{ {\bar{i} \atop \bar{j}\bar{k}} \right\} = \frac{\partial \bar{x}^{\bar{i}}}{\partial x^i} \left(\frac{\partial x^j}{\partial \bar{x}^{\bar{j}}} \frac{\partial x^k}{\partial \bar{x}^{\bar{k}}} \left\{ {i \atop jk} \right\} + \frac{\partial^2 x^i}{\partial \bar{x}^{\bar{j}} \partial \bar{x}^{\bar{k}}} \right)$$

under a coordinate transformation.

A geometrical object Γ^i_{jk} with a similar transformation rule is called the coefficient of the affine connection. The torsion tensor is

$$S^i_{jk} \equiv \Gamma^i_{jk} - \Gamma^i_{kj}.$$

The equation of a geodesic is

$$\frac{d^2 x^i}{ds^2} + \left\{ {i \atop jk} \right\} \frac{dx^j}{ds} \frac{dx^k}{ds} = 0.$$

The covariant derivative of a tensor of weight W with respect to a coefficient of affine connection Γ^i_{jk} is given by

$$T^{j_1 \ldots j_p}_{k_1 \ldots k_q | l} \equiv \partial T^{j_1 \ldots j_p}_{k_1 \ldots k_q} / \partial x^l + \sum_{\nu=1}^{p} T^{j_1 \ldots j_{\nu-1} a j_{\nu+1} \ldots j_p}_{k_1 \ldots \ldots \ldots k_q} \Gamma^{j_\nu}_{al} - \sum_{\mu=1}^{q} T^{j_1 \ldots \ldots \ldots j_p}_{k_1 \ldots k_{\mu-1} a k_{\mu+1} \ldots k_q} \Gamma^a_{l k_\mu} - W T^{j_1 \ldots j_p}_{k_1 \ldots k_q} \Gamma^a_{al}.$$

For the Christoffel symbol, we denote the covariant derivative by $;l$. Then we have the following formulas:

$$g_{jk;l} = 0, \quad g^{jk}{}_{;l} = 0, \quad \delta^j_{k;l} = 0, \quad \sqrt{g}\, \varepsilon_{j_1 \ldots j_n; l} = 0, \quad (1/\sqrt{g})\, \varepsilon^{j_1 \ldots j_n}{}_{;l} = 0.$$

For a scalar f $\quad \operatorname{grad} f = (f_{;j})$,
for a covariant vector v_j $\quad \operatorname{rot} v = (v_{j;k} - v_{k;j}) = (\partial v_j / \partial x^k - \partial v_k / \partial x^j)$,
and for a contravariant vector v^j $\quad \operatorname{div} v = v^j{}_{;j} = \dfrac{1}{\sqrt{g}} \dfrac{\partial(\sqrt{g} v^j}{\partial x^j}.$

Beltrami's differential operator of the first kind is

$$\Delta_1 f \equiv g^{jk} f_{;j} f_{;k}.$$

Beltrami's differential operator of the second kind is

$$\Delta_2 f = \operatorname{div}\operatorname{grad} f = \frac{1}{\sqrt{g}}\frac{\partial\left(\sqrt{g}\, g^{jk}(\partial f/\partial x^k)\right)}{\partial x^j}.$$

For a domain D with sufficiently smooth boundary Γ, we denote the directional derivative along the inner normal by $\partial/\partial n$, the volume element by dV, and the surface element on Γ by dS. Then we have Green's formulas,

$$\int_D (\Delta_1(\varphi\psi) + \psi\Delta_2\varphi)\, dV = -\int_\Gamma \psi\frac{\partial\varphi}{\partial n}\, dS, \qquad \int_D (\varphi\Delta_2\psi - \psi\Delta_2\varphi)\, dV = \int_\Gamma \left(\psi\frac{\partial\varphi}{\partial n} - \varphi\frac{\partial\psi}{\partial n}\right) dS.$$

We denote the curvature tensor with respect to the coefficients of a general affine connection Γ^i_{jk} by B^i_{jkl}, and by R^i_{jkl} when $\Gamma^i_{jk} = \left\{ {i \atop jk} \right\}$. We have the following formulas:

$$B^i_{jkl} \equiv \frac{\partial\Gamma^i_{jl}}{\partial x^k} - \frac{\partial\Gamma^i_{jk}}{\partial x^l} + \Gamma^a_{jl}\Gamma^i_{ak} - \Gamma^a_{jk}\Gamma^i_{al};$$

$$R_{ijkl} \equiv g_{ai}R^a_{jkl} = -R_{ijlk} = R_{klij} = \frac{1}{2}\left[\frac{\partial^2 g_{ik}}{\partial x^j\partial x^l} + \frac{\partial^2 g_{jl}}{\partial x^i\partial x^k} - \frac{\partial^2 g_{jk}}{\partial x^i\partial x^l} - \frac{\partial^2 g_{il}}{\partial x^j\partial x^k}\right]$$

$$+ g_{ab}\left(\left\{ {b \atop ik} \right\}\left\{ {a \atop jl} \right\} - \left\{ {b \atop il} \right\}\left\{ {a \atop jk} \right\}\right);$$

Bianchi's first identity $\quad R^i_{jkl} + R^i_{klj} + R^i_{ljk} = 0,$

$$-\left(B^i_{jkl} + B^i_{klj} + B^i_{ljk}\right) = 2\left(S^i_{jk|l} + S^i_{kl|j} + S^i_{lj|k}\right) + 4\left(S^i_{ja}S^a_{kl} + S^i_{ka}S^a_{lj} + S^i_{la}S^a_{jk}\right);$$

Bianchi's second identity $\quad R^i_{jkl;m} + R^i_{jlm;k} + R^i_{jmk;l} = 0,$

$$B^i_{jkl|m} + B^i_{jlm|k} + B^i_{jmk|l} = -2\left(B^i_{jma}S^a_{kl} + B^i_{jka}S^a_{lm} + B^i_{jla}S^a_{mk}\right);$$

Ricci's tensor $\quad R_{jk} \equiv -R^i_{jki} = R_{kj};$

scalar curvature $\quad R \equiv g^{jk}R_{jk};$

Ricci's formula $\quad T^{j_1\cdots j_p}_{k_1\cdots k_q|s|t} - T^{j_1\cdots j_p}_{k_1\cdots k_q|t|s}$

$$= -\sum_{\nu=1}^{p} T^{j_1\cdots j_{\nu-1}aj_{\nu+1}\cdots j_p}_{k_1\cdots\cdots\cdots k_q}B^{j_\nu}_{ast} + \sum_{\mu=1}^{q} T^{j_1\cdots\cdots\cdots j_p}_{k_1\ldots k_{\mu-1}ak_{\mu+1},\ldots,k_q}B^a_{k_\mu st} + 2T^{j_1\cdots j_p}_{k_1\cdots k_q|l}S^l_{st} + WT^{j_1\cdots j_p}_{k_1\cdots k_q}B^a_{ast},$$

where S and B are the torsion and curvature tensors given above, respectively, and W is the weight of the tensor T.

(3) Special Riemannian Spaces (→ 364 Riemannian Manifolds). In the present Section (3), n means the dimension of the space.

(i) Space of Constant Curvature $R^i_{jkl} = \rho(g_{jl}\delta^i_k - g_{jk}\delta^i_l)$; $\rho = R/n(n-1)$,

(ii) Einstein Space $R_{jk} = \rho g_{jk}$, $\rho = R/n$, for $n \geqslant 3$, where R is a constant.

(iii) Locally Symmetric Riemannian Space $R^i_{jkl;m} = 0$.

(iv) Projectively Flat Space. Weyl's projective curvature tensor is defined by

$$W^i_{jkl} \equiv R^i_{jkl} + \frac{1}{n-1}(R_{jk}\delta^i_l - R_{jl}\delta^i_k).$$

The condition for the space to be projectively flat is given by $W^i_{jkl} = 0$, $R_{jk;l} = R_{jl;k}$.

If $n \geqslant 3$, the latter condition follows from the former condition, and the space reduces to a space of constant curvature. If $n = 2$, the former condition $W = 0$ always holds.

(v) Concircularly Flat Space $Z^i_{jkl} \equiv R^i_{jkl} + \dfrac{R}{n(n-1)}(g_{jk}\delta^i_l - g_{jl}\delta^i_k) = 0$. This space reduces to a space of constant curvature.

(vi) Conformally Flat Space. Weyl's conformal curvature tensor is defined by

$$C^i_{jkl} \equiv R^i_{jkl} + \frac{1}{n-2}(R_{jk}\delta^i_l - R_{jl}\delta^i_k + g_{jk}R^i_l - g_{jl}R^i_k) - \frac{R(g_{jk}\delta^i_l - g_{jl}\delta^i_k)}{(n-1)(n-2)},$$

$$\Pi_{jk} \equiv -\frac{R_{jk}}{(n-2)} + \frac{Rg_{jk}}{2(n-1)(n-2)}.$$

The condition for the space to be conformally flat is given by $C^i_{jkl} = 0$, $\Pi_{jk;l} = \Pi_{jl;k}$.

If $n \geqslant 4$, the latter condition follows from the former condition, and if $n = 3$, the former condition $C = 0$ always holds.

5. Lie Algebras, Symmetric Riemannian Spaces, and Singularities

(I) The Classification of Complex Simple Lie Algebras and Compact Real Simple Lie Algebras
(→ 248 Lie Algebras)

(1) Lie Algebra. The unitary restriction of a noncommutative finite-dimensional complex simple Lie algebra $\mathfrak{g}$ is a compact real simple Lie algebra $\mathfrak{g}_u$, and $\mathfrak{g}$ is given by the complexification $\mathfrak{g}_u^c$ of $\mathfrak{g}_u$. There exists a bijective correspondence between the classifications of these two kinds of Lie algebras. Using Dynkin diagrams, the classification is done as in Fig. 14 (→ 248 Lie Algebras). The system of fundamental roots $(\alpha_1, \ldots, \alpha_l\}$ of a simple Lie algebra $\mathfrak{g}$ is in one-to-one correspondence with the vertices of a Dynkin diagram shown by simple circles in Fig. 14. The number of simple circles coincides with the rank l of $\mathfrak{g}$. The double circle in Fig. 14 means -1 times the highest root θ. Sometimes we mean by the term "Dynkin diagram" the diagram without the double circle and the lines issuing from it. Here we call the diagram with double circle representing $-\theta$ the extended Dynkin diagram. Corresponding to the value of the inner product with respect to the Killing form $-2(\alpha_i, \alpha_j)/(\alpha_j, \alpha_j)$ $(i \neq j)$ (which must be 0, 1, 2, or 3), we connect two vertices representing α_i and α_j as in Fig. 13. When the value is 0, we do not connect α_i and α_j. In Fig. 13, the left circle corresponds to α_i and the right circle to α_j.

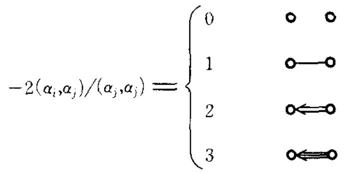


Fig. 13

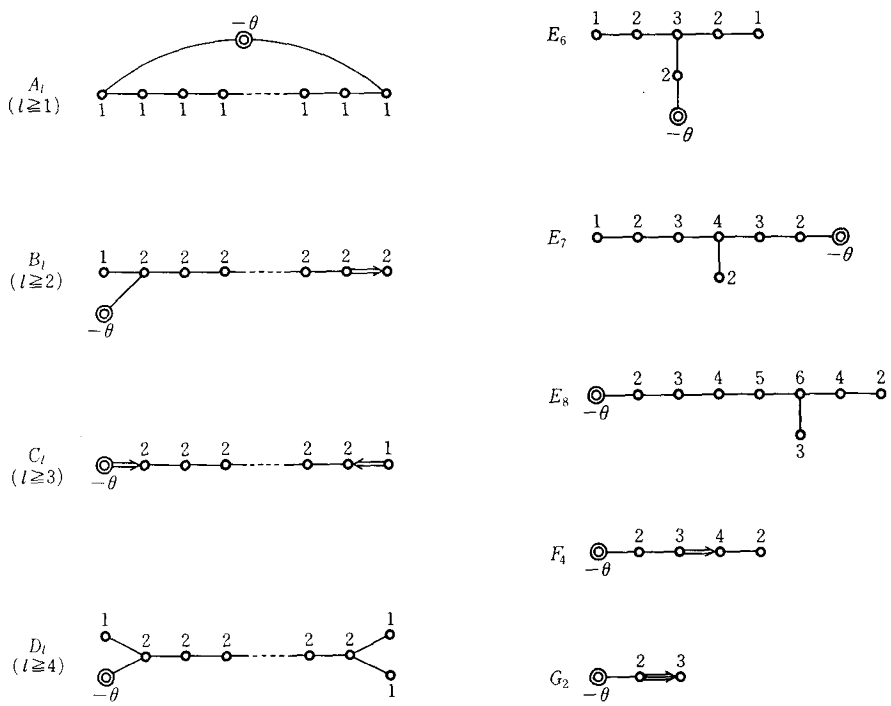


Fig. 14 We have relations $B_1 = C_1 = A_1$, $C_2 = B_2$, and $D_3 = A_3$. ($D_2 = A_1 + A_1$, which is not simple.) In this figure, the number at each vertex means the coefficient m_i in $\theta = \sum m_i \alpha_i$.

From Fig. 14, we have the following information.
(i) The quotient group of the automorphism group $A(\mathfrak{g})$ of $\mathfrak{g}$ with respect to the inner automorphism group $I(\mathfrak{g})$ is isomorphic to the automorphism group of the corresponding Dynkin diagram. The order of the latter group is 2 for A_l $(l \geqslant 2)$ since the diagram is symmetric. It is also 2 for D_l $(l \geqslant 5)$ and for E_6, and it is 6 $(=3!)$ for D_4. For all other cases, the order is 1.

(ii) The order of the center of the simply connected Lie group associated with $\mathfrak{g}$ is equal to the index of the subgroup consisting of elements stabilizing $-\theta$ in the group of automorphisms of the extended Dynkin diagram of $\mathfrak{g}$ (S. Murakami). This index is equal to the order of the fundamental group of the adjoint group of $\mathfrak{g}$ and the number of connected Lie groups, whose Lie algebra is $\mathfrak{g}$.
(iii) Any parabolic Lie subalgebra of $\mathfrak{g}$ is isomorphic to a subalgebra generated by the root vector X_α (and elements of the Cartan subalgebra) such that $\alpha=\sum n_i\alpha_i$, where $\{\alpha_1,\ldots,\alpha_l\}$ is a system of fundamental roots, $n_i\geqslant 0$ $(i=1,\ldots,l)$ or $n_i\leqslant 0$ $(i=1,\ldots,l)$, and $n_j=0$ for α_j belonging to a fixed subset S of $\{\alpha_1,\ldots,\alpha_l\}$.
Hence, isomorphism classes of parabolic Lie subalgebras are in one-to-one correspondence with the set of subsets S of $\{\alpha_1,\ldots,\alpha_l\}$.
(iv) Maximal Lie subalgebra $\mathfrak{k}$ of $\mathfrak{g}$ with the same rank l as $\mathfrak{g}$. The Lie subalgebra $\mathfrak{k}$ is classified by the following rule. First we remove a vertex α_i from the Dynkin diagram. If the number m_i attached to the vertex is 1, $\mathfrak{k}$ is given by the product of the simple Lie algebra corresponding to the Dynkin diagram after removing the vertex α_i and a one-dimensional Lie subalgebra. If $m_i>1$, $\mathfrak{k}$ is given by the diagram after removing α_i from the extended Dynkin diagram.

(2) Lie Groups. The classical complex simple Lie groups of rank n represented by A, B, C, D (in Cartan's symbolism) are the complex special linear group $SL(n+1,\mathbf{C})$, the complex special orthogonal group $SO(2n+1,\mathbf{C})$, the complex symplectic group $Sp(n,\mathbf{C})$, and the complex special orthogonal group $SO(2n,\mathbf{C})$, respectively. The classical compact simple Lie groups of rank n represented by A, B, C, D are the special unitary group $SU(n+1)$, the special orthogonal group $SO(2n+1)$, the unitary-symplectic group $Sp(n)$, and the special orthogonal group $SO(2n)$, respectively (→ 60 Classical Groups).

Cartan's Symbol	Complex Form	Compact Form	Dimension	Rank
A_n	$SL(n+1,\mathbf{C})$	$SU(n+1)$	$(n+1)^2-1$	n
B_n	$SO(2n+1,\mathbf{C})$	$SO(2n+1)$	$2n^2+n$	n
C_n	$Sp(n,\mathbf{C})$	$Sp(n)$	$2n^2+n$	n
D_n	$SO(2n,\mathbf{C})$	$SO(2n)$	$2n^2-n$	n
G_2	$\mathrm{Aut}\,\mathfrak{C}^c$	$\mathrm{Aut}\,\mathfrak{C}$	14	2
F_4	$\mathrm{Aut}\,\mathfrak{J}^c$	$\mathrm{Aut}\,\mathfrak{J}$	52	4
E_6			78	6
E_7			133	7
E_8			248	8

Here $\mathfrak{C}$ is the Cayley algebra over $\mathbf{R}$, $\mathfrak{C}^c$ is the complexification of $\mathfrak{C}$, $\mathfrak{J}$ is the Jordan algebra of Hermitian matrices of order 3 over $\mathfrak{C}$, $\mathfrak{J}^c$ is the complexification of $\mathfrak{J}$, and $\mathrm{Aut}A$ is the automorphism group of A.

(II) Classification of Noncompact Real Simple Lie Algebras

Classical Cases

Cartan's Symbol	Noncompact Real Simple Lie Algebra $\mathfrak{g}$	Maximal Compact Lie Algebra of $\mathfrak{g}$
AI	$\mathfrak{sl}(p+1;\mathbf{R})$	$\mathfrak{su}(p+1)$
AII	$\mathfrak{sl}(n;\mathbf{H})$	$\mathfrak{sp}(n)$
AIII	$\mathfrak{su}(p,q;\mathbf{C})$	$\mathfrak{su}(p)+\mathfrak{su}(q)$
BI	$\mathfrak{so}(p,q;\mathbf{R})$	$\mathfrak{so}(p)+\mathfrak{so}(q)\quad(p+q=2m+1)$
BII	$\mathfrak{so}(1,n-1;\mathbf{R})$	$\mathfrak{so}(n-1)\quad(n=2m+1)$
CI	$\mathfrak{sp}(p;\mathbf{R})$	$\mathfrak{u}(p)$
CII	$\mathfrak{u}(p,q;\mathbf{H})$	$\mathfrak{sp}(p)+\mathfrak{sp}(q)$
DI	$\mathfrak{so}(p,q;\mathbf{R})$	$\mathfrak{so}(p)+\mathfrak{so}(q)\quad(p+q=2m)$
DII	$\mathfrak{so}(1,n-1;\mathbf{R})$	$\mathfrak{so}(n-1)\quad(n=2m)$
DIII	$\mathfrak{so}(p;\mathbf{H})$	$\mathfrak{u}(2p)$

Here the field F is the real field **R**, the complex field **C**, or the quaternion field **H** ($\mathbf{R}\subset\mathbf{C}\subset\mathbf{H}$). **H** is an algebra over **R**. For a quaternion $x=x_0+x_1i+x_2j+x_3k\,(x_0,x_1,x_2,x_3\in\mathbf{R})$, we put

$$\bar{x}=x_0-x_1i-x_2j-x_3k,$$
$$x^*=x_0+x_1i-x_2j+x_3k.$$

Then $\mathfrak{gl}(n;F)=\{$set of all square matrices over F of order $n\}$,

$$\mathfrak{sl}(n;F)=\{A\in\mathfrak{gl}(n;F)\,|\,\mathrm{tr}A=0\},$$
$$\mathfrak{so}(p,q;F)=\{A\in\mathfrak{gl}(p+q;F)\,|\,{}^tA^*I_{p,q}+I_{p,q}A=0\},$$

where $I_{p,q}$ is the symmetric transformation of the Euclidean space $\mathbf{R}^{p+q}$ with respect to $\mathbf{R}^p$, i.e.,

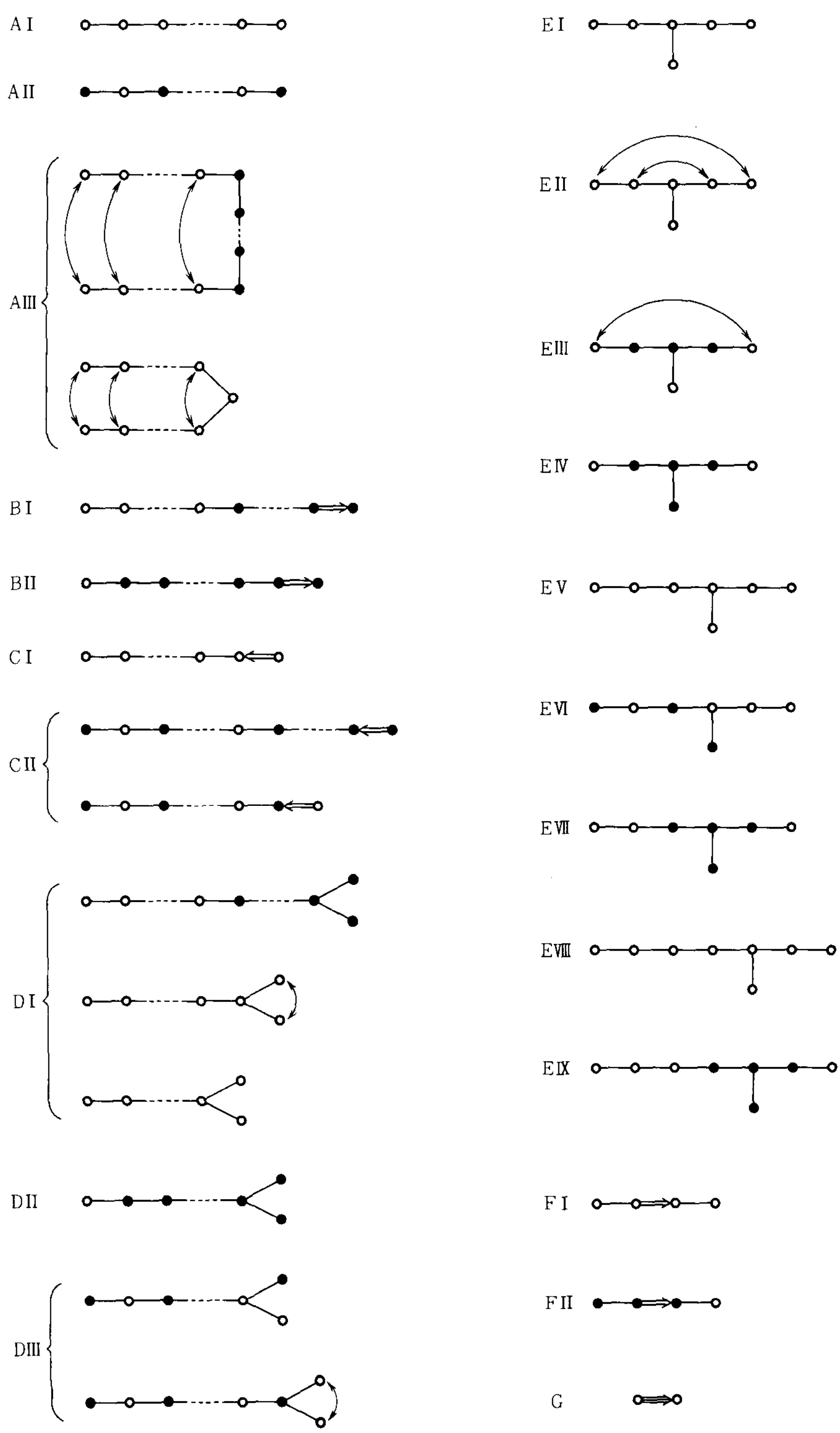


Fig. 15

$I_{p,q}$ is the diagonal sum of the unit matrix I_p of order p and $-I_q$. We have

$$\mathfrak{so}(n;F)=\mathfrak{so}(n,0;F),$$
$$\mathfrak{u}(p,q;F)=\left\{A\in\mathfrak{gl}(p+q;F)\,\middle|\,{}^t\bar{A}I_{p,q}+I_{p,q}A=0\right\},$$
$$\mathfrak{u}(n;F)=\mathfrak{u}(n,0;F),$$
$$\mathfrak{sp}(n;F)=\left\{A\in\mathfrak{gl}(2n;F)\,\middle|\,{}^t\bar{A}J+JA=0\right\},$$

where J is the matrix of an alternating form $\Sigma_{i=1}^{n}(x_i y_{i+n}-x_{i+n}y_i)$ of order $2n$.

A noncompact real simple Lie algebra $\mathfrak{g}$ is classified by the relation of the complex conjugation operator σ with respect to the complexification $\mathfrak{g}^c$ of $\mathfrak{g}$. The results are given by Satake's diagram (Fig. 15).

In the diagram, the fundamental root corresponding to a black circle is multiplied by -1 under σ for a suitable choice of Cartan subalgebra, and the arc with an arrow means that two elements corresponding to both ends of the arc are mutually transformed by a special transformation p such that $\sigma=pw$ $(w\in W)$.

(III) Classification of Irreducible Symmetric Riemannian Spaces (→ 412 Symmetric Riemannian Spaces and Real Forms)

A simply connected irreducible symmetric Riemannian space $M=G/K$ is either a space in the following table or a simply connected compact simple Lie group mentioned in (I). The noncompact forms uniquely corresponding to the compact symmetric Riemannian space are in one-to-one correspondence with the noncompact real simple Lie algebras mentioned in (II).

Cartan's Symbol	$G/K=M$	Dimension	Rank
AI	$SU(n)/SO(n)$ $(n>2)$	$(n-1)(n+2)/2$	$n-1$
AII	$SU(2n)/Sp(n)$ $(n>1)$	$(n-1)(2n+1)$	$n-1$
AIII	$U(p+q)/U(p)\times U(q)$ $(p\geqslant q\geqslant 1)$	$2pq$	q
BDI	$SO(p+q)/SO(p)\times SO(q)$ $(p\geqslant q\geqslant 2, p+q\neq 4)$	pq	q
BDII	$SO(n+1)/SO(n)$ $(n\geqslant 2)$	n	1
DIII	$SO(2l)/U(l)$ $(l\geqslant 4)$	$l(l-1)$	$[l/2]$
CI	$Sp(n)/U(n)$ $(n\geqslant 3)$	$n(n+1)$	n
CII	$Sp(p+q)/Sp(p)\times Sp(q)$ $(p\geqslant q\geqslant 1)$	$4pq$	q
EI	$E_6/Sp(4)$	42	6
EII	$E_6/SU(2)\cdot SU(6)$	40	4
EIII	$E_6/Spin(10)\cdot SO(2)$	32	2
EIV	E_6/F_4	26	2
EV	$E_7/SU(8)$	70	7
EVI	$E_7/Spin(12)\cdot SU(2)$	64	4
EVII	$E_7/E_6\cdot SO(2)$	54	3
EVIII	$E_8/Spin(16)$	128	8
EIX	$E_8/E_7\cdot SU(2)$	112	4
FI	$F_4/Sp(3)\cdot SU(2)$	28	4
FII	$F_4/Spin(9)$	16	1
G	$G_2/SO(4)$	8	2

Notes

The group $G=U(p+q)$ in AIII is not effective, unless it is replaced by $SU(p+q)$. To be precise, $K=Sp(4)$ in EI should be replaced by its quotient group factored by a subgroup of order 2 of its center. K in EII is not a direct product of simple groups; the order of its fundamental group $\pi_1(K)$ is 2. To be precise, K in EV or EVIII should be replaced by its quotient group factored by a subgroup of order 2 of its center. The K's in EII, EIII, EVI, EVII, EIX, and FI are not direct products. The fundamental group $\pi_1(K)$ of K is the infinite cyclic group $\mathbf{Z}$ for EIII, EVII; for all other cases, the order of $\pi_1(K)$ is 2.

In EIII, EVII, the groups E_6, E_7 are adjoint groups of compact simple Lie algebras. In other cases, E_6 and E_7 (E_8, F_4 and G_2 also) are simply connected Lie groups.

The compact symmetric Riemannian space M is a complex Grassmann manifold for AIII, a real Grassmann manifold for BDI, a sphere for BDII, a quaternion Grassmann manifold for CII, and a Cayley projective plane for FII.

(IV) Isomorphic Relations among Classical Lie Algebras

The isomorphic relations among the classical Lie algebras over **R** or **C** are all given in the following table. In the table, we denote, for example, the real form of type AI of the complex Lie algebra with rank 3 by A_3I in Cartan's symbolism. When there are nonisomorphic real forms of the same type and same rank (e.g., in the case of D_3I) we distinguish them by the rank of the corresponding symmetric Riemannian space and denote them by, e.g., D_3I_p, where p is the index of total isotropy of the sesquilinear form which is invariant under the corresponding Lie algebra.

Cartan's Symbol	Isomorphisms among Classical Lie Algebras
$A_1=B_1=C_1$	$\mathfrak{sl}(2,\mathbf{C})\cong\mathfrak{so}(3,\mathbf{C})\cong\mathfrak{sp}(1,\mathbf{C})$; $\mathfrak{su}(2)\cong\mathfrak{so}(3)\cong\mathfrak{sp}(1)$
$B_2=C_2$	$\mathfrak{so}(5,\mathbf{C})\cong\mathfrak{sp}(2,\mathbf{C})$; $\mathfrak{so}(5)\cong\mathfrak{sp}(2)$
$A_3=D_3$	$\mathfrak{sl}(4,\mathbf{C})\cong\mathfrak{so}(6,\mathbf{C})$; $\mathfrak{su}(4)\cong\mathfrak{so}(6)$
$A_1I=A_1III=B_1I=C_1I$	$\mathfrak{sl}(2,\mathbf{R})\cong\mathfrak{su}(1,1;\mathbf{C})\cong\mathfrak{so}(2,1;\mathbf{R})\cong\mathfrak{sp}(1;\mathbf{R})$
$B_2I_2=C_2I$	$\mathfrak{so}(3,2;\mathbf{R})\cong\mathfrak{sp}(2,\mathbf{R})$
$B_2I_1=C_2II$	$\mathfrak{so}(4,1;\mathbf{R})\cong\mathfrak{u}(1,1;\mathbf{H})$
$A_3I=D_3I_3$	$\mathfrak{sl}(4,\mathbf{R})\cong\mathfrak{so}(3,3;\mathbf{R})$
$A_3II=D_3I_1$	$\mathfrak{sl}(2,\mathbf{H})\cong\mathfrak{so}(5,1;\mathbf{R})$
$A_3III_2=D_3I_2$	$\mathfrak{su}(2,2;\mathbf{C})\cong\mathfrak{so}(4,2;\mathbf{R})$
$A_3III_1=D_3III$	$\mathfrak{su}(3,1;\mathbf{C})\cong\mathfrak{so}(3;\mathbf{H})$
$D_4I_2=D_4III$	$\mathfrak{so}(6,2;\mathbf{R})\cong\mathfrak{so}(4,\mathbf{H})$
$D_2=A_1\times A_1$*	$\mathfrak{so}(4,\mathbf{C})\cong\mathfrak{sl}(2,\mathbf{C})\times\mathfrak{sl}(2,\mathbf{C})$; $\mathfrak{so}(4)\cong\mathfrak{su}(2)\times\mathfrak{su}(2)$
$D_2I_2=A_1I\times A_1I$	$\mathfrak{so}(2,2;\mathbf{R})\cong\mathfrak{sl}(2,\mathbf{R})\times\mathfrak{sl}(2;\mathbf{R})$
$D_2III=A_1\times A_1I$*	$\mathfrak{so}(2;\mathbf{H})\cong\mathfrak{su}(2)\times\mathfrak{sl}(2;\mathbf{R})$
$D_2I_1=A_1$*	$\mathfrak{so}(3,1;\mathbf{R})\cong\mathfrak{sl}(2,\mathbf{C})$

Note

(*) In these 3 cases, there are isomorphisms given by the replacement of $\mathfrak{sl}(2,\mathbf{C})$ or $\mathfrak{su}(2)$ by isomorphic Lie algebras of type B_1 or type C_1 due to the isomorphism $A_1\cong B_1\cong C_1$.

(V) Lists of Normal Forms of Singularities with Modulus Number $m=0$, 1, and 2 (→ 418 Theory of Singularities)

Letters $A,\dots,Z$ stand here for stable equivalence classes of function germs (or families of function germs).

(1) Simple Singularities ($m=0$). There are 2 infinite series A, D, and 3 "exceptional" singularities E_6, E_7, E_8:

Notation	Normal form	Restrictions
A_n	$x^{n+1}+y^2+z^2$	$n\geqslant 1$
D_n	$x^{n-1}+xy^2+z^2$	$n\geqslant 4$
E_6	$x^4+y^3+z^2$	
E_7	$x^3y+y^3+z^2$	
E_8	$x^5+y^3+z^2$	

(2) Unimodular Singularities ($m=1$). There are 3 families of parabolic singularities, one series of hyperbolic singularities (with 3 subscripts), and 14 families of exceptional singularities.

The parabolic singularities

Notation	Normal form	Restrictions
$P_8=\tilde{E}_6$	$x^3+y^3+z^3+axyz$	$a^3+27\neq 0$
$X_9=\tilde{E}_7$	$x^4+y^4+z^2+axyz$	$a^4-64\neq 0$
$J_{10}=\tilde{E}_8$	$x^6+y^3+z^2+axyz$	$a^6-432\neq 0$

The hyperbolic singularities

Notation	Normal form	Restrictions
T_{pqr}	$x^p+y^q+z^r+axyz$	$a\neq 0, \frac{1}{p}+\frac{1}{q}+\frac{1}{r}<1$

The 14 exceptional families

Notation	Normal form	Gabrielov numbers	Dolgachëv numbers	Notation	Normal form	Gabrielov numbers	Dolgachëv numbers
K_{12}	$x^3+y^7+z^2+axy^5$	2 3 7	2 3 7	W_{13}	$x^4+xy^4+z^2+ay^6$	2 5 6	3 4 4
K_{13}	$x^3+xy^5+z^2+ay^8$	2 3 8	2 4 5	Q_{10}	$x^3+y^4+yz^2+axy^3$	3 3 4	2 3 9
K_{14}	$x^3+y^8+z^2+axy^6$	2 3 9	3 3 4	Q_{11}	$x^3+y^2z+xz^3+az^5$	3 3 5	2 4 7
Z_{11}	$x^3y+y^5+z^2+axy^4$	2 4 5	2 3 8	Q_{12}	$x^3+y^5+yz^2+axy^4$	3 3 6	3 3 6
Z_{12}	$x^3y+xy^4+z^2+ax^2y^3$	2 4 6	2 4 6	S_{11}	$x^4+y^2z+xz^2+ax^3z$	3 4 4	2 5 6
Z_{13}	$x^3y+y^6+z^2+axy^5$	2 4 7	3 3 5	S_{12}	$x^2y+y^2z+xz^3+az^5$	3 4 5	3 4 5
W_{12}	$x^4+y^5+z^2+ax^2y^3$	2 5 5	2 5 5	U_{12}	$x^3+y^3+z^4+axyz^2$	4 4 4	4 4 4

(3) Bimodular Singularities ($m=2$). There are 8 infinite series and 14 exceptional families. In all the formulas, $\mathbf{a}=a_0+a_1y$.

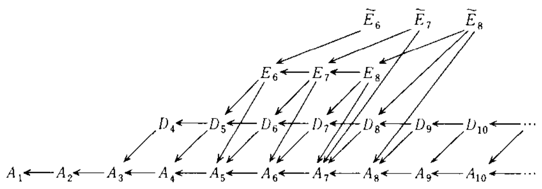
The 8 infinite series of bimodular singularities

Notation	Normal form	Restriction	Milnor number
$J_{3,0}$	$x^3+bx^2y^3+y^9+z^2+cxy^7$	$4b^3+27\neq 0$	16
$J_{3,p}$	$x^3+x^2y^3+z^2+ay^{9+p}$	$p>0, a_0\neq 0$	$16+p$
$Z_{1,0}$	$y(x^3+dx^2y^2+cxy^5+y^6)+z^2$	$4d^3+27\neq 0$	15
$Z_{1,p}$	$y(x^3+x^2y^2+ay^{6+p})+z^2$	$p>0, a_0\neq 0$	$15+p$
$W_{1,0}$	$x^4+ax^2y^3+y^6+z^2$	$a_0^2\neq 4$	15
$W_{1,p}$	$x^4+x^2y^3+ay^{6+p}+z^2$	$p>0, a_0\neq 0$	$15+p$
$W^{\#}_{1,2q-1}$	$(x^2+y^3)^2+axy^{4+q}+z^2$	$q>0, a_0\neq 0$	$15+2q-1$
$W^{\#}_{1,2q}$	$(x^2+y^3)^2+ax^2y^{3+q}+z^2$	$q>0, a_0\neq 0$	$15+2q$
$Q_{2,0}$	$x^3+yz^2+ax^2y^2+xy^4$	$a_0^2\neq 4$	14
$Q_{2,p}$	$x^3+yz^2+x^2y^2+az^{6+p}$	$p>0, a_0\neq 0$	$14+p$
$S_{1,0}$	$x^2z+yz^2+y^5+azy^3$	$a_0^2\neq 4$	14
$S_{1,p}$	$x^2z+yz^2+x^2y^2+ay^{5+p}$	$p>0, a_0\neq 0$	$14+p$
$S^{\#}_{1,2q-1}$	$x^2z+yz^2+zy^3+axy^{2+q}$	$q>0, a_0\neq 0$	$14+2q-1$
$S^{\#}_{1,2q}$	$x^2z+yz^2+zy^3+ax^2y^{2+q}$	$q>0, a_0\neq 0$	$14+2q$
$U_{1,0}$	$x^3+xz^2+xy^3+ay^3z$	$a_0(a_0^2+1)\neq 0$	14
$U_{1,2q-1}$	$x^3+xz^2+xy^3+ay^{1+q}z^2$	$q>0, a_0\neq 0$	$14+2q-1$
$U_{1,2q}$	$x^3+xz^2+xy^3+ay^{3+q}z$	$q>0, a_0\neq 0$	$14+2q$

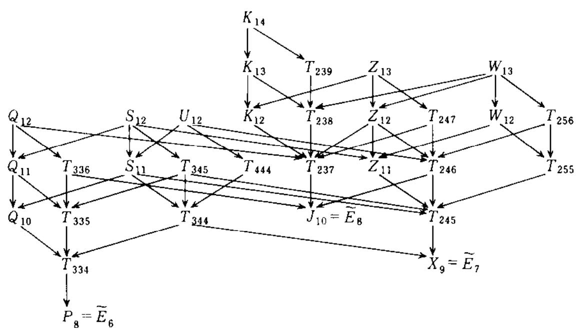
The 14 exceptional families

Notation	Normal form	Notation	Normal form
E_{18}	$x^3+y^{10}+z^2+axy^7$	W_{18}	$x^4+y^7+z^2+ax^2y^4$
E_{19}	$x^3+xy^7+z^2+ay^{11}$	Q_{16}	$x^3+yz^2+y^7+z^2+axy^5$
E_{20}	$x^3+y^{11}+z^2+axy^8$	Q_{17}	$x^3+yz^2+xy^5+z^2+ay^8$
Z_{17}	$x^3y+y^8+z^2+axy^6$	Q_{18}	$x^3+yz^2+y^8+z^2+axy^6$
Z_{18}	$x^3y+xy^6+z^2+ay^9$	S_{16}	$x^2z+yz^2+xy^4+z^2+ay^6$
Z_{19}	$x^3y+y^9+z^2+axy^7$	S_{17}	$x^2z+yz^2+y^6+z^2+azy^4$
W_{17}	$x^4+xy^5+z^2+ay^7$	U_{16}	$x^3+xz^2+y^5+z^2+ax^2y^2$

Adjacency relations between simple and simply elliptic singularities



Adjacency relations among unimodular singularities



References

[1] V. I. Arnol'd, Singularity theory, Lecture note ser. 53, London Math. Soc., 1981.
[2] E. Breiskorn, Die Hierarchie der 1-modularen Singularitäten, Manuscripta Math., 27 (1979), 183–219.
[3] K. Saito, Einfach elliptische Singularitäten, Inventiones Math., 23 (1974), 289–325.

6. Topology

(I) h-Cobordism Groups of Homotopy Spheres and Groups of Differentiable Structures on Combinatorial Spheres

(1) The Structure of the h-Cobordism Group θ_n of n-Dimensional Homotopy Spheres. In the following table, values of θ_n have the following meanings: 0 means that the group consists only of the identity element, an integer l means that the group is isomorphic to the cyclic group of order l, 2^l means that the group is the direct sum of l groups of order 2, + means the direct sum, and ? means that the structure of the group is unknown.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$\theta_n\cong$	0	0	?	0	0	0	28	2	2^3 or $4+2$	6	992	0	3	2	$8128+2$	2	2^4 or $4+2^2$	$8+2$

(2) The Group Γ_n of Differentiable Structures on the n-Dimensional Combinatorial Sphere.

$$\Gamma_n\cong\theta_n \quad (n\neq 3), \quad \Gamma_3=0.$$

(II) Adem's Formula Concerning Steenrod Operators Sq and $\mathscr{P}$ (→ 64 Cohomology Operations)

For the cohomology operators Sq and $\mathscr{P}$, we have

$$Sq^a Sq^b=\sum_{c=0}^{[a/2]}\binom{b-c-1}{a-2c}Sq^{a+b-c}Sq^c \quad (a<2b).$$

$$\mathcal{P}^a\mathcal{P}^b=\sum_{c=0}^{[a/p]}(-1)^{a+c}\binom{(b-c)(p-1)-1}{a-pc}\mathcal{P}^{a+b-c}\mathcal{P}^c \quad (a<pb),$$

$$\mathcal{P}^a\delta\mathcal{P}^b=\sum_{c=0}^{[a/p]}(-1)^{a+c}\binom{(b-c)(p-1)}{a-pc}\delta\mathcal{P}^{a+b-c}\mathcal{P}^c$$
$$+\sum_{c=0}^{[(a-1)/p]}(-1)^{a+c-1}\binom{(b-c)(p-1)-1}{a-pc-1}\mathcal{P}^{a+b-c}\delta\mathcal{P}^c \quad (a\leqslant pb).$$

Several simple cases of the formula above are as follows.

$Sq^1Sq^{2n}=Sq^{2n+1}$, $\quad Sq^1Sq^{2n+1}=0$,

$Sq^2Sq^{4n}=Sq^{4n+2}+Sq^{4n+1}Sq^1$, $\quad Sq^2Sq^{4n+1}=Sq^{4n+2}Sq^1$,

$Sq^2Sq^{4n+2}=Sq^{4n+3}Sq^1$, $\quad Sq^2Sq^{4n+3}=Sq^{4n+5}+Sq^{4n+4}Sq^1$,

$Sq^4Sq^{8n}=Sq^{8n+4}+Sq^{8n+3}Sq^1+Sq^{8n+2}Sq^2$, $\quad Sq^4Sq^{8n+1}=Sq^{8n+4}Sq^1+Sq^{8n+3}Sq^2$,

$Sq^4Sq^{8n+2}=Sq^{8n+4}Sq^2$, $\quad Sq^4Sq^{8n+3}=Sq^{8n+5}Sq^2$,

$Sq^4Sq^{8n+4}=Sq^{8n+7}Sq^1+Sq^{8n+6}Sq^2$, $\quad Sq^4Sq^{8n+5}=Sq^{8n+9}+Sq^{8n+8}Sq^1+Sq^{8n+7}Sq^2$,

$Sq^4Sq^{8n+6}=Sq^{8n+10}+Sq^{8n+8}Sq^2$, $\quad Sq^4Sq^{8n+7}=Sq^{8n+11}+Sq^{8n+9}Sq^2$,

$\mathcal{P}^1\mathcal{P}^n=(n+1)\mathcal{P}^{n+1}$,

$\mathcal{P}^1\delta\mathcal{P}^n=n\cdot\delta\mathcal{P}^{n+1}+\mathcal{P}^{n+1}\delta$.

(III) Cohomology Ring $H^*(\pi,n;\mathbf{Z}_p)$ of Eilenberg-MacLane Complex (→ 70 Complexes)

$\mathbf{Z}$ means the set of integers, and $\mathbf{Z}_p=\mathbf{Z}/p\mathbf{Z}$, where p is a prime number.
(1) The case $p=2$, $\pi=\mathbf{Z}$ or $\mathbf{Z}_{2^f}$ $(f\geqslant 1)$. The degree of a finite sequence $I=(i_r,i_{r-1},\ldots,i_1)$ of positive integers is defined by $d(I)=i_1+i_2+\ldots+i_r$. If such a sequence satisfies $i_{k+1}\geqslant 2i_k$ $(k=1,\ldots,r-1)$, it is called admissible, and we define its excess by

$$e(I)=(i_r-2i_{r-1})+\ldots+(i_2-2r_1)+i_1=2i_r-d(I).$$

Further, we put $Sq^I=Sq^{i_r}Sq^{i_{r-1}}\ldots Sq^{i_1}$. Then we have $H^*(\mathbf{Z}_{2^f},n;\mathbf{Z}_2)=\mathbf{Z}_2[Sq^Iu_n\mid I$ is admissible, $e(I)<n]$, $H^*(\mathbf{Z},n;\mathbf{Z}_2)=\mathbf{Z}_2[Sq^Iu_n\mid I$ is admissible, $e(I)<n$, $i_1>1]$.

Here, $u_n\in H^n(\pi,n;\mathbf{Z}_2)$ is the fundamental cohomology class. $I=\varnothing$ (empty) is also admissible, and for this case we put $n(I)=e(I)=0$, $Sq^I=1$. Due to the Künneth theorem, we have

$$H^*(\pi+\pi',n;\mathbf{Z}_p)=H^*(\pi,n;\mathbf{Z}_p)\otimes H^*(\pi',n;\mathbf{Z}_p)$$

if π is finitely generated. In particular, we have

$$H^*(\mathbf{Z}_2,1;\mathbf{Z}_2)=\mathbf{Z}_2[u_1],$$
$$H^*(\mathbf{Z}_2,2;\mathbf{Z}_2)=\mathbf{Z}_2\left[u_2,Sq^1u_2,Sq^2Sq^1u_2,\ldots,Sq^{2^r}Sq^{2^{r-1}}\ldots Sq^1u_2,\ldots\right],$$
$$H^*(\mathbf{Z}_2,3;\mathbf{Z}_2)=\mathbf{Z}_2\left[u_3,Sq^{2^r}Sq^{2^{r-1}}\ldots Sq^1u_3,Sq^{(2^r+1)2^s}Sq^{(2^r+1)2^{s-1}}\ldots Sq^{2^r+1}Sq^{2^{r-1}}\ldots Sq^1u_3\mid r\geqslant 0,\ s\geqslant 0\right].$$

(2) The case $p\neq 2$, $\pi=\mathbf{Z}$ or $\mathbf{Z}_{p^f}$ $(f\geqslant 1)$. We define the degree of a finite sequence $I=(i_r,i_{r-1},\ldots,i_1,i_0)$ of nonnegative integers by $d(I)=i_r+\ldots+i_1+i_0$. The sequence I is called admissible if it satisfies the following conditions:

$i_k=2\lambda_k(p-1)+\varepsilon_k$ (λ_k is a nonnegative integer, $\varepsilon_k=0$ or 1 $(0\leqslant k\leqslant r)$), and

$i_0=0$ or 1, $\quad i_1\geqslant 2p-2$, $\quad i_{k+1}\geqslant pi_k$ $\quad (1\leqslant k\leqslant r-1)$.

We define its excess by $e(I)=pi_r-(p-1)d(I)$. Further, we put $\mathcal{P}^I=\delta^{\varepsilon_k}\mathcal{P}^{\lambda_k}\ldots\delta^{\varepsilon_1}\mathcal{P}^{\lambda_1}\delta^{\varepsilon_0}$, and assume that $u_n\in H^n(\pi,n;\mathbf{Z})$ is the fundamental cohomology class. Then we have

$$H^*(\mathbf{Z}_{p^f},n;\mathbf{Z}_p)=\mathbf{Z}_p[\mathcal{P}^Iu_n\mid I \text{ is admissible, } e(I)<n(p-1),\ n+d(I) \text{ is even}]$$
$$\otimes\wedge_{\mathbf{Z}_p}(\mathcal{P}^Iu_n\mid I \text{ is admissible, } e(I)<n(p-1),\ n+d(I) \text{ is odd}).$$

$H^*(\mathbf{Z},n;\mathbf{Z})$ is given by the above formula when the admissible sequence is I with $i_0=0$.

(IV) Cohomology Ring of Compact Connected Lie Groups (→ 427 Topology of Lie Groups and Homogeneous Spaces)

(1) General Remarks. Let G be a compact connected Lie group with rank l and dimension n. We have $H^*(G;\mathbf{R})\cong\wedge_{\mathbf{R}}(x_1,\dots,x_l)$, where $\wedge_K(x_1,\dots,x_l)$ means the exterior algebra over K of a linear space $V=Kx_1+\dots+Kx_l$ with the basis $\{x_1,\dots,x_l\}$ over K. We define a new degree in $\wedge_K(x_1,\dots,x_l)$ by putting $\deg x_i=m_i$ (m_i is odd) $(1\leqslant i\leqslant l)$, where $m_1+\dots+m_l=n$. The $\cong$ means isomorphism as graded rings.

(2) Classical Compact Simple Lie Groups. We set $\deg x_i=i$.

$$H^*(U(n);\mathbf{R})\cong\wedge_{\mathbf{R}}(x_1,x_3,\dots,x_{2n-1}),$$
$$H^*(SU(n);\mathbf{R})\cong\wedge_{\mathbf{R}}(x_3,x_5,\dots,x_{2n-1}),$$
$$H^*(Sp(n);\mathbf{R})\cong\wedge_{\mathbf{R}}(x_3,x_7,\dots,x_{4n-1}),$$
$H^*(SO(n);\mathbf{Z}_2)\cong$(Having $x_1,x_2,\dots,x_{n-1}$ as a simple system of generators)
$$\cong\mathbf{Z}_2[x_1,x_3,\dots,x_{2n'-1}]/(x_i^{2^{s(i)}}\mid i=1,\dots,n')$$
($n'=[n/2]$, $s(i)$ is the least integer satisfying $2^{s(i)}(2i-1)\geqslant n$)
$$H^*(SO(2n);K)\cong\wedge_K(x_3,x_7,\dots,x_{4n-5},x_{2n-1}),$$
$H^*(SO(2n-1);K)\cong\wedge_K(x_3,x_7,\dots,x_{4n-5})$, where K is a commutative field whose characteristic is not 2.

For $SO(n)$, $Sq^a(x_i)=\binom{i}{a}x_{i+a}$. For $SU(n)$, $p^a(x_{2i-1})=\binom{i-1}{a}x_{2i-1+2a(p-1)}$.

For $Sp(n)$, $\mathfrak{P}^a(x_{4i-1})=(-1)^{a(p-1)/2}\binom{2i-1}{a}x_{4i-1+2a(p-1)}$.

(3) Exceptional Compact Simple Lie Groups. n and m_i $(1\leqslant i\leqslant l)$ given in (1) are as follows.

G_2:	$n=14$,	$m_i=3,\ 11.$
F_4:	$n=52$,	$m_i=3,\ 11,\ 15,\ 23.$
E_6:	$n=78$,	$m_i=3,\ 9,\ 11,\ 15,\ 17,\ 23.$
E_7:	$n=133$,	$m_i=3,\ 11,\ 15,\ 19,\ 23,\ 27,\ 35.$
E_8:	$n=248$,	$m_i=3,\ 15,\ 23,\ 27,\ 35,\ 39,\ 47,\ 59.$

(4) p-Torsion Groups of Exceptional Groups. The p-torsion groups of exceptional Lie groups are unit groups except when $p=2$ for G_2; $p=2$, 3 for F_4, E_6, E_7; and $p=2$, 3, 5 for E_8. The cohomology ring of $\mathbf{Z}_p$ as a coefficient group in these exceptional cases is given as follows. Here we put $\deg x_i=i$.

$$H^*(G_2;\mathbf{Z}_2)=\mathbf{Z}_2[x_3]/(x_3^4)\otimes\wedge_{\mathbf{Z}_2}(Sq^2x_3);$$
$$H^*(F_4;\mathbf{Z}_2)=\mathbf{Z}_2[x_3]/(x_3^4)\otimes\wedge_{\mathbf{Z}_2}(Sq^2x_3,x_{15},Sq^8x_{15}),$$
$$H^*(F_4;\mathbf{Z}_3)=\mathbf{Z}_3[\delta\mathfrak{P}^1x_3]/((\delta\mathfrak{P}^1x_3)^3)\otimes\wedge_{\mathbf{Z}_3}(x_3,\mathfrak{P}^1x_3,x_{11},\mathfrak{P}^1x_{11});$$
$$H^*(E_6;\mathbf{Z}_2)=\mathbf{Z}_2[x_3]/(x_3^4)\otimes\wedge_{\mathbf{Z}_2}(Sq^2x_3,Sq^4Sq^2x_3,x_{15},Sq^8Sq^4Sq^2x_3,Sq^8x_{15}),$$
$$H^*(E_6;\mathbf{Z}_3)=\mathbf{Z}_3[\delta\mathfrak{P}^1x_3]/((\delta\mathfrak{P}'x_3)^3)\otimes\wedge_{\mathbf{Z}_3}(x_3,\mathfrak{P}^1x_3,x_9,x_{11},\mathfrak{P}^1x_{11},x_{17});$$
$$H^*(E_7;\mathbf{Z}_2)=\mathbf{Z}_2[x_3,Sq^2x_3,Sq^4Sq^2x_3]/(x_3^4,(Sq^2x_3)^4,(Sq^4Sq^2x_3)^4)$$
$$\otimes\wedge_{\mathbf{Z}_2}(x_{15},Sq^8Sq^4Sq^2x_3,Sq^8x_{15},Sq^4Sq^8x_{15}),$$
$$H^*(E_7;\mathbf{Z}_3)=\mathbf{Z}_3[\delta\mathfrak{P}^1x_3]/((\delta\mathfrak{P}^1x_3)^3)\otimes\wedge_{\mathbf{Z}_3}(x_3,\mathfrak{P}^1x_3,x_{11},\mathfrak{P}^1x_{11},\mathfrak{P}^3\mathfrak{P}^1x_3,x_{27},x_{35});$$
$$H^*(E_8;\mathbf{Z}_2)=\mathbf{Z}_2[x_3,Sq^2x_3,Sq^4Sq^2x_3,x_{15}]/(x_3^{16},(Sq^2x_3)^8,(Sq^4Sq^2x_3)^4,x_{15}^4)$$
$$\otimes\wedge_{\mathbf{Z}_2}(Sq^8Sq^4Sq^2x_3,Sq^8x_{15},Sq^4Sq^8x_{15},Sq^2Sq^4Sq^8x_{15}),$$
$$H^*(E_8;\mathbf{Z}_3)=\mathbf{Z}_3[\delta\mathfrak{P}^1x_3,\delta\mathfrak{P}^3\mathfrak{P}^1x_3]/((\delta\mathfrak{P}^1x_3)^3,(\delta\mathfrak{P}^3\mathfrak{P}^1x_3)^3)$$
$$\otimes\wedge_{\mathbf{Z}_3}(x_3,\mathfrak{P}^1x_3,x_{15},\mathfrak{P}^3\mathfrak{P}^1x_3,\mathfrak{P}^3x_{15},x_{35},x_{39},x_{47}),$$
$$H^*(E_8;\mathbf{Z}_5)=\mathbf{Z}_5[\delta\mathfrak{P}^1x_3]/((\delta\mathfrak{P}^1x_3)^5)\otimes\wedge_{\mathbf{Z}_5}(x_3,\mathfrak{P}^1x_3,x_{15},\mathfrak{P}^1x_{15},x_{27},x_{35},x_{39},x_{47}).$$

(V) Cohomology Rings of Classifying Spaces (→ 56 Characteristic Classes, 427 Topology of Lie Groups and Homogeneous Spaces C)

(1) Let $H^*(G;K)=\bigwedge_K(x_1,x_2,\dots,x_n)$. Then the $\deg x_i$ are odd and the x_i may be assumed to be transgressive. y_i being its image, the following formula holds:

$$H^*(BG;K)=K[y_1,y_2,\dots,y_n] \qquad \text{(Borel's theorem)}.$$

(2) $H^*(BU(n))=H^*(BGL(n,\mathbf{C}))=\mathbf{Z}[c_1,c_2,\dots,c_n]$,

$H^*(BSU(n))=H^*(BSL(n,\mathbf{C}))=\mathbf{Z}[c_2,\dots,c_n]$,

$H^*(BSp(n))=\mathbf{Z}[q_1,q_2,\dots,q_n]$,

$H^*(BO(n);K_2)=H^*(BGL(n,\mathbf{R});\mathbf{Z}_2)=K_2[w_1,w_2,\dots,w_n]$,

$H^*(BSO(n);K_2)=H^*(BSL(n,\mathbf{R});\mathbf{Z}_2)=K_2[w_2,\dots,w_n]$,

$H^*(BSO(2m+1);K)=K[p_1,p_2,\dots,p_m]$,

$H^*(BSO(2m);K)=K[p_1,p_2,\dots,p_{m-1},\chi]$.

Here, K denotes a field of characteristic $\neq 2$, and K_2 is the field of characteristic 2. The c_i denote the ith Chern classes and the q_i the ith symplectic Pontryagin classes, the w_i the ith Stiefel-Whitney classes. Moreover, the p_i denote the ith Pontryagin classes, and χ the Euler class. Their degrees are given as follows: $\deg c_i=2i$, $\deg q_i=\deg p_i=4i$, $\deg w_i=i$, and $\deg\chi=2m$.

(3) Wu's Formula. Let $H^2(BSO(n);\mathbf{Z}_2)=\mathbf{Z}_2[w_2,\dots,w_n]$ and $H^*(BU(n),\mathbf{Z}_2)=\mathbf{Z}_2[c_1,\ c_2,\dots,c_n]$. We have

$$Sq^j w_i=\sum_{0\leqslant t\leqslant j}\binom{i-j+t-1}{t}w_{j-t}w_{i+t} \qquad (w_0=1),$$

$$Sq^{2j} c_i=\sum_{0\leqslant t\leqslant j}\binom{i-j+t-1}{t}c_{j-t}c_{i+t} \qquad (c_0=1).$$

Here the symbol $\binom{a}{b}$ denotes the binomial coefficient for $a\geqslant b$; $\binom{a}{0}=1$, and $\binom{a}{b}=0$ otherwise.

(VI) Homotopy Groups of Spheres (→ 202 Homotopy Theory)

Table of the $(n+k)$th Homotopy Group $\pi_{n+k}(S^n)$ of the n-Dimensional Sphere S^n. The table represents Abelian groups. 0 stands for the unit group; integer l the cyclic group of order l; ∞ the infinite cyclic group; 2^l the direct sum of l groups of order 2; and $+$ means the direct sum.

$n \backslash k$	<0	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0	∞	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	∞	∞	2	2	12	2	2	3	15	2	2^2	$12+2$	$84+2^2$	2^2
3	0	∞	2	2	12	2	2	3	15	2	2^2	$12+2$	$84+2^2$	2^2	6
4	0	∞	2	2	$\infty+12$	2^2	2^2	$24+3$	15	2	2^3	$120+12+2$	$84+2^5$	2^6	$24+6+2$
5	0	∞	2	2	24	2	2	2	30	2	2^3	$72+2$	$504+2^2$	2^3	$6+2$
6	0	∞	2	2	24	0	∞	2	60	$24+2$	2^3	$72+2$	$504+4$	240	6
7	0	∞	2	2	24	0	0	2	120	2^3	2^4	$24+2$	$504+2$	0	6
8	0	∞	2	2	24	0	0	2	$\infty+120$	2^4	2^5	$24+24+2$	$504+2$	0	$6+2$
9	0	∞	2	2	24	0	0	2	240	2^3	2^4	$24+2$	$504+2$	0	6
10	0	∞	2	2	24	0	0	2	240	2^2	$\infty+2^3$	$12+2$	504	12	6
11	0	∞	2	2	24	0	0	2	240	2^2	2^3	$6+2$	504	2	$6+2$
12	0	∞	2	2	24	0	0	2	240	2^2	2^3	6	$\infty+504$	2^2	$6+2$
13	0	∞	2	2	24	0	0	2	240	2^2	2^3	6	504	2	6
14	0	∞	2	2	24	0	0	2	240	2^2	2^3	6	504	0	$\infty+3$
$\geqslant 15$	0	∞	2	2	24	0	0	2	240	2^2	2^3	6	504	0	3

Table of the $(n+k)$th Homotopy Group $\pi_{n+k}(S^n)$ of the n-Dimensional Sphere S^n (Continued)

n \ k	14	15	16	17	18	19	20	21	22
1	0	0	0	0	0	0	0	0	0
2	6	30	30	$6+2$	$12+2^2$	$12+2^2$	$132+2$	2^2	2
3	30	30	$6+2$	$12+2^2$	$12+2^2$	$132+2$	2^2	2	210
4	$2520+6+2$	30	$6+6+2$	$24+12+4+2^2$	$120+12+2^5$	$132+2^5$	2^6	$24+2^2$	$9240+6+2$
5	$6+2$	$30+2$	2^2	$4+2^2$	$24+2^2$	$264+2$	$6+2^2$	$6+2$	$90+2^2$
6	$12+2$	$60+2$	$504+2^2$	2^4	$24+6+2$	$1056+8$	$480+12$	6	$180+2^2$
7	$24+4$	$120+2^3$	2^4	2^4	$24+2$	$264+2$	24	$6+2$	$72+2^3$
8	$240+24+4$	$120+2^5$	2^7	$6+2^4$	$504+24+2$	$264+2$	$24+3$	$12+2^3$	$1440+24+2^4$
9	$16+4$	$240+2^3$	2^4	2^4	$24+2$	$264+2$	24	$6+2^2$	$144+2^3$
10	$16+2$	$240+2^2$	$240+2$	2^3	$24+2^2$	$264+6$	$504+24$	$6+2^2$	$144+6+2$
11	$16+2$	$240+2$	2	2^3	$8+4+2$	$264+2^3$	$24+2^2$	2^4	$48+2^2$
12	$48+4+2$	$240+2$	2	2^4	$480+4+4+2$	$264+2^5$	$24+2^5$	$6+2^4$	$2016+12+2^2$
13	$16+2$	$480+2$	2	2^4	$8+8+2$	$264+2^3$	$24+2^3$	$4+2^3$	$16+2^2$
14	$8+2$	$480+2$	$24+2$	2^4	$8+8+2$	$264+4+2$	$240+24$	$4+2^2$	$16+2^2$
15	$4+2$	$480+2$	2^3	2^5	$8+8+2$	$264+2^2$	2^4	2^3	$16+2^3$
16	$2+2$	$\infty+480+2$	2^4	2^6	$24+8+8+2$	$264+2^2$	2^4	2^4	$240+16+2^3$
17	$2+2$	$480+2$	2^3	2^5	$8+8+2$	$264+2^2$	24	2^3	$16+2^3$
18	$2+2$	$480+2$	2^2	$\infty+2^4$	$8+4+2$	$264+2$	$24+12$	2^3	$16+2^2$
19	$2+2$	$480+2$	2^2	2^4	$8+2^2$	$264+2$	$24+2$	2^4	$16+2^2$
20	$2+2$	$480+2$	2^2	2^4	$8+2$	$\infty+264+2$	$24+2^2$	2^4	$16+2^2$
21	$2+2$	$480+2$	2^2	2^4	$8+2$	$264+2$	$24+2$	2^3	$8+2^2$
22	$2+2$	$480+2$	2^2	2^4	$8+2$	$264+2$	24	$\infty+2^2$	$4+2^2$
23	$2+2$	$480+2$	2^2	2^4	$8+2$	$264+2$	24	2^2	2^3
$\geqslant 24$	$2+2$	$480+2$	2^2	2^4	$8+2$	$264+2$	24	2^2	2^2

Remarks

(1) When $n>k+1$ (below the broken line in the table), $\pi_{n+k}(S^n)$ is independent of n and is isomorphic with the kth stable homotopy group G_k.

(2) Let $\iota_n \in \pi_n(S^n)$ be the identity on S^n; $\eta_2 \in \pi_3(S^2)$, $\nu_4 \in \pi_7(S^4)$, $\sigma_8 \in \pi_{15}(S^8)$ be the Hopf mapping $S^3 \to S^2$, $S^7 \to S^4$, $S^{15} \to S^8$ (induced mapping in the homotopy class), respectively; and $[\iota_{2m}, \iota_{2m}] \in \pi_{4m-1}(S^{2m})$ $(m \neq 1,2,4)$ be the Whitehead product of ι_{2m}. These objects generate infinite cyclic groups which are direct factors of $\pi_{n+k}(S^n)$ corresponding to the original mappings.

(3) $\eta_{n+2} = E^n \eta_2$, $\nu_{n+4} = E^n \nu_4$, $\sigma_{n+8} = E^n \sigma_8$ $(n \geqslant 1)$ (E is the suspension) are the generator for $\pi_{n+k}(S^n)$, which contains the mappings.

(4) The orders of the following compositions are 2:gs$_{n+7}$

$$\eta_n \circ \eta_{n+1}\ (n \geqslant 2),\quad \nu_n \circ \nu_{n+3}\ (n \geqslant 5),\quad \sigma_n \circ \sigma_{n+7}\ (n \geqslant 16),\quad \eta_n \circ \nu_{n+1}\ (n=3,4),$$

$$\nu_n \circ \eta_{n+3}\ (n=4,5),\quad \eta_n \circ \sigma_{n+1}\ (n \geqslant 7),\quad \sigma_n \circ \eta_{n+7}\ (n \geqslant 8),\quad \nu_{10} \circ \sigma_{13},\quad \sigma_{11} \circ \nu_{18},$$

$$\eta_n \circ \eta_{n+1} \circ \eta_{n+2}\ (n \geqslant 2),\quad \nu_n \circ \nu_{n+3} \circ \nu_{n+6} (n \geqslant 4),\quad \sigma_n \circ \sigma_{n+7} \circ \sigma_{n+14}\ (n \geqslant 9).$$

(VII) The Homotopy Groups $\pi_k(G)$ of Compact Connected Lie Groups G

Here the group G is one of the following:

$SO(n)$ $(n \geqslant 2)$, $Spin(n)$ $(n \geqslant 3)$, $U(n)$ $(n \geqslant 1)$, $SU(n)$ $(n \geqslant 2)$,
$Sp(n)$ $(n \geqslant 1)$, G_2, F_4, E_6, E_7, E_8.

(1) The Fundamental Group $\pi_1(G)$.

$$\pi_1(G) \cong \begin{cases} \infty & (G=U(n)\ (n \geqslant 1),\ SO(2)), \\ 2 & (G=SO(n)\ (n \geqslant 3)), \\ 0 & (\text{for all other groups } G). \end{cases}$$

(2) Isomorphic Relations $(k \geqslant 2)$.

$\pi_k(U(n)) \cong \pi_k(SU(n))$ $(n \geqslant 2)$,

$\pi_k(U(1)) \cong \pi_k(SO(2)) \cong 0.$

$\pi_k(Spin(n)) \cong \pi_k(SO(n))\ (n \geqslant 3),$

$\pi_k(Spin(3)) \cong \pi_k(Sp(1)) \cong \pi_k(SU(2)) \cong \pi_k(S^3),$

$\pi_k(Spin(4)) \cong \pi_k(Spin(3)) + \pi_k(S^3),$

$\pi_k(Spin(5)) \cong \pi_k(Sp(2)),$

$\pi_k(Spin(6)) \cong \pi_k(SU(4)).$

(3) The Homotopy Group $\pi_k(G)$ $(k \geqslant 2)$.

$\pi_2(G) \cong 0.$

$\pi_3(G) \cong \infty$ $(G \neq SO(2),\ U(1),\ SO(4),\ Spin(4))$, $\pi_3(SO(4)) \cong \infty + \infty$.

$$\pi_4(G) \cong \begin{cases} 2+2 & (G = SO(4),\ Spin(4)), \\ 2 & (G = Sp(n),\ SU(2),\ SO(3),\ SO(5),\ Spin(3),\ Spin(5)), \\ 0 & (G = SU(n)\ (n \geqslant 3),\ SO(n)\ (n \geqslant 6),\ G_2,\ F_4,\ E_6,\ E_7,\ E_8). \end{cases}$$

$$\pi_5(G) \cong \begin{cases} 2+2 & (G = SO(4),\ Spin(4)), \\ 2 & (G = Sp(n),\ SU(2),\ SO(3),\ SO(5)\ Spin(3),\ Spin(5)), \\ \infty & (G = SU(n)\ (n \geqslant 3),\ SO(6),\ Spin(6)), \\ 0 & (G = SO(n),\ Spin(n)\ (n \geqslant 7),\ G_2,\ F_4,\ E_6,\ E_7,\ E_8). \end{cases}$$

$\pi_k(G)$, $k \geqslant 6$.

G \ k	6	7	8	9	10	11	12	13	14	15
$Sp(1)$	12	2	2	3	15	2	2^2	12+2	$84+2^2$	2^2
$Sp(2)$	0	∞	0	0	120	2	2^2	4+2	1680	2
$Sp(3)$	0	∞	0	0	0	∞	2	2	10080	2
$Sp(4)$	0	∞	0	0	0	∞	2	2	0	∞
$SU(2)$	12	2	2	3	15	2	2^2	12+3	$84+2^2$	2^2
$SU(3)$	6	0	12	3	30	4	60	6	84+2	36
$SU(4)$	0	∞	24	2	120+2	4	60	4	1680+2	72+2
$SU(5)$	0	∞	0	∞	120	0	360	4	1680	6
$SU(6)$	0	∞	0	∞	0	∞	720	2	5040+2	6
$SU(7)$	0	∞	0	∞	0	∞	0	∞	5040	0
$SU(8)$	0	∞	0	∞	0	∞	0	∞	0	∞
$SO(5)$	0	∞	0	0	120	2	2^2	4+2	1680	2
$SO(6)$	0	∞	24	2	120+2	4	60	4	1680+2	72+2
$SO(7)$	0	∞	2^2	2^2	8	$\infty+2$	0	2	2520+8+2	2^4
$SO(8)$	0	$\infty+\infty$	2^3	2^3	24+8	$\infty+2$	0	2^2	2520+120+8+2	2^7
$SO(9)$	0	∞	2^2	2^2	8	$\infty+2$	0	2	8+2	$\infty+2^3$
$SO(10)$	0	∞	2	$\infty+2$	4	∞	12	2	8	$\infty+2^2$
$SO(11)$	0	∞	2	2	2	∞	2	2^2	8	$\infty+2$
$SO(12)$	0	∞	2	2	0	$\infty+\infty$	2^2	2^2	24+4	$\infty+2$
$SO(13)$	0	∞	2	2	0	∞	2	2	8	$\infty+2$
$SO(14)$	0	∞	2	2	0	∞	0	∞	4	∞
$SO(15)$	0	∞	2	2	0	∞	0	0	2	∞
$SO(16)$	0	∞	2	2	0	∞	0	0	0	$\infty+\infty$
$SO(17)$	0	∞	2	2	0	∞	0	0	0	∞
G_2	3	0	2	6	0	$\infty+2$	0	0	168+2	2
F_4	0	0	2	2	0	$\infty+2$	0	0	2	∞
E_6	0	0	0	∞	0	∞	12	0	0	∞
E_7	0	0	0	0	0	∞	2	2	0	∞
E_8	0	0	0	0	0	0	0	0	0	∞

(4) Stable Homotopy Groups. For sufficiently large n for fixed k, the homotopy groups for classical compact simple Lie groups $G=Sp(n)$, $SU(n)$, $SO(n)$ become stable. We denote them by the following notations. Here we assume $k \geqslant 2$.

$$\pi_k(Sp)=\pi_k(Sp(n)) \qquad (n \geqslant (k-1)/4),$$
$$\pi_k(U)=\pi_k(U(n)) \cong \pi_k(SU(n)) \qquad (n \geqslant (k+1)/2),$$
$$\pi_k(O)=\pi_k(SO(n)) \qquad (n \geqslant k+2).$$

Bott periodicity theorem

$$\pi_k(Sp) \cong \begin{cases} \infty & (k \equiv 3,7 \ (\mathrm{mod}\, 8)), \\ 2 & (k \equiv 4,5 \ (\mathrm{mod}\, 8)), \\ 0 & (k \equiv 0,1,2,6 \ (\mathrm{mod}\, 8)). \end{cases}$$

$$\pi_k(O) \cong \begin{cases} \infty & (k \equiv 3,7 \ (\mathrm{mod}\, 8)), \\ 2 & (k \equiv 0,1 \ (\mathrm{mod}\, 8)), \\ 0 & (k \equiv 2,4,5,6 \ (\mathrm{mod}\, 8)). \end{cases}$$

$$\pi_k(U) \cong \begin{cases} \infty & (k \equiv 1 \ (\mathrm{mod}\, 2)), \\ 0 & (k \equiv 0 \ (\mathrm{mod}\, 2)). \end{cases}$$

(5) Metastable Homotopy Groups.

(a,b) means the greatest common divisor of two integers a and b.

$$\pi_{2n}(SU(n)) \cong n!.$$

$$\pi_{2n+1}(SU(n)) \cong \begin{cases} 2 & (n \text{ even}), \\ 0 & (n \text{ odd}). \end{cases}$$

$$\pi_{2n+2}(SU(n)) \cong \begin{cases} (n+1)!+2 & (n \text{ even}, \geqslant 4), \\ (n+1)!/2 & (n \text{ odd}). \end{cases}$$

$$\pi_{2n+3}(SU(n)) \cong \begin{cases} (24,n) & (n \text{ even}), \\ (24,\pi+3)/2 & (n \text{ odd}). \end{cases}$$

$$\pi_{2n+4}(SU(n)) \cong \begin{cases} (n+2)!(24,n)/48 & (n \text{ even}, \geqslant 4), \\ (n+2)!(24,n+3)/24 & (n \text{ odd}). \end{cases}$$

$$\pi_{2n+5}(SU(n)) \cong \pi_{2n+5}(U(n+1)).$$

$$\pi_{2n+6}(SU(n)) \cong \begin{cases} \pi_{2n+6}(U(n+1)) & (n \equiv 2,3 \ (\mathrm{mod}\, 4),\ n \geqslant 3), \\ \pi_{2n+6}(U(n+1))+2 & (n \equiv 0,1 \ (\mathrm{mod}\, 4)). \end{cases}$$

$$\pi_{4n+2}(Sp(n)) \cong \begin{cases} (2n+1)! & (n \text{ even}), \\ 2(2n+1)! & (n \text{ odd}). \end{cases}$$

$$\pi_{4n+3}(Sp(n)) \cong 2.$$

$$\pi_{4n+4}(Sp(n)) \cong \begin{cases} 2+2 & (n \text{ even}), \\ 2 & (n \text{ odd}). \end{cases}$$

$$\pi_{4n+5}(Sp(n)) \cong \begin{cases} (24,n+2)+2 & (n \text{ even}), \\ (24,n+2) & (n \text{ odd}). \end{cases}$$

$$\pi_{4n+6}(Sp(n)) \cong \begin{cases} (2n+3)!(24,n+2)/12 & (n \text{ even}), \\ (2n+3)!(24,n+2)/24 & (n \text{ odd}). \end{cases}$$

$$\pi_{4n+7}(Sp(n)) \cong 2.$$
$$\pi_{4n+8}(Sp(n)) \cong 2+2.$$

The homotopy groups $\pi_{n+i}(SO(n))$ for $n \geqslant 16$, $3 \geqslant i \geqslant -1$ are determined by the isomorphism

$$\pi_{n+i}(SO(n)) \cong \pi_{n+i}(O) + \pi_{n+i+1}(V_{i+3+n,i+3}(\mathbf{R}))$$

and the homotopy groups of $V_{m+n,m}(\mathbf{R})$ given below.

(6) Homotopy Groups of Real Stiefel Manifolds $V_{m+n,m}(\mathbf{R}) = O(m+n)/I_m \times O(n)$.

$$\pi_{n+k}(V_{n+1,1}) \cong \pi_{n+k}(S^n).$$

$$\pi_{n-k}(V_{m+n,m}) \cong 0 \quad (k \geqslant 1).$$

$$\pi_n(V_{m+n,m}) \cong \begin{cases} 2 & (n=2s-1,\ m \geqslant 2), \\ \infty & (n=2s). \end{cases}$$

$\pi_{n+k}(V_{m+n,m})$ $(k=1,2,3,4,5)$ are given in the following table.

	$n \backslash m$	1	2	3	4	5	6	$8s-1$	$8s$	$8s+1$	$8s+2$	$8s+3$	$8s+4$	$8s+5$	$8s+6$
	2	0	∞^2	2	$2+\infty$	2	$2+\infty$	2	$2+\infty$	2	$2+\infty$	2	$2+\infty$	2	$2+\infty$
π_{n+1}	$\geqslant 3$	0	∞	0	2^2	2	4	0	2^2	2	4	0	2^2	2	4
	2	∞	2^2	4	2^2	4	2^2	4	2^2	4	2^2	4	2^2	4	2^2
π_{n+2}	3	∞^2	2	$2+\infty$	2^2	$4+\infty$	2	$2+\infty$	2^2	$4+\infty$	2	$2+\infty$	2^2	$4+\infty$	2
	$\geqslant 4$	∞	0	2	2^2	8	0	2	2^2	8	0	2	2^2	8	0
	2	2	2^2	2	$\infty+12+2$	2^2	$24+2$	2^2	$24+2$	2^2	$24+4$	2^2	$24+2$	2^2	$24+2$
	3	2^2	2	2	$\infty+12+4$	2^3	$12+2$	2^2	$24+4$	2^3	$12+2$	2^2	$24+4$	2^3	$12+2$
π_{n+3}	4	2	∞	2	∞^2+12+4	2^2	$12+\infty$	2^2	$24+4+\infty$	2^2	$12+\infty$	2^2	$24+4+\infty$	2^2	$12+\infty$
	$\geqslant 5$	0	0	2	$12+4+\infty$	2	12	2	$24+8$	2	12	2^2	$4+48$	2	12
	2	2	12^2	$\infty+2$	2^2+24	2^2	24	2	24	2	24	2	24	2	24
	3	2^2	0	$\infty+4$	2^4	2^3	2	4	2^2	2^2	2	4	2^2	2^2	2
π_{n+4}	4	2	0	$4+\infty$	2^5	2^2	2	8	2^3	2	2	8	2^3	2	2
	5	∞	0	$4+\infty^2$	2^4	$2+\infty$	2	$8+\infty$	2^2	∞	2	$8+\infty$	2^2	∞	2
	$\geqslant 6$	0	0	$4+\infty$	2^3	2	0	8	2	0	2	16	2	0	0
	2	12	2^2	2	2^3	0	∞	0	0	0	0	0	0	0	0
π_{n+5}	3	12^2	∞	$2+24$	2^4	24	$\infty+2$	24	2	24	2	24	2	24	2
	4	0	∞	2^3	2^5	2	$\infty+4$	2^2	2^2	2	4	2^2	2^2	2	4

(VIII) Immersion and Embedding of Projective Spaces (→ 114 Differential Topology)

We denote immersion by $\subset$, and embedding by $\subseteq$. $\mathbf{P}^n(A)$ is an n-dimensional real or complex projective space where $A = \mathbf{R}$ or $\mathbf{C}$, $k\{\mathbf{P}^n(A)\}$ is the integer k such that $\mathbf{P}^n(A) \subset \mathbf{R}^k$ and $\mathbf{P}^n(A) \not\subset \mathbf{R}^{k-1}$, and $\tilde{k}\{\mathbf{P}^n(A)\}$ is the integer k such that $\mathbf{P}^n(A) \subseteq \mathbf{R}^k$ and $\mathbf{P}^n(A) \not\subseteq \mathbf{R}^{k-1}$.

In the table, for example, numbers 9–11 in the row $k\{\mathbf{P}^n(\mathbf{R})\}$ for $n=6$ mean $\mathbf{P}^6(\mathbf{R}) \not\subset \mathbf{R}^8$, $P^6(\mathbf{R}) \subset \mathbf{R}^{11}$.

n	1	2	3	4	5	6	7	8	9	10	11	12
$k\{\mathbf{P}^n(\mathbf{R})\}$	2	4	5	8	9	9∼11	9∼12	16	17	17∼19	...	...
$\tilde{k}\{\mathbf{P}^n(\mathbf{R})\}$	2	3	4	7	7	7	8	15	15	16	16	17∼19
$k\{\mathbf{P}^n(\mathbf{C})\}$	3	7	9	15	17	22	22∼25	31	33	38	38∼41	...
$\tilde{k}\{\mathbf{P}^n(\mathbf{C})\}$	3	7	8∼9	15	16∼17	22	22∼25	31	32∼33	38	38∼41	...

	2^r	2^r+1	2^r+2	2^r+3	2^r+2^s $(r>s>0)$
$k\{\mathbf{P}^n(\mathbf{R})\}$	$2n$	$2n-1$	$2n-3$∼$2n-1$	...	...
$\tilde{k}\{\mathbf{P}^n(\mathbf{R})\}$	$2n-1$	$2n-3$	$2n-4$	$2n-6$	...
$k\{\mathbf{P}^n(\mathbf{C})\}$	$4n-1$	$4n-3$	$4n-2$	...	$4n-2$
$\tilde{k}\{\mathbf{P}^n(\mathbf{C})\}$	$4n-1$	$4n-4$∼$4n-3$	$4n-2$	...	$4n-2$

7. Knot Theory (→ 235 Knot Theory)

Let k be a projection on a plane of a knot K. We color the domains separated by k, white and black alternatively. The outermost (unbounded) domain determined by k is colored white. In Fig. 16, hatching means black. Take a point (a black point in Fig. 16) in each black domain. The self-intersections of k are represented by white points (Fig. 16). Through each white point we draw a line segment connecting the black points in the black regions meeting at the white point. In Fig. 16, we show this as a broken line. We assign the signature + if the torsion of K at the intersection of k has the orientation of a right-hand screw (as in Fig. 17, left), and the signature − if the orientation is opposite (as in Fig. 17, right). The picture of the line segments with signatures is called the graph corresponding to the projection k of the knot K. Given such a graph, we can reconstruct the original knot K.

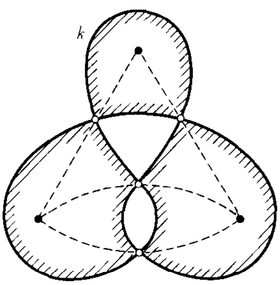


Fig. 16

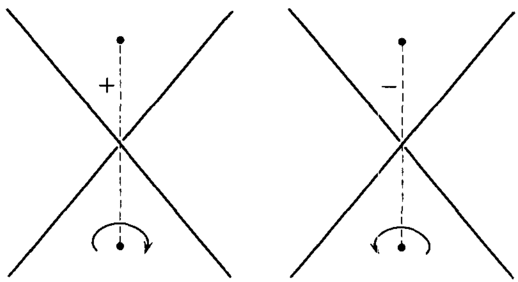


Fig. 17

Fig. 18 shows the classification table of knots for which the numbers of intersections of k are 3 to 8 when we minimize the intersections. The projection of k is described by a solid line, and the graph by broken lines. We omit the signatures since for each graph from 3_1 to 8_{18} they are all + or all −. Such knots are called alternating knots.

8. Inequalities (→ 88 Convex Analysis, 211 Inequalities)

(1) $|a+b| \leqslant |a|+|b|$,

$$|a-b| \geqslant ||a|-|b||.$$

For real a_ν, we have $\Sigma a_\nu^2 \geqslant 0$, and the equality holds only if all $a_\nu = 0$.

(2) $n! < n^n < (n!)^2 \quad (n \geqslant 3)$.

$$e^n \geqslant n^n/n!.$$

$$n^{1/n} < 3^{1/3} \quad (n \neq 3).$$

(3) $2/\pi < (\sin x)/x < 1 \quad (0 < x < \pi/2)$ (Jordan's inequality).

(4) Denote the elementary symmetric polynomials of positive numbers $a_1, \ldots, a_n > 0$ by S_r $(r = 1, \ldots, n)$. Then

$$S_1 \Big/ \binom{n}{1} \geqslant \left[S_2 \Big/ \binom{n}{2} \right]^{1/2} \geqslant \ldots \geqslant \left[S_r \Big/ \binom{n}{r} \right]^{1/r} \geqslant \ldots \geqslant \left[S_n \Big/ \binom{n}{n} \right]^{1/n}.$$

If at least one equality holds, then $a_1 = \ldots = a_n$. In particular, from the two external terms, we have the following inequalities concerning mean values:

$$\frac{1}{n} \sum_{\nu=1}^{n} a_\nu \geqslant \left(\prod_{\nu=1}^{n} a_\nu \right)^{1/n} \geqslant n \Big/ \sum_{\nu=1}^{n} \frac{1}{a_\nu}.$$

For weighted means, we have

$$\sum_{\nu=1}^{n} \lambda_\nu a_\nu \geqslant \prod_{\nu=1}^{n} a_\nu^{\lambda_\nu} \qquad \left(\sum \lambda_\nu = 1, \quad \lambda_\nu > 0 \right).$$

(5) When $a_\nu > 0$, $b_\nu > 0$, $p > 1$, $q > 1$, $(1/p) + (1/q) = 1$,

$$\left[\sum_{\nu=1}^{n} (a_\nu)^p \right]^{1/p} \left[\sum_{\nu=1}^{n} (b_\nu)^q \right]^{1/q} \geqslant \sum_{\nu=1}^{n} a_\nu b_\nu \qquad \text{(Hölder's inequality).}$$

The equality holds only if $(a_\nu)^p = c(b_\nu)^q$ (c is a constant).

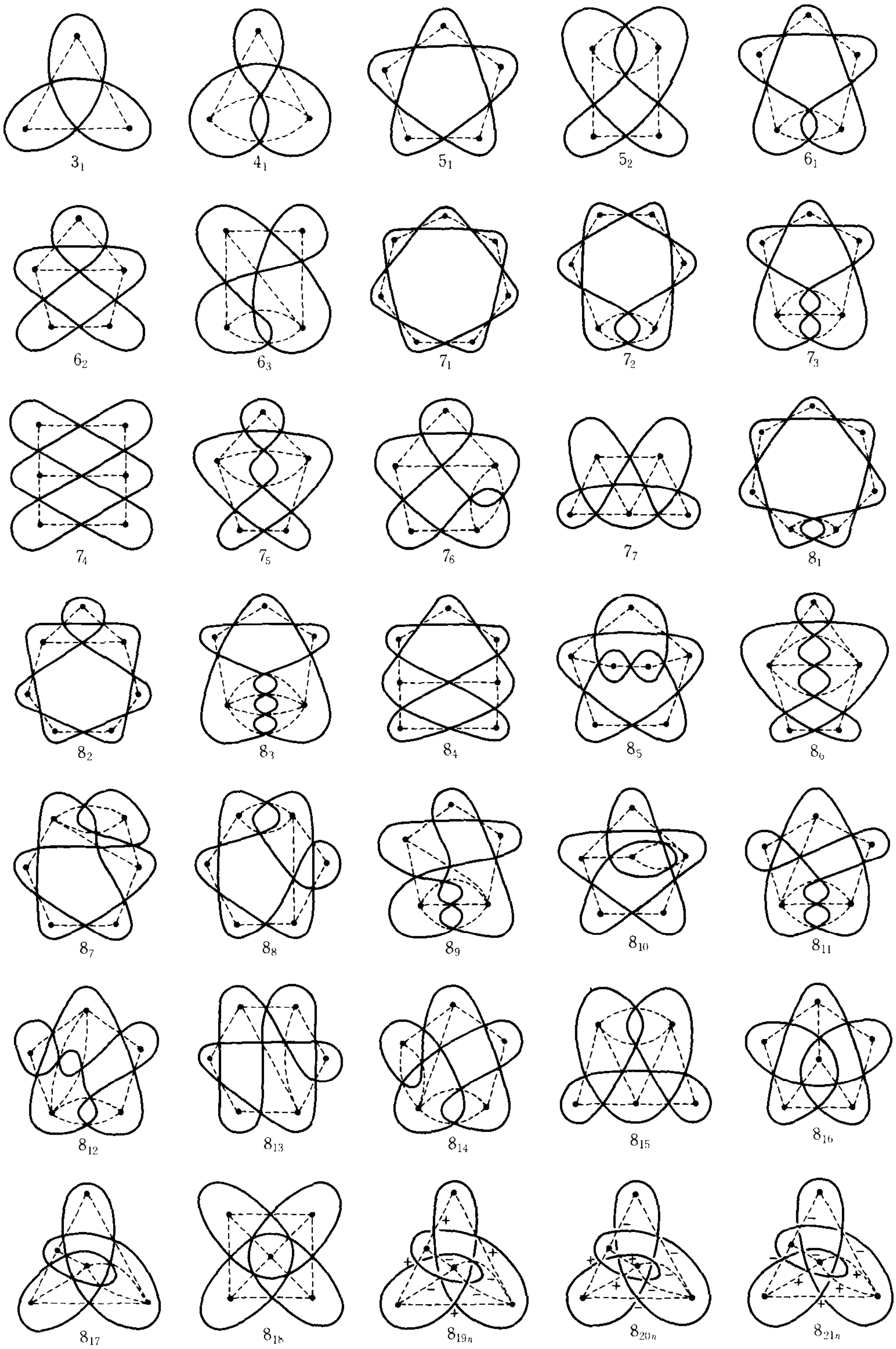


Fig. 18
Classification table of knots. The signatures from 3_1 to 8_{18} are all + or all −.

When $p=q=2$, the inequality is called Cauchy's inequality, the Cauchy-Schwarz inequality, or Bunyakovskiĭ's inequality. As special cases, we have

$$\left(\sum_{\nu=1}^{n} a_\nu\right)\left(\sum_{\nu=1}^{n} \frac{1}{a_\nu}\right) \geqslant n^2 \quad (a_\nu > 0),$$

$$\left(\sum_{\nu=1}^{n} a_\nu\right)^2 \leqslant n\left(\sum_{\nu=1}^{n} a_\nu^2\right) \quad (a_\nu > 0).$$

When $0<p<1$, we have an inequality by reversing the inequality sign in Hölder's inequality.

(6) When $a_\nu > 0$, $b_\nu > 0$, $p > 0$, and $\{a_\nu\}$ and $\{b_\nu\}$ are not proportional, we have

$$\left[\sum_{\nu=1}^{n}(a_\nu+b_\nu)^p\right]^{1/p} \lesseqgtr \left[\sum_{\nu=1}^{n}(a_\nu)^p\right]^{1/p} + \left[\sum_{\nu=1}^{n}(b_\nu)^p\right]^{1/p} \quad (p \gtreqless 1) \qquad \text{(Minkowski's inequality).}$$

The integral inequality corresponding to (5) or (6) has the same name.

(7) If $a_{\mu\nu} > 0$, $\sum_{\mu=1}^{n} a_{\mu\nu} = \sum_{\nu=1}^{n} a_{\mu\nu} = 1$, $b_\nu \geqslant 0$,

$$\prod_{\nu=1}^{n} b_\nu \leqslant \prod_{\mu=1}^{n}\left(\sum_{\nu=1}^{n} a_{\mu\nu} b_\nu\right).$$

In particular, for the determinant $\Delta = \det(a_{\mu\nu})$,

$$|\Delta|^2 \leqslant \prod_{\nu=1}^{n}\left(\sum_{\mu=1}^{n} |a_{\mu\nu}|^2\right).$$

The equality in this holds only if all rows are mutually orthogonal. If all $|a_{\mu\nu}| \leqslant M$, we have

$$|\Delta| \leqslant n^{n/2} M^n \qquad \text{(Hadamard's estimation).}$$

(8) Suppose that a function $f(x)$ is continuous, strictly monotone increasing in $x \geqslant 0$, and $f(0) = 0$. Denote the inverse function of f by f^{-1}. For $a, b > 0$, we have

$$ab \leqslant \int_0^a f(x)\,dx + \int_0^b f^{-1}(x)\,dx \qquad \text{(Young's inequality),}$$

and the equality holds only if $b = f(a)$.

In particular, for $f(x) = x^{p-1}$ $(p > 1)$, we have

$$\frac{a^p}{p} + \frac{b^q}{q} \geqslant ab,$$

where $(1/p) + (1/q) = 1$.

(9) If $p, q > 1$, $(1/p) + (1/q) = 1$, $a_\mu \geqslant 0$, $b_\nu \geqslant 0$,

$$\sum_{\mu,\nu=0}^{\infty} \frac{a_\mu b_\nu}{\mu+\nu+1} \leqslant \frac{\pi}{\sin(\pi/p)}\left[\sum_{\mu=0}^{\infty}(a_\mu)^p\right]^{1/p}\left[\sum_{\nu=0}^{\infty}(b_\nu)^q\right]^{1/q} \qquad \text{(Hilbert's inequality),}$$

and the equality holds only when the right-hand side vanishes.

(10) For a continuous function $f(x) \geqslant 0$ $(0 \leqslant x < \infty)$, we put

$$F(x) = \int_0^x f(t)\,dt,$$

and assume that $p > 1$. Then

$$\int_0^\infty \left[\frac{F(x)}{x}\right]^p dx \leqslant \left(\frac{p}{p-1}\right)^p \int_0^\infty [f(x)]^p\,dx \qquad \text{(Hardy's inequality),}$$

and the equality holds only if $f(x)$ is identically 0.

Further, if $f(x) > 0$,

$$\int_0^\infty \exp\left[\frac{1}{x}\int_0^x \log f(t)\,dt\right]dx < e\int_0^\infty f(x)\,dx \qquad \text{(Carleman's inequality).}$$

(11) Let $a \leqslant x < \xi \leqslant b$, $p > 1$, and

$$\sup_{\xi} \frac{1}{\xi - x}\int_x^\xi f(t)\,dt = \theta(x).$$

Then

$$\int_a^b [\theta(x)]^p\,dx \leqslant 2\left(\frac{p}{p-1}\right)^p \int_a^b |f(x)|^p\,dx \qquad \text{(Hardy-Littlewood supremum theorem).}$$

(12) If $f(x)$ is piecewise smooth in $0 \leqslant x \leqslant \pi$ and $f(0) = f(\pi) = 0$,

$$\int_0^\pi [f'(x)]^2\,dx \geqslant \int_0^\pi [f(x)]^2\,dx \qquad \text{(Wirtinger's inequality),}$$

and the equality holds only if $f(x)$ is a constant multiple of $\sin x$.

9. Differential and Integral Calculus

(I) Derivatives and Primitive Functions (→ 106 Differential Calculus, 216 Integral Calculus)

$F(x) \equiv \int f(x)dx$	$f(x) \equiv F'(x)$
$\alpha\varphi + \beta\psi$ (α, β constants)	$\alpha\varphi' + \beta\psi'$
$\varphi \cdot \psi$	$\varphi'\psi + \varphi\psi'$
φ/ψ $(\psi \neq 0)$	$(\varphi'\psi - \varphi\psi')/\psi^2$
$\log\lvert\varphi\rvert$ $(\varphi \neq 0)$	φ'/φ (logarithmic differentiation)
$\Phi(\varphi)$ (composite)	$(d\Phi/d\varphi)\cdot\varphi'$
c (constant)	0
x^n	nx^{n-1}
$x^{n+1}/(n+1)$	x^n $(n \neq -1)$
$\log\lvert x\rvert$	$1/x$
$\log_a\lvert x\rvert$	$(\log_a e)/x$
$x(\log x - 1)$	$\log x$
$\exp x = e^x$	$\exp x = e^x$
a^x $(a > 0)$	$a^x \log a$
x^x	$x^x(1 + \log x)$
$(x-1)e^x$	xe^x
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\sec^2 x$
$\cot x$	$-\operatorname{cosec}^2 x$
$\sec x$	$\sec x \tan x$
$\operatorname{cosec} x$	$-\operatorname{cosec} x \cot x$
$\sinh x = (e^x - e^{-x})/2$	$\cosh x$
$\cosh x = (e^x + e^{-x})/2$	$\sinh x$
$\tanh x = \sinh x/\cosh x$	$\operatorname{sech}^2 x$
$\coth x = \cosh x/\sinh x$	$-\operatorname{cosech}^2 x$
$\operatorname{sech} x = 1/\cosh x$	$-\operatorname{sech} x \tanh x$
$\operatorname{cosech} x = 1/\sinh x$	$-\operatorname{cosech} x \coth x$
$\arcsin x$ $(\lvert F\rvert < \pi/2)$	$1/\sqrt{1-x^2}$
$\arccos x$ $(0 < F < \pi)$	$-1/\sqrt{1-x^2}$
$\arctan x$ $(\lvert F\rvert < \pi/2)$	$1/(1+x^2)$
$\operatorname{arccot} x$ $(\lvert F\rvert < \pi/2)$	$-1/(1+x^2)$
$\operatorname{arcsec} x$ $(0 < F < \pi)$	$1/\lvert x\rvert\sqrt{x^2-1}$
$\operatorname{arccosec} x$ $(\lvert F\rvert < \pi/2)$	$-1/\lvert x\rvert\sqrt{x^2-1}$
$\operatorname{arcsinh} x = \log(x + \sqrt{x^2+1})$	$1/\sqrt{x^2+1}$
$\operatorname{arccosh} x = \log(x + \sqrt{x^2-1})$	$1/\sqrt{x^2-1}$
$\frac{1}{2}\log\left\lvert\frac{1+x}{1-x}\right\rvert = \begin{cases} \operatorname{arctanh} x & (\lvert x\rvert < 1) \\ \operatorname{arccoth} x & (\lvert x\rvert > 1) \end{cases}$	$\frac{1}{1-x^2}$
$\operatorname{arcsech} x$	$-1/x\sqrt{1-x^2}$
$\operatorname{arccosech} x$	$-1/\lvert x\rvert\sqrt{1+x^2}$

$F(x) \equiv \int f(x)\,dx$	$f(x) \equiv F'(x)$
$\frac{1}{2a}\log\left\lvert\frac{a-x}{a+x}\right\rvert \quad (a>0)$	$\frac{1}{x^2-a^2}$
$(1/a)\arctan(x/a)$	$1/(x^2+a^2)$
$(x\sqrt{1-x^2}+\arcsin x)/2$	$\sqrt{1-x^2}$
$[x\sqrt{x^2\pm 1}\pm\log(x+\sqrt{x^2\pm 1})]/2$	$\sqrt{x^2\pm 1}$
$-\log\lvert\cos x\rvert$	$\tan x$
$\log\lvert\sin x\rvert$	$\cot x$
$\log\lvert\tan x\rvert$	$1/\sin x\cos x$
$\log\lvert\tan[(\pi/4)+(x/2)]\rvert$	$\sec x$
$\log\lvert\tan(x/2)\rvert$	$\operatorname{cosec} x$
$(x/2)-(1/4)\sin 2x$	$\sin^2 x$
$\sin x - x\cos x$	$x\sin x$
$\cos x + x\sin x$	$x\cos x$
$\frac{n\sin mx\sin nx+m\cos mx\cos nx}{n^2-m^2} \quad (n^2\neq m^2)$	$\sin mx\cos nx$
$e^{bx}\frac{b\sin ax-a\cos ax}{a^2+b^2}$	$e^{bx}\sin ax$
$e^{bx}\frac{b\cos ax+a\sin ax}{a^2+b^2}$	$e^{bx}\cos ax$
$x\arcsin x+\sqrt{1-x^2}$	$\arcsin x$
$x\arctan x-(1/2)\log(1+x^2)$	$\arctan x$
$\det(\varphi_{jk})_{j,k=1,\ldots,n}$	$\sum\det(\varphi_{j1}\cdots\varphi_{j\nu-1}\varphi'_{j\nu}\varphi_{j\nu+1}\cdots\varphi_{jn})_{j=1,\ldots,n}$

(II) Recurrence Formulas for Indefinite Integrals

(1) $I_m\equiv\int\frac{dx}{(1+x^2)^m}$ (m is a positive integer).

$$I_m=\frac{1}{2m-2}\frac{x}{(1+x^2)^{m-1}}+\frac{2m-3}{2m-2}I_{m-1}\quad(m\geqslant 2);\quad I_1=\arctan x.$$

(2) $I_m\equiv\int\frac{x^m}{\sqrt{ax^2+bx+c}}\,dx$ (m is an integer, $a\neq 0$).

The case $m<0$ is reduced to the case $m\geqslant 0$ by the change of variable $1/x=t$.

$$I_m=\frac{1}{ma}x^{m-1}\sqrt{ax^2+bx+c}-\frac{(2m-1)b}{2ma}I_{m-1}-\frac{(m-1)c}{ma}I_{m-2}\quad(m\geqslant 1);$$

$$I_0=\begin{cases}(1/\sqrt{a})\log\lvert 2ax+b+2\sqrt{a}\sqrt{ax^2+bx+c}\rvert & (a>0),\\ \frac{-1}{\sqrt{-a}}\arcsin\frac{2ax+b}{\sqrt{b^2-4ac}} & (a<0);\end{cases}$$

In this case, for the integrand to be a real function it is necessary that $b^2-4ac>0$.

(3) $I_m\equiv\int x^m e^x\,dx$ (m is an integer).

$$I_m=x^m e^x-mI_{m-1};\ I_0=e^x,\ I_{-1}=\operatorname{Ei}x,$$

where Ei is the exponential integral function (→ Table 19.II.3, this Appendix).

(4) $I_{m,n}\equiv\int x^m(\log x)^n\,dx$ (m,n are integers, $n\geqslant 0$).

$$I_{m,n}=\frac{x^{m+1}}{m+1}(\log x)^n-\frac{n}{m+1}I_{m,n-1};\ I_{m,0}=\frac{x^{m+1}}{m+1}\quad (m\neq -1),\ I_{-1,n}=(\log x)^{n+1}/(n+1).$$

(5) $I_m\equiv\int x^m\sin x\,dx,\ J_m\equiv\int x^m\cos x\,dx$ (m is a nonnegative integer).

$$I_m=-x^m\cos x+mJ_{m-1}=x^{m-1}(m\sin x-x\cos x)-m(m-1)I_{m-2},$$
$$J_m=x^m\sin x-mI_{m-1}=x^{m-1}(x\sin x+m\cos x)-m(m-1)J_{m-2};$$
$$I_0=-\cos x,$$
$$J_0=\sin x,$$

(6) $I_{m,n}\equiv\int \sin^m x\cdot\cos^n x\,dx$ (m,n are integers).

$$\left.\begin{aligned} I_{m,n}&=\frac{\sin^{m+1}x\cos^{n-1}x}{m+n}+\frac{n-1}{m+n}I_{m,n-2}\\ I_{m,n}&=\frac{-\sin^{m-1}x\cos^{n+1}x}{m+n}+\frac{m-1}{m+n}I_{m-2,n}\end{aligned}\right\}\ (m+n\neq 0),$$

$$I_{m,n}=\frac{-\sin^{m+1}x\cos^{n+1}x}{n+1}+\frac{m+n+2}{n+1}I_{m,n+2}\ (n\neq -1),$$
$$I_{m,n}=\frac{\sin^{m+1}x\cos^{n+1}x}{m+1}+\frac{m+n+2}{m+1}I_{m+2,n}\ (m\neq -1);$$
$$I_{1,1}=(\sin^2 x)/2,\quad I_{1,0}=-\cos x,\quad I_{1,-1}=-\log|\cos x|,\quad I_{0,1}=\sin x,\quad I_{0,0}=x,$$
$$I_{0,-1}=\log|\tan[(x/2)+(\pi/4)]|,\quad I_{-1,1}=\log|\sin x|,$$
$$I_{-1,0}=\log|\tan(x/2)|,\quad I_{-1,-1}=\log|\tan x|.$$
$$I_{m,-m}\equiv\int\tan^m x\,dx=\frac{\tan^{m-1}x}{m-1}-I_{m-2,-(m-2)}(m\neq 1).$$

(III) Derivatives of Higher Order

$f(x)$	$f^{(n)}(x)$
$\varphi\cdot\psi$	$\sum_{\nu=0}^{n}\binom{n}{\nu}\varphi^{(\nu)}\psi^{(n-\nu)}$ (Leibniz's formula)
x^k	$\prod_{\nu=0}^{n-1}(k-\nu)x^{k-n}$
$(x+a)^n$	$n!$
$\exp x$	$\exp x$
$a^x\ (a>0)$	$a^x(\log a)^n$
$\log x$	$(-1)^{n-1}(n-1)!/x^n$
$\sin x$	$\sin[x+(n\pi/2)]$
$\cos x$	$\cos[x+(n\pi/2)]$
$e^{ax}\cos bx$	$r^n e^{ax}\cos(bx+n\theta)$ (where $a=r\cos\theta,\ b=r\sin\theta$)
$\arcsin x$	$\frac{1}{2^{n-1}}\sum_{\nu=0}^{n-1}(-1)^\nu\binom{n-1}{\nu}(2\nu-1)!!(2n-2\nu-3)!!(1+x)^{-(1/2)-\nu}(1-x)^{(1/2)-n+\nu}$ (where $(2\nu-1)!!\equiv 1\cdot 3\cdot 5\cdot\ldots\cdot(2\nu-1),\ (-1)!!\equiv 1$)
$\arctan x$	$(-1)^{n-1}(n-1)!\sin^n\theta\sin n\theta$ (where $x=\cot\theta$)

$$\left(\frac{1}{f}\right)'=-\frac{f'}{f^2},\quad \left(\frac{1}{f}\right)''=\frac{2f'^2-ff''}{f^3},\quad \left(\frac{1}{f}\right)'''=\frac{6ff'f''-6f'^3-f^2f'''}{f^4}.$$

Higher-order derivatives of a composite function $g(t)\equiv f(x_1(t),\ldots,x_n(t))$

$$\frac{dg}{dt}=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\frac{dx_i}{dt},\quad \frac{d^2g}{dt^2}=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\frac{d^2x_i}{dt^2}+\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}\frac{dx_i}{dt}\frac{dx_j}{dt},$$

$$\frac{d^3g}{dt^3}=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}\frac{d^3x_i}{dt^3}+2\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}\frac{d^2x_i}{dt^2}\frac{dx_j}{dt}+\sum_{i,j,k=1}^{n}\frac{\partial^3 f}{\partial x_i\partial x_j\partial x_k}\frac{dx_i}{dt}\frac{dx_j}{dt}\frac{dx_k}{dt}.$$

For a function $z=z\,(x_1,\ldots,x_n)$ determined implicitly by $F(z;x_1,\ldots,x_n)=0$, we have

$$\frac{\partial z}{\partial x_i}=-\frac{F_{x_i}}{F_z},\quad \frac{\partial^2 z}{\partial x_i\partial x_j}=-\frac{F_{x_ix_j}}{F_z}+\frac{F_{x_i}F_{x_jz}+F_{x_j}F_{x_iz}}{F_z^2}-\frac{F_{x_i}F_{x_j}F_{zz}}{F_z^3}.$$

Schwarzian derivative:

$$\{y;x\}\equiv\left(\frac{d^3y}{dx^3}\right)\Big/\left(\frac{dy}{dx}\right)-\frac{3}{2}\left[\left(\frac{d^2y}{dx^2}\right)\Big/\left(\frac{dy}{dx}\right)\right]^2,\quad \{y;x\}=0\Leftrightarrow y=(ax+b)/(cx+d),$$

$$\{y;x\}=\left(\frac{dz}{dx}\right)^2[\{y;z\}-\{x;z\}]=-\left(\frac{dy}{dx}\right)^2\{x;y\},\ \{(ay+b)/(cy+d);x\}=\{y;x\}.$$

(IV) The Taylor Expansion and Remainder

If $f(x)$ is n times continuously differentiable in the interval $[a,b]$ (i.e., of class C^n),

$$f(b)=\sum_{\nu=0}^{n-1}\frac{(b-a)^\nu}{\nu!}f^{(\nu)}(a)+R_n\qquad\text{(Taylor's formula)}.$$

R_n is called the remainder, and is represented as follows:

$$R_n=\frac{1}{(n-1)!}\int_a^b(b-x)^{n-1}f^{(n)}(x)\,dx=\frac{(b-a)^p(b-\xi)^{n-p}}{(n-1)!p}f^{(n)}(\xi)$$

$$=\frac{(b-a)^n}{(n-1)!p}(1-\theta)^{n-p}f^{(n)}(a+\theta(b-a))\quad(n\geqslant p>0,\ 0<\theta<1,\ a<\xi<b,\ \xi=a+\theta(b-a))$$

(Roche-Schlömilch remainder);

$$=\frac{1}{n!}(b-a)^n f^{(n)}(\xi)\qquad\text{(Lagrange's remainder)};$$

$$=\frac{1}{(n-1)!}(b-a)(b-\xi)^{n-1}f^{(n)}(\xi)\qquad\text{(Cauchy's remainder)}.$$

If $f(x,y)$ is m times continuously differentiable in a neighborhood of a point (x_0,y_0),

$$f(x_0+h,y_0+k)=\sum_{\lambda=0}^{m-1}\frac{1}{\lambda!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^\lambda f(x_0,y_0)+R_m$$

$$=\sum_{0\leqslant\mu+\nu\leqslant m-1;\,\mu,\nu\geqslant0}\frac{1}{\mu!\nu!}h^\mu k^\nu\frac{\partial^{\mu+\nu}f(x_0,y_0)}{\partial x^\mu\partial y^\nu}+R_m,$$

$$R_m=\frac{1}{m!}\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^m f(x_0+\theta h,y_0+\theta k)\quad(0<\theta<1).$$

If all partial derivatives up to order $m-1$ are totally differentiable,

$$f(x_1+h_1,\ldots,x_n+h_n)=\sum_{\nu=0}^{m-1}\frac{1}{\nu!}\left(\sum_{\mu=1}^{n}h_\mu\frac{\partial}{\partial x_\mu}\right)^\nu f(x_1,\ldots,x_n)+R_m$$

$$=\sum\frac{1}{\nu_1!\ldots\nu_n!}h_1^{\nu_1}\ldots h_n^{\nu_n}\frac{\partial^{\nu_1+\ldots+\nu_n}f(x_1,\ldots,x_n)}{\partial x_1^{\nu_1}\ldots\partial x_n^{\nu_n}}+R_m,$$

where the Σ means the sum for $\nu_1,\ldots,\nu_n$ in the domain $0\leqslant\nu_1+\ldots+\nu_n\leqslant m-1;\ \nu_1,\ldots,\nu_m\geqslant0$.
The remainder R_m is expressed as

$$R_m=\frac{1}{m!}\left(\sum_{\mu=1}^{n}h_\mu\frac{\partial}{\partial x_\mu}\right)^m f(x_1+\theta h_1,\ldots,x_n+\theta h_n)\quad(0<\theta<1).$$

(V) Definite Integrals [4]

In the following formulas, we assume that m, n are positive integers. δ_{mn} is Kronecker's delta ($\delta_{mn}=0$ or 1 for $m\neq n$ or $m=n$), Γ is the gamma function, B is the beta function, and C is the Euler constant.

For simplicity, we put

$$m!!\equiv\begin{cases}1\cdot3\cdot5\cdot\ldots\cdot(m-2)\cdot m=2^{(m+1)/2}\Gamma[(m/2)+1]/\sqrt{\pi}=m!/2^{(m-1)/2}[(m-1)/2]! & (m\text{ is odd}),\\ 2\cdot4\cdot6\cdot\ldots\cdot(m-2)\cdot m=2^{m/2}\Gamma[(m/2)+1]=2^{m/2}(m/2)! & (m\text{ is even}).\end{cases}$$

In an n-dimensional real space, the volume of the domain

$$|x_1|^p+\ldots+|x_n|^p\leqslant 1\quad(p>0)\qquad\text{is}\qquad\frac{2^n[\Gamma(1/p)]^n}{p^{n-1}n\Gamma(n/p)}.$$

For $p=2$, this is the volume of the unit hypersphere, which is

$$\frac{\pi^{n/2}}{\Gamma[(n/2)+1]}=\begin{cases}(2\pi)^{n/2}/n!! & (n\text{ is even}),\\ 2(2\pi)^{(n-1)/2}/n!! & (n\text{ is odd}).\end{cases}$$

The surface area of the $(n-1)$-dimensional unit hypersphere

$$|x_1|^2+\ldots+|x_n|^2=1\qquad\text{is}\qquad\frac{2\pi^{n/2}}{\Gamma(n/2)}=\begin{cases}(2\pi)^{n/2}/(n-2)!! & (n\text{ is even}),\\ 2(2\pi)^{(n-1)/2}/(n-2)!! & (n\text{ is odd}).\end{cases}$$

$$\int_0^\infty x^{p-1}e^{-x}dx=\int_0^1\left(\log\frac{1}{x}\right)^{p-1}dx\equiv\Gamma(p).$$

$$\int_0^1 x^p\left(\log\frac{1}{x}\right)^q dx=\frac{\Gamma(q+1)}{(p+1)^{q+1}}\quad(\mathrm{Re}\,p,\mathrm{Re}\,q>-1).$$

$$\int_0^1 x^{p-1}(1-x)^{q-1}dx=\int_0^\infty\frac{x^{p-1}}{(1+x)^{p+q}}dx\equiv B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

$$\int_0^\infty\frac{x^a}{(1+x^c)^{1+b}}dx=\frac{1}{c}\,\frac{\Gamma[(a+1)/c]\Gamma[b-\{(a-c+1)/c\}]}{\Gamma(1+b)}$$

$$\left(\mathrm{Re}\,c>0;\mathrm{Re}\,a,\mathrm{Re}\,b>-1;\mathrm{Re}\,b>\mathrm{Re}\frac{a-c+1}{c}\right).$$

$$\int_0^a\frac{(a-x)^{p-1}(x-b)^{q-1}}{|x-c|^{p+q}}dx=\frac{(a-b)^{p+q-1}}{|a-c|^q|b-c|^p}B(p,q)$$

$$(0<c<b<a\quad\text{or}\quad 0<b<a<c;\mathrm{Re}\,p,\mathrm{Re}\,q>0).$$

$$\int_{-\infty}^\infty\frac{1}{(1+x^2)^{n+1}}dx=\frac{\pi(2n)!}{2^{2n}(n!)^2}=\pi\frac{(2n-1)!!}{(2n)!!}.$$

$$\int_{-\infty}^\infty\frac{x^{2m}}{1+x^{2n}}dx=\frac{\pi}{n\sin[(2m+1)\pi/2n]}\quad(2m+1<2n).$$

$$\int_0^\infty\frac{x^{a-1}}{1+x}dx=\Gamma(a)\Gamma(1-a)=\frac{\pi}{\sin a\pi}\quad(0<a<1).$$

$$\int_0^\infty e^{-a^2x^2}dx=\frac{\sqrt{\pi}}{2|a|}.\qquad\int_0^\infty x^pe^{-x^2}dx=\frac{1}{2}\Gamma\left(\frac{p+1}{2}\right)=\begin{cases}(2n-1)!!\sqrt{\pi}/2^{n+1} & (p=2n),\\ n!/2 & (p=2n+1).\end{cases}$$

$$\int_0^\infty(e^{-a^2/x^2}-e^{-b^2/x^2})dx=(b-a)\sqrt{\pi}\quad(a,b\geqslant0).\qquad\int_0^\infty e^{-[x-(1/x)]^2}dx=\frac{\sqrt{\pi}}{2}.$$

$$\int_0^\infty e^{-x^2-(a^2/x^2)}dx=\frac{e^{-2a}\sqrt{\pi}}{2}\quad(a\geqslant0).\qquad\int_0^\infty\frac{1}{e^{ax}+e^{-ax}}dx=\frac{\pi}{4a}.$$

$$\int_0^\infty\frac{x}{e^{ax}-e^{-ax}}dx=\frac{\pi^2}{8a^2}\quad(a>0).$$

$$\int_0^\infty\frac{x}{e^x-e^{-x}}dx=\int_0^1\frac{\log(1/x)}{1-x^2}dx=\frac{\pi^2}{8}.$$

$$\int_0^\infty \frac{x}{e^x+1}\,dx=\int_0^1 \frac{\log(1/x)}{1+x}\,dx=\frac{\pi^2}{12}.\quad \int_0^\infty \frac{x}{e^x-1}\,dx=\int_0^1 \frac{\log(1/x)}{1-x}\,dx=\frac{\pi^2}{6}.$$

$$\int_0^\infty \log\left(\frac{e^x+1}{e^x-1}\right)dx=\int_0^1 \log\left(\frac{1+x}{1-x}\right)\frac{1}{x}\,dx=\frac{\pi^2}{4}.$$

$$\int_0^1 \frac{\log x}{\sqrt{1-x^2}}\,dx=\int_0^{\pi/2}\log\sin x\,dx=-\frac{\pi}{2}\log 2.$$

$$\int_0^1 \frac{\log(1+x)}{1+x^2}\,dx=\frac{\pi}{8}\log 2.\quad \int_0^1 \frac{\log x}{1+x^2}\,dx=-\int_1^\infty \frac{\log x}{1+x^2}\,dx=\sum_{n=0}^{\infty}\frac{(-1)^{n-1}}{(2n+1)^2}=-0.91596\ldots.$$

$$\int_0^\infty \frac{(\log x)^2}{1+x+x^2}\,dx=\frac{16\pi^3}{81\sqrt{3}}.\quad \int_0^1 \frac{x^p-x^q}{\log x}\,dx=\log\frac{p+1}{q+1}\quad (p,q>-1).$$

$$\int_0^1 \log|\log x|\,dx=\int_0^\infty e^{-t}\log t\,dt=-C=-0.57721\ldots.$$

$$\int_0^\pi \sin mx\sin nx\,dx=\int_0^\pi \cos mx\cos nx\,dx=\delta_{mn}\frac{\pi}{2}.$$

$$\int_0^\pi \sin mx\cos nx\,dx=\begin{cases}\left[1-(-1)^{m+n}\right]\dfrac{m}{m^2-n^2} & (m\neq n);\\ 0 & (m=n).\end{cases}$$

$$\int_0^{\pi/2}\sin^p x\cos^q x\,dx=\frac{1}{2}B\left(\frac{p+1}{2},\frac{q+1}{2}\right)\quad \left(\mathrm{Re}\,p,\mathrm{Re}\,q>-\frac{1}{2}\right);$$

$$=\begin{cases}(\pi/2)(p-1)!!(q-1)!!/(p+q)!! & (p,q\text{ are even positive integers}),\\ (p-1)!!(q-1)!!/(p+q)!! & (p,q\text{ are positive integers not both even}).\end{cases}$$

$$\int_0^{\pi/2}\sin^p x\,dx=\int_0^{\pi/2}\cos^p x\,dx=\frac{\sqrt{\pi}}{2}\frac{\Gamma[(p+1)/2]}{\Gamma[(p/2)+1]}\quad (\mathrm{Re}\,p>-1);$$

$$=\begin{cases}(\pi/2)(2n-1)!!/(2n)!! & (p=2n),\\ (2n)!!/(2n+1)!! & (p=2n+1).\end{cases}$$

$$\int_{-\infty}^\infty \sin(x^2)\,dx=\int_{-\infty}^\infty \cos(x^2)\,dx=\int_0^\infty \frac{\sin x}{\sqrt{x}}\,dx=\int_0^\infty \frac{\cos x}{\sqrt{x}}\,dx=\sqrt{\frac{\pi}{2}}\quad \text{(Fresnel integral)}.$$

$$\int_0^\infty \frac{\sin ax}{x}\,dx=\frac{\pi}{2}\quad (a>0).\quad \int_0^\infty \frac{\tan x}{x}\,dx=\frac{\pi}{2}$$

$$\left(\text{take Cauchy's principal value at } x=\left(n+\frac{1}{2}\right)\pi\right).$$

$$\int_0^\infty \frac{\sin^{2n+1}x}{x}\,dx=\frac{\pi}{2}\frac{(2n-1)!!}{(2n)!!}.\quad \int_0^\infty \frac{\sin^2 x}{x^2}\,dx=\frac{\pi}{2}.\quad \int_0^\infty \frac{\sin(x^2)}{x}\,dx=\frac{\pi}{4}.$$

$$\int_0^\infty \frac{\sin qx}{x^p}\,dx=\frac{\pi q^{p-1}}{2\Gamma(p)\sin(p\pi/2)}\quad (0<p<2).$$

$$\int_0^\infty \frac{\cos px}{1+x^2}\,dx=\frac{\pi}{2}e^{-|p|}.\quad \int_0^\infty \frac{\cos^2 ax}{1+x^2}\,dx=\frac{\pi}{4}(1+e^{-2a})\quad (a>0).$$

$$\int_0^\infty \frac{\sin ax}{x(1+x^2)}\,dx=\frac{\pi}{2}(1-e^{-a})\quad (a>0).\quad \int_0^\infty \frac{x\sin ax}{1+x^2}\,dx=\frac{\pi}{2}e^{-a}\quad (a>0).$$

$$\int_0^\infty \frac{\sin^{2m+1}x\cos^{2n}x}{x}\,dx=\int_0^\infty \frac{\sin^{2m+1}x\cos^{2n-1}x}{x}\,dx=\frac{\pi}{2}\frac{(2m-1)!!(2n-1)!!}{(2m+2n)!!}.$$

$$\int_0^\infty \frac{\sin ax\cos bx}{x}\,dx=\begin{cases}\pi/2 & (a>b>0),\\ \pi/4 & (a=b>0),\\ 0 & (b>a>0)\end{cases}\quad \text{(Dirichlet's discontinuous factor)}.$$

$$\int_0^{2\pi}\frac{1}{1+a\cos x}\,dx=\frac{2\pi}{\sqrt{1-a^2}}\quad (|a|<1).\quad \int_0^{\pi/2}\frac{1}{a^2\cos^2 x+b^2\sin^2 x}\,dx=\frac{\pi}{2ab}\quad (ab\neq 0).$$

$$\int_0^\pi \frac{x\sin x}{1+\cos^2 x}\,dx=\frac{\pi^2}{4}.$$

$$\int_0^\pi \frac{\cos nx}{1-2a\cos x+a^2}dx=\begin{cases}\pi a^n/(1-a^2) & (|a|<1),\\ \pi/a^n(a^2-1) & (|a|>1).\end{cases}$$

References

[1] B. O. Peirce, A short table of integrals, Ginn, Boston, second revised edition, 1910.
[2] D. Bierens de Haan, Nouvelles tables d'intégrales définies, Leiden, 1867.
There are several mistakes in this table. For the errata, see
[3] C. F. Lindmann, Examen des nouvelles tables de M. Bierens de Haan, Handlingar Svenska Vetenskaps-Akad., 1891.
[4] E. W. Sheldon, Critical revision of de Haan's tables of definite integrals, Amer. J. Math., 34 (1912), 39–114.

10. Series (→ 379 Series)

(I) Finite Series

(1) $S_k \equiv 1^k+2^k+\ldots+n^k$ (k is an integer). For $k \geqslant 0$, we have

$$S_k=\frac{B_{k+1}(n+1)-B_{k+1}(1)}{k+1}=\sum_{i=0}^{k}(-1)^i\binom{k+1}{i}\frac{B_{2i}(n+1)^{k+1-i}}{k+1},$$

where B_l is a Bernoulli number and $B_l(x)$ is a Bernoulli polynomial. In particular,

$$S_0=n,\quad S_1=n(n+1)/2,\quad S_2=n(n+1)(2n+1)/6,\quad S_3=n^2(n+1)^2/4,$$
$$S_4=n(n+1)(2n+1)(3n^2+3n-1)/30.$$

For $k<0$ and $k=-l$,

$$S_{-l}=c_l-\left[(-1)^l/(l-1)!\right]\left[d^l\log\Gamma(x)/dx^l\right]_{x=n+1}$$
$$=c_l-\frac{1}{(l-1)(n+1)^{l-1}}-\frac{1}{2(n+1)^l}+\sum_{i=1}^{\infty}(-1)^i\frac{B_{2(i+1)}}{(i+1)!}\frac{(l+i-1)!}{(l-1)!}\frac{1}{(n+1)^{l+i}}.$$

For $l=1$, the second term in the latter formula is replaced by $\log[1/(n+1)]$. Here Γ is the gamma function, and the constants c_l are

$$c_l=\begin{cases} C \quad (\text{Euler constant}) & (l=1),\\ \zeta(l) \quad (\zeta \text{ is the Riemann zeta function}) & (l\geqslant 2).\end{cases}$$

(2) $$\sum_{i=1}^{n} i(i+1)\ldots(i+m-1)\equiv\sum_{i=1}^{n}\frac{(i+m-1)!}{(i-1)!}=\frac{1}{m+1}\frac{(n+m)!}{(n-1)!},$$

$$\sum_{i=1}^{n}\frac{(i-1)!}{(i+m-1)!}=\frac{1}{m-1}\left[\frac{1}{(m-1)!}-\frac{n!}{(n+m-1)!}\right]\quad(m\geqslant 2),$$

$$\sum_{i=1}^{n} i!i=(n+1)!-1,\quad \sum_{i=1}^{n} i\binom{n}{i}=n2^{n-1},$$

$$\sum_{i=m}^{n}\binom{i}{m}\binom{n+s-i-1}{n-i}=\binom{n+s}{m+s}\quad(m\leqslant n),$$

$$\sum_{i=0}^{n}\binom{n}{i}\binom{m}{r-i}=\binom{n+m}{r}.$$

$$\sum_{i=1}^{n} a^i=\begin{cases} a(a^n-1)/(a-1) & (a\neq 1)\\ n & (a=1)\end{cases}\quad \text{(geometric progression)}$$

$$\sum_{j=0}^{n}(a+jd)=(n+1)a+\frac{n(n+1)}{2}d=\frac{n+1}{2}(a+a+nd)\quad\text{(arithmetic progression)}$$

$$\sum_{j=0}^{n}\sin(\alpha+j\beta)=\sin\left(\alpha+\frac{n}{2}\beta\right)\sin\frac{(n+1)\beta}{2}\Big/\sin\frac{\beta}{2},$$

$$\sum_{j=0}^{n} \cos(\alpha + j\beta) = \cos\left(\alpha + \frac{n}{2}\beta\right)\sin\frac{(n+1)\beta}{2} \Big/ \sin\frac{\beta}{2},$$

$$\sum_{j=0}^{n} \operatorname{cosec} 2^j\alpha = \cot(\alpha/2) - \cot 2^n\alpha.$$

(II) Convergence Criteria for Positive Series $\sum a_n$

In the present Section II, we assume that $a_n \geqslant 0$.
Cauchy's criterion: The series converges when $\limsup \sqrt[n]{a_n} < 1$ and it diverges when $\limsup \sqrt[n]{a_n} > 1$.
d'Alembert's criterion: The series converges when $\limsup a_{n+1}/a_n < 1$ and diverges when $\liminf a_{n+1}/a_n > 1$.
Raabe's criterion: The series converges when $\liminf n[(a_n/a_{n+1}) - 1] > 1$ and diverges when $\limsup n[(a_n/a_{n+1}) - 1] < 1$.
Kummer's criterion: For a positive divergent series $\sum(1/b_n)$, the series $\sum a_n$ converges when $\liminf[(b_n a_n/a_{n+1}) - b_{n+1}] > 0$ and diverges when $\limsup[(b_n a_n/a_{n+1}) - b_{n+1}] < 0$ diverges.
Gauss's criterion: Suppose $a_n/a_{n+1} = 1 + (k/n) + (\theta_n/n^\lambda)$, where $\lambda > 1$ and $\{\theta_n\}$ is bounded. Then the series $\sum a_n$ converges when $k > 1$; and diverges when $k \leqslant 1$.
Schlömilch's criterion: For a decreasing positive sequence $a_\nu \downarrow 0$, let n_ν be an increasing sequence of positive integers and suppose that $(n_{\nu+2} - n_{\nu+1})/(n_{\nu+1} - n_\nu)$ is bounded. Then the two series $\sum a_n$ and $\sum(n_{\nu+1} - n_\nu)a_{n_\nu}$ converge or diverge simultaneously.
Logarithmic criterion: For a positive integer k, we put

$$\log_k x \equiv \log(\log_{k-1} x), \quad \log_1 x = \log x.$$

Then for sufficiently large n we have
The first logarithmic criterion: If

$$a_n - 1/(n \log_1 n \ldots \log_{k-1} n (\log_k n)^p) \begin{cases} \leqslant 0, & p > 1 \quad \text{then } \sum a_n \text{ converges}, \\ \geqslant 0, & p \leqslant 1 \quad \text{then } \sum a_n \text{ diverges}. \end{cases}$$

The second logarithmic criterion: If

$$\frac{a_{n+1}}{a_n} - \frac{n}{n+1}\,\frac{\log_1 n}{\log_1(n+1)} \cdots \frac{\log_{k-1} n}{\log_{k-1}(n+1)} \left(\frac{\log_k n}{\log_k(n+1)}\right)^p$$

$$\begin{cases} \leqslant 0, & p > 1 \quad \text{then } \sum a_n \text{ converges}, \\ \geqslant 0, & p \leqslant 1 \quad \text{then } \sum a_n \text{ diverges}. \end{cases}$$

(III) Infinite Series

$$\sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i} = \log 2, \quad \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{2i-1} = \frac{\pi}{4} \qquad \text{(Leibniz's formula)},$$

$$\sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{2i(2i+1)(2i+2)} = \frac{\pi - 3}{4},$$

$$\sum_{i=0}^{\infty} \frac{(2i)!}{2^{2i}(i!)^2}\,\frac{1}{2i+1} = \frac{\pi}{2}, \quad \sum_{i=1}^{\infty}\left(\frac{1}{i} - \log\left(1 + \frac{1}{i}\right)\right) = C \quad (C \text{ is Euler's constant}).$$

Putting

$$\zeta(n) = \sum_{i=1}^{\infty} \frac{1}{i^n}, \quad \alpha(n) = \sum_{i=1}^{\infty} \frac{1}{(2i-1)^n}, \quad \beta(n) = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i^n}, \quad \varepsilon(n) = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{(2i-1)^n},$$

we have

$$\zeta(2n)=\frac{(2\pi)^{2n}}{2(2n)!}B_{2n},\quad \alpha(2n)=\frac{(2^{2n}-1)\pi^{2n}}{2(2n)!}B_{2n},$$

$$\beta(2n)=\frac{(2^{2n-1}-1)\pi^{2n}}{(2n)!}B_{2n},\quad \varepsilon(2n+1)=\frac{\pi^{2n+1}}{2^{2n+2}(2n)!}E_{2n}.$$

where B_n is a Bernoulli number, and E_n is an Euler number.

$\zeta(2)=\pi^2/6,\quad \zeta(4)=\pi^4/90,\quad \zeta(6)=\pi^6/945.$

$\alpha(2)=\pi^2/8,\quad \alpha(4)=\pi^4/96,\quad \alpha(6)=\pi^6/960.$

$\beta(2)=\pi^2/12,\quad \beta(4)=7\pi^4/720,\quad \beta(6)=31\pi^6/30240.$

$\varepsilon(1)=\pi/4,\quad \varepsilon(2)=0.91596\,55941\,77219\,01505\,46035\,14932\ldots$ (Catalan's constant),

$\varepsilon(3)=\pi^3/32,\quad \varepsilon(5)=5\pi^5/1536,\quad \varepsilon(7)=61\pi^7/92160.$

(IV) Power Series (→ 339 Power Series)

(1) Binomial Series $(1+x)^\alpha=\sum_{i=0}^{\infty}\binom{\alpha}{i}x^i$. This converges always in $|x|<1$. If $\alpha>0$, it converges in $-1\leqslant x\leqslant 1$, and if $-1<\alpha<0$, it converges in $-1<x\leqslant 1$. When α is 0 or a positive integer, it reduces to a polynomial and converges in $|x|<\infty$.

$$\frac{1}{1+x}=\sum_{i=0}^{\infty}(-1)^i x^i\quad(|x|<1),$$

$$\sqrt{1+x}=\sum_{i=0}^{\infty}\frac{(-1)^{i-1}(2i)!}{(2i-1)2^{2i}(i!)^2}x^i\quad(|x|<1),\qquad \frac{1}{\sqrt{1+x}}=\sum_{i=0}^{\infty}\frac{(-1)^i(2i)!}{2^{2i}(i!)^2}x^i\quad(|x|<1).$$

(2) Elementary Transcendental Functions (→ 131 Elementary Functions).

$$e^x\equiv\exp x=\sum_{i=0}^{\infty}\frac{x^i}{i!}=\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n,\quad a^x=\exp(x\log a)\quad(|x|<\infty).$$

$$\log(1+x)=\sum_{i=1}^{\infty}\frac{(-1)^{i-1}}{i}x^i\quad(-1<x\leqslant 1),\quad \log x=2\sum_{i=0}^{\infty}\frac{1}{2i+1}\left(\frac{x-1}{x+1}\right)^{2i+1}\quad(0<x<\infty).$$

$$\sin x=\sum_{i=0}^{\infty}\frac{(-1)^i}{(2i+1)!}x^{2i+1},\quad \cos x=\sum_{i=0}^{\infty}\frac{(-1)^i}{(2i)!}x^{2i}\quad(|x|<\infty).$$

$$\tan x=\sum_{i=1}^{\infty}\frac{2^{2i}(2^{2i}-1)B_{2i}}{(2i)!}x^{2i-1}\quad\left(|x|<\frac{\pi}{2}\right)\quad(B_i\text{ is a Bernoulli number}),$$

$$\cot x=\frac{1}{x}-\sum_{i=1}^{\infty}\frac{2^{2i}B_{2i}}{(2i)!}x^{2i-1}\quad\left(0<|x|<\frac{\pi}{2}\right),$$

$$\sec x=\sum_{i=0}^{\infty}\frac{E_{2i}}{(2i)!}x^{2i}\quad\left(|x|<\frac{\pi}{2}\right)\quad(E_i\text{ is an Euler number}),$$

$$\operatorname{cosec} x=\frac{1}{x}+\sum_{i=1}^{\infty}\frac{(2^{2i}-2)B_{2i}}{(2i)!}x^{2i-1}\quad(0<|x|<\pi).$$

$$\arcsin x=\sum_{i=0}^{\infty}\frac{(2i)!}{2^{2i}(i!)^2}\frac{x^{2i+1}}{2i+1}\quad(|x|\leqslant 1),\quad \arctan x=\sum_{i=0}^{\infty}\frac{(-1)^i}{2i+1}x^{2i+1}\quad(|x|\leqslant 1).$$

(V) Partial Fractions for Elementary Functions

$$\tan x=\sum_{n=0}^{\infty}\frac{8x}{(2n+1)^2\pi^2-4x^2},\quad \cot x=\frac{1}{x}+2x\sum_{n=1}^{\infty}\frac{1}{x^2-n^2\pi^2},$$

$$\sec x=4\sum_{n=1}^{\infty}\frac{(-1)^n(2n-1)\pi}{(2n-1)^2\pi^2-4x^2},\quad \operatorname{cosec} x=\frac{1}{x}+2x\sum_{n=1}^{\infty}\frac{(-1)^n}{x^2-n^2\pi^2},$$

$$\sec^2 x=\sum_{n=-\infty}^{\infty}\frac{1}{\left[x+\{(2n+1)\pi/2\}\right]^2},\quad \operatorname{cosec}^2 x=\sum_{n=-\infty}^{\infty}\frac{1}{(x+n\pi)^2}.$$

(VI) Infinite Products (→ 379 Series F)

$$\prod_{n=1}^{\infty}\frac{4n^2}{4n^2-1}=\frac{\pi}{2}\ \text{(Wallis formula)},\quad \prod_{n=1}^{\infty}\left(1+\frac{x}{a+n}\right)e^{-x/n}=e^{-Cx}\frac{\Gamma(1+a)}{\Gamma(1+a+x)}$$

(C is Euler's constant).

$$\prod_{n=1}^{\infty}\left(1-\frac{x}{2n-1}\right)\left(1+\frac{x}{2n}\right)=\sqrt{\pi}\Big/\Gamma\left(1+\frac{x}{2}\right)\Gamma\left(\frac{1}{2}-\frac{x}{2}\right).$$

$$\prod_p 1/(1-p^{-s})=\zeta(s)\quad (p \text{ ranges over all prime numbers}, s>1),$$

$$\prod_{n=1}^{\infty}\left(1-\frac{x^2}{n^2\pi^2}\right)=\frac{\sin x}{x},\quad \prod_{n=1}^{\infty}\cos\frac{x}{2^n}=\frac{\sin x}{x},\quad \prod_{n=1}^{\infty}\left(1-\frac{4x^2}{(2n-1)^2\pi^2}\right)=\cos x.$$

For $|q|<1$, putting $q_1\equiv\prod_{n=1}^{\infty}(1+q^{2n})$, $q_2\equiv\prod_{n=1}^{\infty}(1+q^{2n-1})$, $q_3\equiv\prod_{n=1}^{\infty}(1-q^{2n-1})$,

$q_4\equiv\prod_{n=1}^{\infty}(1-q^{2n})$ we have $q_1q_2q_3=1$.

Further, putting $q=e^{i\pi\tau}$, we have the following formulas concerning ϑ-functions (→134 Elliptic Functions):

$$\vartheta_4(0,\tau)=q_4q_3^2,\quad \vartheta_2(0,\tau)=2q^{1/4}q_4q_1^2,\quad \vartheta_3(0,\tau)=q_4q_2^2,\quad \vartheta_1'(0,\tau)=2\pi q^{1/4}q_4^3.$$

11. Fourier Analysis

(I) Fourier Series (→ 159 Fourier Series)

(1) Fourier coefficients $a_0=\frac{1}{a}\int_0^a f(x)\,dx,\quad a_n=\frac{2}{a}\int_0^a f(x)\cos\frac{n\pi x}{a}\,dx$

$$b_n=\frac{2}{a}\int_0^a f(x)\sin\frac{n\pi x}{a}\,dx.$$

Fourier cosine series $a_0+\sum_{n=1}^{\infty}a_n\cos\frac{n\pi x}{a}=\begin{cases} f(x) & (0<x<a),\\ f(-x) & (-a<x<0).\end{cases}$

Fourier sine series $\sum_{n=1}^{\infty}b_n\sin\frac{n\pi x}{a}=\begin{cases} f(x) & (0<x<a),\\ -f(-x) & (-a<x<0).\end{cases}$

The next table shows the Fourier coefficients of the functions $F(x)$ directly in the following manner from a given function $f(x)$ on the interval $[0,a]$. For x in $[-a,0]$ and when the cosine series $\{a_n\}$ is in question, we set $f(x)=f(-x)$, and when the sine series $\{b_n\}$ is in question we set $f(x)=-f(-x)$. Thus $f(x)$ is extended in two ways to functions on $[-a,a]$. The functions $F(x)$ are the periodic continuations of such functions. We remark that the sum of the Fourier series given by the Fourier coefficients in the right hand side has, in general, some singularities (discontinuity of the function or its higher derivatives, for example) at the points given by the integral multiples of a. We assume that μ is not an integer.

$f(x)$	a_0	$a_n \quad (n=1,2,\ldots)$	$b_n \quad (n=1,2,\ldots)$
1	1	0	$[1+(-1)^{n+1}]2a/n\pi$
x	$\frac{a}{2}$	$[1+(-1)^{n+1}]\frac{-2a}{n^2\pi^2}$	$(-1)^{n+1}\frac{2a}{n\pi}$
x^2	$\frac{a^2}{3}$	$(-1)^n\frac{4a^2}{n^2\pi^2}$	$(-1)^{n-1}\frac{2a^2}{n\pi}-[1+(-1)^{n+1}]\frac{4a^2}{n^3\pi^3}$
e^{kx}	$\frac{e^{ka}-1}{ka}$	$\frac{2ka\left[(-1)^n e^{ka}-1\right]}{k^2a^2+n^2\pi^2}$	$\frac{2n\pi\left[1-(-1)^n\right]e^{ka}}{k^2a^2+n^2\pi^2}$
$\cos\frac{\mu\pi x}{a}$	$\frac{\sin\mu\pi}{\mu\pi}$	$(-1)^n\frac{2}{\pi}\frac{\mu\sin\mu\pi}{\mu^2-n^2}$	$\frac{2}{\pi}\frac{\left[(-1)^n\cos\mu\pi-1\right]}{\mu^2-n^2}$
$\sin\frac{\mu\pi x}{a}$	$\frac{1-\cos\mu\pi}{\mu\pi}$	$\frac{2}{\pi}\frac{\mu\left[1-(-1)^n\cos\mu\pi\right]}{\mu^2-n^2}$	$(-1)^n\frac{2}{\pi}\frac{n\sin\mu\pi}{\mu^2-n^2}$
$\frac{1-\lambda^2}{1-2\lambda\cos(\pi x/a)+\lambda^2}$	1	$2\lambda^n \quad (\lvert\lambda\rvert<1)$	
$\frac{\lambda\sin(\pi x/a)}{1-2\lambda\cos(\pi x/a)+\lambda^2}$			$\lambda^n \quad (\lvert\lambda\rvert<1)$
$B_{2m}(x/2a)$	0	$(-1)^{m+1}2(2m)!/(2n\pi)^{2m}$	
$B_{2m+1}(x/a)$			$(-1)^{m+1}2(2m+1)!/(2n\pi)^{2m+1}$
$\log\sin(\pi x/2a)$	$-\log 2$	$-1/n$	
$(1/2)\cot(\pi x/2a)$			$1^{(1)}$

Note

(1)The Fourier series does not converge in the sense of Cauchy, but it is summable, for example, by the Cesàro summation of the first order.

(2) $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\cos nx}{n}=\log\left(2\cos\frac{x}{2}\right) \quad (-\pi<x<\pi), \qquad \sum_{n=1}^{\infty}\frac{\sin nx}{n}=\frac{1}{2}(\pi-x) \quad (0<x<2\pi).$

$$\sum_{n=1}^{\infty}\frac{\cos(2n-1)x}{2n-1}=\frac{1}{2}\log\left|\cot\frac{x}{2}\right| \quad (0<x<2\pi,\quad x\neq\pi),$$

$$\sum_{n=1}^{\infty}\frac{\sin(2n-1)x}{2n-1}=\begin{cases}\pi/4 & (0<x<\pi),\\ -\pi/4 & (\pi<x<2\pi).\end{cases}$$

$$\sum_{n=1}^{\infty}\frac{\cos nx}{n^2}=\frac{1}{4}(x-\pi)^2-\frac{\pi^2}{12} \quad (0\leqslant x\leqslant 2\pi),$$

$$\sum_{n=1}^{\infty}\frac{\sin nx}{n^2}=-x\log 2-\int_0^x\log\left(\sin\frac{t}{2}\right)dt \quad (0\leqslant x<2\pi).$$

$$\sum_{n=1}^{\infty}\frac{a^n}{n!}\cos nx=e^{a\cos x}\cos(a\sin x)-1, \qquad \sum_{n=1}^{\infty}\frac{a^n}{n!}\sin nx=e^{a\cos x}\sin(a\sin x).$$

$$\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\cos nx}{n^2-a^2}=\frac{\pi\cos ax}{2a\sin a\pi}-\frac{1}{2a^2} \quad (-\pi\leqslant x\leqslant\pi),$$

$$\sum_{n=1}^{\infty}(-1)^{n-1}\frac{n\sin nx}{n^2-a^2}=\frac{\pi\sin ax}{2\sin a\pi} \quad (-\pi<x<\pi).$$

In the final two formulas, we assume that a is not an integer.

(II) Fourier Transforms (→ 160 Fourier Transform)

The Fourier transform $\mathscr{F}[f]$ and the inverse Fourier transform $\bar{\mathscr{F}}[g]$ for integrable functions f and g are defined as

$$\mathscr{F}[f(x)]=\mathscr{F}[f](\xi)=(2\pi)^{-n/2}\int_{\mathbf{R}^n} f(x)e^{-ix\xi}dx,$$

$$\bar{\mathscr{F}}[g(\xi)]=\bar{\mathscr{F}}[g](x)=(2\pi)^{-n/2}\int_{\mathbf{R}^n} g(\xi)e^{ix\xi}d\xi, \qquad x\xi=x_1\xi_1+x_2\xi_2+\ldots+x_n\xi_n.$$

In some textbooks the factor $(2\pi)^{-n/2}$ is deleted or the symbols i and $-i$ are switched when defining $\mathscr{F}$ and $\bar{\mathscr{F}}$. However, conversion of the formulas above to ones due to other definitions is straightforward. These transforms are also defined for some nonintegrable functions, or even more generally for †tempered distributions. The Fourier transform $\mathscr{F}$ and the inverse Fourier transform $\bar{\mathscr{F}}$ defined on the space of tempered distributions $\mathscr{S}'=\mathscr{S}'(\mathbf{R}^n)$ are linear homeomorphic mappings from $\mathscr{S}'$ to itself. Useful formulas of these transforms are given in the table below, where $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ $(\alpha_j=0,1,2,\ldots)$, $|\alpha|=\alpha_1+\alpha_2+\ldots+\alpha_n$, C is †Euler's constant, $\lambda\in\mathbf{C}$, and $\mathbf{Z}_+=\{m\in\mathbf{Z}\mid m\geqslant 0\}$.

Case 1. $n=1$. First we explain the meaning of the symbols appearing in the table:

$x_+=\max(x,0)$ (the positive part of x),

$x_-=\max(-x,0)$ (the negative part of x),

x_+^λ and x_-^λ are understood in the sense of finite parts (→ 125 Distributions and Hyperfunctions),

$(x+i\varepsilon)^\lambda=\exp[\lambda\operatorname{Log}(x+i\varepsilon)]$ ($\varepsilon\neq 0$; Log is the principal value of log)

$=(x^2+\varepsilon^2)^{\lambda/2}\exp[i\lambda\operatorname{Arg}(x+i\varepsilon)]$ ($-\pi<\operatorname{Arg}z\leqslant\pi$),

$(x\pm i0)^\lambda=\lim_{\varepsilon\downarrow 0}(x\pm i\varepsilon)^\lambda$ (limit in the sense of distributions).

Then the following formula holds:

$(x\pm i0)^\lambda=x_+^\lambda+e^{\pm i\lambda\pi}x_-^\lambda$

$\operatorname{Pf}x^m=x_+^m+(-1)^m x_-^m$ $(m\in Z)$ (Pf is the finite part).

In the special case $m=-1$, $\operatorname{Pf}x^{-1}$ coincides with Cauchy's principal value p.v.x^{-1}.

$T\ (\in\mathscr{S}')$	$\mathscr{F}[T]\ (\in\mathscr{S}')$
$\delta(x)$	$\sqrt{2\pi}$
$P(x)$ (polynomial)	$\sqrt{2\pi}\ P(id/d\xi)\delta(\xi)$
p.v. $1/x$	$\sqrt{\pi/2}\ i\operatorname{sgn}\xi$
$\operatorname{Pf}x^{-m}$	$\sqrt{\pi/2}\,[(-i)^m/(m-1)!]\xi^{m-1}\operatorname{sgn}\xi\quad (m\in\mathbf{N},\mathbf{Z})$
x_+^λ	$\dfrac{\Gamma(\lambda+1)}{\sqrt{2\pi}}[e^{-i\pi(\lambda+1)/2}\xi_+^{-\lambda-1}+e^{i\pi(\lambda+1)/2}\xi_-^{-\lambda-1}]$
	$\left[=\dfrac{\Gamma(\lambda+1)}{\sqrt{2\pi}}e^{-i\pi(\lambda+1)/2}(\xi+i0)^{-\lambda-1}\right]\quad (\lambda\notin\mathbf{Z})$
x_+^m	$(i^m/\sqrt{2\pi})[\pi\delta^{(m)}-i(-1)^m m!\operatorname{Pf}\xi^{-m-1}]\quad (m\in\mathbf{Z}_+)$
x_+^{-m}	$\dfrac{(-i)^{m-1}}{\sqrt{2\pi}(m-1)!}\left[\left(\sum_{j=1}^{m-1}\dfrac{1}{j}-C\right)\xi^{m-1}-\dfrac{i\pi}{2}\xi^{m-1}\operatorname{sgn}\xi-\xi^{m-1}\log\lvert\xi\rvert\right]\quad (m\in\mathbf{N},\mathbf{Z})$
x_-^λ	$\dfrac{\Gamma(\lambda+1)}{\sqrt{2\pi}}[e^{i\pi(\lambda+1)/2}\xi_+^{-\lambda-1}+e^{-i\pi(\lambda+1)/2}\xi_-^{-\lambda-1}]$
	$\left[=\dfrac{\Gamma(\lambda+1)}{\sqrt{2\pi}}e^{i\pi(\lambda+1)/2}(\xi-i0)^{-\lambda-1}\right]\quad (\lambda\notin\mathbf{Z})$
x_-^m	$\dfrac{(-i)^m}{\sqrt{2\pi}}[\pi\delta^{(m)}+i(-1)^m m!\operatorname{Pf}\xi^{-m-1}]\quad (m\in\mathbf{Z}_+)$
x_-^{-m}	$\dfrac{i^{m-1}}{\sqrt{2\pi}(m-1)!}\left[\left(\sum_{j=1}^{m-1}\dfrac{1}{j}-C\right)\xi^{m-1}+\dfrac{i\pi}{2}\xi^{m-1}\operatorname{sgn}\xi-\xi^{m-1}\log\lvert\xi\rvert\right]\quad (m\in\mathbf{N},\mathbf{Z})$
$(x+i0)^\lambda$	$[\sqrt{2\pi}\ e^{i\pi\lambda/2}/\Gamma(-\lambda)]\xi_+^{-\lambda-1}\quad (\lambda\notin\mathbf{Z}_+)$
$(x-i0)^\lambda$	$[\sqrt{2\pi}\ e^{-i\pi\lambda/2}/\Gamma(-\lambda)]\xi_-^{-\lambda-1}\quad (\lambda\notin\mathbf{Z}_+)$
$(x\pm i0)^m=x^m$	$\sqrt{2\pi}\ i^m\delta^{(m)}\quad (m\in\mathbf{Z}_+)$
$x^{-1}\log\lvert x\rvert$	$\sqrt{\pi/2}\ i\operatorname{sgn}\xi\cdot(C+\log\lvert\xi\rvert)$
e^{-x^2}/α	$\sqrt{a/2}\ e^{-a\xi^2/4}\quad (a>0)$
$\begin{cases}e^{-ax} & (x>0)\\ 0 & (x\leqslant 0)\end{cases}$	$\dfrac{1}{\sqrt{2\pi}}\dfrac{1}{a+i\xi}\quad (a>0)$

$T\,(\in\mathscr{S}')$	$\mathscr{F}[T]\,(\in\mathscr{S}')$
$\dfrac{e^{-a\lvert x\rvert}}{\sqrt{\lvert x\rvert}}$	$\dfrac{\sqrt{a+\sqrt{a^2+\xi^2}}}{\sqrt{a^2+\xi^2}}\quad (a>0)$
$e^{-b\sqrt{x^2+a^2}}$	$\sqrt{\dfrac{2}{\pi}}\dfrac{ab}{\sqrt{\xi^2+b^2}}K_1(a\sqrt{\xi^2+b^2})\quad (a>0, b>0)$
$\log\dfrac{x^2+a^2}{x^2+b^2}$	$-\dfrac{\sqrt{2\pi}}{\lvert\xi\rvert}(e^{-a\lvert\xi\rvert}-e^{-b\lvert\xi\rvert})\quad (a\geqslant 0, b\geqslant 0)$
$\arctan(x/a)$	$-\sqrt{\pi/2}\,i\,\mathrm{sgn}\,a\cdot(e^{-\lvert a\rvert\lvert\xi\rvert}/\xi)\quad (a\in\mathbf{R}, a\neq 0)$
$\dfrac{\sin ax}{\lvert x\rvert^{1-\nu}}$	$-\dfrac{i}{\sqrt{2\pi}}\Gamma(\nu)\cos\dfrac{\nu\pi}{2}\left(\dfrac{1}{\lvert\xi-a\rvert^\nu}-\dfrac{1}{\lvert\xi+a\rvert^\nu}\right)\quad (\nu\notin\mathbf{Z})$
$\dfrac{\cos ax}{\lvert x\rvert^{1-\nu}}$	$\dfrac{1}{\sqrt{2\pi}}\Gamma(\nu)\cos\dfrac{\nu\pi}{2}\left(\dfrac{1}{\lvert\xi-a\rvert^\nu}-\dfrac{1}{\lvert\xi+a\rvert^\nu}\right)\quad (\nu\notin\mathbf{Z})$
$\dfrac{\sin ax}{x}$	$\begin{cases}\sqrt{\pi/2} & (\lvert\xi\rvert<\lvert a\rvert)\\ 0 & (\lvert\xi\rvert>\lvert a\rvert)\end{cases}$
$\begin{cases}1/(a^2-x^2)^{\nu+(1/2)} & (\lvert x\rvert<a)\\ 0 & (\lvert x\rvert>a)\end{cases}$	$\dfrac{1}{\sqrt{2}}\Gamma((1/2)-\nu)\left(\dfrac{\lvert\xi\rvert}{2a}\right)^\nu J_{-\nu}(a\lvert\xi\rvert)\quad (\mathrm{Re}\,\nu<1/2,\ a>0)$
$\begin{cases}0 & (\lvert x\rvert<a)\\ 1/(x^2-a^2)^{\nu+(1/2)} & (\lvert x\rvert>a)\end{cases}$	$-\dfrac{1}{\sqrt{2}}\Gamma((1/2)-\nu)\left(\dfrac{\lvert\xi\rvert}{2a}\right)^\nu N_\nu(a\lvert\xi\rvert)\quad (-1/2<\mathrm{Re}\,\nu<1/2,\ a>0)$
$\dfrac{1}{(x^2+a^2)^{\nu+(1/2)}}$	$\dfrac{\sqrt{2}}{\Gamma(\nu+(1/2))}\left(\dfrac{\lvert\xi\rvert}{2a}\right)^\nu K_\nu(a\lvert\xi\rvert)\quad (\mathrm{Re}\,\nu>-1/2,\ a>0)$

Case 2. $n>1$. Let $r=\sqrt{x_1^2+x_2^2+\ldots+x_n^2}$ and $\rho=\sqrt{\xi_1^2+\xi_2^2+\ldots+\xi_n^2}$, where $x=(x_i)\in\mathbf{R}^n$ and $\xi=(\xi_i)\in\mathbf{R}^n$. If $f\in L_1(\mathbf{R}^n)$ depends only on r, then $\mathscr{F}[f]$ depends only on ρ and is expressed as

$$\mathscr{F}[f](\rho)=\rho^{-(n-2)/2}\int_0^\infty f(r)r^{n/2}J_{(n-2)/2}(r\rho)\,dr.$$

The constant C in the table stands for Euler's number.

$T\,(\in\mathscr{S}')$	$\mathscr{F}[T]\,(\in\mathscr{S}')$
$\delta(x)$	$(2\pi)^{n/2}$
$P(x)$ (polynomial)	$(2\pi)^{n/2}P(i\partial/\partial\xi)\delta(\xi)$
$\mathrm{Pf}\,r^\lambda$	$\dfrac{2^{(n/2)+\lambda}\Gamma((n+\lambda)/2)}{\Gamma(-\lambda/2)}\mathrm{Pf}\,\rho^{-n-\lambda}\quad (\lambda\notin 2\mathbf{Z}_+,\ \lambda\notin -n-2\mathbf{Z}_+)$
r^{2m}	$(2\pi)^{n/2}(-\Delta)^m\delta(\xi)\quad (m\in\mathbf{Z}_+)$
$\mathrm{Pf}\,r^{-n-2m}$	$\dfrac{(-1)^m\rho^{2m}}{2^{(n/2)+2m}\Gamma((n/2)+m)m!}\left[2\log\dfrac{2}{\rho}-C+\sum_{j=1}^m\dfrac{1}{j}+\dfrac{\Gamma'((n/2)+m)}{\Gamma((n/2)+m)}\right]\quad (m\in\mathbf{Z}_+)$
$(1+r^2)^\lambda$	$\dfrac{\rho^{-[(n/2)+\lambda]}K_{(n/2)+\lambda}(\rho)}{2^{-\lambda-1}\Gamma(-\lambda)}\quad (\lambda\notin\mathbf{Z}_+)$
$(1+r^2)^m$	$(2\pi)^{n/2}(1-\Delta)^m\delta(\xi)\quad (m\in\mathbf{Z}_+)$
$\mathrm{Pf}\,r^\lambda\log\lambda$	$\dfrac{2^{(n/2)+\lambda}\Gamma((n+\lambda)/2)}{\Gamma(-\lambda/2)}\mathrm{Pf}\,\rho^{-n-\lambda}\left[\log\dfrac{2}{\rho}+\dfrac{1}{2}\dfrac{\Gamma'((n+\lambda)/2)}{\Gamma((n+\lambda)/2)}+\dfrac{1}{2}\dfrac{\Gamma'(-\lambda/2)}{\Gamma(-\lambda/2)}\right]$ $(\lambda\notin 2\mathbf{Z}_+,\ \lambda\notin -n-2\mathbf{Z}_+)$
$r^{2m}\log r$	$(-1)^{m-1}2^{(n/2)+2m-1}m!\,\Gamma((n/2)+m)\,\mathrm{Pf}\,\rho^{-n-2m}$ $+(2\pi)^{n/2}\left[\log 2-\dfrac{1}{2}C+\dfrac{1}{2}\sum_{j=1}^m\dfrac{1}{j}+\dfrac{\Gamma'((n/2)+m)}{\Gamma((n/2)+m)}\right](-\Delta)^m\delta(\xi)\quad (m\in\mathbf{Z}_+)$
$\mathrm{Pf}\,r^{-n-2m}\log r$	$\dfrac{(-1)^m}{2^{(n/2)+2m}\Gamma((n/2)+m)m!}\rho^{2m}\left[\left\{\log\dfrac{2}{\rho}-\dfrac{1}{2}C+\dfrac{1}{2}\sum_{j=1}^m\dfrac{1}{j}+\dfrac{1}{2}\dfrac{\Gamma'((n/2)+m)}{\Gamma((n/2)+m)}\right\}^2\right.$ $\left.+\dfrac{\pi^2}{24}+\dfrac{1}{4}\sum_{j=1}^m\dfrac{1}{j^2}-\dfrac{1}{4}\dfrac{\Gamma''((n/2)+m)}{\Gamma((n/2)+m)}+\dfrac{\Gamma'((n/2)+m)^2}{\Gamma((n/2)+m)^2}\right]\quad (m\in\mathbf{Z}_+)$
e^{-ar}	$\dfrac{\sqrt{2^n}}{\sqrt{\pi}}\Gamma\left(\dfrac{n+1}{2}\right)\dfrac{a}{(a^2+\rho^2)^{(n+1)/2}}\quad (a>0)$

The Fourier transform mentioned above is a transformation in the family of complex-valued functions or distributions. Similar transformations in the family of real-valued functions are frequently used in applications:

Fourier cosine transform $f_c(u)=\int_0^\infty F(t)\cos ut\,dt.$

Inverse transform $\frac{2}{\pi}\int_0^\infty f_c(u)\cos ut\,du=\begin{cases} F(t) & (t>0),\\ F(-t) & (t<0).\end{cases}$

Fourier sine transform $f_s(u)=\int_0^\infty F(t)\sin ut\,dt.$

Inverse transform $\frac{2}{\pi}\int_0^\infty f_s(u)\sin ut\,du=\begin{cases} F(t) & (t>0),\\ -F(-t) & (t<0).\end{cases}$

The Fourier transform can be expressed in terms of these transforms. For example (in $\mathbf{R}^1$),

$$\mathcal{F}[f](u)=\frac{1}{\sqrt{2\pi}}\int_0^\infty [f(t)+f(-t)]\cos ut\,dt-\frac{i}{\sqrt{2\pi}}\int_0^\infty [f(t)-f(-t)]\sin ut\,dt.$$

$F(t)$	$f_c(u)$	$f_s(u)$
$\begin{cases}1 & (0<t<a)\\ 0 & (a<t)\end{cases}$	$\frac{\sin au}{u}$	$\frac{1-\cos au}{u}$
t^{-1}	(diverges)	$(\pi/2)\operatorname{sgn} u$
$t^{\alpha-1}$ $(0<\alpha<1)$	$\Gamma(\alpha)\cos(\pi\alpha/2)u^{-\alpha}$	$\Gamma(\alpha)\sin(\pi\alpha/2)u^{-\alpha}$
$1/(a^2+t^2)$	$\pi e^{-a\lvert u\rvert}/2a$	$[e^{-au}\mathrm{Ei}(au)-e^{au}\mathrm{Ei}(-au)]/a^{(2)}$
e^{-at}	$a/(a^2+u^2)$	$u/(a^2+u^2)$
$e^{-\lambda t^2}$ $(\mathrm{Re}\,\lambda>0)$	$\sqrt{\pi/4\lambda}\,e^{-u^2/4\lambda}$	$e^{-u^2/4\lambda}\varphi(u/2\sqrt{\lambda})/\sqrt{\lambda}^{\,(3)}$
$e^{-\lambda t^2}t$		$\sqrt{\pi/4\lambda}\,(u/4\lambda)e^{-u^2/4\lambda}$
$\frac{\sin at}{t}$ $(a>0)$	$\begin{cases}\pi/2 & (0<u<a)\\ 0 & (a<u)\end{cases}$	$\frac{1}{2}\log\left\lvert\frac{a+u}{a-u}\right\rvert$
$\tanh(\pi t/2)$		$\operatorname{cosech} u$
$\operatorname{sech}(\pi t/2)$	$\operatorname{sech} u$	
$J_\nu(t)$ $(\mathrm{Re}\,\nu>-1)$	$\begin{cases}\frac{\cos(\nu\arcsin u)}{\sqrt{1-u^2}}\\ -\frac{(u-\sqrt{u^2-1})^\nu}{\sqrt{u^2-1}}\sin\frac{\nu\pi}{2}\end{cases}$	$\frac{\sin(\nu\arcsin u)}{\sqrt{1-u^2}}$ $(0<u<1)$ $\frac{(u-\sqrt{u^2-1})^\nu}{\sqrt{u^2-1}}\cos\frac{\nu\pi}{2}$ $(1<u)$
$J_0(at)$	$\begin{cases}1/\sqrt{a^2-u^2}\\ 0\end{cases}$	0 $(0\leqslant u<a)$ $1/\sqrt{u^2-a^2}$ $(a<u)$
$N_0(t)$	$\begin{cases}0\\ -\frac{1}{\sqrt{u^2-1}}\end{cases}$	$\frac{2}{\pi}\frac{\arcsin u}{\sqrt{1-u^2}}$ $(0<u<1)$ $\frac{2}{\pi}\frac{\log(u-\sqrt{u^2-1})}{\sqrt{u^2-1}}$ $(1<u)$
$K_0(t)$	$\pi/2\sqrt{1+u^2}$	$(\operatorname{arcsinh} u)/\sqrt{1+u^2}$

Notes

(2) Ei is the exponential integral function (→ Table 19.II.3, this Appendix).

(3) We put $\varphi(x)=\int_0^x e^{t^2}dt$.

12. Laplace Transforms and Operational Calculus

(I) Laplace Transforms (→ 240 Laplace Transform)

Laplace transform $V(p)=\int_0^\infty e^{-pt}F(t)\,dt$ $(\mathrm{Re}\,p>0)$.

Inverse transform (Bromwich integral) $\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{pt}V(p)\,dp=\begin{cases}F(t) & (t>0),\\ 0 & (t<0).\end{cases}$

$F(t)$	$V(p)$	$F(t)$	$V(p)$
1	$1/p$	$J_\nu(t)\quad(\mathrm{Re}\,\nu>-1)$	$\dfrac{\left(\sqrt{1+p^2}-p\right)^\nu}{\sqrt{1+p^2}}$
$1(t-a)=\begin{cases}0 & (0\leqslant t<a)\\ 1 & (a\leqslant t)\end{cases}$	e^{-ap}/p	$\dfrac{1}{t}J_\nu(at)\quad(\mathrm{Re}\,\nu>0)$	$\dfrac{\left(\sqrt{a^2+p^2}-p\right)^\nu}{\nu a^\nu}$
$[x/a]$ (integral part)	$1/p(e^{ap}-1)$		
$t^{\alpha-1}\quad(\mathrm{Re}\,\alpha>0)$	$\Gamma(\alpha)/p^\alpha$	$t^\nu J_\nu(at)\left(\mathrm{Re}\,\nu>-\dfrac{1}{2}\right)$	$\dfrac{(2a)^\nu\Gamma[\nu+(1/2)]}{\sqrt{\pi}\,(p^2+a^2)^{\nu+(1/2)}}$
e^{-at}	$1/(p+a)$		
$e^{-at}t^{\alpha-1}\quad(\mathrm{Re}\,\alpha>0,\ a>0)$	$(p+a)^{-\alpha}\Gamma(\alpha)$	$t^{\nu/2}J_\nu(x\sqrt{t})$ $(\mathrm{Re}\,\nu>-1)$	$\dfrac{x^\nu}{2^\nu p^{\nu+1}}e^{-x^2/4p}$
$e^{-at}F(t)\ (a>0)$	$V(p+a)$		
$(1-e^{-t})/t$	$\log(1+p^{-1})$	$J_0(t)$	$(1+p^2)^{-1/2}$
$(\pi t)^{-1/2}e^{-x^2/4t}$	$p^{-1/2}e^{-x\sqrt{p}}\ (x>0)$	$J_0(x\sqrt{t})$	$e^{-x^2/4p}/p$
$\log t$	$-(\log p+C)/p^{(1)}$	$N_0(t)$	$\dfrac{2}{\pi}\dfrac{\log\left(\sqrt{1+p^2}-p\right)}{\sqrt{1+p^2}}$
$\sin at$	$a/(p^2+a^2)$	$L_n(t)^{(2)}$	$\dfrac{1}{p}\left(\dfrac{p-1}{p}\right)^n$
$\cos at$	$p/(p^2+a^2)$		
$\sin(x\sqrt{t})$	$\dfrac{\sqrt{\pi}}{2}\dfrac{x}{p^{3/2}}e^{-x^2/4p}$	$t^\alpha L_n^{(\alpha)}(t)$	$\dfrac{\Gamma(\alpha+n+1)}{n!}\dfrac{1}{p^{\alpha+1}}\left(\dfrac{p-1}{p}\right)^n$
$t^{-1/2}\cos(x\sqrt{t})$	$\sqrt{\pi/p}\,e^{-x^2/4p}$	$H_{2n+1}(\sqrt{t})^{(3)}$	$\sqrt{\dfrac{\pi}{2}}\,(2n+1)!!\,\dfrac{(1-p)^n}{p^{n+(3/2)}}\ ^{(4)}$
$t^{-1}\sin xt$	$\arctan(x/p)$		
$x^{-1}(1-\cos ax)$	$\frac{1}{2}\log[1+(a^2/p^2)]$		
$\sinh at$	$a/(p^2-a^2)$	$\dfrac{H_{2n}(\sqrt{t})}{\sqrt{t}}$	$\sqrt{\pi}\,(2n-1)!!\,\dfrac{(1-p)^n}{p^{n+(1/2)}}$
$\cosh at$	$p/(p^2-a^2)$		

Notes
(1) C is Euler's constant.
(2) $L_n(t)$ is a Laguerre polynomial.
(3) $H_n(t)$ is a Hermite polynomial.
(4) $(2n+1)!!=(2n+1)(2n-1)(2n-3)\ldots5\cdot3\cdot1$.

(II) Operational Calculus (→ 306 Operational Calculus)

Heaviside function (unit function) $\mathbf{1}(t)=\begin{cases}0 & (t<0)\\ 1 & (t\geqslant 0).\end{cases}$

Dirac delta function (unit impulse function) $\delta(t)=\lim_{\varepsilon\to0}\dfrac{1}{2\varepsilon}[\mathbf{1}(t+\varepsilon)-\mathbf{1}(t-\varepsilon)]$.

When an operator $\Omega(p)$ operates on $\mathbf{1}(t)$ and the result is $A(t)$ we write $\Omega(p)\mathbf{1}(t)=A(t)$.
In the following table (i) of general formulas, we assume the relations $\Omega_i(p)\mathbf{1}(t)=A_i(t)$ $(i=1,2)$.

Carson's integral $\Omega(p)=p\displaystyle\int_0^\infty e^{-pt}A(t)\,dt\quad(\mathrm{Re}\,p>0)$.

Laplace transform $V(p)=\dfrac{\Omega(p)}{p}=\displaystyle\int_0^\infty e^{-pt}A(t)\,dt$.

(i) General Formulas		(ii) Examples	
$\Omega(p)$	$A(t)$	$\Omega(p)=pV(p)$	$A(t)$
$\Omega_1(p)+\Omega_2(p)$	$A_1(t)+A_2(t)$	p	$\delta(t)$
$a\Omega_1(p)$	$aA_1(t)$	$1/p^n$ $(n=0,1,2,\ldots)$	$(t^n/n!)\mathbf{1}(t)$
$p\Omega_1(p)$	$A_1(0)\delta(t)+A_1'(t)$	$p/(p+a)$	$(e^{-at})\mathbf{1}(t)$
		$p^2/(p^2+a^2)$	$(\cos at)\mathbf{1}(t)$
$\frac{1}{p}\Omega_1(p)$	$\int_0^t A_1(\tau)d\tau$	$ap/(p^2+a^2)$	$(\sin at)\mathbf{1}(t)$
$\Omega_1(ap)$	$A_1(t/a)$		
$[p/(p+a)]\Omega_1(p+a)$	$e^{-at}A_1(t)$ $(\mathrm{Re}\,a\geqslant 0)$	$a_0+\frac{a_1}{p}+\frac{a_2}{p^2}+\ldots$	$\left(a_0+a_1\frac{t}{1!}+a_2\frac{t^2}{2!}+\ldots\right)\mathbf{1}(t)$
$\frac{1}{p}\Omega_1(p)\Omega_2(p)$	$\int_0^t A_1(\tau)A_2(t-\tau)d\tau$ $=\int_0^t A_1(t-\tau)A_2(\tau)d\tau$	$\sum_{k=1}^{n}\frac{B_k}{p-p_k}$	$\sum_{k=1}^{n}\frac{B_k}{p_k}(e^{p_kt}-1)\mathbf{1}(t)$ $=\Omega(0)\mathbf{1}(t)+\sum_{k=1}^{n}\frac{B_k}{p_k}e^{p_kt}$

13. Conformal Mappings

(→ 77 Conformal Mappings)

Original Domain	Image Domain	Mapping Function
$\lvert z\rvert<1$ (unit disk)	$\lvert w\rvert<1$	$w=\varepsilon\frac{z-z_0}{1-\bar{z}_0z}$, $\lvert z_0\rvert<1$, $\lvert\varepsilon\rvert=1$ (general form)
$\mathrm{Im}\,z>0$ (upper half-plane)	$\lvert w\rvert<1$	$w=\varepsilon\frac{z-z_0}{z-\bar{z}_0}$, $\mathrm{Im}\,z_0>0$, $\lvert\varepsilon\rvert=1$ (general form)
$\mathrm{Im}\,z>0$ (upper half-plane)	$\mathrm{Im}\,w>0$	$w=\frac{az+b}{cz+d}$, a,b,c,d are real; $ad-bc>0$ (general form)
$0<\arg z<\alpha$ (angular domain)	$\mathrm{Im}\,w>0$	$w=z^{\pi/\alpha}$
$\lvert z\rvert<1, \mathrm{Im}\,z>0$ (upper semidisk)	$\mathrm{Im}\,w>0$	$w=\left(\frac{1+z}{1-z}\right)^2$
$0<\arg z<\alpha$, $\lvert z\rvert<1$ (fan shape)	$\mathrm{Im}\,w>0$	$w=\left(\frac{1+z^{\pi/\alpha}}{1-z^{\pi/\alpha}}\right)^2$
$\alpha<\arg\frac{z-p}{z-q}<\beta$ (circular triangle)	$0<\arg w<\gamma$	$w=\left(e^{-i\alpha}\frac{z-p}{z-q}\right)^{\frac{\gamma}{\beta-\alpha}}$
$0<\mathrm{Im}\,z<\eta$ (parallel strip)	$\mathrm{Im}\,w>0$	$w=e^{\pi z/\eta}$
$\mathrm{Re}\,z<0$, $0<\mathrm{Im}\,z<\eta$ (semiparallel strip)	$\mathrm{Im}\,w>0$, $\lvert w\rvert<1$	$w=e^{\pi z/\eta}$
$y^2>4c^2(x+c^2)$, $z=x+iy$, $c>0$ (exterior of a parabola)	$\mathrm{Im}\,w>0$	$w=\sqrt{z}-ic$
$y^2<4c^2(x+c^2)$, $z=x+iy$, $c>0$ (interior of a parabola)	$\mathrm{Im}\,w>0$	$w=i\sec\frac{\pi\sqrt{z}}{2ic}$

Original Domain	Image Domain	Mapping Function
$\frac{x^2}{[c+(1/c)]^2} + \frac{y^2}{[c-(1/c)]^2} > 1$, $z=x+iy$, $c>1$ (exterior of an ellipse)	$\lvert w\rvert > c$	$w=\frac{z+\sqrt{z^2-4}}{2}$, $z=w+\frac{1}{w}$
$\frac{x^2}{\cos^2\alpha} - \frac{y^2}{\sin^2\alpha} < 4$, $z=x+iy$, $0<\alpha<\pi/2$ (exterior of a hyperbola)	$\operatorname{Im} w > 0$	$w=\left(e^{-i\alpha}\frac{z+\sqrt{z^2-4}}{2}\right)^{\pi(\pi-2\alpha)}$
$\frac{x^2}{\cos^2\alpha} - \frac{y^2}{\sin^2\alpha} > 4$, $x>0$, $z=x+iy$, $0<\alpha<\pi/2$ (right-hand side interior of a hyperbola)	$\operatorname{Im} w > 0$	$w=\frac{1}{2i}\left[\left(\frac{z+\sqrt{z^2-4}}{2}\right)^{\frac{\pi}{2\alpha}} + \left(\frac{z-\sqrt{z^2-4}}{2}\right)^{\frac{\pi}{2\alpha}}\right]$
$\lvert z\rvert<1$	Slit domain with boundary $\lvert \operatorname{Re} w\rvert \leqslant 2$, $\operatorname{Im} w=0$	$w=z+\frac{1}{z}$
$\lvert z\rvert<1$	Slit domain with boundary $\lvert w\rvert \geqslant 1/4$, $\arg w=\lambda$	$w=\frac{z}{(1+e^{-i\lambda}z)^2}$
$\lvert z\rvert<1$	Slit domain with boundary $\lvert w\rvert \geqslant 1/4^{1/p}$ $\arg w=\lambda+(2j\pi/p)$, $j=0,\dots,p-1$	$w=\frac{z}{(1+e^{-ip\lambda}z^p)^{2/p}}$
$-\pi/2<\operatorname{Re} z<\pi/2$ (parallel strip)	Slit domain with boundary $\lvert \operatorname{Re} w\rvert \geqslant 1$, $\operatorname{Im} w=0$	$w=\sin z$
$-\pi<\operatorname{Im} z<\pi$ (parallel strip)	Slit domain with boundary $\operatorname{Re} w \leqslant -1$, $\operatorname{Im} w=\pm\pi$	$w=z+e^z$
Arbitrary circle or half plane	Interior of an n-gon	$w=c\int^z \prod_{j=1}^{n} (t-z_j)^{\alpha_j-1}dt+c'$ ($c\neq 0$, c' are constants), where the inverse image of the vertex with the inner angle $\alpha_j\pi$ ($j=1,\dots,n$) is $z=z_j$. When $z_n=\infty$, we omit the factor $(t-z_n)^{\alpha_n-1}$ (Schwarz-Christoffel transformation)
Arbitrary circle or half-plane	Exterior of an n-gon	$w=c\int^z (t-p)^{-2}\prod_{j=1}^{n} (t-z_j)^{1-\alpha_j}dt+c'$ ($c\neq 0$, c' are constants), where the inverse image of the vertex with the inner angle $\alpha_j\pi$ ($j=1,\dots,n$) is $z=z_j$, and the inverse image of ∞ is $z=p$
$\operatorname{Im} z>0$	Interior of an equilateral triangle	$w=\int_0^z \frac{1}{\sqrt[3]{t^2(1-t)^2}}dt$
$\operatorname{Im} z>0$	Interior of an isosceles right triangle	$w=\int_0^z \frac{1}{\sqrt[5]{t^2(1-t)^3}}dt$

Original Domain	Image Domain	Mapping Function
$\operatorname{Im} z > 0$	Interior of a right triangle with one angle $\pi/6$	$w = \int_0^z \frac{1}{\sqrt[6]{t^3(1-t)^4}}\, dt$
$\|z\| < 1$	Interior of a regular n-gon	$w = \int_0^z (1-t^n)^{-2/n}\, dt$
$0 < \operatorname{Re} z < \omega_1$, $0 < \operatorname{Im} z < \omega_3/i$ (rectangle)	$\operatorname{Im} w > 0$	$w = \wp(z \| 2\omega_1, 2\omega_3)$ ($\wp$ is the Weierstrass $\wp$-function)
$-K < \operatorname{Re} z < K$, $0 < \operatorname{Im} z < K'$ (rectangle)(1)	$\operatorname{Im} w > 0$	$w = \operatorname{sn}(z,k)$, $z = \int_0^w \frac{1}{\sqrt{(1-t^2)(1-k^2t^2)}}\, dt$ (sn is Jacobi's sn function)
$v < \|z\| < 1$ $\operatorname{Im} z < 0$ (upper half-ring domain)	$\log q < \operatorname{Re} w < 0$, $0 < \operatorname{Im} w < \pi$ (rectangle)	$w = \log z$
$\|z\| < 1$	$\frac{u^2}{A^2} + \frac{v^2}{B^2} < 1$, $w = u + v$, $A > B > 0$; (interior of an ellipse)	$w = \sqrt{A^2 - B^2}\, \sin\left(\frac{\pi}{2K} \int_0^{2z/k(1+z^2)} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} \right)$, $\frac{1}{\pi} \log \frac{A+B}{A-B} = \frac{K'}{K}$ (1)
$\operatorname{Im} z > 0$	interior of a circular polygon	$\{w; z\} \equiv \frac{w'''}{w'} - \frac{3}{2}\left(\frac{w''}{w'}\right)^2 = R(z)$ ($R(z)$ is a rational function)
$\|z\| < 1$	Interior of an equilateral circular triangle with inner angle π/k, $1 < k \leqslant \infty$	$w = \dfrac{z \int_0^1 t^{-\frac{1}{2} - \frac{1}{2k}} (1-t)^{-\frac{1}{6} + \frac{1}{2k}} (1 - z^3 t)^{-\frac{5}{6} + \frac{1}{2k}}\, dt}{\int_0^1 t^{-\frac{1}{2} - \frac{1}{2k}} (1-t)^{-\frac{5}{6} + \frac{1}{2k}} (1 - z^3 t)^{-\frac{1}{6} + \frac{1}{2k}}\, dt}$ The vertices are the images of $z = 1$, $e^{2\pi i/3}$, and $e^{4\pi i/3}$; and $\left[\frac{dw}{dz}\right]_{z=0} = \frac{\Gamma[(5/6) + (1/2k)]\Gamma(2/3)}{\Gamma[(1/6) + (1/2k)]\Gamma(4/3)}$
$\operatorname{Im} z > 0$	Interior of a circular triangle with inner angles $\pi\alpha$, $\pi\beta$, $\pi\gamma$, $\alpha + \beta + \gamma < 1$(2)	$w = \dfrac{\int_0^1 t^{-\frac{1+\alpha+\beta+\gamma}{2}} (1-t)^{-\frac{1+\alpha-\beta-\gamma}{2}} (1-zt)^{-\frac{1-\alpha+\beta-\gamma}{2}}\, dt}{\int_0^1 t^{-\frac{1+\alpha+\beta+\gamma}{2}} (1-t)^{-\frac{1-\alpha-\beta+\gamma}{2}} (1-t+zt)^{-\frac{1-\alpha+\beta-\gamma}{2}}\, dt}$
$\|\tau\| > 1$, $-1/2 < \operatorname{Re} \tau < 0$	$\operatorname{Im} J > 0$	$J = J(\tau)$, $\tau = \omega_3/\omega_1$, $J = g_2^3/(g_2^3 - 27 g_3^2)$ (the absolute invariant of the elliptic modular function); $J(e^{2\pi i/3}) = 0$, $J(i) = 1$, $J(\infty) = \infty$
$\|\tau + 1/2\| < 1/2$, $-1 < \operatorname{Re} \tau < 0$	$\operatorname{Im} \lambda < 0$	$\lambda = \lambda(\tau)$, $\tau = \omega_3/\omega_1$, $\lambda = (e_2 - e_3)/(e_1 - e_3)$; $J(\tau) \equiv \frac{4}{27} \frac{[\lambda(\tau)^2 - \lambda(\tau) + 1]^2}{\lambda(\tau)^2[\lambda(\tau) - 1]^2}$, $\lambda(-1) = \infty$, $\lambda(0) = 1$, $\lambda(\infty) = 0$

Notes

(1) K, K', k' are the usual notations in the in the theory of elliptic integrals:

$$K \equiv \int_0^1 \frac{1}{\sqrt{(1-t^2)(1-k^2t^2)}}\, dt, \qquad K' = \int_0^1 \frac{1}{\sqrt{(1-t^2)(1-k'^2t^2)}}\, dt, \qquad k^2 + k'^2 = 1.$$

(2) When $\alpha + \beta + \gamma = 1$, the circular triangle is mapped into the ordinary linear triangle by a suitable linear transformation, and we can apply the Schwarz-Christoffel transformation. When $\alpha + \beta + \gamma > 1$, we have a similar mapping function replacing the integral representations of hypergeometric functions in the formula by the corresponding integral representations of the hypergeometric functions converging at α, β, and γ.

14. Ordinary Differential Equations

(I) Solution by Quadrature

$a,b,c,\ldots$ are integral constants.

(1) Solution of the First-Order Differential Equations (→ 313 Ordinary Differential Equations).
(i) Separated type $dy/dx = X(x)Y(y)$. The general solution is

$$\int^{y}\frac{dy}{Y(y)}=\int^{x}X(x)\,dx+c.$$

(ii) Homogeneous ordinary differential equation $dy/dx=f(y/x)$. Putting $y=ux$, we have $du/dx=[f(u)-u]x$, and the equation reduces to type (i). The general solution is

$$x=c\exp\left[\int^{u}\frac{du}{f(u)-u}\right]\quad\left(u=\frac{y}{x}\right).$$

(iii) Linear ordinary differential equation of the first order. $dy/dx+p(x)y+q(x)=0$. The general solution is

$$y=\left[c-\int q(x)P(x)\,dx\right]\Big/P(x),$$

where

$$P(x)\equiv\exp\left[\int p(x)\,dx\right].$$

(iv) Bernoulli's differential equation $dy/dx+p(x)y+q(x)y^{\alpha}=0$ $(\alpha\neq 0,1)$. Putting $z=y^{1-\alpha}$, the equation is transformed into

$$dz/dx+(1-\alpha)p(x)z+(1-\alpha)q(x)=0,$$

which reduces to (iii).
(v) Riccati's differential equation $dy/dx+ay^2=bx^m$. If $m=-2,\ 4k/(1-2k)$ (k an integer), this is solved by quadrature. In general, it is reduced to Bessel's differential equation by $ay=u'/u$.
(vi) Generalized Riccati differential equation $dy/dx+p(x)y^2+q(x)y+r(x)=0$. If we know one, two, or three special solutions $y=y_i(x)$, the general solution is represented as follows. When $y_1(x)$ is one known special solution,

$$y=y_1(x)+P(x)\Big/\left[\int p(x)P(x)\,dx+c\right],$$

where

$$P(x)\equiv\exp\left[-\int\{q(x)+2p(x)y_1(x)\}\,dx\right].$$

When $y_1(x)$, $y_2(x)$ are the known solutions,

$$\frac{y-y_1(x)}{y-y_2(x)}=c\exp\left[\int p(x)\{y_2(x)-y_1(x)\}\,dx\right].$$

When $y_1(x)$, $y_2(x)$, $y_3(x)$ are known solutions,

$$\frac{y-y_1(x)}{y-y_2(x)}=c\frac{y_3(x)-y_1(x)}{y_3(x)-y_2(x)}.$$

(vii) Exact differential equation $P(x,y)dx+Q(x,y)dy=0$. If the left-hand side is an exact differential form, the condition is $\partial P/\partial y=\partial Q/\partial x$. The general solution is

$$\int P\,dx+\int\left(Q-\frac{\partial}{\partial y}\int P\,dx\right)dy=c.$$

(viii) Integrating factors. A function $M(x,y)$ is called an integrating factor of a differential equation $P(x,y)dx+Q(x,y)dy=0$, if $M(x,y)[P(x,y)dx+Q(x,y)dy]$ is an exact differential form $d\varphi(x,y)$. If we know an integrating factor, the general solution is given by $\varphi(x,y)=c$. If we know two independent integrating factors M and N, the general solution is given by $M/N=c$.
(ix) Clairaut's differential equation $y=xp+f(p)$ $(p=dy/dx)$. The general solution is the family of straight lines $y=cx+f(c)$, and the singular solution is the envelope of this family, which is given by eliminating p from the original equation and $x+f'(p)=0$.
(x) Lagrange's differential equation $y=x\varphi(p)+\psi(p)$ $(p\equiv dy/dx)$. Differentiation with respect to x reduces the equation to a linear differential equation $[\varphi(p)-p](dx/dp)+\varphi'(p)x+\psi'(p)=0$

with respect to x, p (see (iii)). The general solution of the original equation is given by eliminating p from the original equation and the solution of the latter linear equation. The parameter p may represent the solution. If the equation $p=\varphi(p)$ has a solution $p=p_0$, we have a solution $y=p_0x+\psi(p_0)$ (straight line). This solution is sometimes the singular solution.

(xi) Singular solutions. The singular solution of $f(x,y,p)=0$ is included in the equation resulting from eliminating p from $f=0$ and $\partial f/\partial p=0$, though the eliminant may contain various curves that are not the singular solutions.

(xii) System of differential equations.

eq. (1) $\quad dx:dy:dz=P:Q:R.$

A function $M(x,y,z)$ is called a Jacobi's last multiplier for eq. (1) if M is a solution of a partial differential equation $(\partial MP/\partial x)+(\partial MQ/\partial y)+(\partial MR/\partial z)=0$. If we know two independent last multipliers M and N, then $M/N=c$ is a solution of eq. (1). If we know a last multiplier M and a solution $f=a$ of eq. (1), we may find another solution of (1) as follows: solving $f=a$ with respect to z and inserting the solution into eq. (1), we see that $M(Qdx-Pdy)/f_z$ is an exact differential form $dG(x,y,a)$ in three variables x, y, and a. Then $G(x,y,f(x,y,z))=b$ is another solution of eq. (1).

(2) Solutions of Higher-Order Ordinary Differential Equations. The following (i)–(iv) are several examples of depression.

(i) $f(x,y^{(k)},y^{(k+1)},\ldots,y^{(n)})=0$ $(0<k\leqslant n)$. Set $y^{(k)}=z$; the equation reduces to one of the $(n-k)$th order in z.

(ii) $f(y,y',y'',\ldots,y^{(n)})=0$. This is reduced to $(n-1)$st order if we consider $y'=p$ as a variable dependent on y.

(iii) $y''=f(y)$. The general solution is given by

$$x=a\pm\int\left[2\int f(y)\,dy+b\right]^{-1/2}dy.$$

We have a similar formula for $y^{(n)}=f(y^{(n-2)})$.

(iv) Homogeneous ordinary differential equation of higher order. If the left-hand side of $F(x,y,y',\ldots,y^{(n)})=0$ satisfies the homogeneity relation $F(x,\rho y,\rho y',\ldots,\rho y^{(n)})=\rho^{\alpha}F(x,y,y',\ldots,y^{(n)})$, the equation is reduced to one of the $(n-1)$st order in u by $u=y'/y$.

If F satisfies $F(\rho x,\rho^l y,\rho^{l-1}y',\ldots,\rho^{l-n}y^{(n)})=\rho^{\alpha}F(x,y,y',\ldots,y^{(n)})$, then $u=y/x^l$, $t=\log x$ reduces the equation to one of type (ii) not containing t.

(v) Euler's linear ordinary differential equation.

$$p_n(x)x^n y^{(n)}+p_{n-1}(x)x^{n-1}y^{(n-1)}+\ldots+p_1(x)xy'+p_0(x)y=q(x)$$

is reduced to a linear equation by $t=\log x$.

(vi) Linear ordinary differential equations of higher order (exact equations). A necessary and sufficient condition that $L[y]\equiv\sum_{j=0}^{n}p_j(x)y^{(j)}=X(x)$ is an exact differential form is $\sum_{j=0}^{n}(-1)^j p_j^{(j)}=0$, and then the first integral of the equation is given by

$$\sum_{j=0}^{n-1}\sum_{k=0}^{n-j-1}(-1)^k p_{k+j+1}^{(k)}y^{(j)}=\int X(x)\,dx+c.$$

(vii) Linear ordinary differential equation of higher order (depression).

$$L[y]\equiv\sum_{j=0}^{n}p_j(x)y^{(j)}=X(x).$$

If we know mutually independent special solutions $y_1(x),\ldots,y_m(x)$ for the homogeneous linear ordinary differential equation $L[y]=0$, the equation is reduced to the $(n-m)$th linear ordinary differential equation with respect to z by a transformation $z=A(y)$, where $A(y)=0$ is the mth linear ordinary differential equation with solutions $y_1(x),\ldots,y_m(x)$. For example, if $m=1$, the equation is reduced to the $(n-1)$st linear ordinary differential equation with respect to z by the transformation

$$y(x)=y_1(x)\int z(x)\,dx.$$

Also, if $n=m=2$, the general solution is

$$y=c_1y_1+c_2y_2-y_1\int Ty_2\,dx+y_2\int Ty_1\,dx,$$

where $T(x)\equiv X(x)/[y_1(x)y_2'(x)-y_2(x)y_1'(x)]$. The denominator of the last expression is the Wronskian of y_1 and y_2.

(viii) Regular singularity. For a linear ordinary differential equation of higher order,

eq. (1) $\quad x^n y^{(n)} + x^{n-1} p_1(x) y^{(n-1)} + \ldots + p_n(x) y = 0,$

the point $x=0$ is its regular singularity if $p_1(x), \ldots, p_n(x)$ are analytic at $x=0$.

We put $p_0 = 1$ and

$$\sum_{\nu=0}^{\infty} f_\nu(\rho) x^\nu \equiv \sum_{j=0}^{n} p_{n-j}(x) \rho(\rho-1)\ldots(\rho-j+1).$$

If ρ is a root of the characteristic equation $f_0(\rho)=0$ and $\rho+1, \rho+2, \ldots$ are not roots, we can determine the coefficients c_ν uniquely from

eq. (2) $\quad \sum_{\nu=0}^{m} c_\nu f_{m-\nu}(\rho+\nu) = 0 \qquad (m=1,2,\ldots),$

starting from a fixed value $c_0\ (\neq 0)$, and the series $y = x^\rho \sum_{\nu=0}^{\infty} c_\nu x^\nu$ converges and represents a solution of eq. (1). If the differences of all pairs of roots of the determining equation are not integers, we have n linearly independent solutions of eq. (1) applying the process for each characteristic root.

If there are roots whose differences are integers (including multiple roots), we denote such a system of roots by $\rho_1, \ldots, \rho_l$. We arrange them in increasing order, and denote the multiplicities of the roots by $e_1, \ldots, e_l$, respectively. Put $q_k = \rho_k - \rho_1$ $(k=1,2,\ldots,l; 0=q_1<q_2<\ldots<q_l)$. Take $N \geqslant q_l$ and a constant $c\ (\neq 0)$. Let λ be a parameter, and starting from $c_0 = c_0(\lambda) \equiv c\prod_{k=1}^{N} f_0(\lambda+k)$, we determine $c_\nu = c_\nu(\lambda)$ uniquely by the relation (2). Putting

$m_k \equiv e_k + e_{k+1} + \ldots + e_l \quad (k=1,\ldots,l) \qquad$ for h in $\quad m_{k+1} \leqslant h \leqslant m_k - 1,$

the series

eq. (3) $\quad y = \left[\dfrac{\partial^h}{\partial \lambda^h} x^\lambda \sum_{\nu=0}^{\infty} c_\nu(\lambda) x^\nu \right]_{\lambda=\rho_k} = x^{\rho_k} \sum_{\nu=0}^{\infty} x^\nu \left[\sum_{j=0}^{h} \binom{h}{j} c_\nu^{(j)}(\rho_k)(\log x)^{h-j} \right]$

converges and gives e_k independent solutions of eq. (1). Hence for $k=1,\ldots,l$, we may have $\sum_{k=1}^{l} e_k = m_1$ mutually independent solutions of (1). Applying this process to every characteristic root, we have finally n independent solutions of (1) (Frobenius method).

In the practical computation of the solution, since it is known to have the expression (3), we often determine its coefficients successively by the method of undetermined coefficients.

(3) Solution of Linear Ordinary Differential Equations with Constant Coefficient (→ 252 Linear Ordinary Differential Equations). Let α_i, α_{jk} be constants. We consider the following linear ordinary differential equation of higher order (eq. (1)) and system of linear ordinary differential equations (eq. (2)).

eq. (1) $\quad \sum_{i=0}^{n} \alpha_i y^{(i)} = X(x).$

eq. (2) $\quad y_j' = \sum_{k=1}^{n} \alpha_{jk} y_k + X_j(x) \quad (j=1,\ldots,n).$

(i) The general solution of the homogeneous equation (cofactor) is given by the following formulas:

for eq. (1) $\quad y = x^j \exp \lambda_k x \quad (j=0,1,,\ldots,e_k-1; k=1,\ldots,m),$

for eq. (2) $\quad y_j(x) = \sum_{k=1}^{m} p_{jk}(x) \exp \lambda_k x \quad (j=1,\ldots,n),$

where $\lambda_1, \ldots, \lambda_m$ are the roots of the characteristic equation of eq. (1) or eq. (2) given by

eq. (1′) $\quad \sum_{i=0}^{n} \alpha_i \lambda^i = 0,$

eq. (2′) $\quad \det(\alpha_{jk} - \lambda \delta_{jk}) = 0,$

respectively. We denote the multiplicities of the roots by $e_1, \ldots, e_m$ $(e_1 + \ldots + e_m = n)$; $p_{jk}(x)$ is a polynomial of degree at most $e_k - 1$ containing e_k arbitrary constants.

If all the coefficients in the original equation are real, and the root $\lambda_k = \mu_k + i\nu_k$ is imaginary, then $\bar{\lambda}_k = \mu_k - i\nu_k$ is also a root with the same multiplicity. Then we may replace $\exp \lambda_k x$ and $\exp \bar{\lambda}_k x$ by $\exp \mu_k x \cos \nu_k x$ and $\exp \mu_k x \sin \nu_k x$, and in this way we can represent the solution using real functions.

(ii) Inhomogenous equation. The solution of an inhomogeneous linear ordinary differential equation is given by the method of variation of parameters or by the method described in Section (2)(vii).

We explain the method of variation of parameters for eq. (2). First we use (i) to find a fundamental system of n independent solutions $y_j = \varphi_{jk}(x)$ $(k=1,\ldots,n)$ by (i)). Inserting $y_j = \sum_{k=1}^{n} c_k(x)\varphi_{jk}(x)$ into eq. (2), we have a system of linear equations in the $c_k'(x)$. Solving for the $c_k'(x)$ and integrating, we have $c_k(x)$.

Special forms of $X(x)$ or $X_j(x)$ determine the form of the solutions, and the parameters may be found by the method of undetermined coefficients. The following table shows some examples of special solutions for eq. (1). In the table, α, k, a, b, c, are constants, p_r, q_r are polynomials of degree r, and I_a is the operator defined by

$$I_a \cdot F = \frac{1}{a}\left[\sin ax \int \cos ax \cdot F(x)\,dx - \cos ax \int \sin ax \cdot F(x)\,dx\right] \quad (a \neq 0).$$

$X(x)$	Condition	Special Solution
$p_r(x)$	$\lambda=0$ is an m-tuple root of(1′)	$x^m q_r(x)$
$ke^{\alpha x}$	$\lambda=\alpha$ is an m-tuple root of (1′)	$cx^m e^{\alpha x}$
$e^{\alpha x}p_r(x)$	$\lambda=\alpha$ is an m-tuple root of (1′)	$x^m q_r(x)e^{\alpha x}$
$\left.\begin{matrix}\cos(ax+b)\\ \sin(ax+b)\end{matrix}\right\}$	$\left\{\begin{matrix}(1')\equiv\varphi(\lambda^2)+\lambda\psi(\lambda^2), \text{ and}\\ \varphi(-a^2)+a^2\psi(-a^2)\neq 0\end{matrix}\right\}$	$c_1\cos(ax+b)+c_2\sin(ax+b)$
$\left.\begin{matrix}\cos(ax+b)\\ \sin(ax+b)\end{matrix}\right\}$	$\left\{\begin{matrix}(1')=g(\lambda)/f(\lambda^2) \text{ and } f(\lambda^2)\\ \text{is divisible by } (\lambda^2+a^2)^m\\ \text{(but not by } (\lambda^2+a^2)^{m+1})\end{matrix}\right\}$	$c(I_a)^m\left\{\begin{matrix}\cos(ax+b)\\ \sin(ax+b)\end{matrix}\right.$

(II) Riemann's P-Function and Special Functions (→ 253 Linear Ordinary Differential Equations (Global Theory))

(1) Some Examples Expressed by Elementary Functions. A,B are integral constants.

$$P\left\{\begin{matrix} a & b & c & \\ 0 & \mu & -\mu & x \\ 1 & \mu' & -\mu' & \end{matrix}\right\} = \left\{\begin{matrix} A\left(\dfrac{x-b}{x-c}\right)^{\mu} + B\left(\dfrac{x-b}{x-c}\right)^{\mu'} & (\mu\neq\mu'), \\ \left(\dfrac{x-b}{x-c}\right)^{\mu}\left[A+B\log\left(\dfrac{x-b}{x-c}\right)\right] & (\mu=\mu'). \end{matrix}\right.$$

$$P\left\{\begin{matrix} a & b & c & \\ \lambda & \mu & \nu & x \\ 0 & 0 & 0 & \end{matrix}\right\} = A + B\int (x-a)^{\lambda-1}(x-b)^{\mu-1}(x-c)^{\nu-1}dx \quad (\lambda+\mu+\nu=1).$$

These are for finite a, b, c. If $c=\infty$, $x-c$ should be replaced by 1.

$$P\left\{\begin{matrix} \overbrace{\quad\infty\quad} & 0 & \\ \alpha \quad 0 & 0 & x \\ \alpha' \quad 0 & 1 & \end{matrix}\right\} = \left\{\begin{matrix} Ae^{\alpha x} + Be^{\alpha' x} & (\alpha\neq\alpha'), \\ e^{\alpha x}(Ax+B) & (\alpha=\alpha'). \end{matrix}\right.$$

$$P\left\{\begin{matrix} \overbrace{\quad\infty\quad} & 0 & \\ \alpha \quad 1-\sigma & \sigma & x \\ 0 \quad 0 & 0 & \end{matrix}\right\} = A + B\int e^{\alpha x}x^{\sigma-1}dx \quad (\alpha\neq 0).$$

$$P\left\{\begin{matrix} \overbrace{\quad\infty\quad} & 0 & \\ \alpha \quad -\sigma & \sigma & x \\ \alpha' \quad 1-\sigma' & \sigma' & \end{matrix}\right\} = x^{\sigma}e^{\alpha x}\left[A + B\int e^{(\alpha'-\alpha)x}x^{\sigma'-\sigma-1}dx\right] \quad (\alpha\neq\alpha').$$

Riemann's P-function is reduced to Gauss's hypergeometric function with parameters $\alpha=\lambda+\mu+\nu$, $\beta=\lambda+\mu+\nu'$, $\gamma=1+\lambda+\lambda'$ by transforming a, b, c to 0, 1, ∞ by a suitable linear transformation and by putting $z=x^{-\lambda}(x-1)^{-\mu}y$.

(2) Representation of Special Functions by Riemann's P-function.
(i) Gauss's hypergeometric differential equation $x(1-x)y''+[\gamma-(\alpha+\beta+1)x]y'-\alpha\beta y=0$.

$$y = P\left\{\begin{matrix} 0 & 1 & \infty & \\ 0 & 0 & \alpha & x \\ 1-\gamma & \gamma-\alpha-\beta & \beta & \end{matrix}\right\}.$$

A special solution is $F(\alpha,\beta,\gamma;x)$ (→ 206 Hypergeometric Functions).

(ii) Confluent hypergeometric differential equation $xy'' - (x-\mu)y' - \lambda y = 0$.

$$y = P\begin{Bmatrix} & \overbrace{\infty} & & 0 & \\ 0 & \lambda & & 0 & x \\ 1 & \mu-\lambda & & 1-\mu & \end{Bmatrix}.$$

A special solution is

$${}_1F_1(\lambda, \mu; x) \equiv \sum_{k=0}^{\infty} \frac{\Gamma(\lambda+k)}{\Gamma(\lambda)} \frac{\Gamma(\mu)}{\Gamma(\mu+k)} \frac{x^k}{k!}.$$

(iii) Whittaker's differential equation $y'' + \left[-\frac{1}{4} + \frac{k}{x} + \frac{(1/4)-n^2}{x^2} \right] y = 0$.

$$y = P\begin{Bmatrix} & \overbrace{\infty} & & 0 & \\ 1/2 & k & & 1/2+n & x \\ -1/2 & -k & & 1/2-n & \end{Bmatrix}.$$

Special solutions are $M_{k,n}(x), W_{k,n}(x)$.

(iv) Bessel's differential equation $x^2y'' + xy' + (x^2 - \nu^2)y = 0$.

$$y = P\begin{Bmatrix} & \overbrace{\infty} & & 0 & \\ i & 1/2 & & \nu & x \\ -i & 1/2 & & -\nu & \end{Bmatrix}.$$

Special solutions are $J_\nu(x)$, $N_\nu(x)$ (→ 39 Bessel Functions). When $m = 0, 1, 2, \ldots, J_{m-1/2}(x) = (-1)^m 2^{m+1/2} \pi^{-1/2} x^{m-1/2} d^m(\cos x)/d(x^2)^m$.

(v) Hermite's different equation (parabolic cylindrical equation) $y'' - 2xy' + 2ny = 0$.

$$y = P\begin{Bmatrix} & \overbrace{\infty} & & 0 & \\ 0 & -n/2 & & 0 & x^2 \\ 1 & (n+1)/2 & & 1/2 & \end{Bmatrix}.$$

When $n = 0, 1, 2, \ldots,$ the Hermite polynomial $H_n(x) = (-1)^n 2^{-n/2} e^{x^2} d^n(e^{-x^2})/dx^n$ is the solution.

(vi) Laguerre's differential equation $xy'' + (l - x + 1)y' + ny = 0$.

$$y = P\begin{Bmatrix} & \overbrace{\infty} & & 0 & \\ 0 & -n & & 0 & x \\ 1 & l+n+1 & & -l & \end{Bmatrix}.$$

When $n = 0, 1, 2, \ldots,$ the Laguerre polynomial $L_n^l(x) = (1/n!) x^{-l} e^x d^n(x^{n+l} e^{-x})/dx^n$ is the solution.

(vii) Jacobi's differential equation $x(1-x)y'' + [q - (p+1)x]y' + n(n+p)y = 0$.

$$y = P\begin{Bmatrix} 0 & 1 & \infty & \\ 1-q & q-p & p+n & x \\ 0 & 0 & -n & \end{Bmatrix}.$$

When $n = 0, 1, 2, \ldots,$ the Jacobi polynomial

$$G_n(p, q; x) = \frac{\Gamma(q) x^{1-q} (1-x)^{q-p}}{\Gamma(n+q)} \frac{d^n\left[x^{q+n-1}(1-x)^{p+n-q}\right]}{dx^n}$$

is the solution.

(viii) Legendre's differential equation $(1-x^2)y'' - 2xy' + n(n+1)y = 0$.

$$y = P\begin{Bmatrix} 1 & -1 & \infty & \\ 0 & 0 & n+1 & x \\ 0 & 0 & -n & \end{Bmatrix}.$$

When $n = 0, 1, 2, \ldots,$ the general solution is

$$\frac{d^n}{dx^n}\left[A(x^2-1)^n + B(x^2-1)^n \int \frac{dx}{(x^2-1)^{n+1}} \right].$$

The Legendre polynomial $P_n(x) = [d^n\{(x^2-1)^n\}/dx^n]/2^n n!$ is a special solution.

(3) Solution by Cylindrical Functions of Ordinary Linear Differential Equations of the Second Order. We denote cylindrical functions by $C_\nu(x)$ (→ 39 Bessel Functions).

Equation	Solution
$y''+\frac{1-2\alpha}{x}y'+\left[(\beta\gamma x^{\gamma-1})^2+\frac{\alpha^2-\nu^2\gamma^2}{x^2}\right]y=0$	$y=x^\alpha C_\nu(\beta x^\gamma)$
$y''+\left[\frac{1-2\alpha}{x}-2\beta\gamma i x^{\gamma-1}\right]y'+\left[\frac{\alpha^2-\nu^2\gamma^2}{x^2}-\beta\gamma(\gamma-2\alpha)ix^{\gamma-2}\right]y=0$	$y=x^\alpha\exp(i\beta x^\gamma)C_\nu(\beta x^\gamma)$
$y''+\left[\frac{1}{x}-2u(x)\right]y'+\left[1-\frac{\nu^2}{x^2}+u(x)^2-u'(x)-\frac{u(x)}{x}\right]y=0$	$y=\exp\left[\int u(x)\,dx\right]C_\nu(x)$
$y''+\alpha^2\nu^2x^{2\nu-2}y=0$	$y=\sqrt{x}\,C_{1/2\nu}(\alpha x^\nu)$
$y''+(e^{2x}-\nu^2)y=0$	$y=C_\nu(e^x)$
$x^2y''+xy'+(\beta^2x^2-\nu^2)y=0$	$y=C_\nu(\beta x)$
$x^2y''+xy'-(x^2+\nu^2)y=0$	$y=C_\nu(ix)$ (modified Bessel function)

(III) Transformation Groups and Invariants

Let $U\equiv\xi\frac{\partial}{\partial x}+\eta\frac{\partial}{\partial y}$ be the infinitesimal transformation of a given continuous transformation group of two variables, and $U'\equiv\xi\frac{\partial}{\partial x}+\eta\frac{\partial}{\partial y}+\zeta\frac{\partial}{\partial p}$ be that of its extended group.

We have

$$\zeta=\frac{\partial\eta}{\partial x}+p\left(\frac{\partial\eta}{\partial y}-\frac{\partial\xi}{\partial x}\right)-p^2\frac{d\xi}{\partial x}.$$

We put

$$p\equiv\frac{dy}{dx},\qquad r\equiv\frac{d^2y}{dx^2}.$$

Let α, β and γ be invariants of the 0th, first, and second order, respectively. The general form of the diffferential equation of the first or of the second order invariant under U is given by $\Phi(\alpha,\beta)=0$ (or $\beta=F(\alpha)$), and $\Psi(\alpha,\beta,\gamma)=0$ (or $\gamma=G(\alpha,\beta)$), respectively, where F, Φ, Ψ, G denote arbitrary functions of the corresponding variables.

Group With Infinitesimal Transformation U			Invariants			Note
ξ	η	ζ	0th	1st	2nd	
0	1	0	x	p	r	(1)
1	0	0	y	p	r	(1)
$-y$	x	$1+p^2$	x^2+y^2	$(y-xp)/(x+yp)$	$r/(1+p^2)^{3/2}$	(2)
0	y	p	x	p/y	r/y	(3)
x	0	$-p$	y	xp	x^2r	(3)
x	y	0	y/x	p	xr	
x	$-y$	$-2p$	xy	x^2p	x^3r	(4)
μx	νy	$(\nu-\mu)p$	y^μ/x^ν	$x^{1-\nu/\mu}p$ or px/y	$r\mu/x^{\nu-\mu-1}$	
μ	ν	0	$\nu x-\mu y$	p	r	
0	$h(x)$	$h'(x)$	x	$h(x)p-h'(x)y$	$h(x)r-h''(x)y$	
$k(y)$	0	$-k'(y)p^2$	y	$\frac{1}{p}-\frac{k'(y)}{k(y)}x$	$\frac{r}{p^3}+\frac{k''(y)}{k'(y)p}$	(5)
0	$k(y)$	$k'(y)p$	x	$\frac{p}{k(y)}$	$\frac{r}{k(y)}-\frac{k'(y)p^2}{[k(y)]^2}$	(6)
$h(x)$	0	$-h'(x)p$	y	$h(x)p$	$(h(x))^2r+h(x)h'(x)p$	

Group With Infinitesimal Transformation U			Invariants			Note
ξ	η	ζ	0th	1st	2nd	
0	$h(x)k(y)$	$h'(x)k(y) + h(x)k'(y)p$	x	$\frac{p}{k(y)}-\frac{h'(x)}{h(x)}\int^{y}\frac{dy}{k(y)}$	—	(7)
$xh(x)$	$yh(x)$	$h'(x)(y-xp)$	$\frac{y}{x}$	$\left(p-\frac{y}{x}\right)h(x)$	$\left(\frac{x^2r}{xp-y}-1\right)h(x) + h'(x)$	
y	x	$1-p^2$	x^2-y^2	$\frac{1-p}{1+p}\frac{x+y}{x-y}$ or $(x-yp)/(1+p)(x-y)$	$\frac{r}{(1-p^2)^{3/2}}$	

Notes

(1) Parallel translation.
(2) Rotation.
(3) Affine transformation.
(4) Similar transformation; the equation is a homogeneous differential equation.
(5) Linear differential equation.
(6) Separated variable type.
(7) When $k(y)=y^n$, the equation is Bernoulli's differential equation.

Reference

[1] A. R. Forsyth, A treatise on differential equations, Macmillan, fourth edition, 1914.

15. Total and Partial Differential Equations

(I) Total Differential Equations (→ 428 Total Differential Equations)

Suppose we are given a system of total differential equations

$$dz_j=\sum_{k=1}^{n} P_{jk}(x;z)\,dx_k \quad (j=1,2,\ldots,m).$$

A condition for complete integrability is given by

$$\frac{\partial P_{jk}(x;z)}{\partial x_l}+\sum_i \frac{\partial P_{jk}(x;z)}{\partial z_i}P_{il}(x;z)=\frac{\partial P_{jl}(x;z)}{\partial x_k}+\sum_i \frac{\partial P_{jl}(x;z)}{\partial z_i}P_{ik}(x;z).$$

Under this condition, the solution with the initial condition $(x_1^0,\ldots,x_n^0;z_1^0,\ldots,z_m^0)$ is obtained as follows: First, solve the system of differential equations $dz_j/dx_1=P_{j1}(x_1,x_2^0,\ldots,x_n^0;z)$ in x_1 with the initial condition $z_j(x_1^0)=z_j^0$, and denote the solution by $z_j=\varphi_j(x_1)$. Next, considering x_1 as a parameter, solve the system of differential equations $dz_j/dx_2=P_{j2}(x_1,x_2,x_3^0,\ldots,x_n^0;z)$ in x_2 with the initial condition $z_j(x_2^0)=\varphi_j(x_1)$, and denote the solution by $z_j=\varphi_j(x_1,x_2)$. Repeat the process, until we finally have $z_j=\varphi_j(x_1,\ldots,x_n)$, which is the solution of the original equation. Or, if we have m independent first integrals $f_j(x;z)=c_j$ of the equation $dz_j/dx_1=P_{j1}(x;z)$, we may transform the equation into $du_j=\sum_{k=1}^{n}Q_{jk}(x;u)dx_k$ by the transformation $u_j=f_j(x;z)$. Since the $Q_{jk}(x;u)$ do not involve x_1 and the equation is a completely integrable total differential equation, we have reduced the number of variables. We obtain the general solution by repeating this process n times.

For

$$P(x,y,z)\,dx+Q(x,y,z)\,dy+R(x,y,z)\,dz=0$$

$(n=3,m=1)$, the complete integrability condition is

$$P\left(\frac{\partial Q}{\partial z}-\frac{\partial R}{\partial y}\right)+Q\left(\frac{\partial R}{\partial x}-\frac{\partial P}{\partial z}\right)+R\left(\frac{\partial P}{\partial y}-\frac{\partial Q}{\partial x}\right)=0.$$

(II) Solution of Partial Differential Equations of First Order (→ 322 Partial Differential Equations (Methods of Integration), 324 Partial Differential Equations of First Order)

Let z be a function of x and y, and

$$p \equiv \partial z/\partial x, \quad q \equiv \partial z/\partial y, \quad r \equiv \partial^2 z/\partial x^2, \quad s \equiv \partial^2 z/\partial x \partial y, \quad t \equiv \partial^2 z/\partial y^2.$$

We consider a partial differential equation of the first order $F(x,y,z,p,q)=0$.

(1) The Lagrange-Charpit Method. We consider the auxiliary equation

$$\frac{dx}{F_p}=\frac{dy}{F_q}=\frac{dz}{pF_p+qF_q}=\frac{-dp}{F_x+pF_z}=\frac{-dq}{F_y+qF_z},$$

which is a system of ordinary differential equations. Let $G(x,y,z,p,q)=a$ be the solution of the auxiliary equation. Using this together with the original equation $F=0$, we obtain $p=P(x,y,z,a)$, $q=Q(x,y,z,a)$, and the complete solution by integrating $dz=P\,dx+Q\,dy$. If we know another solution of the auxiliary equation $H(x,y,z,p,q)=b$ independent of $G=a$, we have the complete solution $z=\Phi(x,y,a,b)$ by eliminating p and q from $F=0$, $G=a$, and $H=b$.

(2) Solution of Various Standard Forms of Partial Differential Equations of the First Order. The integration constants are a, b.
(i) $f(p,q)=0$. The complete solution is $z=ax+\varphi(a)y+b$, where the function $t=\varphi(a)$ is defined by $f(t,a)=0$.
(ii) $f(px,q)=0$, $f(x,qy)=0$, $f(p/z,q/z)=0$. These equations reduce to (i) if $x=e^X$, $y=e^Y$, $z=e^Z$, respectively.
(iii) $f(x,p,q)=0$. If we can solve for $p=F(x,q)$, the complete solution is $z=\int F(x,a)dx+ay+b$. A similar procedure applies to $f(y,p,q)=0$.
(iv) $f(z,p,q)=0$. Solve $f(z,t,at)=0$ for $t=F(z,a)$. The complete solution is then given by $x+ay+b=\int dz/F(z,a)$. If we eliminate a and b from the complete solution $\Phi(x,y,z,a,b)=0$ and $\partial\Phi/\partial a=\partial\Phi/\partial b=0$, we have the singular solution of the original equation.
(v) Separated variable type $f(x,p)=g(y,q)$. Solve the two ordinary differential equations $f(x,p)=a$ and $g(y,q)=a$ for the solutions $p=P(x,a)$ and $q=Q(y,a)$, respectively. Then the complete solution is $z=\int P(x,a)dx+\int Q(y,a)dy+b$.
(vi) Lagrange's partial differential equation $Pp+Qq=R$. Here P, Q, R are functions of x, y, and z. Denote the solutions of the system of differential equations $dx:dy:dz=P:Q:R$ by $u(x,y,z)=a$, $v(x,y,z)=b$. Then the general solution is $\Phi(u,v)=0$, where Φ is an arbitrary function. A similar method is applicable to

$$\sum_{j=1}^{n} P_j(x_1,\ldots,x_n)\frac{\partial z}{\partial x_j}=R(x_1,\ldots,x_n).$$

If we have n independent solutions $u_j(x)=a_j$ of a system of n differential equations $dx_j/P_j=dz/R$ $(j=1,\ldots,n)$, the general solution is given by $\Phi(u_1,\ldots,u_n)=0$.
(vii) Clairaut's partial differential equation $z=px+qy+f(p,q)$. The complete solution is given by the family of planes $z=ax+by+f(a,b)$. The singular solution as the envelope of the family of planes is given by eliminating p and q from the original equation and $x=-\partial f/\partial p$ and $y=-\partial f/\partial q$.

(III) Solutions of Partial Differential Equations of Second Order (→ 322 Partial Differential Equations (Methods of Integration))

(1) Quadrature. Here φ and ψ are arbitrary functions.
(i) $r=f(x)$. The general solution is $z=\int\int f(x)dx\,dx+\varphi(y)x+\psi(y)$. A similar rule applies to $t=f(y)$.
(ii) $s=f(x,y)$. The general solution is $z=\int\int f(x,y)dx\,dy+\varphi(x)+\psi(y)$.
(iii) Wave equation. $r-t=0$. The general solution is $z=\varphi(x+y)+\psi(x-y)$.
(iv) Laplace's differential equation. $r+t=0$. Let $x+iy=\zeta$ and φ, ψ be complex analytic functions of ζ. The general solution is $z=\varphi(\zeta)+\psi(\bar\zeta)$, and a real solution is $z=\varphi(\zeta)+\bar\varphi(\bar\zeta)$.
(v) $r+Mp=N$, where M and N are functions of x and y. The general solution is given by $z=\int[\int L(x,y)N(x,y)dx+\varphi(y)]/L(x,y)dx+\psi(y)$, $L(x,y)=\exp\int[M(x,y)dx]$. In the integration, y is considered a constant.
A similar method is applicable to $s+Mp=N$, $s+Mq=N$, and $t+Mq=N$.
(vi) Monge-Ampère partial differential equation. $Rr+Ss+Tt+U(rt-s^2)=V$, where R, S, T, U, V are functions of x, y, z, p, q.

First, in the case $U=0$, we take auxiliary equations

eq. (1) $\quad R\,dy^2+T\,dx^2-S\,dx\,dy=0,$

eq. (2) $\quad R\,dp\,dy+T\,dq\,dx=V\,dx\,dy.$

Equation (1) is decomposed into two linear differential forms $X_i dx+Y_i dy=0$ $(i=1,2)$. The combination with (2) gives a solution $u_i(x,y,z,p,q)=a_i$, $v_i(x,y,z,p,q)=b_i$ $(i=1,2)$, and we have intermediate integrals $F_i(u_i,v_i)=0$ $(i=1,2)$ for an arbitrary function F_i. We have the solution of the original equation by solving the intermediate integrals. If $S^2\neq 4RT$, two intermediate integrals are distinct, and hence we can solve them in the form $p=P(x,y,z)$, $q=Q(x,y,z)$, and then we may integrate $dz=P\,dx+Q\,dy$.

Next, in the case $U\neq 0$, let λ_1 and λ_2 be the solutions of $U^2\lambda^2+US\lambda+TR+UV=0$. We have two auxiliary equations

$$\begin{cases}\lambda_1 U\,dy+T\,dx+U\,dp=0,\\ \lambda_2 U\,dx+R\,dy+U\,dq=0,\end{cases} \quad\text{or}\quad \begin{cases}\lambda_2 U\,dy+T\,dx+U\,dp=0,\\ \lambda_1 U\,dx+R\,dy+U\,dq=0,\end{cases}$$

and from the solutions $u_i=a_i$, $v_i=b_i$ $(i=1,2)$, we have intermediate integrals $F_i(u_i,v_i)=0$ $(i=1,2)$. If $4(TR+UV)\neq S^2$, $\lambda_1\neq\lambda_2$, we have two different intermediate integrals $F_i=0$. Solving the simultaneous equations $F_i=0$ in $p=P(x,y,z)$, $q=Q(x,y,z)$, we may also find the solution by integrating $dz=P\,dx+Q\,dy$.

(vii) Poisson's differential equation. $P=(rt-s^2)^n Q$, where $P=P(p,q,r,s,t)$ is homogeneous with respect to r, s, t and we assume that $Q=Q(x,y,z)$ satisfies $\partial Q/\partial z\neq\infty$ for x, y, z when $rt=s^2$. The equation $P(p,\varphi(p),r,r\varphi'(p),r\{\varphi'(p)\}^2)=0$ is then an ordinary differential equation in φ as a function of p. We first solve this for φ, and then solve a partial differential equation of the first order $q=\varphi(p)$ by the method (II)(2)(i).

(2) Intermediate Integrals. Let $f(x,y,z,p,q,r,s,t)$ be polynomials with respect to r, s, t. Suppose that $f(x,y,z,p,q,r,s,t)=0$ has the first integral $u(x,y,z,p,q)=0$. We insert

$$r=-\left(\frac{\partial u}{\partial x}+p\frac{\partial u}{\partial z}+s\frac{\partial u}{\partial q}\right)\Big/\frac{\partial u}{\partial p},\qquad t=-\left(\frac{\partial u}{\partial y}+q\frac{\partial u}{\partial z}+s\frac{\partial u}{\partial p}\right)\Big/\frac{\partial u}{\partial q}$$

into the original equation, and replace all the coefficients that are polynomials of s by 0. We thus obtain a system of differential equations in u. If u and v are two independent solutions of this system, an intermediate integral of the original equation is given in the form $\Phi(u,v)=0$.

(3) Initial Value Problem for a Hyperbolic Partial Differential Equation $L[u]\equiv u_{xy}+au_x+bu_y+cu=h$.

$$u(\xi,\eta)=[(uR)_A+(uR)_B]/2+\iint_\Delta R(x,y;\xi,\eta)h(x,y)\,dx\,dy$$
$$+\int_A^B\left[\frac{1}{2}\left(u\frac{\partial R}{\partial n'}-R\frac{\partial u}{\partial n'}\right)-\{a\cos(n,x)+b\cos(n,y)\}uR\right]ds,$$

where Δ is the hatched region in Fig. 19, and the conormal n' is the mirror image of the normal n with respect to $x=y$.

$$u(\xi,\eta)=(uR)_C+\int_C^A R(u_y+au)\,dy+\int_C^B R(u_x+bu)\,dx+\iint_\Box R(x,y;\xi,\eta)h(x,y)\,dx\,dy$$

(characteristic initial value problem).

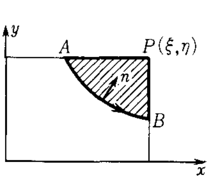


Fig. 19

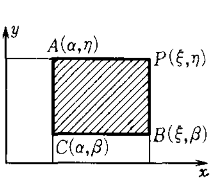


Fig. 20

Here $\Box$ is the hatched rectangular region in Fig. 20. $R(x,y;\xi,\eta)$ is the Riemann function; it satisfies

$M[R(x,y;\xi,\eta)]=0,$

$R_x-bR=0$ (on $x=\xi$),

$R_y-aR=0$ (on $y=\eta$),

$R(\xi,\eta;\xi,\eta)=1.$

Example (i). $u_{xy}=h(x,y)$. $R(x,y;\xi,\eta)=1$.

$$u(\xi,\eta)=\frac{1}{2}[u_A+u_B]+\frac{1}{2}\int_A^B\left[u_y\cos(n,x)+u_x\cos(n,y)\right]ds+\iint_\Delta h(x,y)\,dx\,dy.$$

Example (ii). Telegraph equation $u_{xy}+cu=0$ $(c>0)$. $R(x,y;\xi,\eta)=J_0\left(2\sqrt{c(x-\xi)(y-\eta)}\right)$.

Example (iii). $u_{xy}+\dfrac{n}{x+y}(u_x+u_y)=0$ ($n=$ a constant >0).

$$R(x,y;\xi,\eta)=\left(\frac{x+y}{\xi+\eta}\right)^n F\left(1-n,n;1;-\frac{(x-\xi)(y-\eta)}{(x+y)(\xi+\eta)}\right).$$

(IV) Contact Transformations (→ 82 Contact Transformations)

We consider a transformation $(x_1,\ldots,x_n;z)\to(X_1,\ldots,X_n;Z)$. We put $p_j\equiv\partial z/\partial x_j$, $P_j\equiv\partial Z/\partial X_j$ $(j=1,\ldots,n)$. The transformation is called a contact transformation if there exists a function $\rho(x,z,p)\neq 0$ satisfying $dZ-\Sigma P_j dX_j=\rho(x,z,p)(dz-\Sigma p_j dx_j)$.

A transformation given by $(2n+1)$ equations $\Omega=0$, $\partial\Omega/\partial X_j+P_j\partial\Omega/\partial Z=0$, $\partial\Omega/\partial x_j+p_j\partial\Omega/\partial z=0$ generated by a generating function $\Omega(x,z,X,Z)$ is a contact transformation.

Generating Function	ρ	Transformation	Name
Σx_jX_j+z+Z	-1	$X_j=-p_j,\quad P_j=-x_j,$ $Z=\Sigma p_jx_j-z$	Legendre's transformation
$\Sigma X_j^2+Z^2-\Sigma x_jX_j-zZ$	$Z/(2Z-z)$	$X_j=-p_jZ,$ $P_j=-(2X_j-x_j)/(2Z-z)$	Pedal transformation
$\Sigma(X_j-x_j)^2+(Z-z)^2-a^2$	1	$X_j=x_j-ap_j(1+\Sigma p_j^2)^{-1/2},$ $P_j=p_j,$ $Z=z_j+a(1+\Sigma p_j^2)^{-1/2}$	Similarity
$\Sigma(X_j-x_j)^2-Z^2-z^2$	$-\dfrac{1}{\sqrt{\Sigma p_j^2-1}}$	$X_j=x_j-p_jz,$ $P_j=-p_j(\Sigma p_j^2-1)^{-1/2},$ $Z=z(\Sigma p_j^2-1)^{1/2}$	

(V) Fundamental Solutions (→ 320 Partial Differential Equations H)

A function (or a generalized function such as a distribution) T satisfying $LT=\delta$ (δ is the Dirac delta function) for a linear differential operator L is called the fundamental (or elementary) solution of L. In the following table, we put

$$\Delta\equiv\sum_{i=1}^n\frac{\partial^2}{\partial x_i^2},\quad \square\equiv\frac{\partial^2}{\partial x_n^2}-\sum_{i=1}^{n-1}\frac{\partial^2}{\partial x_i^2},\quad r^2\equiv\sum_{i=1}^n x_i^2,\quad \mathbf{1}(x)=\begin{cases}1 & (x>0)\\ 0 & (x\leqslant 0)\end{cases}\quad\text{(Heaviside function).}$$

J_ν is the Bessel function of the first kind; K_ν and I_ν are the modified Bessel functions. (→Table 19.IV, this Appendix.)

$$s=\begin{cases}\sqrt{x_n^2-x_1^2-\ldots-x_{n-1}^2} & \text{(if } x_n>0 \text{ and the quantity under the radical sign is positive),}\\ 0 & \text{(otherwise).}\end{cases}$$

(For Pf (finite part) → 125 Distributions and Hyperfunctions.)

Operator	Fundamental Solution
d/dx	$\mathbf{1}(x)$
$\dfrac{d^m}{dx^m}$	$\begin{cases}x^{m-1}/(m-1)! & (x>0)\\ 0 & (x\leqslant 0)\end{cases}$
$\partial^n/\partial x_1\partial x_2\ldots\partial x_n$	$\mathbf{1}(x_1)\mathbf{1}(x_2)\ldots\mathbf{1}(x_n)$
$\dfrac{\partial}{\partial x}+i\dfrac{\partial}{\partial y}$	$\dfrac{1}{2\pi}\dfrac{1}{x+iy}\quad(i\equiv\sqrt{-1})$

Operator	Fundamental Solution
Δ	$\begin{cases} -\left[\Gamma\left(\frac{n}{2}\right)/2(n-2)\pi^{n/2}\right]\frac{1}{r^{n-2}} & (n\geq 3) \\ (1/2\pi)\log r & (n=2) \end{cases}$
Δ^m	$\begin{cases} \left[\Gamma\left(\frac{n}{2}\right)/2^m(m-1)!\pi^{n/2}\prod_{k=1}^{m}(2k-n)\right]\frac{1}{r^{n-2m}} \begin{pmatrix} n-2m \text{ is a positive integer} \\ \text{or a negative odd integer} \end{pmatrix} \\ \left[\Gamma\left(\frac{n}{2}\right)/2^m(m-1)!\pi^{n/2}\prod_{k=1,\ k\neq m-h}^{m}(2k-n)\right]\frac{\log r}{r^{n-2m}} \begin{pmatrix} n-2m=-2h, \\ h=0,1,2,\dots \end{pmatrix} \end{cases}$
$\left(1-\frac{\Delta}{4\pi^2}\right)^m$	$\frac{2\pi^m}{(m-1)!}r^{m-(n/2)}K_{(n/2)-m}(2\pi r)$
$\Box^m$	$(\mathrm{Pf}\, s^{2m-n})/\pi^{(n/2)-1}2^{2m-1}(m-1)!\Gamma[m+1-(n/2)]$
$(\Box-\lambda)^m$ (λ is real and $\neq 0$)	$\frac{\lvert\lambda\rvert^{(n/4)-(m/2)}}{\pi^{(n/2)-1}(m-1)!2^{m-1+(n/2)}}\mathrm{Pf}\, s^{m-(n/2)}\begin{cases} I_{m-(n/2)}(\sqrt{\lambda s}) & (\lambda>0) \\ J_{m-(n/2)}(\sqrt{\lvert\lambda\rvert s}) & (\lambda<0) \end{cases}$
$\left(\frac{\partial}{\partial x_n}-\sum_{i=1}^{n-1}\frac{\partial^2}{\partial x_i^2}\right)^m$	$\begin{cases} \frac{x_n^{m-1}}{(m-1)!}\left(\frac{1}{2\sqrt{\pi x_n}}\right)^{n-1}\exp\left(-\sum_{i=1}^{n-1}x_i^2/4x_n\right) & (x_n>0) \\ 0 & (x_n\leq 0) \end{cases}$

(VI) Solution of Boundary Value Problems (→ 188 Green's Functions, 323 Partial Differential Equations of Elliptic Type, 327 Partial Differential Equations of Parabolic Type)

$$L[u]=Au_{xx}+2Bu_{xy}+Cu_{yy}+Du_x+Eu_y+Fu,$$

$$M[v]=(Av)_{xx}+2(Bv)_{xy}+(Cv)_{yy}-(Dv)_x-(Ev)_y+Fv.$$

Green's formula $\int\int_D\{vL[u]-uM[v]\}\,dx\,dy=\int_C\left\{P\left(u\frac{\partial v}{\partial n'}-v\frac{\partial u}{\partial n'}\right)+Quv\right\}ds.$

eq. (1) $\begin{cases} A\cos(n,x)+B\cos(n,y)=P\cos(n',x), \\ B\cos(n,x)+C\cos(n,y)=P\cos(n',y). \end{cases}$

$$Q=(A_x+B_y-D)\cos(n,x)+(B_x+C_y-E)\cos(n,y).$$

The integration contour C is the boundary of the domain D (Fig. 21), n is the inner normal of C, and n', called the conormal, is given by (1).

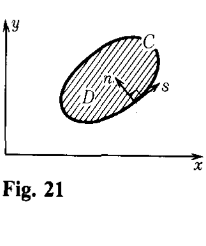


Fig. 21

(1) Elliptic Partial Differential Equation $L[u]\equiv u_{xx}+u_{yy}+au_x+bu_y+cu=h$.

$$u(\xi,\eta)=-\int_C u(s)\frac{\partial G}{\partial n}ds+\int\int_D G(x,y;\xi,\eta)h(x,y)\,dx\,dy.$$

Here $G(x,y;\xi,\eta)$ is Green's function, which satisfies $M(G(x,y;\xi,\eta))=0$ in the interior of D except at $(x,y)\neq(\xi,\eta)$, and

$$G(x,y;\xi,\eta)=-(1/2\pi)\log\sqrt{(x-\xi)^2+(y-\eta)^2}\ +\text{ a regular function},$$

$$G(x,y;\xi,\eta)=0\quad((x,y)\in C).$$

(2) Laplace's Differential Equation in the 2-Dimensional Case $\dfrac{\partial^2 u}{\partial x^2}+\dfrac{\partial^2 u}{\partial y^2}=0.$

$u(x,y)\equiv\tilde u(r,\varphi)=\operatorname{Re}f(z)\quad(z\equiv x+iy=re^{i\theta}).$

(i) Interior of a disk ($r\leqslant 1$).

$$f(z)=\frac{1}{2\pi}\int_0^{2\pi}\tilde u(1,\varphi)\frac{e^{i\varphi}+z}{e^{i\varphi}-z}d\varphi,\qquad \tilde u(r,\theta)=\frac{1}{2\pi}\int_0^{2x}\tilde u(1,\varphi)\frac{1-r^2}{1-2r\cos(\theta-\varphi)+r^2}d\varphi$$

(Poisson's integration formula).

(ii) Annulus ($0<q\leqslant r\leqslant 1$).

$$f(z)=\frac{\omega_1}{\pi^2 i}\left\{\int_0^{2\pi}\tilde u(1,\varphi)\zeta_1(w)\,d\varphi-\int_0^{2\pi}\tilde u(q,\varphi)\zeta_3(w)\,d\varphi-a\log z\right\}$$

(Villat's integration formula).

$$w=\frac{\omega_1}{\pi}(i\log z+\varphi),\quad a=\left(\frac{1}{2\omega_3}-\frac{\eta_1}{\pi i}\right)\int_0^{2\pi}\{\tilde u(1,\varphi)-\tilde u(q,\varphi)\}\,d\varphi,\quad \frac{\omega_3}{\omega_1}=-\frac{i}{\pi}\log q.$$

Here ζ_1 and ζ_3 are the Weierstrass ζ-functions (→ 134 Elliptic Functions) with the fundamental periods $2\omega_1$ and $2\omega_3$.

(iii) Half-plane ($y\geqslant 0$). $f(z)=\dfrac{i}{\pi}\displaystyle\int_{-\infty}^{\infty}\frac{u(t,0)}{z-t}dt,\quad u(x,y)=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{u(t,0)y}{(x-t)^2+y^2}dt.$

(3) Laplace's Differential Equation in the 3-Dimensional Case.

(i) Interior of a sphere ($r\leqslant 1$).

$$\tilde u(r,\varphi,\theta)=\frac{1}{4\pi}\int_0^{\pi}\int_0^{2\pi}\tilde u(1,\Phi,\Theta)\frac{1-r^2}{(1-2r\cos\gamma+r^2)^{3/2}}\sin\Theta\,d\Theta\,d\Phi,$$

where

$$\cos\gamma=\cos\Theta\cos\theta+\sin\Theta\sin\theta\cos(\Phi-\varphi).$$

(ii) Half-space ($z\geqslant 0$).

$$u(x,y,z)=\frac{z}{2\pi}\int\int_{-\infty}^{\infty}\frac{u(\xi,\eta,0)}{\{(x-\xi)^2+(y-\eta)^2+z^2\}^{3/2}}d\xi\,d\eta.$$

(4) Equation of Oscillation (Helmholtz Differential Equation) $\Delta u+k^2u=0$. Let u_n be the normalized eigenfunction with the same boundary condition for the eigenvalue k_n. Green's function is

$$G(P,Q)=\sum\frac{u_n(P)u_n^*(Q)}{k^2-k_n^2}$$

Domain	Boundary Condition	Eigenvalue	Eigenfunction
rectangle $0\leqslant x\leqslant a,\ 0\leqslant y\leqslant b$	$u=0$	$k_{nm}=\pi\sqrt{\dfrac{n^2}{a^2}+\dfrac{m^2}{b^2}}$ $(n,\ m=1,2,\ldots)$	$\sin n\pi\dfrac{x}{a}\sin m\pi\dfrac{y}{b}$
circle $0\leqslant r\leqslant a$	$u=0$	k_{nm} is the root of $J_m(kx)=0$	$J_m(k_{nm}r)e^{\pm im\varphi}$
annulus $b\leqslant r\leqslant a$	$u=0$	k_{nm} is the root of $J_m(ka)N_m(kb)-J_m(kb)N_m(ka)=0$	$\left\{\dfrac{J_m(k_{nm}r)}{J_m(k_{nm}a)}-\dfrac{N_m(k_{nm}r)}{N_m(k_{nm}a)}\right\}e^{\pm im\varphi}$
fan shape $0\leqslant r\leqslant a,\ 0\leqslant\varphi\leqslant\alpha$	$u=0$	k_{nm} is the root of $J_\mu(ka)=0$ $(\mu=m\pi/\alpha)$	$J_\mu(k_{nm}r)\sin\mu\varphi$
rectangular parallelepiped $0\leqslant x\leqslant a,\ 0\leqslant y\leqslant b,\ 0\leqslant z\leqslant c$	$\dfrac{\partial u}{\partial n}=0$	$k_{nml}=\pi\sqrt{\dfrac{n^2}{a^2}+\dfrac{m^2}{b^2}+\dfrac{l^2}{c^2}}$	$\cos n\pi\dfrac{x}{a}\cos m\pi\dfrac{y}{b}\cos l\pi\dfrac{z}{c}$
sphere $0\leqslant r\leqslant a$	$\dfrac{\partial u}{\partial n}=0$	k_{nl} is the root of $\psi_n'(ka)=0$, where $\psi_n(\rho)\equiv\sqrt{\pi/2}\,J_{n+(1/2)}(\rho)$	$\psi_n(k_{nl}r)P_n^m(\cos\theta)^{\pm im\varphi}$

(5) Heat Equation. $\dfrac{\partial u}{\partial t}=\kappa\Delta u\left(\Delta=\dfrac{\partial^2}{\partial x_1^2}+\ldots+\dfrac{\partial^2}{\partial x_m^2};\ \kappa\text{ is a positive constant}\right)$. Boundary condi-

tion: $hu - k\,\partial u/\partial n = \varphi$, where h and k are nonnegative constants with $h+k=1$, and φ is a given function.

$$u(P,t)=\int_V G(P,Q,t)u(Q,0)\,dV_Q+\kappa\int_0^t d\tau\int_S\left\{\frac{\partial G(P,Q,t-\tau)}{\partial n_Q}+G(P,Q,t-\tau)\right\}\varphi(Q,\tau)\,dS_Q.$$

Here V is the domain, and S is its boundary. $G(P,Q,t)$ is the elementary solution that satisfies $\partial G/\partial t=\kappa\Delta G$ in V and $k\partial G/\partial n = hG$ on S, and further in the neighborhood of $P=Q$, $t=0$, it has the form $G(P,Q,t)=(4\pi\kappa t)^{-m/2}e^{-R^2/4\kappa t}$ + terms of lower degree ($R=\overline{PQ}$).

(i) $-\infty<x<\infty$, $G=U(x-\xi,t)$, where $U(x,t)=e^{-x^2/4\kappa t}/\sqrt{4\pi\kappa t}$ (similar in the following case (ii)).

(ii) $0\leqslant x<\infty$. $u(0,t)=0$: $\quad G=U(x-\xi,t)-U(x+\xi,t)$.

$\dfrac{\partial u}{\partial x}=hu$: $\quad G=U(x-\xi,t)+U(x+\xi,t)-2he^{h\xi}\displaystyle\int_{-\infty}^{-\xi}e^{h\eta}U(x-\eta,t)\,d\eta.$

(iii) $0\leqslant x\leqslant l$.

$u(0,t)=u(l,t)=0$: $$G=\vartheta\left(\frac{x-\xi}{2l}\middle|\tau\right)-\vartheta\left(\frac{x+\xi}{2l}\middle|\tau\right).$$

$\dfrac{\partial u}{\partial x}(0,t)=\dfrac{\partial u}{\partial x}(l,t)=0$: $$G=\vartheta\left(\frac{x-\xi}{2l}\middle|\tau\right)+\vartheta\left(\frac{x+\xi}{2l}\middle|\tau\right).$$

$u(0,t)=\dfrac{\partial u}{\partial x}(l,t)=0$: $$G=\vartheta\left(\frac{x-\xi}{4l}\middle|\tau\right)-\vartheta\left(\frac{x+\xi}{4l}\middle|\tau\right)+\vartheta\left(\frac{x+\xi-2l}{4l}\middle|\tau\right)-\vartheta\left(\frac{x-\xi-2l}{4l}\middle|\tau\right).$$

Here ϑ is the elliptic theta function: $\vartheta(x|\tau)=\vartheta_3(x,\tau)\equiv 1+2\sum e^{i\pi\tau n^2}\cos 2n\pi x$.

(iv) $0\leqslant x<\infty$, $\quad 0\leqslant y<\infty$. $\quad u(x,0,t)=u(0,y,t)=0$:

$$G=\left(e^{-(x-\xi)^2/4\kappa t}-e^{-(x+\xi)^2/4\kappa t}\right)\left(e^{-(y-\eta)^2/4\kappa t}-e^{-(y+\eta)^2/4\kappa t}\right)/4\pi\kappa t.$$

(v) $0\leqslant x\leqslant a$, $\quad 0\leqslant y\leqslant b$. $\quad u=0$ on the boundary:

$$G=\frac{4}{ab}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\exp\left\{-\kappa\pi^2 t\left(\frac{m^2}{a^2}+\frac{n^2}{b^2}\right)\right\}\sin\frac{m\pi x}{a}\sin\frac{m\pi\xi}{a}\sin\frac{n\pi y}{b}\sin\frac{n\pi\eta}{b}.$$

(vi) $0\leqslant x\leqslant a$, $0\leqslant y\leqslant b$, $0\leqslant z\leqslant c$. $u=0$ on the boundary:

$$G=\frac{8}{abc}\sum_{l=1}^{\infty}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\exp\left\{-\kappa\pi^2 t\left(\frac{l^2}{a^2}+\frac{m^2}{b^2}+\frac{n^2}{c^2}\right)\right\}$$
$$\times\sin\frac{l\pi x}{a}\sin\frac{l\pi\xi}{a}\sin\frac{m\pi y}{b}\sin\frac{m\pi\eta}{b}\sin\frac{n\pi z}{c}\sin\frac{n\pi\zeta}{c}.$$

(vii) $0\leqslant r<\infty$. Spherically symmetric. $\quad |x|=r$, $|\xi|=r'$:

$$G=\left(e^{-(r-r')^2/4\kappa t}-e^{-(r+r')^2/4\kappa t}\right)/8\pi rr'(\pi\kappa t)^{1/2}.$$

(viii) $0\leqslant r\leqslant a$. Spherically symmetric. $\quad u=0$ on the boundary:

$$G=\frac{1}{2\pi arr'}\sum_{n=1}^{\infty}e^{-\kappa n^2\pi^2 t/a^2}\sin\frac{n\pi r}{a}\sin\frac{n\pi r'}{a}.$$

(ix) $a\leqslant r<\infty$. Spherically symmetric. $\quad k\partial u/\partial r-hu=0$ on the boundary:

$$G=\frac{1}{8\pi rr'(\pi\kappa t)^{1/2}}\left[e^{-(r-r')^2/4\kappa t}+e^{-(r+r'-2a)^2/4\kappa t}-\frac{ah+k}{ak}(4\pi\kappa t)^{1/2}\right.$$
$$\times\exp\left\{\kappa t\left(\frac{ah+k}{ak}\right)^2+(r+r'-2a)\frac{ah+k}{ak}\right\}$$
$$\left.\times\operatorname{erfc}\left\{\frac{r+r'-2a}{2\sqrt{\kappa t}}+\frac{ah+k}{ak}\sqrt{\kappa t}\right\}\right]\qquad\left(\operatorname{erfc}x\equiv\int_x^\infty e^{-t^2}dt\right).$$

(x) $0\leqslant r<\infty$. Axially symmetric: $\quad G=e^{-(r^2+r'^2)/4\kappa t}I_0(rr'/2\kappa t)/4\pi\kappa t$.

(xi) $0\leqslant r\leqslant a$. Axially symmetric. $\quad k\dfrac{\partial u}{\partial r}-hu=0$ on the boundary:

$$G=\frac{1}{\pi a^2}\sum_{n=1}^{\infty}\frac{J_0(r\alpha_n)J_0(r'\alpha_n)}{\{J_0(a\alpha_n)\}^2+\{J_1(a\alpha_n)\}^2}e^{-\kappa\alpha_n t/a},$$

where α_n is given by $k\alpha_n J_1(a\alpha_n)-h\,J_0(a\alpha_n)=0$.

16. Elliptic Integrals and Elliptic Functions

(I) Elliptic Integrals (→ 134 Elliptic Functions)

(1) Legendre-Jacobi Standard Form.

Elliptic integral of the first kind

$$F(k,\varphi)\equiv\int_0^{\varphi}\frac{d\psi}{\sqrt{1-k^2\sin^2\psi}}=\int_0^{\sin\varphi}\frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}\qquad (k \text{ is the modulus}).$$

Elliptic integral of the second kind

$$E(k,\varphi)\equiv\int_0^{\varphi}\sqrt{1-k^2\sin^2\psi}\;d\psi=\int_0^{\sin\varphi}\sqrt{\frac{1-k^2t^2}{1-t^2}}\;dt.$$

Elliptic integral of the third kind

$$\Pi(\varphi,n,k)=\int_0^{\varphi}\frac{d\psi}{(1+n\sin^2\psi)\sqrt{1-k^2\sin^2\psi}}=\int_0^{\sin\varphi}\frac{dt}{(1+nt^2)\sqrt{(1-t^2)(1-k^2t^2)}}.$$

When $\varphi=\pi/2$, elliptic integrals of the first and the second kinds are called complete elliptic integrals:

$$K(k)\equiv F\left(k,\frac{\pi}{2}\right)=\int_0^{\pi/2}\frac{d\psi}{\sqrt{1-k^2\sin^2\psi}}=\int_0^1\frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}=\frac{\pi}{2}F\left(\frac{1}{2},\frac{1}{2};1;k^2\right),$$

$$E(k)\equiv E\left(k,\frac{\pi}{2}\right)=\int_0^{\pi/2}\sqrt{1-k^2\sin^2\psi}\;d\psi=\int_0^1\sqrt{\frac{1-k^2t^2}{1-t^2}}\;dt=\frac{\pi}{2}F\left(-\frac{1}{2},\frac{1}{2};1;k^2\right),$$

where F is the hypergeometric function.

$$K(k')=K(\sqrt{1-k^2})\equiv K'(k),\qquad E(k')=E(\sqrt{1-k^2})\equiv E'(k)\qquad (k'^2=1-k^2;\ k' \text{ is the complementary modulus}).$$

$$EK'+E'K-KK'=\frac{\pi}{2}\ \text{(Legendre's relation)}.\qquad K\left(\frac{1}{\sqrt{2}}\right)=\frac{\Gamma(1/4)^2}{4\sqrt{\pi}}.$$

$$\frac{\partial F}{\partial k}=\frac{1}{k'^2}\left(\frac{E-k'^2F}{k}-\frac{\sin\varphi\cos\varphi}{\sqrt{1-k^2\sin^2\varphi}}\right),\qquad \frac{\partial E}{\partial k}=\frac{E-F}{k}.$$

(2) Change of Variables.

$$\tan(\psi-\varphi)=k'\tan\varphi:\quad F\left(\frac{1-k'}{1+k'},\psi\right)=(1+k')F(k,\varphi),$$

$$E\left(\frac{1-k'}{1+k'},\psi\right)=\frac{2}{1+k'}[E(k,\varphi)+k'F(k,\varphi)]-\frac{1-k'}{1+k'}\sin\psi.$$

$$\sin\chi=\frac{(1+k)\sin\varphi}{1+k\sin^2\varphi}:\quad F\left(\frac{2\sqrt{k}}{1-k},\chi\right)=(1+k)F(k,\varphi),$$

$$E\left(\frac{2\sqrt{k}}{1+k},\chi\right)=\frac{1}{1+k}\left[2E(k,\varphi)-k'^2F(k,\varphi)+2k\frac{\sin\varphi\cos\varphi}{1+k\sin^2\varphi}\sqrt{1-k^2\sin^2\varphi}\right].$$

k_1	$\sin\varphi_1$	$\cos\varphi_1$	$F(k_1,\varphi_1)$	$E(k_1,\varphi_1)$
$i\frac{k}{k'}$	$k'\frac{\sin\varphi}{\sqrt{1-k^2\sin^2\varphi}}$	$\frac{\cos\varphi}{\sqrt{1-k^2\sin^2\varphi}}$	$k'F(k,\varphi)$	$\frac{1}{k'}\left[E(k,\varphi)-k^2\frac{\sin\varphi\cos\varphi}{\sqrt{1-k^2\sin^2\varphi}}\right]$
k'	$-i\tan\varphi$	$\frac{1}{\cos\varphi}$	$-iF(k,\varphi)$	$i[E(k,\varphi)-F(k,\varphi)-\sqrt{1-k^2\sin^2\varphi}\ \tan\varphi]$
$\frac{1}{k}$	$k\sin\varphi$	$\sqrt{1-k^2\sin^2\varphi}$	$kF(k,\varphi)$	$\frac{1}{k}[E(k,\varphi)-k'^2F(k,\varphi)]$

(3) Transformation into Standard Form.

(i) The following are reducible to elliptic integrals of the first kind (we assume $a>b>0$ for parameters).

$AF(k,\varphi)$	A	k	φ
$\int_1^x \frac{dt}{\sqrt{t^3-1}}$	$\frac{1}{\sqrt[4]{3}}$	$\frac{\sqrt{3}-1}{2\sqrt{2}}$	$\arccos\frac{\sqrt{3}+1-x}{\sqrt{3}-1+x}$
$\int_1^x \frac{dt}{\sqrt{1-t^3}}$	$\frac{1}{\sqrt[4]{3}}$	$\frac{\sqrt{3}+1}{2\sqrt{2}}$	$\arccos\frac{\sqrt{3}-1+x}{\sqrt{3}+1-x}$
$\int_0^x \frac{dt}{\sqrt{1+t^4}}$	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\arccos\frac{1-x^2}{1+x^2}$
$\int_0^x \frac{dt}{\sqrt{(a^2-t^2)(b^2-t^2)}}$	$\frac{1}{a}$	$\frac{b}{a}$	$\arcsin\frac{x}{b}$
$\int_b^x \frac{dt}{\sqrt{(a^2-t^2)(t^2-b^2)}}$	$\frac{1}{a}$	$\sqrt{1-(b/a)^2}$	$\arcsin\sqrt{\frac{1-(b/x)^2}{1-(b/a)^2}}$
$\int_a^x \frac{dt}{\sqrt{(t^2-a^2)(t^2-b^2)}}$	$\frac{1}{a}$	$\frac{b}{a}$	$\arcsin\sqrt{\frac{1-(a/x)^2}{1-(b/x)^2}}$
$\int_0^x \frac{dt}{\sqrt{(a^2+t^2)(b^2+t^2)}}$	$\frac{1}{a}$	$\sqrt{1-(b/a)^2}$	$\arctan\frac{x}{b}$
$\int_0^x \frac{dt}{\sqrt{(a^2-t^2)(b^2+t^2)}}$	$\frac{1}{\sqrt{a^2+b^2}}$	$\frac{a}{\sqrt{a^2+b^2}}$	$\arcsin\sqrt{\frac{1+(b/a)^2}{1+(b/x)^2}}$
$\int_b^x \frac{dt}{\sqrt{(a^2+t^2)(t^2-b^2)}}$	$\frac{1}{\sqrt{a^2+b^2}}$	$\frac{a}{\sqrt{a^2+b^2}}$	$\arccos\frac{b}{x}$

(ii) The following are reducible to elliptic integrals of the second kind (we assume $a>b>0$ for parameters).

$AE(k,\varphi)$	A	k	φ
$\int_0^x \sqrt{\frac{a^2-t^2}{b^2-t^2}}\,dt$	a	$\frac{b}{a}$	$\arcsin\frac{x}{b}$
$\int_x^a \sqrt{\frac{b^2+t^2}{a^2-t^2}}\,dt$	$\sqrt{a^2+b^2}$	$\frac{a}{\sqrt{a^2+b^2}}$	$\arccos\frac{x}{a}$
$\int_b^x \frac{1}{t^2}\sqrt{\frac{t^2+a^2}{t^2-b^2}}\,dt$	$\frac{\sqrt{a^2+b^2}}{b^2}$	$\frac{a}{\sqrt{a^2+b^2}}$	$\arccos\frac{b}{x}$
$\int_\infty^x t^2\sqrt{\frac{t^2+a^2}{t^2-b^2}}\,dt$	$\sqrt{a^2+b^2}$	$\frac{a}{\sqrt{a^2+b^2}}$	$\arcsin\sqrt{\frac{1+(b/a)^2}{1+(x/a)^2}}$
$\int_0^x \sqrt{\frac{a^2+t^2}{(b^2+t^2)^3}}\,dt$	$\frac{a}{b^2}$	$\frac{\sqrt{a^2-b^2}}{a}$	$\arctan\frac{x}{b}$
$\int_b^x \frac{dt}{t^2\sqrt{(t^2-b^2)(a^2-t^2)}}$	$\frac{1}{ab^2}$	$\frac{\sqrt{a^2-b^2}}{a}$	$\arcsin\sqrt{\frac{1-(b/x)^2}{1-(b/a)^2}}$

(II) Elliptic Theta Functions

(1) For $\operatorname{Im}\tau>0$, we put $q\equiv e^{i\pi\tau}$ and define

$$\vartheta_0(u,\tau)\equiv\vartheta_4(u,\tau)\equiv 1+2\sum_{n=1}^{\infty}(-1)^n q^{n^2}\cos 2n\pi u,$$

$$\vartheta_1(u,\tau)\equiv 2\sum_{n=0}^{\infty}(-1)^n q^{[n+(1/2)]^2}\sin(2n+1)\pi u,$$

$$\vartheta_2(u,\tau)\equiv 2\sum_{n=0}^{\infty} q^{[n+(1/2)]^2}\cos(2n+1)\pi u,$$

$$\vartheta_3(u,\tau)\equiv 1+2\sum_{n=1}^{\infty} q^{n^2}\cos 2n\pi u.$$

Each of the four functions ϑ_j $(j=0,1,2,3)$ as a function of two variables u and τ satisfies the following partial differential equation

$$\frac{\partial^2\vartheta(u,\tau)}{\partial u^2}=4\pi i\frac{\partial\vartheta(u,\tau)}{\partial\tau}.$$

(2) Mutual Relations.

$$\vartheta_0^4(u)+\vartheta_2^4(u)=\vartheta_1^4(u)+\vartheta_3^4(u),\qquad \vartheta_0^2(u)=k\vartheta_1^2(u)+k'\vartheta_3^2(u),$$

$$\vartheta_2^2(u)=-k'\vartheta_1^2(u)+k\vartheta_3^2(u),\qquad \vartheta_1^2(u)=k\vartheta_0^2(u)-k'\vartheta_2^2(u),$$

where k is the modulus such that $iK'(k)/K(k)=\tau$, and k' is the corresponding complementary modulus.

$$k=\vartheta_2^2(0)/\vartheta_3^2(0),\qquad k'=\vartheta_0^2(0)/\vartheta_3^2(0).$$

$$\vartheta_1'(0)=\pi\vartheta_2(0)\vartheta_3(0)\vartheta_0(0),\qquad \frac{\vartheta_1'''(0)}{\vartheta_1'(0)}=\frac{\vartheta_2''(0)}{\vartheta_0(0)}+\frac{\vartheta_3''(0)}{\vartheta_3(0)}+\frac{\vartheta_0''(0)}{\vartheta_0(0)}.$$

(3) Pseudoperiodicity. In the following table, the only variables in ϑ are u and τ. m and n are integers.

Increment of u	ϑ_0	ϑ_1	ϑ_2	ϑ_3	Exponential Factor
$m+n\tau$	$(-1)^n\vartheta_0$	$(-1)^{m+n}\vartheta_1$	$(-1)^m\vartheta_2$	ϑ_3	$\exp[-n\pi i \times(2u+n\tau)]$
$m-\frac{1}{2}+n\tau$	ϑ_3	$(-1)^{m+1}\vartheta_2$	$(-1)^{m+n}\vartheta_1$	$(-1)^n\vartheta_0$	
$m+\left(n+\frac{1}{2}\right)\tau$	$(-1)^n i\vartheta_1$	$(-1)^{m+n} i\vartheta_0$	$(-1)^m\vartheta_3$	ϑ_2	$\exp\left[-\left(n+\frac{1}{2}\right)\pi i \times\left\{2u+\left(n+\frac{1}{2}\right)\tau\right\}\right]$
$m-\frac{1}{2}+\left(n+\frac{1}{2}\right)\tau$	ϑ_2	ϑ_3	$(-1)^{m+n} i\vartheta_0$	$(-1)^n i\vartheta_1$	
Zeros $u=$	$m+\left(n+\frac{1}{2}\right)\tau$	$m+n\tau$	$m+\frac{1}{2}+n\tau$	$m+\frac{1}{2}+\left(n+\frac{1}{2}\right)\tau$	

(4) Expansion into Infinite Products. We put $Q_0 \equiv \prod_{n=1}^{\infty}(1-q^{2n})$. Then we have

$$\vartheta_0(u)=Q_0\prod_{n=1}^{\infty}(1-2q^{2n-1}\cos 2\pi u+q^{4n-2}),$$

$$\vartheta_1(u)=2Q_0 q^{1/4}\sin\pi u\prod_{n=1}^{\infty}(1-2q^{2n}\cos 2\pi u+q^{4n}),$$

$$\vartheta_2(u)=2Q_0 q^{1/4}\cos\pi u\prod_{n=1}^{\infty}(1+2q^{2n}\cos 2\pi u+q^{4n}),$$

$$\vartheta_3(u)=Q_0\prod_{n=1}^{\infty}(1+2q^{2n-1}\cos 2\pi u+q^{4n-2}).$$

(III) Jacobi's Elliptic Functions

(1) We express the modulus k and the complementary modulus as follows.

$$k=\frac{\vartheta_2^2(0)}{\vartheta_3^2(0)},\qquad k'=\frac{\vartheta_0^2(0)}{\vartheta_3^2(0)},\qquad k^2+k'^2=1.$$

Then we have

$$K(k)=K=\frac{\pi}{2}\vartheta_3^2(0),\qquad K'(k)=K'=-i\tau K.$$

The relation between q and k is

$$q=e^{i\pi\tau}=e^{-\pi(K'/K)},$$

$$q^{1/4}=\left(\frac{k}{4}\right)^{1/2}\left[1+2\left(\frac{k}{4}\right)^2+15\left(\frac{k}{4}\right)^4+150\left(\frac{k}{4}\right)^6+1707\left(\frac{k}{4}\right)^8+\dots\right],$$

$$q=\frac{1}{2}L+\frac{2}{2^5}L^5+\frac{15}{2^9}L^9+\frac{150}{2^{13}}L^{13}+\frac{1707}{2^{17}}L^{17}+\dots,\quad \text{where}\quad L=\frac{1-\sqrt[4]{1-k^2}}{1+\sqrt[4]{1-k^2}}.$$

(2) Functions sn, cn, dn; Addition Theorem.

$$\operatorname{sn}(u,k)\equiv\frac{1}{\sqrt{k}}\frac{\vartheta_1(u/2K)}{\vartheta_0(u/2K)},\qquad \operatorname{cn}(u,k)\equiv\sqrt{\frac{k'}{k}}\frac{\vartheta_2(u/2K)}{\vartheta_0(u/2K)},\qquad \operatorname{dn}(u,k)\equiv\sqrt{k'}\frac{\vartheta_3(u/2K)}{\vartheta_0(u/2K)}.$$

$$\operatorname{sn}^2 u+\operatorname{cn}^2 u=1,\qquad \operatorname{dn}^2 u+k^2\operatorname{sn}^2 u=1.$$

$$\operatorname{sn}(u+v)=\frac{\operatorname{sn}u\operatorname{cn}v\operatorname{dn}v+\operatorname{sn}v\operatorname{cn}u\operatorname{dn}u}{1-k^2\operatorname{sn}^2u\operatorname{sn}^2v},\qquad \operatorname{cn}(u+v)=\frac{\operatorname{cn}u\operatorname{cn}v-\operatorname{sn}u\operatorname{dn}u\operatorname{sn}v\operatorname{dn}v}{1-k^2\operatorname{sn}^2u\operatorname{sn}^2v},$$

$$\operatorname{dn}(u+v)=\frac{\operatorname{dn}u\operatorname{dn}v-k^2\operatorname{sn}u\operatorname{cn}u\operatorname{sn}v\operatorname{cn}v}{1-k^2\operatorname{sn}^2u\operatorname{sn}^2v}.$$

$$\frac{d\operatorname{sn}u}{du}=\operatorname{cn}u\operatorname{dn}u,\qquad \frac{d\operatorname{cn}u}{du}=-\operatorname{sn}u\operatorname{dn}u,\qquad \frac{d\operatorname{dn}u}{du}=-k^2\operatorname{sn}u\operatorname{cn}u.$$

(3) Periodicity. In the next table, m and n are integers.

Increment of u	$\operatorname{sn} u$	$\operatorname{cn} u$	$\operatorname{dn} u$
$2mK+2niK'$	$(-1)^m \operatorname{sn} u$	$(-1)^{m+n} \operatorname{cn} u$	$(-1)^n \operatorname{dn} u$
$(2m-1)K+2niK'$	$(-1)^{m+1}\dfrac{\operatorname{cn} u}{\operatorname{dn} u}$	$(-1)^{m+n}k'\dfrac{\operatorname{sn} u}{\operatorname{dn} u}$	$(-1)^n k'\dfrac{1}{\operatorname{dn} u}$
$2mK+(2n+1)iK'$	$(-1)^m k^{-1}\dfrac{1}{\operatorname{sn} u}$	$(-1)^{m+n+1} i k^{-1}\dfrac{\operatorname{dn} u}{\operatorname{sn} u}$	$i(-1)^{n+1}\dfrac{\operatorname{cn} u}{\operatorname{sn} u}$
$(2m-1)K+(2n+1)iK'$	$(-1)^{m+1}k^{-1}\dfrac{\operatorname{dn} u}{\operatorname{cn} u}$	$(-1)^{m+n} i k' k^{-1}\dfrac{1}{\operatorname{cn} u}$	$(-1)^n i k'\dfrac{\operatorname{sn} u}{\operatorname{cn} u}$
Zeros $u=$	$2nK+2miK'$	$(2n+1)K+2miK'$	$(2n+1)K+(2m+1)iK'$
Poles $u=$	$2nK+(2m+1)iK'$	$2nK+(2m+1)iK'$	$2nK+(2m+1)iK'$
Fundamental periods	$4K,\quad 2iK'$	$4K,\quad 2K+2iK'$	$2K,\quad 4iK'$

(4) Change of Variables. In the next table, the second column, for example, means the relation $\operatorname{sn}(ku, 1/k) = k\operatorname{sn}(u,k)$.

u	k	sn	cn	dn
ku	$1/k$	$k\operatorname{sn}$	dn	cn
iu	k'	$i\dfrac{\operatorname{sn}}{\operatorname{cn}}$	$\dfrac{1}{\operatorname{cn}}$	$\dfrac{\operatorname{dn}}{\operatorname{cn}}$
$k'u$	$i\dfrac{k}{k'}$	$k'\dfrac{\operatorname{sn}}{\operatorname{dn}}$	$\dfrac{\operatorname{cn}}{\operatorname{dn}}$	$\dfrac{1}{\operatorname{dn}}$
iku	$i\dfrac{k'}{k}$	$ik\dfrac{\operatorname{sn}}{\operatorname{dn}}$	$\dfrac{1}{\operatorname{dn}}$	$\dfrac{\operatorname{cn}}{\operatorname{dn}}$
$ik'u$	$\dfrac{1}{k'}$	$ik'\dfrac{\operatorname{sn}}{\operatorname{cn}}$	$\dfrac{\operatorname{dn}}{\operatorname{cn}}$	$\dfrac{1}{\operatorname{cn}}$
$(1+k)u$	$\dfrac{2\sqrt{k}}{1+k}$	$\dfrac{(1+k)\operatorname{sn}}{1+k\operatorname{sn}^2}$	$\dfrac{\operatorname{cn}\operatorname{dn}}{1+k\operatorname{sn}^2}$	$\dfrac{1-k\operatorname{sn}^2}{1+k\operatorname{sn}^2}$ (Gauss's transformation)
$(1+k')u$	$\dfrac{1-k'}{1+k'}$	$(1+k')\dfrac{\operatorname{sn}\operatorname{cn}}{\operatorname{dn}}$	$\dfrac{1-(1+k')\operatorname{sn}^2}{\operatorname{dn}}$	$1-(1-k')\operatorname{sn}^2$ (Landen's transformation)
$\dfrac{(1+k')^2 u}{2}$	$\left(\dfrac{1-\sqrt{k'}}{1+\sqrt{k'}}\right)^2$	$\dfrac{1+\sqrt{k'}}{1-\sqrt{k'}}\dfrac{k^2\operatorname{sn}\operatorname{cn}}{(1+\operatorname{dn})(k'+\operatorname{dn})}$	$\dfrac{\operatorname{dn}-\sqrt{k'}}{1-\sqrt{k'}}\times\sqrt{\dfrac{2(1+k')}{(1+\operatorname{dn})(k'+\operatorname{dn})}}$	$\dfrac{\sqrt{2(1+k')}}{1+\sqrt{k'}}\dfrac{\operatorname{dn}+\sqrt{k'}}{\sqrt{(1+\operatorname{dn})(k'+\operatorname{dn})}}$

Jacobi's transformation. $\operatorname{sn}(iu,k) = i\dfrac{\operatorname{sn}(u,k')}{\operatorname{cn}(u,k')}$, $\quad \operatorname{cn}(iu,k) = \dfrac{1}{\operatorname{cn}(u,k')}$,

$$\operatorname{dn}(iu,k) = \frac{\operatorname{dn}(u,k')}{\operatorname{cn}(u,k')}.$$

(5) Amplitude.
The inverse function $\varphi = \operatorname{am}(u,k)$ of $u(k,\varphi) \equiv F(k,\varphi) = \displaystyle\int_0^{\varphi}\frac{d\psi}{\sqrt{1-k^2\sin^2\psi}}$

is called the amplitude.

$\mathrm{sn}(u,k)=\sin\varphi=\sin\mathrm{am}(u,k),\qquad \mathrm{cn}(u,k)=\cos\varphi=\cos\mathrm{am}(u,k),$

$\mathrm{dn}(u,k)=\sqrt{1-k^2\sin^2\varphi}=\sqrt{1-k^2\mathrm{sn}^2(u,k)}\ .$

$$u(k,x)=\int_0^x\frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}},\qquad x=\mathrm{sn}(u,k).$$

$$\mathrm{am}(u,k)=\frac{\pi u}{2K}+\sum_{n=1}^{\infty}\frac{2q^n}{n(1+q^{2n})}\sin\left(n\pi\frac{u}{2K}\right)\qquad (q=e^{i\pi\tau}=e^{-\pi(K'/K)}).$$

$\mathrm{am}(\theta,1)=\mathrm{gd}\,\theta$ (Gudermann function).

(IV) Weierstrass's Elliptic Functions

(1) Weierstrass's $\wp$-function. For the fundamental periods $2\omega_1$, $2\omega_3$, we have

$$\wp(u)\equiv\frac{1}{u^2}+\sum_{n,m}{}'\left[\frac{1}{(u-2n\omega_1-2m\omega_3)^2}-\frac{1}{(2n\omega_1+2m\omega_3)^2}\right]$$

$$=\frac{1}{u^2}+\frac{g_2}{20}u^2+\frac{g_3}{28}u^4+\frac{g_2^2}{1200}u^6+\frac{3g_2g_3}{6160}u^8+\dots.$$

$$g_2\equiv 60\sum_{n,m}{}'\frac{1}{(2n\omega_1+2m\omega_3)^4},\qquad g_3\equiv 140\sum_{n,m}{}'\frac{1}{(2n\omega_1+12m\omega_3)^6},$$

where Σ' means the sum over all integers except $m=n=0$.

$\wp(-u)=\wp(u)$. Putting $\omega_2\equiv-(\omega_1+\omega_3)$, $e_j\equiv\wp(\omega_j)$ $(j=1,2,3)$ we have

$$e_1+e_2+e_3=0,\qquad e_1e_2+e_2e_3+e_3e_1=-g_2/4,\qquad e_1e_2e_3=g_3/4.$$

$$\wp'(u)\equiv d\wp/du=-2\sum_{m,n}\frac{1}{(u-2n\omega_1-2m\omega_3)^3}.$$

$$\wp'^2(u)=4[\wp(u)-e_1][\wp(u)-e_2][\wp(u)-e_3]=4\wp^3(u)-g_2\wp(u)-g_3.$$

Addition theorem

$$\wp(u+v)=-\wp(u)-\wp(v)+\frac{1}{4}\left[\frac{\wp'(u)-\wp'(v)}{\wp(u)-\wp(v)}\right]^2,$$

$$\wp(u+\omega_j)=e_j+\frac{(e_j-e_k)(e_j-e_l)}{\wp(u)-e_j}\qquad (j,k,l)=(1,2,3).$$

Using theta functions corresponding to $\tau=\omega_3/\omega_1$,

$$\wp(u)=-\frac{\eta_1}{\omega_1}-\frac{d^2\log\vartheta_1(u/2\omega_1)}{du^2}\qquad\left(\eta_1=\zeta(\omega_1)=-\frac{1}{12\omega_1}\frac{\vartheta_1'''(0)}{\vartheta_1'(0)}\right),$$

$$\wp'(u)=-\frac{1}{4\omega_1^3}\frac{\vartheta_1'^3(0)\vartheta_2(u/2\omega_1)\vartheta_3(u/2\omega_1)\vartheta_0(u/2\omega_1)}{\vartheta_2(0)\vartheta_0(0)\vartheta_1^3(u/2\omega_1)}.$$

The relations to Jacobi's elliptic functions are

$q\equiv\exp(i\pi\omega_3/\omega_1)$.

$$\wp\left(\frac{u}{\sqrt{e_1-e_3}}\right)=e_1+(e_1-e_3)\frac{\mathrm{cn}^2u}{\mathrm{sn}^2u}=e_2+(e_1-e_3)\frac{\mathrm{dn}^2u}{\mathrm{sn}^2u}=e_3+(e_1-e_3)\frac{1}{\mathrm{sn}^2u},$$

where the modulus is $k=\sqrt{\dfrac{e_2-e_3}{e_1-e_3}}$, $\qquad K(k)=\omega_1\sqrt{e_1-e_3}$.

(2) ζ-function.

$$\zeta(u)\equiv\frac{1}{u}+\sum_{n,m}{}'\left[\frac{1}{u-2n\omega_1-2m\omega_3}+\frac{u}{(2n\omega_1+2m\omega_3)^2}+\frac{1}{2n\omega_1+2m\omega_3}\right]$$

$$=\frac{1}{u}-\frac{g_2}{60}u^3-\frac{g_3}{140}u^5-\frac{g_2^2}{8400}u^7-\frac{g_2g_3}{18480}u^9-\dots$$

$$=(\zeta_1/\omega_1)u+d\log\vartheta_1(u/2\omega_1)/du.$$

$\zeta'(u)=-\wp(u)$.

Pseudoperiodicity. Putting $\eta_j \equiv \zeta(\omega_j)$ $(j=1,2,3)$ we have

$$\zeta(u+2n\omega_1+2m\omega_3)=\zeta(u)+2n\eta_1+2m\eta_3 \qquad (n,m=0,\pm1,\pm2,\dots),$$

$$\eta_1=-\frac{1}{12\omega_1}\frac{\vartheta_1'''(0)}{\vartheta_1'(0)}, \qquad \eta_1+\eta_2+\eta_3=0,$$

$$\eta_1\omega_3-\eta_3\omega_1=\eta_2\omega_1-\eta_1\omega_2=\eta_3\omega_2-\eta_2\omega_3=\pi i/2 \text{ (Legendre's relation).}$$

Addition theorem $\zeta(u+v)=\zeta(u)+\zeta(v)+\dfrac{1}{2}\dfrac{\zeta''(u)-\zeta''(v)}{\zeta'(u)-\zeta'(v)}$.

(3) σ-function.

$$\sigma(u)\equiv u\prod_{n,m}{}'\left(1-\frac{u}{2n\omega_1+2m\omega_3}\right)\exp\left[\frac{u}{2n\omega_1+2m\omega_3}+\frac{1}{2}\left(\frac{u}{2n\omega_1+2m\omega_3}\right)^2\right] \qquad \left[\begin{matrix}n,m=0,\pm1,\\ \pm2,\dots,\\ (n,m)\neq(0,0)\end{matrix}\right]$$

$$=u-\frac{g_2}{2^4\cdot3\cdot5}u^5-\frac{g_3}{2^3\cdot3\cdot5\cdot7}u^7-\frac{g_2^2}{2^9\cdot3^2\cdot5\cdot7}u^9-\dots$$

$$=2\omega_1\left(\exp\frac{\eta_1u^2}{2\omega_1}\right)\frac{\vartheta_1(u/2\omega_1)}{\vartheta_1'(0)}.$$

$$\zeta(u)=\sigma'(u)/\sigma(u). \qquad \sigma(-u)=-\sigma(u).$$

Pseudoperiodicity. $\sigma(u+2n\omega_1+2m\omega_3)=(-1)^{n+m+mn}[\exp(2n\eta_1+2m\eta_3)(u+n\omega_1+m\omega_3)]\sigma(u)$.

(4) Cosigma functions σ_1, σ_2, σ_3.

$$\sigma_j(u)\equiv -e^{\eta_j u}\frac{\sigma(u+\omega_j)}{\sigma(\omega_j)}=\left(\exp\frac{\eta_1u^2}{2\omega_1}\right)\frac{\vartheta_{j+1}(u/2\omega_1)}{\vartheta_{j+1}(0)} \qquad (j=1,2,3;\ \vartheta_4\equiv\vartheta_0).$$

$$\wp(u)-e_j=\left[\frac{\sigma_j(u)}{\sigma(u)}\right]^2, \qquad \wp'(2u)=-\frac{2\sigma_1(u)\sigma_2(u)\sigma_3(u)}{\sigma^3(u)}=-\frac{\sigma(2u)}{\sigma^4(u)}.$$

$$\operatorname{sn}u=\alpha\frac{\sigma(u/\alpha)}{\sigma_3(u/\sigma)}, \qquad \operatorname{cn}u=\frac{\sigma_1(u/\alpha)}{\sigma_3(u/\alpha)}, \qquad \operatorname{dn}u=\frac{\sigma_2(u/\alpha)}{\sigma_3(u/\alpha)}, \qquad \text{where } \alpha=\sqrt{e_1-e_3}=\frac{K}{\omega_1}.$$

References

[1] W. F. Magnus, F. Oberhettinger, and R. P. Soni, Formulas and theorems for the special functions of mathematical physics, Springer, third enlarged edition, 1966.
[2] Y. L. Luke, The special functions and their approximations I, II, Academic Press, 1969.
[3] A. Erdelyi, Higher transcendental functions I, II, III (Bateman manuscript project) McGraw-Hill, 1953, 1955.
In particular, for hypergeometric functions of two variables see
[4] P. Appell, Sur les fonctions hypergéométriques de plusieurs variables, Mémor. Sci. Math., Gauthier-Villars, 1925.
Also → references to 39 Bessel Functions, 134 Elliptic Functions, 174 Gamma Function, 389 Special Functions.

17. Gamma Functions and Related Functions

(I) Gamma Functions and Beta Functions (→ 174 Gamma Function)

In this Section (I), C means Euler's constant, B_n means a Bernoulli number, ζ means the Riemann zeta function.

(1) Gamma function. $\Gamma(z)\equiv\displaystyle\int_0^\infty e^{-t}t^{z-1}\,dt \qquad (\operatorname{Re}z>0)$

$$=\frac{1}{e^{2\pi iz}-1}\int_\infty^{(0+)}e^{-t}t^{z-1}\,dt.$$

In the last integral, the integration contour goes once around the positive real axis in the positive direction.

$$\Gamma(n+1)=n! \qquad (n=0,1,2,\ldots), \qquad \Gamma(1/2)=\sqrt{\pi}.$$

$$\Gamma(z+1)=z\Gamma(z), \qquad \Gamma(z)\Gamma(1-z)=\frac{\pi}{\sin\pi z}, \qquad \prod_{j=0}^{n-1}\Gamma\left(z+\frac{j}{n}\right)=(2\pi)^{(n-1)/2}n^{(1/2)-nz}\Gamma(nz),$$

$$\frac{1}{\Gamma(z)}=ze^{Cz}\prod_{n=1}^{\infty}\left(1+\frac{z}{n}\right)e^{-z/n}, \qquad \log\Gamma(1+z)=-\frac{1}{2}\log\frac{\sin\pi z}{\pi z}-Cz-\sum_{n=1}^{\infty}\frac{\zeta(2n+1)}{2n+1}z^{2n+1}$$
$$(|z|<1).$$

$$\left|\frac{\Gamma(x+iy)}{\Gamma(x)}\right|^2=\prod_{n=0}^{\infty}1\Big/\left(1+\frac{y^2}{(x+n)^2}\right) \qquad (x,y \text{ are real and } x>0).$$

Asymptotic expansion (Stirling formula).

$$\Gamma(z)\approx e^{-z}z^{z-1}\sqrt{2\pi z}\,\exp\left[\sum_{n=1}^{\infty}\frac{(-1)^{n-1}B_{2n}z^{1-2n}}{2n(2n-1)}\right] \qquad (|\arg z|<\pi)$$

$$=e^{-z}z^{z-(1/2)}\sqrt{2\pi}\left[1+\frac{1}{12z}+\frac{1}{288z^2}-\frac{139}{51840z^3}-\frac{571}{2488320z^4}+O(z^{-5})\right].$$

(2) Beta Function. $B(x,y)\equiv\int_0^1 t^{x-1}(1-t)^{y-1}dt \qquad (\mathrm{Re}\,x, \quad \mathrm{Re}\,y>0)$

$$=\Gamma(x)\Gamma(y)/\Gamma(x+y).$$

(3) Incomplete Gamma Function.

$$\gamma(\nu,x)\equiv\int_0^x t^{\nu-1}e^{-t}dt=\Gamma(\nu)-x^{(\nu-1)/2}e^{-x/2}W_{(\nu-1)/2,\nu/2}(x) \qquad (\mathrm{Re}\,\nu>0).$$

(4) Incomplete Beta Function. $B_\alpha(x,y)\equiv\int_0^\alpha t^{x-1}(1-t)^{y-1}dt \qquad (0<\alpha\leqslant 1).$

(5) Polygamma Functions. $\psi(z)\equiv\frac{d}{dz}\log\Gamma(z)$

$$=\frac{\Gamma'(z)}{\Gamma(z)}=\int_0^\infty\left[\frac{e^{-t}}{t}-\frac{e^{-zt}}{1-e^{-t}}\right]dt=-C+\sum_{n=0}^{\infty}\left(\frac{1}{n+1}-\frac{1}{z+n}\right).$$

$$\psi'(z)=\sum_{n=0}^{\infty}\frac{1}{(z+n)^2}, \qquad \psi^{(k)}(z)=\sum_{n=0}^{\infty}\frac{(-1)^{k+1}k!}{(z+n)^{k+1}} \qquad (k=1,2,\ldots).$$

(II) Combinatorial Problems (→ 330 Permutations and Combinations)

Factorial $n!=n(n-1)(n-2)\ldots3\cdot2\cdot1. \quad 0!=1.$

Binomial coefficient $\binom{\alpha}{r}=\frac{\alpha(\alpha-1)\ldots(\alpha-r+1)}{r!}.$

(1) Number of Permutations of n Elements Taken r at a Time.

$${}_nP_r=n(n-1)\ldots(n-r+1)=n!/(n-r)!.$$

Number of combinations of n elements taken r at a time

$${}_nC_r=\frac{{}_nP_r}{r!}=\frac{n!}{r!(n-r)!}=\binom{n}{r}.$$

$${}_nC_r={}_nC_{n-r}, \qquad {}_nC_r={}_{n-1}C_r+{}_{n-1}C_{r-1}.$$

Number of multiple permutations $\quad {}_n\Pi_r=n^r.$

Number of multiple combinations $\quad {}_nH_r={}_{n+r-1}C_r=\frac{(n+r-1)!}{r!(n-1)!}.$

Number of circular permutations $\quad {}_nP_r/r.$

(2) Binomial Theorem. $(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r.$

Multinomial theorem $(a_1+\ldots+a_m)^n = \sum \frac{n!}{p_1!\ldots p_m!} a_1^{p_1}\ldots a_m^{p_m}.$

The latter summation runs over all nonnegative integers satisfying $p_1+\ldots+p_m=n$.

References

See references to Table 16, this Appendix.

18. Hypergeometric Functions and Spherical Functions

(I) Hypergeometric Function (→ 206 Hypergeometric Functions)

(1) Hypergeometric Function. $F(a,b;c;z) \equiv \sum_{n=0}^{\infty} \frac{\Gamma(a+n)}{\Gamma(a)} \frac{\Gamma(b+n)}{\Gamma(b)} \frac{\Gamma(c)}{\Gamma(c+n)} \frac{z^n}{n!}.$

The fundamental system of solutions of the hypergeometric differential equation $z(1-z)\frac{d^2u}{dz^2}+[c-(a+b+1)z]\frac{du}{dz}-abu=0$ at $z=0$ is given by

$$u_1=F(a,b;c;z), \qquad u_2=z^{1-c}F(a-c+1,\, b-c+1;\, 2-c;\, z) \qquad (c\neq 0,-1,-2,\ldots).$$

$$F(a,b;c;z)=F(b,a;c;z). \qquad dF/dz=(ab/c)F(a+1,\, b+1;\, c+1;\, z).$$

$$F(a,b;c;1)=\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \qquad (\mathrm{Re}(a+b-c)<0).$$

$$F(a,b;c;z)=\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\,dt \qquad (\mathrm{Re}\,c>\mathrm{Re}\,b>0, \quad |z|<1),$$

$$F(a,b;c;z)=\frac{1}{2\pi i}\frac{\Gamma(c)}{\Gamma(a)\Gamma(b)}\int_{-i\infty}^{i\infty}\frac{\Gamma(a+s)\Gamma(b+s)\Gamma(-s)}{\Gamma(c+s)}(-z)^s\,ds.$$

(2) Transformations of the Hypergeometric Function.

$$\begin{aligned}
F(a,b;\, c;\, z)&=(1-z)^{-a}F\left(a,c-b;\, c;\, \frac{z}{z-1}\right)\\
&=(1-z)^{c-a-b}F(c-a,c-b;\, c;\, z)\\
&=(1-z)^{-a}\frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)}F\left(a,c-b;\, a-b+1;\, \frac{1}{1-z}\right)\\
&\quad+(1-z)^{-b}\frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)}F\left(b,c-a;\, b-a+1;\, \frac{1}{1-z}\right)\\
&=\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}F(a,b;\, a+b-c+1;\, 1-z)\\
&\quad+(1-z)^{c-a-b}\frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}F(c-a,c-b;\, c-a-b+1;\, 1-z)\\
&=\frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)}(-z)^{-a}F\left(a,1-c+a;\, 1-b+a;\, \frac{1}{z}\right)\\
&\quad+\frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)}(-z)^{-b}F\left(b,1-c+b;\, 1-a+b;\, \frac{1}{z}\right).
\end{aligned}$$

(3) Riemann's Differential Equation (→ Table 14.II, this Appendix).

$$\frac{d^2u}{dz^2}+\left[\frac{1-\alpha-\alpha'}{z-a}+\frac{1-\beta-\beta'}{z-b}+\frac{1-\gamma-\gamma'}{z-c}\right]\frac{du}{dz}$$

$$+\left[\frac{\alpha\alpha'(a-b)(a-c)}{z-a}+\frac{\beta\beta'(b-c)(b-a)}{z-b}+\frac{\gamma\gamma'(c-a)(c-b)}{z-c}\right]\frac{u}{(z-a)(z-b)(z-c)}=0.$$

Here we have $\alpha+\alpha'+\beta+\beta'+\gamma+\gamma'=1$ (Fuchsian relation). The solution of this equation is given by Riemann's P-function

$$u=P\left\{\begin{matrix} a & b & c & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{matrix}\right\}$$

$$=\left(\frac{z-a}{z-b}\right)^{\alpha}\left(\frac{z-c}{z-b}\right)^{\gamma}F\left(\alpha+\beta+\gamma,\alpha+\beta'+\gamma;\ 1+\alpha-\alpha';\ \frac{(c-b)(z-a)}{(c-a)(z-b)}\right)$$

$(\alpha-\alpha',\beta-\beta',\gamma-\gamma'\neq\text{integer}).$

We have 24 representations of the above function by interchanging the parameters a, b, c; α, α'; β, β'; γ, γ' in the right-hand side.

(4) Barnes's Extended Hypergeometric Function.

$${}_pF_q(\alpha_1,\ldots,\alpha_p;\beta_1,\ldots,\beta_q;z)\equiv\sum_{n=0}^{\infty}\frac{(\alpha_1)_n\cdots(\alpha_p)_n}{(\beta_1)_n\cdots(\beta_q)_n}\frac{z^n}{n!},\qquad\text{where}\quad(\alpha)_n=\alpha(\alpha+1)\ldots(\alpha+n-1)$$

$=\Gamma(\alpha+n)/\Gamma(\alpha)$. $\quad F(a,b;c;z)={}_2F_1(a,b;c;z)$. $\quad {}_0F_0(x)=e^x$, $\quad {}_1F_0(\alpha;x)=(1-x)^{-\alpha}$.

(5) Appell's Hypergeometric Functions of Two Variables.

$$F_1(\alpha;\beta,\beta';\gamma;x,y)=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\alpha)_{m+n}(\beta)_m(\beta')_n}{m!n!(\gamma)_{m+n}}x^m y^n,$$

$$F_2(\alpha;\beta,\beta';\gamma,\gamma';x,y)=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\alpha)_{m+n}(\beta)_m(\beta')_n}{m!n!(\gamma)_m(\gamma')_n}x^m y^n,$$

$$F_3(\alpha,\alpha';\beta,\beta';\gamma;x,y)=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\alpha)_m(\alpha')_n(\beta)_m(\beta')_n}{m!n!(\gamma)_{m+n}}x^m y^n,$$

$$F_4(\alpha;\beta;\gamma,\gamma';x,y)=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\alpha)_{m+n}(\beta)_{m+n}}{m!n!(\gamma)_m(\gamma')_n}x^m y^n.$$

(6) Representation of Various Special Functions by Hypergeometric Functions.

$$(1-x)^{\nu}=F(-\nu,b;b;x),\qquad e^{-nx}=\left(\frac{\operatorname{sech}x}{2}\right)^n(\tanh x)F\left(1+\frac{n}{2},\frac{1+n}{2};\ 1+n;\ \operatorname{sech}^2x\right).$$

$$\log(1+x)=xF(1,1;\ 2;\ -x),\qquad \frac{1}{2}\log\frac{1+x}{1-x}=xF\left(\frac{1}{2},1;\ \frac{3}{2};\ x^2\right),$$

$$\sin nx=n(\sin x)F\left(\frac{1+n}{2},\frac{1-n}{2};\ \frac{3}{2};\sin^2x\right),$$

$$\cos nx=F\left(\frac{n}{2},-\frac{n}{2};\ \frac{1}{2};\ \sin^2x\right)=(\cos x)F\left(\frac{1+n}{2},\frac{1-n}{2};\ \frac{1}{2};\ \sin^2x\right),$$

$$\arcsin x=xF\left(\frac{1}{2},\frac{1}{2};\ \frac{3}{2};\ x^2\right),\qquad \arctan x=xF\left(\frac{1}{2},1;\ \frac{3}{2};\ -x^2\right).$$

$$P_{2n}(x)=(-1)^n\frac{(2n-1)!!}{(2n)!!}F\left(-n,n+\frac{1}{2};\ \frac{1}{2};\ x^2\right),$$

$$P_{2n+1}(x)=(-1)^n\frac{(2n+1)!!}{(2n)!!}xF\left(-n,n+\frac{3}{2};\ \frac{3}{2};\ x^2\right)\qquad\text{(spherical function)},$$

where

$$n=0,1,2,\ldots; \qquad m!!=\begin{cases} m(m-2)\ldots4\cdot2 & (m \text{ even}),\\ m(m-2)\ldots3\cdot1 & (m \text{ odd}),\end{cases} \qquad 0!!=(-1)!!=1.$$

$$K(x)=\frac{\pi}{2}F\left(\frac{1}{2},\frac{1}{2};1;x^2\right), \qquad E(x)=\frac{\pi}{2}F\left(-\frac{1}{2},\frac{1}{2};1;x^2\right) \quad \text{(complete elliptic integral).}$$

$$J_\nu(x)=\frac{x}{2\Gamma(\nu+1)}{}_0F_1\left(\nu+1;\,-\frac{x^2}{4}\right)=\frac{x^\nu e^{-ix}}{2^\nu\Gamma(\nu+1)}{}_1F_1\left(\nu+\frac{1}{2};\,2\nu+1;\,2ix\right).$$

$$e^x=\lim_{b\to\infty}F(a,b;a;x/b)={}_1F_1(a;a;x)={}_0F_0(x).$$

(II) Legendre Function (→ 393 Spherical Functions)

(1) Legendre Functions. The generalized spherical function corresponding to the rotation group of 3-dimensional space is the solution of the following differential equation.

$$(1-z^2)\frac{d^2u}{dz^2}-2z\frac{du}{dz}+\left[\nu(\nu+1)-\frac{\mu^2}{1-z^2}\right]u=0.$$

When $\mu=0$, the equation is Legendre's differential equation, and the fundamental system of solutions is given by the following two kind of functions.

Legendre function of the first kind $\quad \mathfrak{P}_\nu(z)\equiv P_\nu(z)\equiv {}_2F_1\left(-\nu,\ \nu+1;\ 1;\ \frac{1-z}{2}\right).$

Legendre function of the second kind

$$\mathfrak{Q}_\nu(z)\equiv\frac{\Gamma(\nu+1)\sqrt{\pi}}{2^{\nu+1}\Gamma[\nu+(3/2)]}z^{-\nu-1}{}_2F_1\left(\frac{\nu+2}{2},\ \frac{\nu+1}{2};\ \nu+\frac{3}{2};\ \frac{1}{z^2}\right).$$

$$Q_\nu(x)\equiv\frac{1}{2}[\mathfrak{Q}_\nu(x+i0)+\mathfrak{Q}_\nu(x-i0)]$$

$$=\pi\frac{(\cos\nu\pi)P_\nu(x)-P_\nu(-x)}{2\sin\nu\pi} \qquad (\nu\neq\text{integer};\ -1<x<1).$$

Recurrence formulas:

$$\mathfrak{P}_\nu(z)=\mathfrak{P}_{-\nu-1}(z). \qquad \mathfrak{Q}_\nu(z)-\mathfrak{Q}_{-\nu-1}(z)=\pi(\cot\nu\pi)\mathfrak{P}_\nu(z) \qquad (\nu\neq\text{integer}).$$

$$\mathfrak{P}_\nu(-z)=e^{\pm\nu\pi i}\mathfrak{P}_\nu(z)-(2/\pi)(\sin\nu\pi)\mathfrak{Q}_\nu(z), \qquad \mathfrak{Q}_\nu(-z)=-e^{\pm\nu\pi i}\mathfrak{Q}_\nu(z) \qquad (\pm=\operatorname{sgn}(\operatorname{Im}z)).$$

$$(z^2-1)d\mathfrak{P}_\nu(z)/dz=(\nu+1)[\mathfrak{P}_{\nu+1}(z)-z\mathfrak{P}_\nu(z)],$$

$$(2\nu+1)z\mathfrak{P}_\nu(z)=(\nu+1)\mathfrak{P}_{\nu+1}(z)+\nu\mathfrak{P}_{\nu-1}(z),$$

$$(z^2-1)d\mathfrak{Q}_\nu(z)/dz=(\nu+1)[\mathfrak{Q}_{\nu+1}(z)-z\mathfrak{Q}_\nu(z)],$$

$$(2\nu+1)z\mathfrak{Q}_\nu(z)=(\nu+1)\mathfrak{Q}_{\nu+1}(z)+\nu\mathfrak{Q}_{\nu-1}(z).$$

$$\mathfrak{P}_\nu(z)=\pi^{-1/2}2^{-\nu-1}\tan\nu\pi\frac{\Gamma(\nu+1)}{\Gamma[\nu+(3/2)]}z^{-\nu-1}{}_2F_1\left(\frac{\nu}{2}+1,\ \frac{\nu+1}{2};\ \nu+\frac{3}{2};\ \frac{1}{z^2}\right)$$

$$+\pi^{-1/2}2^\nu\frac{\Gamma[\nu+(1/2)]}{\Gamma(\nu+1)}z^\nu{}_2F_1\left(\frac{1-\nu}{2},\ \frac{-\nu}{2};\ \frac{1}{2}-\nu;\ \frac{1}{z^2}\right).$$

$$P_\nu(\cos\theta)=\frac{\sin\nu\pi}{\pi}\sum_{n=0}^{\infty}(-1)^n\left(\frac{1}{\nu-n}-\frac{1}{\nu+n+1}\right)P_n(\cos\theta) \qquad (\nu\neq\text{integer};\ 0\leqslant\theta<\pi).$$

Estimation: $\quad |P_\nu(\cos\theta)|\leqslant\dfrac{2}{\sqrt{\nu\pi\sin\theta}}, \qquad |Q_\nu(\cos\theta)|\leqslant\dfrac{\sqrt{\pi}}{\sqrt{\nu\sin\theta}} \qquad (0<\theta<\pi;\ \nu>1).$

$$P_\nu(1)=1, \qquad P_\nu(0)=-\frac{\sin\nu\pi}{2\pi^{3/2}}\Gamma\left(\frac{\nu+1}{2}\right)\Gamma\left(\frac{-\nu}{2}\right),$$

$$Q_\nu(0)=\frac{1}{4\sqrt{\pi}}(1-\cos\nu\pi)\Gamma\left(\frac{\nu+1}{2}\right)\Gamma\left(\frac{-\nu}{2}\right).$$

(2) The Case $\nu=n$ $(=0,1,2,\ldots)$. In the following, the symbol !! means

$$m!!\equiv\begin{cases} m(m-2)\ldots4\cdot2 & (m \text{ even}),\\ m(m-2)\ldots5\cdot3\cdot1 & (m \text{ odd}).\end{cases}$$

The function P_n is a polynomial of degree n (Legendre polynomial) and is represented as follows:

$$P_n(z)\equiv\frac{1}{2^n n!}\frac{d^n}{dz^n}(z^2-1)^n$$

$$=\frac{(2n)!}{2^n(n!)^2}z^n{}_2F_1\left(-\frac{n}{2},\ \frac{1-n}{2};\ \frac{1}{2}-n;\ \frac{1}{z^2}\right)$$

$$=\frac{(2n-1)!!}{n!}\left[z^n-\frac{n(n-1)}{(2n-1)}z^{n-2}+\frac{n(n-1)(n-2)(n-3)}{2\cdot 4\cdot(2n-1)(2n-3)}z^{n-4}\mp\dots\right].$$

$$P_{2m}(z)=\sum_{j=0}^{m}(-1)^{m-j}\frac{(2m+2j-1)!!}{(2j)!(2m-2j)!!}z^{2j},$$

$$P_{2m+1}(z)=\sum_{j=0}^{m}(-1)^{m-j}\frac{(2m+2j+1)!!}{(2j+1)!(2m-2j)!!}z^{2j+1}.$$

$$P_n(\cos\theta)=\frac{(2n)!}{2^{2n}(n!)^2}e^{\mp in\theta}{}_2F_1\left(\frac{1}{2},\ -n;\ \frac{1}{2}-n;\ e^{\pm 2i\theta}\right)$$

$$=\frac{2(2n-1)!!}{(2n)!!}\left[\cos n\theta+\frac{1}{1}\frac{n}{(2n-1)}\cos(n-2)\theta+\frac{1\cdot 3}{1\cdot 2}\frac{n(n-1)}{(2n-1)(2n-3)}\cos(n-4)\theta\right.$$

$$\left.+\frac{1\cdot 3\cdot 5}{1\cdot 2\cdot 3}\frac{n(n-1)(n-2)}{(2n-1)(2n-3)(2n-5)}\cos(n-6)\theta+\dots\right]$$

$$+\begin{cases}\left[\dfrac{(n-1)!!}{n!!}\right]^2 & (n\text{ even}),\\ 0 & (n\text{ odd}).\end{cases}$$

$$=\frac{4}{\pi}\frac{(2n)!!}{(2n+1)!!}\left[\sin(n+1)\theta+\frac{1\cdot(n+1)}{1\cdot(2n+3)}\sin(n+3)\theta\right.$$

$$\left.+\frac{1\cdot 3\cdot(n+1)(n+2)}{1\cdot 2\cdot(2n+3)(2n+5)}\sin(n+5)\theta+\dots(\text{ad infinitum})\right]\qquad(0<\theta<\pi).$$

Laplace-Mehler integral representation

$$P_n(\cos\theta)=\frac{1}{\pi}\int_0^{\pi}(\cos\theta+i\sin\theta\cos\varphi)^n\,d\varphi$$

$$=\frac{\sqrt{2}}{\pi}\int_0^{\theta}\frac{\cos[n+(1/2)]\varphi}{\sqrt{\cos\varphi-\cos\theta}}\,d\varphi=\frac{\sqrt{2}}{\pi}\int_{\theta}^{\pi}\frac{\sin[n+(1/2)]\varphi}{\sqrt{\cos\theta-\cos\varphi}}\,d\varphi.$$

$$P_n(x)=\frac{(-1)^n}{n!}r^{n+1}\frac{\partial^n}{\partial z^n}\left(\frac{1}{r}\right)\qquad\left(x=\frac{z}{r},\ r=\sqrt{z^2+\rho^2}\ \right).$$

$$P_n(1)=1,\qquad P_n(-1)=(-1)^n,\qquad P_{2n+1}(0)=0,$$

$$P_{2n}(0)=(-1)^n\frac{(2n)!}{2^{2n}(n!)^2}=\frac{(-1)^n(2n-1)!!}{(2n)!!}.$$

Recurrence formulas: $nP_n(z)-(2n-1)zP_{n-1}(z)+(n-1)P_{n-2}(z)=0,$

$$(z^2-1)\frac{dP_n}{dz}=n(zP_n-P_{n-1})=\frac{n(n+1)}{2n+1}(P_{n+1}-P_{n-1})=(n+1)(P_{n+1}-zP_n).$$

$$\mathfrak{Q}_n(z)=\frac{1}{2^n n!}\frac{d^n}{dz^n}\left[(z^2-1)^n\log\frac{z+1}{z-1}\right]-\frac{1}{2}P_n(z)\log\frac{z+1}{z-1}$$

$$=2^n n!\int_z^{\infty}\dots\int_z^{\infty}\frac{(dz)^{n+1}}{(z^2-1)^{n+1}}$$

$$=2^n\int_z^{\infty}\frac{(t-z)^n}{(t^2-1)^{n+1}}\,dt$$

$$=(-1)^n\frac{1}{(2n-1)!!}\frac{d^n}{dz^n}\left[(z^2-1)^n\int_z^{\infty}\frac{dt}{(t^2-1)^{n+1}}\right]\qquad(\operatorname{Re}z>1).$$

$$Q_n(\cos\theta)=\frac{2\cdot(2n)!!}{(2n+1)!!}\left[\cos(n+1)\theta+\frac{1\cdot(n+1)}{1\cdot(2n+3)}\cos(n+3)\theta\right.$$
$$\left.+\frac{1\cdot 3\cdot(n+1)(n+2)}{1\cdot 2\cdot(2n+3)(2n+5)}\cos(n+5)\theta+\dots\right]\qquad(0<\theta<\pi).$$

$$Q_n(x)=\frac{1}{2^n n!}\frac{d^n}{dx^n}\left[(x^2-1)^n\log\frac{1+x}{1-x}\right]-\frac{1}{2}P_n(x)\log\frac{1+x}{1-x}$$
$$=\frac{1}{2}P_n(x)\log\frac{1+x}{1-x}-\sum_{j=1}^{n}\frac{1}{j}P_{j-1}(x)P_{n-j}(x).$$

$$Q_0(x)=\frac{1}{2}\log\frac{1+x}{1-x},\qquad Q_1(x)=\frac{x}{2}\log\frac{1+x}{1-x}-1,\qquad Q_2(x)=\frac{1}{4}(3x^2-1)\log\frac{1+x}{1-x}-\frac{3}{2}x.$$

(3) Generating Functions.

$$\frac{1}{\sqrt{1-2hz+h^2}}=\begin{cases}\sum_{n=0}^{\infty}h^nP_n(z) & \left(|h|<\min|z\pm\sqrt{z^2-1}\,|\right),\\ \sum_{n=0}^{\infty}\frac{1}{h^{n+1}}P_n(z) & \left(|h|>\max|z\pm\sqrt{z^2-1}\,|\right).\end{cases}$$

(If $-1\leqslant z\leqslant 1$, the right-hand side is equal to 1.)

$$\frac{1}{z-t}=\sum_{n=0}^{\infty}(2n+1)\mathfrak{P}_n(t)\mathfrak{Q}_n(z)\qquad\left(|t+\sqrt{t^2-1}\,|<|z+\sqrt{z^2-1}\,|\right).$$

$$\frac{1}{\sqrt{1-2tz+z^2}}\log\frac{z-t+\sqrt{1-2tz+z^2}}{\sqrt{z^2-1}}=\sum_{n=0}^{\infty}t^n\mathfrak{Q}_n(z)\qquad(\mathrm{Re}\,z>1,\quad |t|<1).$$

$$r=\sqrt{x^2+y^2+z^2}\,,\qquad \cos\theta=z/r,\qquad x,y\ \text{real},$$

$$\frac{1}{r}+\frac{1}{2}+\sum_{m=1}^{\infty}\left[\frac{1}{\sqrt{(2mi\pi+z)^2+x^2+y^2}}+\frac{1}{\sqrt{(2mi\pi-z)^2+x^2+y^2}}\right]$$

(Here the square root of a complex number is taken so that its real part is positive.)

$$=\begin{cases}1+\sum_{n=1}^{\infty}e^{-nz}J_0\left(n\sqrt{x^2+y^2}\,\right) & (\mathrm{Re}\,z>0),\\ \frac{1}{r}+\frac{1}{2}+\sum_{n=1}^{\infty}\frac{(-1)^nB_{2n}}{(2n)!}r^{2n-1}P_{2n-1}(\cos\theta) & (0<\theta<2\pi;\ z\ \text{real}).\end{cases}$$

(4) Integrals of Legendre Polynomials.

Orthogonal relations: $\int_{-1}^{+1}P_n(z)P_m(z)\,dz=\delta_{nm}\frac{2}{2n+1}$.

$$\int_{-1}^{+1}z^kP_n(z)\,dz=0\qquad(k=0,1,\dots,n-1).$$

$$\int_0^1 z^\lambda P_n(z)\,dz=\begin{cases}\frac{\lambda(\lambda-2)\dots(\lambda-n+2)}{(\lambda+n+1)(\lambda+n-1)\dots(\lambda+1)} & (n\ \text{even}),\\ \frac{(\lambda-1)(\lambda-3)\dots(\lambda-n+2)}{(\lambda+n+1)(\lambda+n-1)\dots(\lambda+2)} & (n\ \text{odd})\qquad(\mathrm{Re}\,\lambda>-1).\end{cases}$$

$$\int_0^\pi P_n(\cos\theta)\sin m\theta\,d\theta=\begin{cases}\frac{2(m-n+1)(m-n+3)\dots(m+n-1)}{(m-n)(m-n+2)\dots(m+n)} & (m>n;\ m+n\ \text{is odd}),\\ 0 & (\text{otherwise}).\end{cases}$$

(5) Conical Function (Kegelfunktion). This is the Legendre function corresponding to the case $\nu=-(1/2)+i\lambda$ (λ is a real parameter),

$$P_{-(1/2)+i\lambda}(\cos\theta)=1+\frac{4\lambda^2+1^2}{2^2}\sin^2\frac{\theta}{2}+\frac{(4\lambda^2+1^2)(4\lambda^2+3^2)}{2^2\cdot 4^2}\sin^4\frac{\theta}{2}+\dots.$$

$$P_{-(1/2)+i\lambda}(x)\equiv P_{-(1/2)-i\lambda}(x).$$

(III) Associated Legendre Functions (→ 393 Spherical Functions)

(1) Associated Legendre Functions. The fundamental system of solutions of the differential equation in (II) (1) is given by the following two kind of functions when $\mu\neq 0$.

Associated Legendre function of the first kind:

$$\mathfrak{P}_\nu^\mu(z)\equiv\frac{1}{\Gamma(1-\mu)}\left(\frac{z+1}{z-1}\right)^{\mu/2}{}_2F_1\left(-\nu,\nu+1;1-\mu;\frac{1-z}{2}\right),$$

where we take the branch satisfying $\arg[(z+1)/(z-1)]^{\mu/2}=0$ for $z>1$ in the expression raised to the $(\mu/2)$th power.

Associated Legendre function of the second kind:

$$\mathfrak{Q}_\nu^\mu(z)\equiv\frac{e^{i\mu\pi}}{2^{\nu+1}}\frac{\Gamma(\nu+\mu+1)\sqrt{\pi}}{\Gamma[\nu+(3/2)]}(z^2-1)^{\mu/2}z^{-\nu-\mu-1}{}_2F_1\left(\frac{\nu+\mu+2}{2},\frac{\nu+\mu+1}{2};\nu+\frac{3}{2};\frac{1}{z^2}\right),$$

where we take the branch satisfying $\arg(z^2-1)^{\mu/2}=0$ for $z>1$ in $(z^2-1)^{\mu/2}$, and $\arg z^{-\nu-\mu-1}=0$ for $z>0$ in $z^{-\nu-\mu-1}$, respectively.

$$P_\nu^\mu(x)\equiv e^{i\mu\pi/2}\mathfrak{P}_\nu^\mu(x+i0)=e^{i\mu\pi/2}\mathfrak{P}_\nu^\mu(x-i0)$$

$$=\frac{1}{\Gamma(1-\mu)}\left(\frac{1+x}{1-x}\right)^{\mu/2}{}_2F_1\left(-\nu,\nu+1;1-\mu;\frac{1-x}{2}\right)\qquad(-1\leqslant x\leqslant 1).$$

$$Q_\nu^\mu(x)\equiv e^{-i\mu\pi}\left[e^{-i\mu\pi/2}\mathfrak{Q}_\nu^\mu(x+i0)+e^{i\mu\pi/2}\mathfrak{Q}_\nu^\mu(x-i0)\right]/2$$

$$=\frac{\pi}{2\sin\mu\pi}\left[P_\nu^\mu(x)\cos\mu\pi-\frac{\Gamma(\nu+\mu+1)}{\Gamma(\nu-\mu+1)}P_\nu^{-\mu}(x)\right]\qquad(-1\leqslant x\leqslant 1).$$

Integral representations:

$$\mathfrak{P}_\nu^{-\mu}(z)=\frac{(z^2-1)^{\mu/2}}{2\mu\sqrt{\pi}\,\Gamma[\mu+(1/2)]}\int_{-1}^{+1}\frac{(1-t^2)^{\mu-1/2}}{\left(z+t\sqrt{z^2-1}\right)^{\mu-\nu}}dt$$

$$\left(\operatorname{Re}\mu>-\frac{1}{2},\ |\arg(z\pm 1)|<\pi\right).$$

$$\mathfrak{P}_\nu^{-\mu}(z)=\frac{(z^2-1)^{\mu-2}}{2^\nu\Gamma(\mu-\nu)\Gamma(\nu+1)}\int_0^\infty\frac{(\sinh t)^{2\nu+1}}{(z+\cosh t)^{\mu+\nu+1}}dt$$

$$(\operatorname{Re}z>-1,\ |\arg(z\pm 1)|<\pi,\ \operatorname{Re}\nu>-1,\ \operatorname{Re}(\mu-\nu)>0).$$

$$\mathfrak{P}_\nu^{-\mu}(z)=\sqrt{\frac{2}{\pi}}\frac{\Gamma[\mu+(1/2)](z^2-1)^{\mu-2}}{\Gamma(\mu+\nu+1)\Gamma(\mu-\nu)}\int_0^\infty\frac{\cosh[\nu+(1/2)]t}{(z+\cosh t)^{\mu+(1/2)}}dt$$

$$(\operatorname{Re}z>-1,\ |\arg(z\pm 1)|<\pi,\ \operatorname{Re}(\mu+\nu)>-1,\ \operatorname{Re}(\mu-\nu)>0).$$

$$\mathfrak{P}_\nu^\mu(\cosh\alpha)=\sqrt{\frac{2}{\pi}}\frac{(\sinh\alpha)^\mu}{\Gamma[-\mu+(1/2)]}\int_0^\alpha\frac{\cosh[\{\nu+(1/2)\}t]\,dt}{(\cosh\alpha-\cosh t)^{\mu+(1/2)}}\qquad\left(\alpha>0,\ \operatorname{Re}\mu<\frac{1}{2}\right).$$

$$P_\nu^\mu(\cos\theta)=\sqrt{\frac{2}{\pi}}\frac{(\sin\theta)^\mu}{\Gamma[-\mu+(1/2)]}\int_0^\theta\frac{\cos[\{\nu+(1/2)\}\varphi]\,d\varphi}{(\cos\varphi-\cos\theta)^{\mu+(1/2)}}\qquad\left(0<\theta<\pi,\ \operatorname{Re}\mu<\frac{1}{2}\right).$$

$$P_\nu^{-\mu}(\cos\theta)=\frac{\Gamma(2\mu+1)2^{-\mu}(\sin\theta)^\mu}{\Gamma(\mu+1)\Gamma(\mu+\nu+1)\Gamma(\mu-\nu)}\int_0^\infty\frac{t^{\mu+\nu}dt}{(1+2t\cos\theta+t^2)^{\mu+(1/2)}}$$

$$(\operatorname{Re}(\mu+\nu)>-1,\ \operatorname{Re}(\mu-\nu)>0).$$

$$P_\nu^{-\mu}(\cos\theta)=\frac{1}{\Gamma(\nu+\mu+1)}\int_0^\infty e^{-t\cos\theta}J_\mu(t\sin\theta)t^\nu dt\qquad\left(0<\theta<\frac{\pi}{2},\operatorname{Re}(\mu+\nu)>-1\right).$$

$$\mathfrak{Q}_\nu^\mu(z)=\frac{e^{i\mu\pi}}{2^{\nu+1}}\frac{\Gamma(\nu+\mu+1)}{\Gamma(\nu+1)}(z^2-1)^{\mu/2}\int_{-1}^{+1}(1-t)^\nu(z-1)^{-\nu-\mu-1}dt$$

$$(\operatorname{Re}(\nu+\mu)>-1,\operatorname{Re}\nu>-1,|\arg(z\pm 1)|<\pi).$$

$$\mathfrak{Q}_\nu^\mu(\cosh\alpha)=\sqrt{\frac{\pi}{2}}\frac{e^{i\mu\pi}(\sinh\alpha)^\mu}{\Gamma[-\mu+(1/2)]}\int_\alpha^\infty\frac{e^{-[\nu+(1/2)]t}dt}{(\cosh t-\cosh\alpha)^{\mu+(1/2)}}$$

$$(\alpha>0,\ \operatorname{Re}\mu<1/2,\ \operatorname{Re}(\nu+\mu)>-1).$$

Recurrence formulas:

$$(z^2-1)d\mathfrak{P}_\nu^\mu(z)/dz=(\nu-\mu+1)\mathfrak{P}_{\nu+1}^\mu(z)-(\nu+1)z\mathfrak{P}_\nu^\mu(z),$$

$$(2\nu+1)z\mathfrak{P}_\nu^\mu(z)=(\nu-\mu+1)\mathfrak{P}_{\nu+1}^\mu(z)+(\nu+\mu)\mathfrak{P}_{\nu-1}^\mu(z),$$

$$\mathfrak{P}_{-\nu-1}^\mu(z)=\mathfrak{P}_\nu^\mu(z),$$

$$\mathfrak{Q}_\nu^{-\mu}(z)=e^{-2i\mu\pi}\frac{\Gamma(\nu-\mu+1)}{\Gamma(\nu+\mu+1)}\mathfrak{Q}_\nu^\mu(z),$$

$$(1-x^2)dP_\nu^\mu(x)/dx=(\nu+1)xP_\nu^\mu(x)-(\nu-\mu+1)P_{\nu+1}^\mu(x).$$

The case when μ is an integer $m(m=0,1,2,\ldots)$ and ν is also an integer n:

$$P_n^{m+2}(x)+2(m+1)x(1-x^2)^{-1/2}P_n^{m+1}(x)+(n-m)(n+m+1)P_n^m(x)=0.$$

$$(2n+1)xP_n^m(x)-(n-m+1)P_{n+1}^m(x)-(n+m)P_{n-1}^m(x)=0 \qquad (0\leqslant m\leqslant n-2),$$

$$(x^2-1)dP_n^m(x)/dx-(n-m+1)P_{n+1}^m(x)+(n+1)xP_n^m(x)=0,$$

$$P_{n-1}^m(x)-P_{n+1}^m(x)=(2n+1)\sqrt{1-x^2}\;P_n^{m-1}(x).$$

$$\mathfrak{P}_\nu^{-\mu}(z)=\frac{\Gamma(\nu-\mu+1)}{\Gamma(\nu+\mu+1)}\left[\mathfrak{P}_\nu^\mu(z)-\frac{2}{\pi}e^{-i\mu\pi}(\sin\mu\pi)\mathfrak{Q}_\nu^\mu(z)\right],$$

$$\mathfrak{Q}_\nu^\mu(z)\sin[(\nu+\mu)\pi]-\mathfrak{Q}_{-\nu-1}^\mu(z)\sin[(\nu-\mu)\pi]=\pi e^{i\mu\pi}(\cos\nu\pi)\mathfrak{P}_\nu^\mu(z),$$

$$\mathfrak{P}_\nu^\mu(-z)=e^{\mp i\nu\pi}\mathfrak{P}_\nu^\mu(z)-(2/\pi)[\sin(\nu+\mu)\pi]e^{-i\mu\pi}\mathfrak{Q}_\nu(z) \qquad (\mp=-\operatorname{sgn}(\operatorname{Im}z)),$$

$$\mathfrak{Q}_\nu^\mu(-z)=-e^{\pm i\nu\pi}\mathfrak{Q}_\nu^\mu(z) \qquad (\pm=\operatorname{sgn}(\operatorname{Im}z)).$$

$$e^{-i\mu\pi}\mathfrak{Q}_\nu^\mu(\cosh\alpha)=\frac{\pi\Gamma(1+\mu+\nu)}{\sqrt{2\pi\sinh\alpha}}\mathfrak{P}_{-\mu-1/2}^{-\nu-1/2}(\coth\alpha) \qquad (\operatorname{Re}\cosh\alpha>0).$$

$$e^{-i\mu\pi}\mathfrak{Q}_\nu^\mu(x\pm i0)=e^{\pm i\mu\pi/2}[Q_\nu^\mu(x)\mp(i\pi/2)P_\nu^\mu(x)].$$

$$Q_{-\nu-1}^\mu(x)=\frac{\sin(\nu+\mu)\pi}{\sin(\nu-\mu)\pi}Q_\nu^\mu(x)-\frac{\pi\cos\nu\pi\cos\mu\pi}{\sin(\nu-\mu)\pi}P_\nu^\mu(x),$$

$$P_\nu^{-\mu}(x)=\frac{\Gamma(\nu-\mu+1)}{\Gamma(\nu+\mu+1)}\left[\cos\mu\pi P_\nu^\mu(x)-\frac{2}{\pi}\sin\mu\pi Q_\nu^\mu(x)\right],$$

$$P_\nu^\mu(-x)=[\cos(\nu+\mu)\pi]P_\nu^\mu(x)-(2/\pi)[\sin(\nu+\mu)\pi]Q_\nu^\mu(x),$$

$$Q_\nu^\mu(-x)=-[\cos(\nu+\mu)\pi]Q_\nu^\mu(x)+(\pi/2)[\sin(\nu+\mu)\pi]P_\nu^\mu(x).$$

$$\mathfrak{P}_\nu^m(z)=\frac{\Gamma(1+\nu+m)(z^2-1)^{m/2}}{\Gamma(1+\nu-m)m!2^m}{}_2F_1\left(m-\nu,\,m+\nu+1;\,m+1;\,\frac{1-z}{2}\right)=(z^2-1)^{m/2}\frac{d^m\mathfrak{P}_\nu(z)}{dz^m},$$

$$\mathfrak{P}_\nu^{-m}(z)=(z^2-1)^{-m/2}\int_1^z\cdots\int_1^z P_\nu(z)(dz)^m.$$

$$\mathfrak{Q}_\nu^m(z)=(z^2-1)^{m/2}\frac{d^m\mathfrak{Q}_\nu(z)}{dx^m}, \qquad \mathfrak{Q}_\nu^{-m}(z)=(-1)^m(z^2-1)^{-m/2}\int_z^\infty\cdots\int_z^\infty\mathfrak{Q}_\nu(z)(dz)^m.$$

$$P_\nu^m(x)=(-1)^m\frac{\Gamma(1+\nu+m)(1-x^2)^{m/2}}{\Gamma(1+\nu-m)m!2^m}{}_2F_1\left(m-\nu,\,m+\nu+1;\,m+1;\,\frac{1-x}{2}\right)$$

$$=(-1)^m(1-x^2)^{m/2}\frac{d^mP_\nu(x)}{dx^m},$$

$$P_\nu^{-m}(x)=(1-x^2)^{-m/2}\int_x^1\cdots\int_x^1 P_\nu(x)(dx)^m=(-1)^m\frac{\Gamma(\nu-m+1)}{\Gamma(\nu+m+1)}P_\nu^m(x).$$

$$Q_\nu^m(x)=(-1)^m(1-x^2)^{m/2}\frac{d^mQ_\nu(x)}{dx^m}, \qquad Q_\nu^{-m}(x)=(-1)^m\frac{\Gamma(\nu-m+1)}{\Gamma(\nu+m+1)}Q_\nu^m(x).$$

The values at the origin are

$$P_\nu^\mu(0)=\frac{\sqrt{\pi}\,2^\mu}{\Gamma[(\nu-\mu)/2+1]\Gamma[(-\nu-\mu+1)/2]},$$

$$\frac{dP_\nu^\mu(0)}{dx}=\frac{2^{\mu+1}\sin[\pi(\nu+\mu)/2]\Gamma[(\nu+\mu+2)/2]}{\Gamma[(\nu-\mu+1)/2]\sqrt{\pi}}=\frac{\sqrt{\pi}\,2^{\mu+1}}{\Gamma[(\nu-\mu+1)/2]\Gamma[(-\nu-\mu)/2]},$$

$$Q_\nu^\mu(0)=-2^{\mu-1}\sqrt{\pi}\,\sin\left(\frac{\nu+\mu}{2}\pi\right)\frac{\Gamma[(\nu+\mu+1)/2]}{\Gamma[(\nu-\mu+2)/2]},$$

$$\frac{dQ_\nu^\mu(0)}{dx}=2^\mu\sqrt{\pi}\,\cos\left(\frac{\nu+\mu}{2}\pi\right)\frac{\Gamma[(\nu+\mu+2)/2]}{\Gamma[(\nu-\mu+1)/2]}.$$

(2) Generating Functions.

$$(\cos\theta+i\sin\theta\sin\varphi)^n=P_n(\cos\theta)+2\sum_{m=1}^{n}(-i)^m\frac{n!}{(n+m)!}(\cos m\varphi)P_n^m(\cos\theta).$$

$$P_\nu^{-\mu}(\cos\theta)=\frac{\sin\nu\pi}{\pi}\sum_{n=0}^{\infty}(-1)^n\left(\frac{1}{\nu-n}-\frac{1}{\nu+n+1}\right)P_n^{-\mu}(\cos\theta)\qquad(0<\theta<\pi,\ \mu\geqslant 0).$$

(3) Orthogonal Relations.

$$\int_{-1}^{+1}P_n^m(x)P_{n'}^m(x)\,dx=\frac{2}{2n+1}\frac{(n+m)!}{(n-m)!}\delta_{nn'}.$$

$$\int_0^{2\pi}d\varphi\int_0^{\pi}\sin\theta e^{\pm i(m-m')\varphi}P_n^m(\cos\theta)P_{n'}^{m'}(\cos\theta)\,d\theta=\frac{4\pi}{2n+1}\frac{(n+m)!}{(n-m)!}\delta_{nn'}\delta_{mm'}.$$

(4) Addition Theorems.

$$\mathfrak{P}_\nu\left(z\zeta-\sqrt{z^2-1}\,\sqrt{\zeta^2-1}\,\cos\varphi\right)=\mathfrak{P}_\nu(z)\mathfrak{P}_\nu(\zeta)+2\sum_{m=1}^{\infty}(-1)^m\mathfrak{P}_\nu^m(z)\mathfrak{P}_\nu^{-m}(\zeta)\cos m\varphi$$

$$(\mathrm{Re}\,z>0,\ \mathrm{Re}\,\zeta>0,\ |\arg(z-1)|<\pi,\ |\arg(\zeta-1)|<\pi).$$

$$\mathfrak{Q}_\nu\left(tt'-\sqrt{t^2-1}\,\sqrt{t'^2-1}\,\cos\varphi\right)=\mathfrak{Q}_\nu(t)\mathfrak{P}_\nu(t')+2\sum_{m=1}^{\infty}(-1)^m\mathfrak{Q}_\nu^m(t)\mathfrak{P}_\nu^{-m}(t')\cos m\varphi$$

$$(t,t'\text{ real},\ 1<t'<t,\ \nu\neq\text{negative integer},\ \varphi\text{ real}).$$

$$P_\nu(\cos\theta\cos\theta'+\sin\theta\sin\theta'\cos\varphi)=P_\nu(\cos\theta)P_\nu(\cos\theta')+2\sum_{m=1}^{\infty}(-1)^mP_\nu^{-m}(\cos\theta')P_\nu^m(\cos\theta')\cos m\varphi$$

$$=P_\nu(\cos\theta)P_\nu(\cos\theta')+2\sum_{m=1}^{\infty}\frac{\Gamma(\nu-m+1)}{\Gamma(\nu+m+1)}P_\nu^m(\cos\theta)P_\nu^m(\cos\theta')\cos m\varphi$$

$$(0\leqslant\theta<\pi,\ 0\leqslant\theta'<\pi,\ \theta+\theta'<\pi,\ \varphi\text{ real}).$$

$$Q_\nu(\cos\theta\cos\theta'+\sin\theta\sin\theta'\cos\varphi)=P_\nu(\cos\theta')Q_\nu(\cos\theta)+2\sum_{m=1}^{\infty}(-1)^mP_\nu^{-m}(\cos\theta')Q_\nu^m(\cos\theta)\cos m\varphi$$

$$(0<\theta'<\pi/2,\ 0<\theta<\pi,\ \theta+\theta'<\pi,\ \varphi\text{ real}).$$

$$\mathfrak{Q}_n\left(\tau\tau'+\sqrt{\tau^2+1}\,\sqrt{\tau'^2+1}\,\cosh\alpha\right)=\sum_{m=n+1}^{\infty}\frac{1}{(m-n-1)!(m+n)!}\mathfrak{Q}_n^m(i\tau)\mathfrak{Q}_n^m(i\tau')e^{-m\alpha}$$

$$(\tau,\ \tau',\ \alpha>0).$$

(5) Asymptotic Expansions.

$$\mathfrak{P}_\nu^\mu(z)=\left[\frac{2^\nu\Gamma[\nu+(1/2)]}{\sqrt{\pi}\,\Gamma(\nu-\mu+1)}z^\nu+\frac{2^{-\nu-1}\Gamma[-\nu-(1/2)]}{\sqrt{\pi}\,\Gamma(-\mu-\nu)}z^{-\nu-1}\right]\left[1+O(z^{-2})\right]$$

$$(\nu+(1/2)\neq\text{integer},\ |\arg z|<\pi,\ |z|\gg 1).$$

$$\mathfrak{Q}_\nu^\mu(z)=\frac{\sqrt{\pi}\,e^{i\mu\pi}}{2^{\nu+1}}\frac{\Gamma(\nu+\mu+1)}{\Gamma[\nu+(3/2)]}z^{-\nu-1}\left[1+O(z^{-2})\right]$$

$$(\nu+(1/2)\neq\text{negative integer},\ |\arg z|<\pi,\ |z|\gg 1).$$

$$P_\nu^\mu(\cos\theta)=\frac{2}{\sqrt{\pi}}\frac{\Gamma(\nu+\mu+1)}{\Gamma[\nu+(3/2)]}\frac{\cos[\{\nu+(1/2)\}\theta-(\pi/4)+(\mu\pi/2)]}{\sqrt{2\sin\theta}}\left[1+O(\nu^{-1})\right]$$

$$(\varepsilon\leqslant\theta\leqslant\pi-\varepsilon,\ \varepsilon>0,\ |\nu|\gg 1/\varepsilon).$$

$$P_\nu^\mu(\cos\theta)=\frac{2\Gamma(\nu+\mu+1)}{\sqrt{\pi}\,\Gamma[\nu+(3/2)]}$$
$$\times\sum_{l=0}^{\infty}\frac{\Gamma\left(\frac{1}{2}+\mu+l\right)\Gamma\left(\frac{1}{2}-\mu+l\right)\Gamma\left(\nu+\frac{3}{2}\right)}{\Gamma\left(\frac{1}{2}+\mu\right)\Gamma\left(\frac{1}{2}-\mu\right)\Gamma\left(\nu+l+\frac{3}{2}\right)l!}\,\frac{\cos\left[\left(\nu+\frac{2l+1}{2}\right)\theta-\frac{(2l+1)\pi}{4}+\frac{\mu\pi}{2}\right]}{(2\sin\theta)^{l+(1/2)}},$$

$$Q_\nu^\mu(\cos\theta)=\sqrt{\pi}\,\frac{\Gamma(\nu+\mu+1)}{\Gamma[\nu+(3/2)]}$$
$$\times\sum_{l=0}^{\infty}(-1)^l\frac{\Gamma\left(\frac{1}{2}+\mu+l\right)\Gamma\left(\frac{1}{2}-\mu+l\right)\Gamma\left(\nu+\frac{3}{2}\right)}{\Gamma\left(\frac{1}{2}+\mu\right)\Gamma\left(\frac{1}{2}-\mu\right)\Gamma\left(\nu+l+\frac{3}{2}\right)l!}\,\frac{\cos\left[\left(\nu+\frac{2l+1}{2}\right)\theta+\frac{(2l+1)\pi}{4}+\frac{\mu\pi}{2}\right]}{(2\sin\theta)^{l+(1/2)}}.$$

(In the final two formulas the series converges when $\nu+\mu\neq$negative integer, $\nu+(1/2)\neq$negative integer, $\pi/6<\theta<5\pi/6$.)

$$\left[\left(\nu+\frac{1}{2}\right)\cos\frac{\theta}{2}\right]^\mu P_\nu^{-\mu}(\cos\theta)=J_\mu(\eta)+\sin^2\frac{\theta}{2}\left[\frac{J_{\mu+1}(\eta)}{2\eta}-J_{\mu+2}(\eta)+\frac{\eta}{6}J_{\mu+3}(\eta)\right]+O\left(\sin^4\frac{\theta}{2}\right)$$
$$(\eta=(2\nu+1)\sin(\theta/2)).$$

(6) Estimation. When $\nu\geqslant 1,\quad \nu-\mu+1>0,\quad \mu\geqslant 0,$

$$|P_\nu^{\pm\mu}(\cos\theta)|<\frac{\Gamma(\nu\pm\mu+1)}{\Gamma(\nu+1)}\left(\frac{8}{\nu\pi\sin\theta}\right)^{1/2}\frac{1}{(\sin\theta)^\mu},$$

$$|Q_\nu^{\pm\mu}(\cos\theta)|<\frac{\Gamma(\nu\pm\mu+1)}{\Gamma(\nu+1)}\left(\frac{2\pi}{\nu\sin\theta}\right)^{1/2}\frac{1}{(\sin\theta)^\mu},$$

$$|P_\nu^{\pm m}(\cos\theta)|<\frac{\Gamma(\nu\pm m+1)}{\Gamma(\nu+1)}\left(\frac{4}{\nu\pi\sin\theta}\right)^{1/2}\frac{1}{(\sin\theta)^m},$$

$$|Q_\nu^{\pm m}(\cos\theta)|\leqslant\frac{\Gamma(\nu\pm m+1)}{\Gamma(\nu+1)}\left(\frac{4}{\nu\sin\theta}\right)^{1/2}\frac{1}{(\sin\theta)^m}.$$

(7) Torus Functions. These are solutions of the differential equation

$$\frac{d^2u}{d\eta^2}+\coth\eta\frac{du}{d\eta}-\left(n^2-\frac{1}{4}+\frac{m^2}{\sinh^2\eta}\right)u=0.$$

The fundamental system of solutions is given by

$\mathfrak{P}_{n-(1/2)}^m(\cosh\eta)$, $\mathfrak{Q}_{n-(1/2)}^m(\cosh\eta)$.

The asymptotic expansion when $m=0$ is

$$\mathfrak{P}_{n-(1/2)}(\cosh\eta)$$
$$=\frac{(n-1)!\,e^{n-(1/2)\eta}}{\Gamma[n+(1/2)]\sqrt{\pi}}\left[\frac{2\Gamma^2[n+(1/2)]}{\pi n!(n-1)!}(\log 4+\eta)e^{-2m\eta}{}_2F_1\left(\frac{1}{2},n+\frac{1}{2};n+1;e^{-2\eta}\right)+A+B\right].$$

Here

$$A=1+\frac{(1/2)[n-(1/2)]}{1\cdot(n-1)}e^{-2\eta}+\frac{(1/2)(3/2)[n-(1/2)][n-(3/2)]}{1\cdot 2\cdot(n-1)(n-2)}e^{-4\eta}+\dots$$
$$+\frac{(2n-3)!!(2n-1)!!}{[(2n-2)!!]^2}e^{-2(n-1)\eta},$$

$$B=\frac{\Gamma[n+(1/2)]}{\pi^{3/2}(n-1)!}\sum_{l=1}^{\infty}\frac{\Gamma[l+(1/2)]\Gamma[n+l+(1/2)]}{(n+l)!l!}(u_{n+l}+u_l-v_{l-(1/2)}-v_{n+l-(1/2)})e^{-2(l+n)\eta},$$

where

$$u_r\equiv 1+\frac{1}{2}+\dots+\frac{1}{r},\qquad v_{r-(1/2)}\equiv\frac{2}{1}+\frac{2}{3}+\frac{2}{5}+\dots+\frac{2}{2r-1}=2u_{2r}-u_r.$$

References

See references to Table 16, this Appendix.

19. Functions of Confluent Type and Bessel Functions

(I) Hypergeometric Function of Confluent Type (→ 167 Functions of Confluent Type)

(1) Kummer Functions.

$$v(z)={}_1F_1(a;\ c;\ z)\equiv\sum_{n=0}^{\infty}\frac{\Gamma(a+n)}{\Gamma(a)}\frac{\Gamma(c)}{\Gamma(c+n)}\frac{z^n}{n!}$$

$$=\frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}z^{1-c}\int_0^z e^t t^{a-1}(z-t)^{c-a-1}dt \qquad (0<\operatorname{Re}a<\operatorname{Re}c)$$

$$=\frac{\Gamma(c)2^{1-c}}{\Gamma(a)\Gamma(c-a)}e^{z/2}\int_{-1}^{+1}e^{zt/2}(1-t)^{c-a-1}(1+t)^{a-1}dt \qquad (0<\operatorname{Re}a<\operatorname{Re}c).$$

The fundamental system of solutions of the confluent hypergeometric differential equation (Kummer's differential equation)

$$z\frac{d^2v}{dz^2}+(c-z)\frac{dv}{dz}-av=0,$$

when $c\neq 0,-1,-2,\ldots$, is given by

$$\mathring{v}_1(z)\equiv{}_1F_1(a;c;z),\qquad \mathring{v}_2(z)\equiv z^{1-c}{}_1F_1(a-c+1;\ 2-c;\ z).$$

$$d\,{}_1F_1(a;\ c;\ z)/dz=(a/c)\,{}_1F_1(a+1;c+1;z),$$

$${}_1F_1(a;c;z)=e^z\,{}_1F_1(c-a;c;-z),$$

$$a\,{}_1F_1(a+1;c+1;z)=(a-c)\,{}_1F_1(a;c+1;z)+c\,{}_1F_1(a;c;z),$$

$$a\,{}_1F_1(a+1;c;z)=(z+2a-c)\,{}_1F_1(a;c;z)+(c-a)\,{}_1F_1(a-1;c;z).$$

Putting $(a)_n=a(a+1)\ldots(a+n-1)=\Gamma(a+n)/\Gamma(a)$ we have

$$\lim_{c\to -n}\frac{1}{\Gamma(c)}{}_1F_1(a;c;z)=\frac{z^{n+1}(a)_{n+1}}{(n+1)!}{}_1F_1(a+n+1;n+2;z)\qquad (n=0,1,2,\ldots).$$

Asymptotic expansion:

$$\mathring{v}_1\approx A_1z^{-a}\sum_{n=0}^{\infty}\frac{(a)_n(a-c+1)_n}{n!}(-z)^{-n}+B_1e^z z^{a-c}\sum_{n=0}^{\infty}\frac{(c-a)_n(1-a)_n}{n!}z^{n},$$

$$\mathring{v}_2\approx A_2z^{-a}\sum_{n=0}^{\infty}\frac{(a)_n(a-c+1)_n}{n!}(-z)^{-n}+B_2e^z z^{a-c}\sum_{n=0}^{\infty}\frac{(c-a)_n(1-a)_n}{n!}z^{n}$$

$(|z|\gg|a|,\ |z|\gg|c|,\ -3\pi/2<\arg z<\pi/2, c\neq\text{integer})$,

where

$$A_1=e^{-i\pi a}\Gamma(c)/\Gamma(c-a),\qquad B_1=\Gamma(c)/\Gamma(a),$$

$$A_2=e^{-i\pi(a-c+1)}\Gamma(2-c)/\Gamma(1-a),\qquad B_2=\Gamma(2-c)/\Gamma(a-c+1).$$

(2) The fundamental system of solutions at $z=0$ of the hypergeometric differential equation of confluent type

$$\frac{d^2u}{dz^2}+\frac{du}{dz}+\left[\frac{\kappa}{z}+\frac{(1/4)-\mu^2}{z^2}\right]u=0$$

is given by

$$z^{(1/2)\pm\mu}e^{-z}{}_1F_1[(1/2)\pm\mu-\kappa;\ \pm2\mu+1;z].$$

(II) Whittaker Functions (→ 167 Functions of Confluent Type)

(1) A pair of linearly independent solutions of Whittaker's differential equation

$$\frac{d^2W}{dz^2}+\left[-\frac{1}{4}+\frac{\kappa}{z}+\frac{(1/4)-\mu^2}{z^2}\right]W=0$$

is given by $M_{\kappa,\pm\mu}(z)=z^{\pm\mu+(1/2)}e^{-z/2}{}_1F_1[\pm\mu-\kappa+(1/2);\pm 2\mu+1;z]$.

Whittaker functions:

$$W_{\kappa,\mu}(z)\equiv\frac{\Gamma(-2\mu)}{\Gamma[(1/2)-\mu-\kappa]}M_{\kappa,\mu}(z)+\frac{\Gamma(2\mu)}{\Gamma[(1/2)+\mu-\kappa]}M_{\kappa,-\mu}(z)=W_{\kappa,-\mu}(z).$$

When 2μ is an integer, the above definition of $W_{\kappa,\mu}(z)$ loses meaning, but by taking the limit with respect to μ we can define it in terms of the following integrals.

$$W_{\kappa,\mu}(z)=\frac{z^{\mu+(1/2)}e^{-z/2}}{\Gamma[\mu+(1/2)-\kappa]}\int_0^\infty e^{-z\tau}\tau^{\mu-\kappa-(1/2)}(1+\tau)^{\mu+\kappa-(1/2)}d\tau$$

$$=\frac{z^{\kappa}e^{-z/2}}{\Gamma[\mu+(1/2)-\kappa]}\int_0^\infty t^{\mu-\kappa-(1/2)}e^{-t}\left(1+\frac{t}{z}\right)^{\mu+\kappa-(1/2)}dt$$

$(\mathrm{Re}[\mu+(1/2)-\kappa]>0,\quad |\arg z|<\pi)$.

$$W_{\kappa,\mu}(z)=\frac{e^{-z/2}}{2\pi i}\int_{-i\infty}^{+i\infty}\frac{\Gamma(s-\kappa)\Gamma[-s-\mu+(1/2)]\Gamma[-s+\mu+(1/2)]}{\Gamma[-\kappa+\mu+(1/2)]\Gamma[-\kappa-\mu+(1/2)]}z^s\,ds.$$

$$M_{l+\mu+(1/2),\mu}(z)=(-1)^l z^{\mu+(1/2)}e^{-z/2}(2\mu+1)_l\,{}_1F_1(-l;2\mu+1;z)\qquad (l=0,1,2,\ldots).$$

$$M_{\kappa,\mu}(z)=e^{-i\pi[\mu+(1/2)]}M_{-\kappa,\mu}(e^{i\pi}z).$$

$$M_{\kappa,\mu}(z)=\frac{\Gamma(2\mu+1)}{\Gamma[\mu+(1/2)-\kappa]}e^{i\pi\kappa}W_{-\kappa,\mu}(e^{i\pi}z)+\frac{\Gamma(2\mu+1)}{\Gamma[\mu+(1/2)+\kappa]}e^{i\pi[\kappa-\mu-(1/2)]}W_{\kappa,\mu}(z)$$

$(-3\pi/2<\arg z<\pi/2,\quad 2\mu\neq-1,-2,\ldots)$.

$$M_{\kappa,\mu}(z)=\frac{\Gamma(2\mu+1)}{\Gamma[\mu+(1/2)-\kappa]}e^{-i\pi\kappa}W_{-\kappa,\mu}(e^{-i\pi}z)+\frac{\Gamma(2\mu+1)}{\Gamma[\mu+(1/2)+\kappa]}e^{-i\pi[\kappa-\mu-(1/2)]}W_{\kappa,\mu}(z)$$

$(-\pi/2<\arg z<3\pi/2,\quad 2\mu\neq-1,-2,\ldots)$.

$$W_{\kappa,\mu}(z)=z^{1/2}W_{\kappa-(1/2),\mu-(1/2)}(z)+[(1/2)-\kappa+\mu]W_{\kappa-1,\mu}(z)$$

$$=z^{1/2}W_{\kappa-(1/2),\mu+(1/2)}(z)+[(1/2)-\kappa-\mu]W_{\kappa-1,\mu}(z).$$

$$z\,dW_{\kappa,\mu}(z)/dz=[\kappa-(z/2)]W_{\kappa,\mu}(z)-\left[\mu^2-\{\kappa-(1/2)\}^2\right]W_{\kappa-1,\mu}(z).$$

When κ is sufficiently large we have

$$M_{\kappa,\mu}(z)\sim\pi^{-1/2}\Gamma(2\mu+1)\kappa^{-\mu-(1/4)}z^{1/4}\cos\left[2(z\kappa)^{1/2}-\mu\pi-(\pi/4)\right],$$

$$W_{\kappa,\mu}(z)\sim-(4z/\kappa)^{1/4}\exp(-\kappa+\kappa\log\kappa)\sin\left[2(z\kappa)^{1/2}-\pi\kappa-(\pi/4)\right],$$

$$W_{-\kappa,\mu}(z)\sim(z/4\kappa)^{1/4}\exp\left(\kappa-\kappa\log\kappa-2(z\kappa)^{1/2}\right).$$

Asymptotic expansion:

$$W_{\kappa,\mu}(z)\approx e^{-z/2}z^{\kappa}$$

$$\times\left(1+\sum_{n=1}^{\infty}\frac{[\mu^2-\{\kappa-(1/2)\}^2][\mu^2-\{\kappa-(3/2)\}^2]\ldots[\mu^2-\{\kappa-n+(1/2)\}^2]}{n!z^n}\right).$$

(2) Representation of Various Special Functions by Whittaker Functions.

(i) Probability integral (error function) $\quad \mathrm{erf}\,x\equiv\Phi(x)\equiv\dfrac{2}{\sqrt{\pi}}\displaystyle\int_0^x e^{-t^2}dt$

$$=1-\pi^{-1/2}x^{-1/2}e^{-x^2/2}W_{-1/4,1/4}(x^2)$$

$$=\frac{2x}{\sqrt{\pi}}{}_1F_1\left(\frac{1}{2};\frac{3}{2};-x^2\right)=\frac{2}{\sqrt{\pi}}\left(x-\frac{x^3}{1!3}+\frac{x^5}{2!5}-\frac{x^7}{3!7}\pm\ldots\right).$$

Asymptotic expansion:

$$\frac{\sqrt{\pi}}{2}[1-\Phi(x)]\approx\frac{e^{-x^2}}{2x}\left(1-\frac{1}{2x^2}+\frac{1\cdot3}{(2x^2)^2}-\frac{1\cdot3\cdot5}{(2x^2)^3}\pm\ldots\right).$$

$$\frac{1}{1+i}\Phi\left(x\frac{1+i}{2}\sqrt{\pi}\right)=C(x)-iS(x),$$

where $C(x)$, $S(x)$ are the following Fresnel integrals.

$$C(x)\equiv\int_0^x\cos\frac{\pi}{2}t^2dt=\frac{1}{2}+\frac{1}{\pi x}\sin\frac{\pi}{2}x^2+O\left(\frac{1}{x^2}\right),$$

$$S(x)\equiv\int_0^x\sin\frac{\pi}{2}t^2dt=\frac{1}{2}-\frac{1}{\pi x}\cos\frac{\pi}{2}x^2+O\left(\frac{1}{x^2}\right).$$

(ii) Logarithmic integral

$$\mathrm{Li}\,z\equiv\int_0^z\frac{dt}{\log t}\qquad\text{(When } z>1\text{, take Cauchy's principal value at } t=1.)$$

$$=-(\log 1/z)^{-1/2}z^{1/2}W_{-1/2,0}(-\log z).$$

$\mathrm{Li}\,z$ is sometimes written as $\mathrm{li}\,z$.

(3) Exponential Integral

$$\mathrm{Ei}\,x\equiv\int_{-\infty}^x\frac{e^t}{t}dt\quad\text{(When } x>0\text{, take the Cauchy's principal value at } t=0 \text{ while integrating.)}$$

$$=C+\log|x|+\sum_{n=1}^{\infty}\frac{x^n}{n\cdot n!}\qquad(x\text{ real, }\neq 0)$$

$$=e^x\sum_{n=1}^{N}\frac{(n-1)!}{t^n}+N!\sum_{\substack{n=0\\n\neq N}}^{\infty}\frac{t^{n-N}}{n!(n-N)}-\left(1+\frac{1}{2}+\cdots+\frac{1}{N}\right)+C+\log|x|.$$

Cosine integral $\quad \mathrm{Ci}\,x\equiv-\int_x^\infty\frac{\cos t}{t}dt=C+\log x-\int_0^x\frac{1-\cos t}{t}dt.$

Sine integral $\quad \mathrm{Si}\,x\equiv\int_0^x\frac{\sin t}{t}dt,$

$$\mathrm{si}\,x\equiv-\int_x^\infty\frac{\sin t}{t}dt=\mathrm{Si}\,x-\frac{\pi}{2}.$$

Asymptotic expansion $\quad \mathrm{Ei}\,ix=\mathrm{Ci}\,x+i\,\mathrm{si}\,x\approx e^{ix}\left(\frac{1}{ix}+\frac{1!}{(ix)^2}+\frac{2!}{(ix)^3}+\frac{3!}{(ix)^4}+\cdots\right).$

(III) Bessel Functions (→ 39 Bessel Functions)

(1) Cylindrical Functions. A cylindrical function Z_ν is a solution of Bessel's differential equation

$$\frac{d^2Z_\nu}{dz^2}+\frac{1}{z}\frac{dZ_\nu}{dz}+\left(1-\frac{\nu^2}{z^2}\right)Z_\nu=0.$$

Recurrence formulas:

$$Z_{\nu-1}(z)+Z_{\nu+1}(z)=(2\nu/z)Z_\nu(z),\qquad Z_{\nu-1}(z)-Z_{\nu+1}(z)=2dZ_\nu(z)/dz.$$

$$\int z^{\nu+1}Z_\nu(z)dz=z^{\nu+1}Z_{\nu+1}(z),\qquad \int z^{-\nu}Z_{\nu+1}(z)dz=-z^{-\nu}Z_\nu(z).$$

As special solutions, we have the following three kinds of functions.

(i) Bessel function (Bessel function of the first kind).

$$J_\nu(z)\equiv\left(\frac{z}{2}\right)^\nu\sum_{l=0}^{\infty}\frac{(-1)^l}{l!\Gamma(\nu+l+1)}\left(\frac{z}{2}\right)^{2l}=\frac{M_{0,\nu}(2iz)}{(2iz)^{1/2}2^{2\nu}i^\nu\Gamma(\nu+1)}\qquad(|\arg z|<\pi).$$

$$J_\nu(e^{im\pi}z)=e^{im\nu\pi}J_\nu(z).$$

$$J_{-n}(z)=(-1)^nJ_n(z).$$

$$J_{n+(1/2)}(z)=\sqrt{\frac{2}{\pi}}\,z^{n+(1/2)}\left(-\frac{1}{z}\frac{d}{dz}\right)^n\left(\frac{\sin z}{z}\right)\qquad(n=0,1,2,\ldots).$$

(ii) Neumann function (Bessel function of the second kind).

$$N_\nu(z)\equiv\frac{1}{\sin\nu\pi}\left[(\cos\nu\pi)J_\nu(z)-J_{-\nu}(z)\right]\qquad(\nu\neq\text{integer};\ |\arg z|<\pi),$$

$$N_n(z)\equiv\frac{2}{\pi}J_n(z)\left(C+\log\frac{z}{2}\right)-\frac{1}{\pi}\left(\frac{z}{2}\right)^n\sum_{l=0}^{\infty}\frac{(-1)^l}{l!(n+l)!}\left(\frac{z}{2}\right)^{2l}[\varphi(l)+\varphi(l+n)]$$

$$-\frac{1}{\pi}\left(\frac{z}{2}\right)^{-n}\sum_{l=0}^{n-1}\frac{(n-l-1)!}{l!}\left(\frac{z}{2}\right)^{2l}\qquad\left(\varphi(l)\equiv\sum_{m=1}^{l}\frac{1}{m}\right),$$

$N_{-n}(z) \equiv (-1)^n N_n(z) \qquad (n=0,1,2,\ldots; \qquad |\arg z|<\pi).$

$N_\nu(e^{im\pi}z) = e^{-im\nu\pi}N_\nu(z) + 2i(\sin m\nu\pi \cot \nu\pi)J_\nu(z).$

$N_{n+(1/2)}(z) = (-1)^{n+1}J_{-[n+(1/2)]}(z).$

(iii) Hankel function (Bessel function of the third kind).

$H_\nu^{(1)}(z) \equiv J_\nu(z) + iN_\nu(z),$

$H_\nu^{(2)}(z) \equiv J_\nu(z) - iN_\nu(z).$

$H_\nu^{(1)}(iz/2) = -2ie^{-i\nu\pi/2}(\pi z)^{-1/2}W_{0,\nu}(z).$

$H_{-\nu}^{(1)}(z) = e^{i\nu\pi}H_\nu^{(1)}(z), \quad H_{-\nu}^{(2)}(z) = e^{-i\nu\pi}H_\nu^{(2)}(z). \quad \overline{H_\nu^{(2)}(x)} = H_\nu^{(1)}(x) \quad (x, \nu \text{ real}).$

(2) Integral Representation.

Hansen-Bessel formula
$$J_n(z) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} e^{iz\cos t}e^{in[t-(\pi/2)]}dt = \frac{i^{-n}}{\pi}\int_0^\pi e^{iz\cos t}\cos nt\,dt = \frac{1}{\pi}\int_0^\pi \cos(z\sin t - nt)dt \qquad (n=0,1,2,\ldots).$$

Mehler's formula
$$J_0(x) = \frac{2}{\pi}\int_0^\infty \sin(x\cosh t)dt,$$
$$N_0(x) = -\frac{2}{\pi}\int_0^\infty \cos(x\cosh t)dt \qquad (x>0).$$

Poisson's formula
$$J_\nu(z) = \frac{2(z/2)^\nu}{\sqrt{\pi}\,\Gamma[\nu+(1/2)]}\int_0^{\pi/2}\cos(z\cos t)\sin^{2\nu}t\,dt \qquad \left(\mathrm{Re}\,\nu > -\frac{1}{2}\right),$$
$$N_\nu(z) = \frac{2(z/2)^\nu}{\sqrt{\pi}\,\Gamma[\nu+(1/2)]}\left[\int_0^{\pi/2}\sin(z\sin t)\cos^{2\nu}t\,dt - \int_0^\infty e^{-z\sinh t}\cosh^{2\nu}t\,dt\right]$$
$$(\mathrm{Re}\,z>0, \quad \mathrm{Re}\,\nu > -1/2).$$

Schläfli's formula
$$J_\nu(z) = \frac{1}{\pi}\int_0^\pi \cos(z\sin t - \nu t)dt - \frac{\sin\nu\pi}{\pi}\int_0^\infty e^{-z\sinh t}e^{-\nu t}dt \quad (\mathrm{Re}\,z>0),$$
$$N_\nu(z) = \frac{1}{\pi}\int_0^\pi \sin(z\sin t - \nu t)dt - \frac{1}{\pi}\int_0^\infty e^{-z\sinh t}[e^{\nu t} + (\cos\nu\pi)e^{-\nu t}]dt \quad (\mathrm{Re}\,z>0).$$

$$J_\nu(z) = \frac{z^\nu}{2\pi i}\int_{c-i\infty}^{c+i\infty}\exp\left[\frac{1}{2}\left(t - \frac{z^2}{t}\right)\right]t^{-\nu-1}dt \quad (c>0, \quad |\arg z|<\pi, \quad \mathrm{Re}\,\nu > -1).$$

$$J_\nu(x) = \frac{2(x/2)^{-\nu}}{\sqrt{\pi}\,\Gamma[(1/2)-\nu]}\int_1^\infty \frac{\sin xt}{(t^2-1)^{\nu+(1/2)}}dt,$$

$$N_\nu(x) = -\frac{2(x/2)^{-\nu}}{\sqrt{\pi}\,\Gamma[(1/2)-\nu]}\int_1^\infty \frac{\cos xt}{(t^2-1)^{\nu+(1/2)}}dt \quad \left(x>0, \quad -\frac{1}{2} < \mathrm{Re}\,\nu < \frac{1}{2}\right).$$

$$J_\nu(z) = \frac{1}{2\pi i}\int_{-\infty}^{(0+)} e^{z[t-(1/t)]/2}t^{-\nu-1}dt \quad (\mathrm{Re}\,z>0).$$

(The contour goes once around the negative real axis in the positive direction.)

Sommerfeld's formula
$$J_\nu(z) = \frac{1}{2\pi}\int_{-\eta+i\infty}^{2\pi-\eta+i\infty} e^{iz\cos t}e^{i\nu[t-(\pi/2)]}dt,$$

$$H_\nu^{(1)}(z) = \frac{1}{\pi}\int_{-\eta+i\infty}^{\eta-i\infty} e^{iz\cos t}e^{i\nu[t-(\pi/2)]}dt,$$

$$H_\nu^{(2)}(z) = \frac{1}{\pi}\int_{\eta-i\infty}^{2\pi-\eta+i\infty} e^{iz\cos t}e^{i\nu[t-(\pi/2)]}dt \qquad (-\eta < \arg z < \pi-\eta, \qquad 0<\eta<\pi).$$

$$H_\nu^{(1)}(z) = -\frac{2i}{\pi}e^{-i\nu\pi/2}\int_0^\infty e^{iz\cosh t}\cosh\nu t\,dt \quad (0<\arg z<\pi;\ \text{when } \nu=0, \text{ it holds also at } z=0).$$

$$H_\nu^{(1)}(z) = -\frac{2ie^{-i\nu\pi}(z/2)^\nu}{\sqrt{\pi}\,\Gamma[\nu+(1/2)]}\int_0^\infty e^{iz\cosh t}\sinh^{2\nu}t\,dt$$
$$(0<\arg z<\pi, \quad \mathrm{Re}\,\nu > -1/2;\ \text{when } z=0,\ -1/2<\mathrm{Re}\,\nu<1/2).$$

$$H_\nu^{(1)}(z) = -i\frac{e^{-i\nu\pi/2}}{\pi}\int_0^\infty e^{iz[t-(1/t)]/2}t^{-\nu-1}\,dt \qquad (0<\arg z<\pi;\ \text{when } \arg z=0,\ -1<\operatorname{Re}\nu<1).$$

(3) Generating Function.

$$\exp\left[\frac{z(t-t^{-1})}{2}\right] = J_0(z) + \sum_{n=1}^{\infty}[t^n + (-t)^{-n}]J_n(z),$$

$$\exp(iz\cos\theta) = \sum_{n=-\infty}^{\infty} i^n J_n(z)e^{in\theta} = J_0(z) + 2\sum_{n=1}^{\infty} i^n J_n(z)\cos n\theta.$$

$$\int J_\nu(z)\,dz = 2\sum_{n=0}^{\infty} J_{\nu+2n+1}(z).$$

Kapteyn's series $\displaystyle \frac{1}{1-z} = 1 + 2\sum_{n=1}^{\infty} J_n(nz),$

$$\frac{1}{2}\frac{z^2}{1-z^2} = \sum_{n=1}^{\infty} J_{2n}(2nz) \qquad \left(\left|\frac{z\exp\sqrt{1-z^2}}{1+\sqrt{1-z^2}}\right|<1\right).$$

Schlömilch's series. Supposing that $f(x)$ is twice continuously differentiable with respect to the real variable x in $0\leqslant x\leqslant\pi$, we have

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n J_0(nx) \quad (0<x<\pi),$$

where $\displaystyle a_0 \equiv 2f(0) + \frac{2}{\pi}\int_0^\pi du\int_0^{\pi/2} f'(u\sin\varphi)\,d\varphi,$

$$a_n \equiv \frac{2}{\pi}\int_0^\pi du\int_0^{\pi/2} uf'(u\sin\varphi)\cos n\varphi\,d\varphi.$$

$$1 = J_0(z) + 2\sum_{n=1}^{\infty} J_{2n}(z) = [J_0(z)]^2 + 2\sum_{n=1}^{\infty}[J_n(z)]^2.$$

(4) Addition Theorem. For the cylindrical function Z_ν, we have

$$e^{i\nu\psi}Z_\nu(kR) = \sum_{n=-\infty}^{\infty} J_n(k\rho)Z_{\nu+n}(kr)e^{in\varphi}$$

$$\left(R=\sqrt{r^2+\rho^2-2r\rho\cos\varphi},\quad 0<\psi<\frac{\pi}{2},\qquad e^{2i\psi}=\frac{r-\rho e^{-i\varphi}}{r-\rho e^{i\varphi}},\quad 0<\rho<r,\right.$$

k is an arbitrary complex number),

$$\frac{Z_\nu(kR)}{R^\nu} = 2^\nu k^{-\nu}\Gamma(\nu)\sum_{m=0}^{\infty}(\nu+m)\frac{J_{\nu+m}(k\rho)}{\rho^\nu}\,\frac{Z_{\nu+m}(kr)}{r^\nu}\,C_m^{(\nu)}(\cos\varphi)$$

($\nu\neq$ negative integer).

$$\frac{\exp\left[(-1)^{\iota+1}ikR\right]}{R} = \frac{\pi}{2}\frac{(-1)^{\iota+1}i}{\sqrt{r\rho}}\sum_{m=0}^{\infty}(2m+1)^{\iota}J_{m+(1/2)}(k\rho)H_{m+(1/2)}^{(\iota)}(kr)P_m(\cos\varphi)$$

$(\iota=1,2)$.

$$e^{ik\rho\cos\varphi} = \left(\frac{\pi}{2k\rho}\right)^{1/2}\sum_{m=0}^{\infty} i^m(2m+1)J_{m+(1/2)}(k\rho)P_m(\cos\varphi)$$

$$= 2^\nu\Gamma(\nu)\sum_{m=0}^{\infty}(\nu+m)i^m J_{\nu+m}(k\rho)(k\rho)^{-\nu}C_m^{(\nu)}(\cos\varphi) \qquad (\nu\neq 0,-1,-2,\ldots),$$

where P_m is a Legendre polynomial, and $C_m^{(\nu)}$ is a Gegenbauer polynomial.

(5) Infinite Products and Partial Fractions. Let $j_{\nu,n}$ be the zeros of $z^{-\nu}J_\nu(z)$ in ascending order with respect to the real part. We have

$$J_\nu(z) = \frac{(z/2)^\nu}{\Gamma(\nu+1)}\prod_{n=1}^{\infty}\left(1-\frac{z^2}{j_{\nu,n}^2}\right) \qquad (\nu\neq -1,-2,-3,\ldots).$$

Note that if ν is real and greater than -1, all zeros are real.

Kneser-Sommerfeld formula

$$\frac{\pi J_\nu(xz)}{4J_\nu(z)}[J_\nu(z)N_\nu(Xz)-N_\nu(z)J_\nu(Xz)]=\sum_{n=1}^{\infty}\frac{J_\nu(j_{\nu,n}x)J_\nu(j_{\nu,n}X)}{(z^2-j_{\nu,n}^2)J_{\nu,n}'^2(j_{\nu,n})}$$

$$(0<x<X<1,\quad \operatorname{Re} z>0).$$

(6) Definite Integrals.

$$\int_0^{\pi/2}J_\nu(z\cos\theta)\cos\theta\,d\theta=\frac{1}{2z}\int_0^{2z}J_{2\nu}(t)\,dt,\qquad \int_0^z J_\mu(t)\,dt=\frac{1}{\pi}\int_0^\pi\frac{\sin(z\sin\theta)}{\sin\theta}\cos\mu\theta\,d\theta.$$

$$\int_0^\infty e^{-at}J_\nu(bt)t^{\mu-1}\,dt=\frac{(b/2a)^\nu\Gamma(\mu+\nu)}{a^\mu\Gamma(\nu+1)}{}_2F_1\left(\frac{\mu+\nu}{2},\ \frac{\mu+\nu+1}{2};\ \nu+1;\ \frac{-b^2}{a^2}\right)$$

$$(\operatorname{Re}(a+ib)>0,\quad \operatorname{Re}(a-ib)>0,\quad \operatorname{Re}(\mu+\nu)>0).$$

$$\int_0^\infty e^{-at}J_\nu(bt)t^\nu\,dt=\frac{(2x)^\nu\Gamma[\nu+(1/2)]}{(a^2+b^2)^{\nu+(1/2)}\sqrt{\pi}}\qquad\left(\operatorname{Re}\nu>-\frac{1}{2},\quad \operatorname{Re}a>|\operatorname{Im}b|\right).$$

$$\int_0^\infty e^{-at}J_\nu(bt)\frac{dt}{t}=\frac{(\sqrt{a^2+b^2}-a)^\nu}{\nu b^\nu}\qquad(\operatorname{Re}\nu>0,\quad \operatorname{Re}a>|\operatorname{Im}b|).$$

Sommerfeld's formula
$$\int_0^\infty J_0(\tau r)e^{-|x|\sqrt{\tau^2-k^2}}\frac{\tau\,d\tau}{\sqrt{\tau^2-k^2}}=\frac{e^{ik\sqrt{r^2+k^2}}}{\sqrt{r^2+x^2}}$$

$$(r,x\text{ real};\quad -\pi/2\leqslant\arg\sqrt{\tau^2-k^2}<\pi/2,\quad 0\leqslant\arg k<\pi).$$

Weyrich's formula
$$\frac{i}{2}\int_{-\infty}^{+\infty}e^{i\tau x}H_0^{(1)}(r\sqrt{k^2-\tau^2})\,d\tau=\frac{e^{ik\sqrt{r^2+x^2}}}{\sqrt{r^2+x^2}}$$

$$(r,x\text{ real};\quad 0\leqslant\arg\sqrt{k^2-\tau^2}<\pi,\quad 0\leqslant\arg k<\pi).$$

Weber-Sonine formula

$$\int_0^\infty J_\nu(at)e^{-p^2t^2}t^{\mu-1}\,dt=\frac{(a/2p)^\nu\Gamma[(\nu+\mu)/2]}{2p^\mu\Gamma(\nu+1)}{}_1F_1\left(\frac{\nu+\mu}{2};\ \nu+1;\ \frac{-a^2}{4p^2}\right)$$

$$(\operatorname{Re}(\mu+\nu)>0,\quad |\arg p|<\pi/4,\quad a>0),$$

$$\int_0^\infty J_\nu(at)e^{-p^2t^2}t^{\nu+1}\,dt=\frac{a^\nu}{(2p^2)^{\nu+1}}e^{-a^2/4p^2}\qquad(\operatorname{Re}\nu>-1,\quad |\arg p|<\pi/4).$$

Sonine-Schafheitlin formula

$$\int_0^\infty J_\mu(at)J_\nu(bt)t^{-\lambda}\,dt=\frac{a^\mu\Gamma[(\mu+\nu-\lambda+1)/2]}{2^\lambda b^{\mu-\lambda+1}\Gamma[(-\mu+\nu+\lambda+1)/2]\Gamma(\mu+1)}$$

$$\times{}_2F_1\left(\frac{\mu+\nu-\lambda+1}{2},\ \frac{\mu-\nu-\lambda+1}{2};\ \mu+1;\ \frac{a^2}{b^2}\right)$$

$$(\operatorname{Re}(\mu+\nu-\lambda+1)>0,\quad \operatorname{Re}\lambda>-1,\quad 0<a<b).$$

(7) Asymptotic Expansion.
(i) Hankel's asymptotic representation. We put

$$(\nu,m)\equiv\frac{[4\nu^2-1^2][4\nu^2-3^2]\ldots[4\nu^2-(2m-1)^2]}{2^{2m}m!}\qquad(m=1,2,3,\ldots);\qquad(\nu,0)\equiv1.$$

For $|z|\gg|\nu|$, $|z|\gg1$,

$$J_\nu(z)=\sqrt{\frac{2}{\pi z}}\cos\left(z-\frac{\nu\pi}{2}-\frac{\pi}{4}\right)\left[\sum_{m=0}^{M-1}(-1)^m\frac{(\nu,2m)}{(2z)^{2m}}+O(|z|^{-2M})\right]$$

$$-\sqrt{\frac{2}{\pi z}}\sin\left(z-\frac{\nu\pi}{2}-\frac{\pi}{4}\right)\left[\sum_{m=0}^{M-1}(-1)^m\frac{(\nu,2m+1)}{(2z)^{2m+1}}+O(|z|^{-2M-1})\right]$$

$$(-\pi<\arg z<\pi),$$

$$N_\nu(z)=\sqrt{\frac{2}{\pi z}}\,\sin\left(z-\frac{\nu\pi}{2}-\frac{\pi}{4}\right)\left[\sum_{m=0}^{M-1}(-1)^m\frac{(\nu,2m)}{(2z)^{2m}}+O(|z|^{-2M})\right]$$

$$+\sqrt{\frac{2}{\pi z}}\,\cos\left(z-\frac{\nu\pi}{2}-\frac{\pi}{4}\right)\left[\sum_{m=0}^{M-1}(-1)^m\frac{(\nu,2m+1)}{(2z)^{2m+1}}+O(|z|^{-2M-1})\right]$$

$$(-\pi<\arg z<\pi),$$

$$H_\nu^{(1)}(z)=\sqrt{\frac{2}{\pi z}}\,\exp\left[i\left(z-\frac{\nu\pi}{2}-\frac{\pi}{4}\right)\right]\left[\sum_{m=0}^{M-1}\frac{(\nu,m)}{(-2iz)^m}+O(|z|^{-M})\right]\qquad(-\pi<\arg z<2\pi),$$

$$H_\nu^{(2)}(z)=\sqrt{\frac{2}{\pi z}}\,\exp\left[-i\left(z-\frac{\nu\pi}{2}-\frac{\pi}{4}\right)\right]\left[\sum_{m=0}^{M-1}\frac{(\nu,m)}{(2iz)^m}+O(|z|^{-M})\right]\qquad(-2\pi<\arg z<\pi).$$

(ii) Debye's asymptotic representation.

$\nu \doteqdot x,\ 1-(\nu/x)>\varepsilon,\ \nu/x=\sin\alpha$, when $1-(\nu/x)>(3/x)\nu^{1/2}$,

$$H_\nu^{(1)}(x)\sim\frac{1}{\sqrt{\pi}}\exp\left[ix\left\{\cos\alpha+\left(\alpha-\frac{\pi}{2}\right)\sin\alpha\right\}\right]$$

$$\times\left[\frac{e^{i\pi/4}}{X}+\left(\frac{1}{8}+\frac{5}{24}\tan^2\alpha\right)\frac{3e^{3\pi i/4}}{2X^3}\right.$$

$$\left.+\left(\frac{3}{128}+\frac{77}{576}\tan^2\alpha+\frac{385}{3456}\tan^4\alpha\right)\frac{3\cdot 5e^{5\pi i/4}}{2^2X^5}+\ldots\right]$$

$$(X=[-x\cos(\alpha/2)]^{1/2}).$$

$\nu\doteqdot x,\quad (\nu/x)-1>\varepsilon,\quad \nu/x=\cosh\sigma,\quad$ when $\quad|\nu^2-x^2|^{1/2}\gg 1,\quad |\nu^2-x^2|^{3/2}\nu^{-2}\gg 1$

$$H_\nu^{(1)}(x)\sim\frac{1}{\sqrt{\pi}}\exp[x(\sigma\cosh\sigma-\sinh\sigma)]$$

$$\times\left[\frac{1}{X}+\left(\frac{1}{8}-\frac{5}{24}\coth^2\sigma\right)\frac{3}{2X^3}+\left(\frac{3}{128}-\frac{77}{576}\coth^2\sigma+\frac{385}{3456}\coth^4\sigma\right)\frac{3\cdot 5}{2^2X^5}+\ldots\right]$$

$$(X=[-x\sinh(\sigma/2)]^{1/2}).$$

When $\nu\doteqdot x,\quad |x-\nu|\ll x^{1/3},\quad x\gg 1,\quad x-\nu=\delta,$

$$H_\nu^{(2)}(x)\sim\frac{6^{1/3}e^{i\pi/3}}{3^{1/2}\pi}\left[\frac{\Gamma(1/3)}{x^{1/3}}-6^{1/3}e^{i\pi/3}\delta\frac{\Gamma(2/3)}{x^{2/3}}+\left(\frac{2}{5}\delta-\delta^3\right)\frac{\Gamma(4/3)}{x^{4/3}}\right.$$

$$\left.+\left(\frac{3}{140}-\frac{\delta^2}{4}+\frac{\delta^4}{4}\right)6^{1/3}e^{i\pi/3}\frac{\Gamma(5/3)}{x^{5/3}}+\ldots\right]$$

(iii) Watson-Nicholson formula. When $x,\nu>0,\quad w=[(x/\nu)^2-1]^{1/2}$,

$$H_\nu^{(\iota)}(x)=3^{-1/2}w\exp[(-1)^{\iota+1}i\{(\pi/6)+\nu(w-(w^3/3)-\arctan w)\}]H_{1/3}^{(\iota)}(\nu w^3/3)+O|\nu^{-1}|$$

$$(\iota=1,2).$$

(IV) Functions Related to Bessel Functions

(1) Modified Bessel Functions.

$$I_\nu(z)\equiv e^{-i\nu\pi/2}J_\nu(e^{i\pi/2}z)$$

$$=\sum_{n=0}^{\infty}\frac{(z/2)^{\nu+2n}}{n!\Gamma(\nu+n+1)},$$

$$K_\nu(z)\equiv\frac{i\pi}{2}e^{i\nu\pi/2}H_\nu^{(1)}(e^{i\pi/2}z)=-\frac{i\pi}{2}e^{-i\nu\pi/2}H_\nu^{(2)}(e^{-i\pi/2}z)$$

$$=\frac{\pi}{2}\,\frac{I_{-\nu}(z)-I_\nu(z)}{\sin\nu\pi}=\left(\frac{\pi}{2z}\right)^{1/2}W_{0,\nu}(2z).$$

Recurrence formulas:

$$I_{\nu-1}(z)-I_{\nu+1}(z)=(2\nu/z)I_\nu(z),$$

$$I_{\nu-1}(z)+I_{\nu+1}(z)=2I_\nu'(z).$$

$$K_{\nu-1}(z)-K_{\nu+1}(z)=-(2\nu/z)K_\nu(z),$$

$$K_{\nu-1}(z)+K_{\nu+1}(z)=-2K_\nu'(z),$$

$$K_{-\nu}(z)=K_\nu(z).$$

Airy's integral: $\int_0^\infty \cos(t^3-tx)\,dt=\frac{\pi}{3}\sqrt{\frac{x}{3}}\left[J_{1/3}\left(\frac{2x\sqrt{x}}{3\sqrt{3}}\right)+J_{-1/3}\left(\frac{2x\sqrt{x}}{3\sqrt{3}}\right)\right],$

$$\int_0^\infty \cos(t^3+tx)\,dt=\frac{1}{3}\sqrt{x}\,K_{1/3}\left(\frac{2x\sqrt{x}}{3\sqrt{3}}\right)\qquad(x>0).$$

H. Weber's formula: $\frac{1}{2p^2}e^{-(a^2+b^2)/4p^2}I_\nu\left(\frac{ab}{2p^2}\right)=\int_0^\infty e^{-p^2t^2}J_\nu(at)J_\nu(bt)\,t\,dt$

$$(\operatorname{Re}\nu>-1,\quad |\arg p|<\pi/4;\quad a,b>0).$$

Watson's formula: $J_\mu(z)N_\nu(z)-J_\nu(z)N_\mu(z)=\frac{4\sin(\mu-\nu)\pi}{\pi^2}\int_0^\infty K_{\nu-\mu}(2z\sinh t)e^{(\mu+\nu)t}\,dt$

$$(\operatorname{Re}z>0,\quad \operatorname{Re}(\mu-\nu)<1),$$

$$J_\nu(z)\frac{\partial N_\nu(z)}{\partial\nu}-N_\nu(z)\frac{\partial J_\nu(z)}{\partial\nu}=-\frac{4}{\pi}\int_0^\infty K_0(2z\sinh t)e^{-2\nu t}\,dt\qquad(\operatorname{Re}z>0).$$

Nicholson's formula: $J_\nu^2(z)+N_\nu^2(z)=\frac{8}{\pi^2}\int_0^\infty K_0(2z\sinh t)\cosh 2\nu t\,dt\qquad(\operatorname{Re}z>0).$

Dixon-Ferrar formula: $J_\nu^2(z)+N_\nu^2(z)=\frac{8\cos\nu\pi}{\pi^2}\int_0^\infty K_{2\nu}(2z\sinh t)\,dt$

$$\left(\operatorname{Re}z>0;\qquad -\tfrac{1}{2}<\operatorname{Re}\nu<\tfrac{1}{2}\right).$$

(2) Kelvin Functions.

$$\operatorname{ber}_\nu(z)\pm i\operatorname{bei}_\nu(z)\equiv J_\nu(e^{\pm 3\pi i/4}z),$$

$$\operatorname{her}_\nu(z)\pm i\operatorname{hei}_\nu(z)\equiv H_\nu^{(1)}(e^{\pm 3\pi i/4}z),$$

$$\operatorname{ker}_\nu(z)\equiv-(\pi/2)\operatorname{hei}_\nu(z),$$

$$\operatorname{kei}_\nu(z)\equiv(\pi/2)\operatorname{her}_\nu(z).$$

When ν is an integer n, $\operatorname{ber}_n(x)-i\operatorname{bei}_n(x)=(-1)^nJ_n(\sqrt{i}\,x),$

$$\operatorname{her}_n(x)-i\operatorname{hei}_n(x)=(-1)^{n+1}H_n^{(1)}(\sqrt{i}\,x)\qquad(x\text{ real}).$$

(3) Struve Function. $\boldsymbol{H}_\nu(x)\equiv\frac{2(z/2)^\nu}{\Gamma[\nu+(1/2)]\sqrt{\pi}}\int_0^{\pi/2}\sin(z\cos\theta)\sin^{2\nu}\theta\,d\theta$

$$=\sum_{m=0}^{\infty}\frac{(-1)^m(z/2)^{\nu+2m+1}}{\Gamma[m+(3/2)]\Gamma[\nu+m+(3/2)]}.$$

Anger function: $\boldsymbol{J}_\nu(z)\equiv\frac{1}{\pi}\int_0^\pi\cos(\nu\theta-z\sin\theta)\,d\theta.$

H. F. Weber function: $\boldsymbol{E}_\nu(z)\equiv\frac{1}{\pi}\int_0^\pi\sin(\nu\theta-z\sin\theta)\,d\theta.$

Putting $\nabla_\nu\equiv z^2\frac{d^2}{dz^2}+z\frac{d}{dz}+z^2-\nu^2,$

$$\nabla_\nu\boldsymbol{H}_\nu(z)=\frac{4(z/2)^{\nu+1}}{\Gamma[\nu+(1/2)]\sqrt{\pi}},\quad \nabla_\nu\boldsymbol{J}_\nu(z)=\frac{(z-\nu)\sin\nu\pi}{\pi},$$

$$\nabla_\nu\boldsymbol{E}_\nu(z)=-\frac{z+\nu}{\pi}-\frac{(z-\nu)\cos\nu\pi}{\pi}.$$

When ν is an integer n, $J_n(z)=J_n(z)$.

$$\int_0^z J_0(t)\,dt = zJ_0(z)+\frac{\pi z}{2}[J_1(z)\boldsymbol{H}_0(z)-J_0(z)\boldsymbol{H}_1(z)],$$

$$\int_0^z N_0(t)\,dt = zN_0(z)+\frac{\pi z}{2}[N_1(z)\boldsymbol{H}_0(z)-N_0(z)\boldsymbol{H}_1(z)].$$

(4) Neumann Polynomials. $O_n(t)=\sum_{j=0}^{[n/2]}\frac{n(n-j-1)!}{j!(t/2)^{n-2j+1}}$ (n is a positive integer),

$$O_0(t)=1/t.$$

$$\frac{1}{t-z}\equiv 1+2\sum_{n=1}^{\infty}O_n(t)J_n(z) \qquad (|t|>|z|).$$

Schläfli polynomials: $S_n(t)\equiv\frac{2}{n}\left[tO_n(t)-\cos^2\frac{n\pi}{2}\right]$ (n is a positive integer),

$$S_0(t)\equiv 0.$$

$\nabla_n S_n(x)=2n+2(x-n)\sin^2(n\pi/2)$ (∇_n is the same operator defined in (3)).

Lommel polynomials: $R_{m,\nu}(z)\equiv\frac{\Gamma(\nu+m)}{\Gamma(\nu)(z/2)^m}\,{}_2F_3\left(\frac{1-m}{2},\ \frac{-m}{2};\ \nu,\ -m,\ 1-\nu-m;\ -z^2\right)$

$$=(\pi z/2\sin\nu\pi)[J_{\nu+m}(z)J_{-\nu+1}(z)+(-1)^m J_{-\nu-m}(z)J_{\nu-1}(z)]$$

(m is a nonnegative integer).

References

See references to Table 16, this Appendix.

20. Systems of Orthogonal Functions (→ 317 Orthogonal Functions)

$$\int_a^b p_n(x)p_m(x)\varphi(x)\,dx=\delta_{nm}A_n$$

Name	Notation $p_n(x)$	Interval (a,b)	Weight $\varphi(x)$	Norm A_n
Legendre	$P_n(x)$	$(-1,+1)$	1	$2/(2n+1)$
Gegenbauer	$C_n^\nu(x)$	$(-1,+1)$	$(1-x^2)^{\nu-(1/2)}$	$2\pi\Gamma(2\nu+n)/2^{2\nu}(n+\nu)n![\Gamma(\nu)]^2$
Chebyshev	$T_n(x)$	$(-1,+1)$	$(1-x^2)^{-1/2}$	$\pi(n=0)$; $\pi/2\ (n\geqslant 1)$
Hermite	$H_n(x)$	$(-\infty,+\infty)$	e^{-x^2}	$\sqrt{\pi}\cdot n!$
Jacobi	$G_n(\alpha,\gamma;x)$	$(0,1)$	$x^{\gamma-1}(1-x)^{\alpha-\gamma}$	$\dfrac{n![\Gamma(\gamma)]^2\Gamma(\alpha+n-\gamma+1)}{(\alpha+2n)\Gamma(\alpha+n)\Gamma(\gamma+n)}$
Laguerre	$L_n^\alpha(x)$	$(0,\infty)$	$x^\alpha e^{-x}$	$\Gamma(\alpha+n+1)/n!$

For Legendre polynomials $P_n(x)$ → Table 18.II, this Appendix.

(I) Gegenbauer Polynomials (Gegenbauer Functions)

$$C_n^\nu(t)\equiv\frac{\Gamma(n+2\nu)}{n!\Gamma(2\nu)}\,{}_2F_1\left(n+2\nu,\ -n;\ \nu+\frac{1}{2};\ \frac{1-t}{2}\right)$$

$$=\frac{\Gamma(2\nu+n)\Gamma[\nu+(1/2)]}{\Gamma(2\nu)n!}\left[\frac{1}{4}(t^2-1)\right]^{(1/4)-(\nu/2)}\mathfrak{P}_{n+\nu-(1/2)}^{(1/2)-\nu}(t).$$

Generating function $(1-2\alpha t+\alpha^2)^{-\nu}\equiv\sum_{n=0}^{\infty}C_n^{\nu}(t)\alpha^n.\quad C_{n-l}^{l+(1/2)}(x)=\frac{1}{(2l-1)!!}\frac{d^lP_n(t)}{dt^l}.$

Orthogonal relation $\int_0^{\pi}(\sin^{2\nu}\theta)C_m^{\nu}(\cos\theta)C_n^{\nu}(\cos\theta)d\theta=\frac{\pi\Gamma(2\nu+n)}{2^{2\nu-1}(\nu+n)n![\Gamma(\nu)]^2}\delta_{nm}.$

(II) Chebyshev (Tschebyscheff) Polynomials

(1) Chebyshev Polynomial (Chebyshev Function of the First Kind)

$$T_n(x)\equiv\cos(n\arccos x)$$
$$=(1/2)\left[\left(x+i\sqrt{1-x^2}\right)^n+\left(x-i\sqrt{1-x^2}\right)^n\right]$$
$$=F(n,-n;\ 1/2;\ (1-x)/2)$$
$$=\sum_{j=0}^{[n/2]}(-1)^j\binom{n}{2j}x^{n-2j}(1-x^2)^j$$
$$=\frac{(-1)^n(1-x^2)^{1/2}}{(2n-1)!!}\frac{d^n(1-x^2)^{n-(1/2)}}{dx^n},$$

Chebyshev function of the second kind

$$U_n(x)\equiv\sin(n\arccos x)$$
$$=(1/2i)\left[\left(x+i\sqrt{1-x^2}\right)^n-\left(x-i\sqrt{1-x^2}\right)^n\right]$$
$$=\frac{(-1)^{n-1}n}{(2n-1)!!}\frac{d^{n-1}(1-x^2)^{n-(1/2)}}{dx^{n-1}}.$$

$T_n(x)$, $U_n(x)$ are mutually linearly independent solutions of Chebyshev's differential equation $(1-x^2)y''-xy'+n^2y=0$. Recurrence relations are

$$T_{n+1}(x)-2xT_n(x)+T_{n-1}(x)=0,\quad U_{n+1}(x)-2xU_n(x)+U_{n-1}(x)=0.$$

Generating function:

$$\frac{1-t^2}{1-2tx+t^2}=T_0(x)+2\sum_{n=0}^{\infty}T_n(x)t^n,\quad\frac{1}{1-2tx+t^2}=\frac{1}{\sqrt{1-x^2}}\sum_{n=0}^{\infty}U_{n+1}(x)t^n.$$

Orthogonal relation:

$$\int_{-1}^{+1}\frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}dx=\begin{cases}0 & (m\neq n),\\ \pi/2 & (m=n\neq 0),\\ \pi & (m=n=0);\end{cases}\quad\int_{-1}^{+1}\frac{U_m(x)U_n(x)}{\sqrt{1-x^2}}dx=\begin{cases}\frac{\pi}{2} & (m=n\neq 0),\\ 0 & (\text{otherwise}).\end{cases}$$

Orthogonality in finite sums. Let $u_0, u_1, \ldots, u_k$ be the zeros of $T_{k+1}(x)$. All zeros are real and situated in the interval $(-1,1)$. Then we have

$$\sum_{i=0}^{k}T_m(u_i)T_n(u_i)=\begin{cases}0 & (m\neq n,\text{ or } m=n=k+1),\\ (k+1)/2 & (1\leqslant m=n\leqslant k),\\ k+1 & (m=n=0).\end{cases}$$

Let $p_n(x)$ be the best approximation of x^n in $-1\leqslant x\leqslant 1$ by polynomials of degree at most $n-1$. Then we have $x^n-p_n(x)=2^{-n+1}T_n(x)$.

(2) Expansions by $T_n(x)$.

$$e^{ax}=I_0(a)+2\sum_{n=1}^{\infty}I_n(a)T_n(x),$$

$$\sin ax=2\sum_{n=0}^{\infty}(-1)^nJ_{2n+1}(a)T_{2n+1}(x),$$

$$\cos ax=J_0(a)+2\sum_{n=1}^{\infty}(-1)^nJ_{2n}(a)T_{2n}(x),$$

$$\log(1+x\sin 2\alpha)=2\log\cos\alpha-2\sum_{n=1}^{\infty}\frac{1}{n}(-\tan\alpha)^n T_n(x),$$

$$\arctan x=2\sum_{n=1}^{\infty}(-1)^n\frac{(\sqrt{2}-1)^{2n+1}}{2n+1}T_{2n+1}(x).$$

(III) Parabolic Cylinder Functions (Weber Functions) (→ 167 Functions of Confluent Type)

Parabolic cylinder functions:

$$D_\nu(z)\equiv 2^{(1/4)+(\nu/2)}z^{-1/2}W_{(1/4)+(\nu/2),\,-1/4}(z^2/2)$$

$$=\sqrt{\pi}\,2^{(1/4)+(\nu/2)}z^{-1/2}\left[\frac{M_{(1/4)+(\nu/2),\,-1/4}(z^2/2)}{\Gamma[(1-\nu)/2]}+\frac{M_{(1/4)-(\nu/2),\,-1/4}(z^2/2)}{\Gamma(-\nu/2)}\right]$$

$$=2^{\nu/2}e^{-z^2/4}\sqrt{\pi}\left[\frac{1}{\Gamma[(1-\nu)/2]}{}_1F_1\left(\frac{-\nu}{2};\frac{1}{2};\frac{z^2}{2}\right)-\frac{\sqrt{2}\,z}{\Gamma(-\nu/2)}{}_1F_1\left(\frac{1-\nu}{2};\frac{3}{2};\frac{z^2}{2}\right)\right].$$

The solutions of Weber's differential equation

$$\frac{d^2u}{dz^2}+\left(\nu+\frac{1}{2}-\frac{z^2}{4}\right)u=0$$

are given by

$$D_\nu(z),\ D_\nu(-z),\ D_{-\nu-1}(iz),\ D_{-\nu-1}(-iz),$$

and the following relations hold among them.

$$D_\nu(z)=\left[\Gamma(\nu+1)/\sqrt{2\pi}\right]\left[e^{i\nu\pi/2}D_{-\nu-1}(iz)+e^{-i\nu\pi/2}D_{-\nu-1}(-iz)\right]$$

$$=e^{-i\nu\pi}D_\nu(-z)+\left[\sqrt{2\pi}/\Gamma(-\nu)\right]e^{-i(\nu+1)\pi/2}D_{-\nu-1}(iz)$$

$$=e^{i\nu\pi}D_\nu(-z)+\left[\sqrt{2\pi}/\Gamma(-\nu)\right]e^{i(\nu+1)\pi/2}D_{-\nu-1}(-iz).$$

Integral representation:

$$D_\nu(z)=\frac{e^{-z^2/4}}{\Gamma(-\nu)}\int_0^\infty e^{-zt-(t^2/2)}t^{-\nu-1}dt\quad(\mathrm{Re}\,\nu<0).$$

$$e^{-(z^2/4)-zt-(t^2/2)}=\sum_{n=0}^{\infty}\frac{(-t)^n}{n!}D_n(z)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}t^\nu\Gamma(-\nu)D_\nu(z)d\nu\quad(c<0,\ |\arg t|<\pi/4).$$

Recurrence formula:

$$D_{\nu+1}(z)-zD_\nu(z)+\nu D_{\nu-1}(z)=0,\quad dD_\nu(z)/dz+(1/2)zD_\nu(z)-\nu D_{\nu-1}(z)=0.$$

$$D_\nu(0)=\frac{2^{\nu/2}\sqrt{\pi}}{\Gamma[(1-\nu)/2]},\quad D_\nu'(0)=-\frac{2^{(\nu+1)/2}\sqrt{\pi}}{\Gamma(-\nu/2)}.$$

Asymptotic expansion:

$$D_\nu(z)\approx e^{-z^2/4}z^\nu\left(1-\frac{\nu(\nu-1)}{2z^2}+\frac{\nu(\nu-1)(\nu-2)(\nu-3)}{2\cdot 4z^4}\mp\ldots\right)\quad\left(|\arg z|<\frac{3}{4}\pi\right).$$

$$D_{-1}(z)=e^{z^2/4}\sqrt{\frac{\pi}{2}}\left[1-\mathrm{erf}\left(\frac{z}{\sqrt{2}}\right)\right],\quad \mathrm{erf}(x)\equiv\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt\quad\text{(error function)}.$$

(IV) Hermite Polynomials

For the parabolic cylinder functions, when ν is an integer n, we have

$$D_n(z)=(-1)^n e^{z^2/4}d^n(e^{-z^2/2})/dz^n=e^{-z^2/4}H_n(z/\sqrt{2}),$$

where $H_n(x)$ is the Hermite polynomial

$$H_n(x) \equiv 2^{-n/2}(-1)^n e^{x^2} d^n (e^{-x^2})/dx^n = e^{x^2/2} D_n(\sqrt{2}\, x).$$

A Hermite polynomial is more often defined by the following function $\mathrm{He}_n(x)$ (e.g., in W.F. Magnus, F. Oberhettinger, and R. P. Soni [1]).

$$\mathrm{He}_n(x) \equiv (-1)^n e^{x^2/2} d^n (e^{-x^2/2})/dx^n = e^{x^2/4} D_n(x) = H_n(x/\sqrt{2}\,).$$

The function $y = H_n(x)$ is a solution of Hermite's differential equation

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

$H_n(x)$ is a polynomial in x of degree n, and is an even or odd function according to whether n is even or odd.

$$H_{2n}(x) = (-1)^n (2n-1)!!\,{}_1F_1(-n; 1/2; x^2),$$

$$H_{2n+1}(x) = (-1)^n (2n+1)!!\sqrt{2}\, x\,{}_1F_1(-n; 3/2; x^2).$$

Recurrence formula:

$$H_{n+1}(x) = \sqrt{2}\, xH_n(x) - nH_{n-1}(x) = \sqrt{2}\, xH_n(x) - H_n'(x)/\sqrt{2}\,,$$

$$H_n'(x) = \sqrt{2}\, nH_{n-1}(x).$$

$$H_{2n}(0) = \frac{(-1)^n (2n)!}{2^n n!} = (-1)^n (2n-1)!!, \quad H_{2n+1}(0) = 0.$$

Generating function:

$$e^{\sqrt{2}\, tx - (t^2/2)} = \sum_{n=0}^{\infty} H_n(x) t^n / n!.$$

Orthogonal relation:

$$\int_{-\infty}^{+\infty} H_n(x) H_m(x) e^{-x^2} dx = \delta_{nm} n! \sqrt{\pi}\,.$$

(V) Jacobi Polynomials

$$G_n(\alpha, \gamma; x) \equiv F(-n, \alpha+n; \gamma; x)$$

$$= x^{1-\gamma}(1-x)^{\gamma-\alpha} \frac{\Gamma(\gamma+n)}{\Gamma(\gamma)} \frac{d^n}{dx^n}\left[x^{\gamma+n-1}(1-x)^{\alpha+n-\gamma}\right].$$

These satisfy Jacobi's differential equation $x(1-x)y'' + [\gamma - (\alpha+1)x]y' + n(\alpha+n)y = 0$.

Orthogonal relation:

$$\int_0^1 x^{\gamma-1}(1-x)^{\alpha-\gamma} G_m(\alpha, \gamma; x) G_n(\alpha, \gamma; x)\, dx = \frac{n!\Gamma(\alpha+n-\gamma+1)\Gamma(\gamma)^2}{(\alpha+2n)\Gamma(\alpha+n)\Gamma(\gamma+n)} \delta_{mn}$$

$(\mathrm{Re}\gamma > 0, \quad \mathrm{Re}(\alpha-\gamma) > -1).$

Representation of other functions:

$$P_n(x) = G\left(1, 1; \frac{1-x}{2}\right), \quad T_n(x) = G\left(0, \frac{1}{2}; \frac{1-x}{2}\right),$$

$$C_n^{\nu}(x) = (-1)^n \frac{\Gamma(2\nu+n)}{\Gamma(2\nu)\cdot n!} G_n\left(2\nu, \nu+\frac{1}{2}; \frac{1+x}{2}\right).$$

(VI) Laguerre Functions

(1) Laguerre Functions.

$$L_\nu^{(\alpha)}(z) \equiv \frac{\Gamma(\alpha+\nu+1)}{\Gamma(\alpha+1)\Gamma(\nu+1)} z^{-(\alpha+1)/2} e^{z/2} M_{[(\alpha+1)/2]+\nu,\alpha/2}(z)$$

$$= \frac{\Gamma(\alpha+\nu+1)}{\Gamma(\alpha+1)\Gamma(\nu+1)} {}_1F_1(-\nu;\alpha+1;z),$$

These satisfy Laguerre's differential equation

$$z d^2[L_\nu^{(\alpha)}(z)]/dz^2 + (\alpha+1-z) d[L_\nu^{(\alpha)}(z)]/dz + \nu L_\nu^{(\alpha)}(z) = 0.$$

(2) Laguerre Polynomials. When ν is an integer n ($n=0,1,2,\ldots$), the function $L_n^{(\alpha)}(x)$ reduces to a polynomial of degree n as follows.

Laguerre polynomials:

$$L_n^{(\alpha)}(x) = \frac{e^x x^{-\alpha}}{n!}\frac{d^n}{dx^n}(e^{-x}x^{n+\alpha}) = \sum_{j=0}^{n}\binom{n+\alpha}{n-j}\frac{(-x)^j}{j!}.$$

$$L_n^{(0)}(x)=1, \quad L_0^{(m)}(x)=1, \quad L_{n+m}^{(-m)}(x) = \frac{(-1)^m n!}{(n+m)!} x^m L_n^{(m)}(x) \quad (m=0,1,2,\ldots).$$

Recurrence formulas:

$$nL_n^{(\alpha)}(x) = (-x+2n+\alpha-1)L_{n-1}^{(\alpha)}(x) - (n+\alpha-1)L_{n-2}^{(\alpha)}(x),$$

$$x d[L_n^{(\alpha)}(x)]/dx = nL_n^{(\alpha)}(x) - (n+\alpha)L_{n-1}^{(\alpha)}(x) \quad (n=2,3,\ldots).$$

Generating function:

$$\frac{e^{-xt/(1-t)}}{(1-t)^{\alpha+1}} = \sum_{n=0}^{\infty} L_n^{(\alpha)}(x) t^n \quad (|t|<1).$$

Orthogonal relations:

$$\int_0^\infty e^{-x} x^\alpha L_m^{(\alpha)}(x) L_n^{(\alpha)} dx = \delta_{mn}\Gamma(\alpha+n+1)/n! = \delta_{mn}\Gamma(1+\alpha)\binom{n+\alpha}{n}.$$

$$H_{2n}(x) = (-2)^n n! L_n^{(-1/2)}(x^2), \quad H_{2n+1}(x) = (-2)^n n! \sqrt{2}\, x L_n^{(1/2)}(x^2).$$

(3) Sonine Polynomials.

$$S_n^{(\alpha)}(x) \equiv \frac{(-1)^n}{\Gamma(\alpha+n+1)} L_n^{(\alpha)}(x).$$

(VII) Orthogonal Polynomials

$$P_{n,m}(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k}\binom{n+k}{k} \frac{x(x-1)\ldots(x-k+1)}{m(m-1)\ldots(m-k+1)}$$

(where n, m are positive integers and $n \leqslant m$).

We have the same polynomials if we replace x^k in $P_n(1-2x)$ by

$$x(x-1)\ldots(x-k+1)/m(m-1)\ldots(m-k+1) \quad (k=0,1,\ldots,n).$$

Orthogonality in finite sums:

$$\sum_{k=0}^{m} P_{n,m}(k) P_{l,m}(k) = \delta_{nl} \frac{(m+n+1)!(m-n)!}{(2n+1)(m!)^2}.$$

Chebyshev's q functions:

$$q_n(m,x) = \frac{(-1)^n (m-1)!}{2^n (m-n-1)!} P_{n,m-1}(x), \quad \xi_{n,m}(x) = [2^n (n!)^2/(2n)!]\, q_n(m,x-1).$$

For given data y_k at m points $x_k = x_1 + (k-1)h$ ($k=1,\ldots,m$) that are equally spaced with step h, the least square approximation among the polynomials $Q(x)$ of degree $n(<m)$, i.e., the polynomial that minimizes the square sum of the residues $S = \sum_{k=1}^{m} [y_k - Q(x_k)]^2$ is given by the

following formula (→ 19 Analog Computation):

$$Q(x)=\sum_{k=0}^{n}\frac{B_k}{S_k}\xi_{k,m}\left(\frac{x-x_1}{h}+1\right),\quad S=\sum_{k=1}^{m}(y_k)^2-\sum_{k=0}^{m}\frac{B_k^2}{S_k},$$

$$B_k=\sum_{i=1}^{m}y_i\xi_{k,m}(i),\quad S_k=\sum_{i=1}^{m}[\xi_{k,m}(i)]^2.$$

References

See references to Table 16, this Appendix.

21. Interpolation (→ 223 Interpolation)

(1) Lagrange's Interpolation Polynomial.

$$f(x)=\sum_{s=0}^{n}f(x_s)\frac{(x-x_0)(x-x_1)\dots(x-x_{s-1})(x-x_{s+1})\dots(x-x_n)}{(x_s-x_0)(x_s-x_1)\dots(x_s-x_{s-1})(x_s-x_{s+1})\dots(x_s-x_n)}.$$

Aitken's interpolation scheme. The interpolation polynomial $f(x)$ corresponding to the value $y_s=f(x_s)$ $(s=0,1,\dots,n)$ is given inductively by the following procedure. The order of $x_0, x_1, \dots, x_s$ is quite arbitrary.

$$p_{s,0}(x)=y_s\quad(s=0,1,\dots,n),$$
$$p_{s,k+1}(x)=\left[(x_s-x)p_{k,k}(x)-(x_k-x)p_{s,k}(x)\right]/(x_s-x_k)\quad(s=k+1,k+2,\dots,n),$$
$$f(x)\equiv p_{n,n}(x).$$

(2) Interpolation for Equally Spaced Points. When the points x_k lie in the order of their subscripts at a uniform distance h $(x_s=x_0+sh)$, we make the following difference table $(\Delta x=h)$.
Forward difference:

$$\Delta_i\equiv\Delta_i^1\equiv f_{i+1}-f_i=f(x_{i+1})-f(x_i),\quad \Delta_i^s\equiv\Delta_{i+1}^{s-1}-\Delta_i^{s-1}.$$

Variable	Value of Function	Difference (1st)	(2nd)	(3rd)	(4th)	...
...	...					
$x_0-2\Delta x$	f_{-2}	...	...			
$x_0-\Delta x$	f_{-1}	Δ_{-2}	Δ_{-2}^2	...	...	
x_0	f_0	Δ_{-1}	Δ_{-1}^2	Δ_{-2}^3	Δ_{-2}^4	...
$x_0+\Delta x$	f_1	Δ_0	Δ_0^2	Δ_{-1}^3	Δ_{-1}^4	...
$x_0+2\Delta x$	f_2	Δ_1	Δ_1^2	Δ_0^3	...	
$x_0+3\Delta x$	f_3	Δ_2	...	...		
...	...	...				

Backward difference:

$$\bar{\Delta}_i^s\equiv\bar{\Delta}_i^{s-1}-\bar{\Delta}_{i-1}^{s-1}=\Delta_{s-i}^s.$$

Central difference:

$$\delta_i^s=\delta_{i+(1/2)}^{s-1}-\delta_{i-(1/2)}^{s-1},\quad \delta_{i+(s/2)}^s=\Delta_i^s.$$

Newton interpolation formula (forward type):

$$f(x_0+u\Delta x)=f(x_0)+\frac{u}{1!}\Delta_0+\frac{u(u-1)}{2!}\Delta_0^2+\frac{u(u-1)(u-2)}{3!}\Delta_0^3$$
$$+\frac{u(u-1)(u-2)(u-3)}{4!}\Delta_0^4+\dots.$$

Gauss's interpolation formula (forward type):

$$f(x_0+u\Delta x)=f(x_0)+\frac{u}{1!}\Delta_0+\frac{u(u-1)}{2!}\Delta_{-1}^2+\frac{u(u-1)(u+1)}{3!}\Delta_{-1}^3$$
$$+\frac{u(u-1)(u+1)(u-2)}{4!}\Delta_{-2}^4+\dots.$$

Stirling's interpolation formula:

$$f(x_0+u\Delta x)=f(x_0)+\frac{u}{1!}\frac{\Delta_{-1}+\Delta_0}{2}+\frac{u^2}{2!}\Delta^2_{-1}+\frac{u(u^2-1)}{3!}\frac{\Delta^3_{-2}+\Delta^3_{-1}}{2}+\frac{u^2(u^2-1)}{4!}\Delta^4_{-2}+\dots .$$

Bessel's interpolation formula:

$$f\left(\frac{x_0+x_1}{2}+v\Delta x\right)=\frac{f(x_0)+f(x_1)}{2}+\frac{v}{1!}\Delta_0+\frac{1}{2!}\left(v^2-\frac{1}{4}\right)\frac{\Delta^2_{-1}+\Delta^2_0}{2}+\frac{v}{3!}\left(v^2-\frac{1}{4}\right)\Delta^3_{-1}$$

$$+\frac{1}{4!}\left(v^2-\frac{1}{4}\right)\left(v^2-\frac{9}{4}\right)\frac{\Delta^4_{-2}+\Delta^4_{-1}}{2}+\dots .$$

Everett's interpolation formula:

$$f(x_0+u\Delta x)=f(x_1-\xi\Delta x)=\xi f(x_0)+\frac{\xi(\xi^2-1)}{3!}\Delta^2_{-1}+\frac{\xi(\xi^2-1)(\xi^2-4)}{5!}\Delta^4_{-2}+\dots$$

$$+uf(x_1)+\frac{u(u^2-1)}{3!}\Delta^2_0+\frac{u(u^2-1)(u^2-4)}{5!}\Delta^4_{-1}+\dots . \quad (\xi=1-u)$$

(3) Interpolation for Functions of Two Variables. Let $x_m=x_0+m\Delta x$, $y_n\equiv y_0+n\Delta y$ (m and n are integers). We define the finite differences as follows:

$$\Delta_x(x_0,y_0)\equiv f(x_1,y_0)-f(x_0,y_0),$$

$$\Delta_y(x_0,y_0)\equiv f(x_0,y_1)-f(x_0,y_0),$$

$$\Delta^2_x(x_0,y_0)\equiv\Delta_x(x_1,y_0)-\Delta_x(x_0,y_0)\equiv\delta^2_x(x_1,y_0),$$

$$\Delta_{xy}(x_0,y_0)\equiv\Delta_y(x_1,y_0)-\Delta_y(x_0,y_0)=\Delta_x(x_0,y_1)-\Delta_x(x_0,y_0),$$

$$\Delta^2_y(x_0,y_0)\equiv\Delta_y(x_0,y_1)-\Delta_y(x_0,y_0)\equiv\delta^2_y(x_0,y_1), \quad \dots .$$

Newton's formula:

$$f(x_0+u\Delta x,y_0+v\Delta y)=f(x_0,y_0)+(u\Delta_x+v\Delta_y)(x_0,y_0)$$

$$+(1/2!)\left[u(u-1)\Delta^2_x+2uv\Delta_{xy}+v(v-1)\Delta^2_y\right](x_0,y_0)+\dots .$$

Everett's formula. Putting $s\equiv 1-u, t\equiv 1-v$ we have

$$f(x_0+u\Delta x,y_0+v\Delta y)=stf(x_0,y_0)+svf(x_0,y_1)+utf(x_1,y_0)+uvf(x_1,y_1)$$

$$-(1/6)\left[us(1+s)\left\{t\delta^2_x(x_0,y_0)+v\delta^2_x(x_0,y_1)\right\}+us(1+u)\left\{t\delta^2_x(x_1,y_0)+v\delta^2_x(x_1,y_1)\right\}\right.$$

$$\left.+vt(1+t)\left\{s\delta^2_y(x_0,y_0)+u\delta^2_y(x_1,y_0)\right\}+vt(1+v)\left\{s\delta^2_y(x_0,y_1)+u\delta^2_y(x_1,y_1)\right\}\right]+\dots .$$

References

[1] F. J. Thompson, Table of the coefficients of Everett's central-difference interpolation formula, Tracts for computers, no. V, Cambridge Univ. Press, 1921.
[2] M. Lindow, Numerische Infinitesimalrechnung, Dummler, Berlin, 1928.
[3] H. T. Davis, Tables of the higher mathematical functions I, Principia Press, Bloomington, 1933.
[4] K. Hayashi and S. Moriguti, Table of higher transcendental functions (in Japanese), Iwanami, second revised edition, 1967.

22. Distribution of Typical Random Variables

(→ 341 Probability Measures, 374 Sampling Distributions)

In the following table, Nos. 1–13 are 1-dimensional continuous distributions, and Nos. 20–21 are k-dimensional continuous distributions, for which the distribution density is the one with respect to Lebesgue measure. Nos. 14–19 are 1-dimensional discrete distributions, and Nos. 22–24 are k-dimensional discrete distributions, where the density function $P(x)$ means the probability at the point x.

The characteristic function, average, and variance are given only for those represented in a simple form.

Distribution of Typical Random Variables

No.	Name	Symbol	Density Function	Domains
1	Normal	$N(\mu,\sigma^2)$	$\frac{1}{(2\pi\sigma^2)^{1/2}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$	$-\infty<x<\infty$
2	Logarithmic normal		$\frac{1}{(2\pi\sigma^2)^{1/2}}\frac{1}{y}\exp\left[-\frac{(\log y-\mu)^2}{2\sigma^2}\right]$	$0<y<\infty$
3	Gamma	$\Gamma(p,\sigma)$	$[\Gamma(p)]^{-1}\sigma^{-p}x^{p-1}e^{-x/\sigma}$	$0<x<\infty$
4	Exponential	$e(\mu,\sigma)$	$(1/\sigma)\exp(-(x-\mu)/\sigma)$	$\mu<x<\infty$
5	Two-sided exponential		$(1/2\sigma)e^{-\lvert x\rvert/\sigma}$	$-\infty<x<\infty$
6	Chi square	$\chi^2(n)$	$2^{-n/2}[\Gamma(n/2)]^{-1}x^{(n/2)-1}e^{-x/2}$	$0<x<\infty$
7	Beta	$B(p,q)$	$[B(p,q)]^{-1}x^{p-1}(1-x)^{q-1}$	$0<x<1$
8	F	$F(m,n)$	$2K_F x^{(m/2)-1}[1+(mx/n)]^{-(m+n)/2}$, $K_F\equiv[B(m/2,n/2)]^{-1}(m/n)^{m/2}$	$0<x<\infty$
9	z	$z(m,n)$	$K_F e^{mz}[1+(me^{2z}/n)]^{-(m+n)/2}$, $K_F\equiv[B(m/2,n/2)]^{-1}(m/n)^{m/2}$	$-\infty<z<\infty$
10	t	$t(n)$	$[\sqrt{n}\,B(n/2,1/2)]^{-1}[1+(t^2/n)]^{-(n+1)/2}$	$-\infty<t<\infty$
11	Cauchy	$C(\mu,\sigma)$	$(\pi\sigma)^{-1}\left[1+\frac{(x-\mu)^2}{\sigma^2}\right]^{-1}$	$-\infty<x<\infty$
12	One-side stable for exponent 1/2		$c(2\pi)^{-1/2}x^{-3/2}\exp(-c^2/2x)$	$0<x<\infty$
13	Uniform rectangular	$U(\alpha,\beta)$	$1/(\beta-\alpha)$	$\alpha<x<\beta$
14	Binomial	$Bin(n,p)$	$\binom{n}{x}p^x q^{n-x}$	$x=0,1,2,\ldots,n$
15	Poisson	$P(\lambda)$	$e^{-\lambda}\lambda^x/x!$	$x=0,1,2,\ldots$
16	Hypergeometric	$H(N,n,p)$	$\binom{Np}{x}\binom{Nq}{n-x}\Big/\binom{N}{n}$	x integer $0\leq x\leq Np$, $0\leq n-x\leq Nq$
17	Negative binomial	$NB(m,p)$	$\Gamma(m+x)[\Gamma(m)x!]^{-1}p^m q^x$	$x=0,1,2,\ldots$
18	Geometric	$G(p)$	pq^x	$x=0,1,2,\ldots$
19	Logarithmic		$K_L q^x/x$, $K_L\equiv-1/\log p$	$x=1,2,3,\ldots$
20	Multidimensional normal	$N(\mu,\Sigma)$	$(2\pi)^{-k/2}\lvert\Sigma\rvert^{-1/2}$ $\times\exp[-(x-\mu)\Sigma^{-1}(x-\mu)'/2]$, $x=(x_1,\ldots,x_k)$, $\mu=(\mu_1,\ldots,\mu_k)$, $\Sigma=(\sigma_{ij})$	$-\infty<x_1,\ldots,x_k<\infty$
21	Dirichlet		$\frac{\Gamma(\nu_1+\ldots+\nu_{k+1})}{\Gamma(\nu_1)\ldots\Gamma(\nu_{k+1})}x_1^{\nu_1-1}\ldots x_{k+1}^{\nu_{k+1}-1}$, $x_{k+1}=1-(x_1+\ldots+x_k)$	$x_1,\ldots,x_k>0$, $x_1+\ldots+x_k<1$
22	Multinomial	$M(n,(p_i))$	$n!(x_1!\ldots x_{k+1}!)^{-1}p_1^{x_1}\ldots p_{k+1}^{x_{k+1}}$, $x_{k+1}=n-(x_1+\ldots+x_k)$	$x_1,\ldots,x_k$ $=0,1,\ldots,n$, $x_1+\ldots+x_k\leq n$
23	Multidimensional hypergeometric	$H(N,n,(p_i))$	$\binom{Np_1}{x_1}\cdots\binom{Np_{k+1}}{x_{k+1}}\Big/\binom{N}{n}$, $x_{k+1}=n-(x_1+\ldots+x_k)$	$x_1,\ldots,x_k$ integers $0\leq x_i\leq Np_i$ $(i=1,\ldots,k+1)$
24	Negative polynomial		$\frac{\Gamma(m+x_1+\ldots+x_k)}{\Gamma(m)x_1!\ldots x_k!}p_0^m p_1^{x_1}\ldots p_k^{x_k}$,	$x_1,\ldots,x_k$ $=0,1,2,\ldots$

Conditions for Parameters	Characteristic Function	Mean	Variance	No.
$-\infty<\mu<\infty$, $\sigma>0$	$\exp\left(i\mu t-\frac{\sigma^2t^2}{2}\right)$	μ	σ^2	1
$-\infty<\mu<\infty$, $\sigma>0$		$e^{\mu+(\sigma^2/2)}$	$e^{2\mu}(e^{2\sigma^2}-e^{\sigma^2})$	2
$p,\sigma>0$	$(1-i\sigma t)^{-p}$	σp	σ^2p	3
$-\infty<\mu<\infty$, $\sigma>0$	$e^{i\mu t}(1-i\sigma t)^{-1}$	$\mu+\sigma$	σ^2	4
$\sigma>0$	$(1+\sigma^2t^2)^{-1}$	0	$2\sigma^2$	5
n positive integer	$(1-2it)^{-n/2}$	n	$2n$	6
$p,q>0$		$\frac{p}{p+q}$	$\frac{pq}{(p+q)^2(p+q+1)}$	7
m,n positive integers		$\frac{n}{n-2}\ (n>2)$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}\ (n>4)$	8
m, n positive integers				9
n positive integer		$0\ (n>1)$	$n/(n-2)\ (n>2)$	10
$-\infty<\mu<\infty$, $\sigma>0$	$\exp(i\mu t-\sigma\|t\|)$	none	none	11
$0<c<\infty$	$\exp[-c\|t\|^{1/2}(1-it/\|t\|)]$	none	none	12
$-\infty<\alpha<\beta<\infty$	$(e^{i\beta t}-e^{i\alpha t})/it(\beta-\alpha)$	$(\alpha+\beta)/2$	$(\beta-\alpha)^2/12$	13
$p+q=1, p,q>0$, n positive integer	$(pe^{it}+q)^n$	np	npq	14
$\lambda>0$	$\exp[-\lambda(1-e^{it})]$	λ	λ	15
$p+q=1, p,q>0$, N, Np, n positive integers $N\geqslant n$	$(Nq)^{[n]}(N^{[n]})^{-1}$ $\times F(-n,-Np;\ Nq-n+1;\ e^{it})$, $m^{[n]}\equiv m!/(m-n)!$	np	$\frac{npq(N-n)}{N-1}$	16
$p+q=1, p,q>0$, $m>0$	$\frac{p^m}{(1-qe^{it})^m}$	$\frac{mq}{p}$	$\frac{mq}{p^2}$	17
$p+q=1, p,q>0$	$\frac{p}{1-qe^{it}}$	$\frac{q}{p}$	$\frac{q}{p^2}$	18
$p+q=1, p,q>0$	$-K_L\log(1-qe^{it})$	K_Lq/p	$K_Lq(1-K_Lq)/p^2$	19
$-\infty<\mu_1,\ldots,\mu_k<\infty$, Σ symmetric positive definite quadratic form	$\exp\left(i\mu t'-\frac{t\Sigma t'}{2}\right)$, $t=(t_1,\ldots,t_k)$	$E(x_i)=\mu_i$	$V(x_i)=\sigma_{ii}$, $\mathrm{Cov}(x_i,x_j)=\sigma_{ij}$	20
$\nu_1,\ldots,\nu_{k+1}>0$		$E(x_i)=\frac{\nu_i}{\nu_1+\ldots+\nu_{k+1}}$	$V(x_i)=C\nu_i(\nu_1+\ldots+\nu_{k+1}-\nu_i)$, $\mathrm{Cov}(x_i,x_j)=-C\nu_i\nu_j$, $C\equiv(\nu_1+\ldots+\nu_{k+1})^{-2}\times(\nu_1+\ldots+\nu_{k+1}+1)^{-1}$	21
$p_1+\ldots+p_{k+1}=1$, $p_1,\ldots,p_{k+1}>0$ n positive integer	$(p_1e^{it_1}+\ldots+p_ke^{it_k}+p_{k+1})^n$	$E(x_i)=np_i$	$V(x_i)=np_i(1-p_i)$, $\mathrm{Cov}(x_i,x_j)=-np_ip_j$	22
$p_1+\ldots+p_{k+1}=1$, $p_1,\ldots,p_{k+1}>0$, $N, Np_1,\ldots,Np_k, n$ positive integers		$E(x_i)=np_i$	$V(x_i)=Cnp_i(1-p_i)$, $\mathrm{Cov}(x_i,x_j)=-Cnp_ip_j$, $C\equiv\frac{N-n}{N-1}$	23
$p_0+p_1+\ldots+p_k=1, p_0,p_1,\ldots,p_k>0$, $m>0$	$p_0^m(1-p_1e^{it_1}-\ldots-p_ke^{it_k})^m$	$E(x_i)=\frac{mp_i}{p_0}$	$V(x_i)=mp_i(p_0+p_i)/p_0^2$, $\mathrm{Cov}(x_i,x_j)=mp_ip_j/p_0^2$	24

Remarks
1. Reproducing property with respect to μ, σ^2.
2. $X = \log Y : N(\mu, \sigma^2)$.
3. Reproducing property with respect to p.
4. $e(0,\sigma) = \Gamma(1,\sigma)$.
6. n is the number of degrees of freedom; reproducing property with respect to n.
8. m and n are the numbers of degrees of freedom.
9. $e^{2z} = F(m,n)$.
10. n is the number of degrees of freedom.
11. $C(0,1) = t(1)$; reproducing property with respect to μ and σ.
14. Reproducing property with respect to n.
15. Reproducing property with respect to λ.
17. Reproducing property with respect to m.
18. $G(p) = NB(1,p)$.
20. Generalization of normal distribution; reproducing property with respect to μ and Σ.
22. Generalization of binomial distribution; reproducing property with respect to n.
23. Generalization of hypergeometric distribution.
24. Generalization of negative binomial distribution; reproducing property with respect to m.

23. Statistical Estimation and Statistical Hypothesis Testing

Listed below are some frequently used and well-investigated statistical procedures. (Concerning main probability distributions → 398 Statistical Decision Functions, 399 Statistical Estimation, 400 Statistical Hypothesis Testing). The following notations and conventions are adopted, unless otherwise stated.

Immediately after the heading number, the distribution is indicated by the symbol as defined in Table 22, this Appendix. It is to be understood that a random sample $(x_1, x_2, \ldots, x_n)$ is observed from this distribution. Where two distributions are involved, samples $(x_1, \ldots, x_{n_1})$ and $(y_1, \ldots, y_{n_2})$ are understood to be observed from the respective distributions.

Next, a necessary and sufficient statistic based on the sample is marked with * when it is complete, and # otherwise. Then appears the sampling distribution of this statistic. For those statistics consisting of several independent components, the distribution of these are shown. Greek lowercase letters except α and χ denote unknown parameters. Italic lowercase letters denote constants, each taking arbitrary real values. Italic capital letters denote constants whose values are specified in each procedure; repeated occurrences of the same letter under the same heading number specify a certain common real value.

Problems of point estimation, interval estimation, and hypothesis testing are presented, with corresponding estimators, confidence intervals, and tests (critical regions) as their solutions. All the confidence intervals here are those constructed from UMP unbiased tests, having $1-\alpha$ as confidence levels. Alternative hypotheses are understood to be the negations of corresponding null hypotheses. Significance levels of all the tests are α. The following symbols are attached to each procedure to describe its properties.

For estimators:
UMV: uniformly minimum variance unbiased.
ML: maximum likelihood.
AD: admissibility with respect to quadratic loss function.
IAD: inadmissibility with respect to quadratic loss function.

For tests:
UMP: uniformly most powerful.
UMPU: uniformly most powerful unbiased.
UMPI( ): uniformly most powerful invariant with respect to the product of transformation groups shown in ( ).
LR: likelihood ratio.
O: group of orthogonal transformations.
L: group of shift transformations.
S: group of change of scales.
AD: admissibility with respect to simple loss function.

IAD: inadmissibility with respect to simple loss function. (Note that UMPU implies AD.)

The following symbols denote $100(1-\alpha)\%$ points of respective distributions, α being sufficiently small.

$u(\alpha)$: standard normal distribution.

$t_f(\alpha)$: t-distribution with f degrees of freedom.

$\chi_f^2(\alpha)$: χ^2 distribution with f degrees of freedom.

$F_{f_2}^{f_1}(\alpha)$: F-distribution with (f_1,f_2) degrees of freedom.

(1) $N(\mu,b^2)$. $\quad \Sigma x_i$*. $\quad N(n\mu,nb^2)$.

Point estimation of μ. $\bar{x}=\frac{1}{2}\Sigma x_i$: UMV, ML, AD.

Interval estimation of μ. $\left(\bar{x}\pm u(\alpha/2)\frac{b}{\sqrt{n}}\right)$.

Hypothesis $[\mu\leqslant k]$. $\bar{x}>k+u(\alpha)\frac{b}{\sqrt{n}}$: UMP, LR.

Hypothesis $[h\leqslant\mu\leqslant l]$. $\bar{x}<h-C$ or $\bar{x}>l+C$: UMPU, LR.

(2) $N(a,\sigma^2)$. $\Sigma(x_i-a)^2$*. $\sigma^2\chi_n^2$. ($\sigma^2\chi_n^2$ is the σ^2-multiplication of a random variable obeying the $\chi^2(n)$ distribution. We use similar notations in the following.)

Point estimation of σ^2. $\frac{\Sigma(x_i-a)^2}{n}$: UMV, ML, IAD.

Interval estimation of σ^2. $(A\Sigma(x_i-a)^2,\ B\Sigma(x_i-a)^2)$.

Hypothesis $[\sigma^2\leqslant k]$. $\Sigma(x_i-a)^2>\chi_n^2(\alpha)k$: UMP, LR.

Hypothesis $[\sigma^2=k]$. $\Sigma(x_i-a)^2<Ak$ or $\Sigma(x_i-a)^2>Bk$: UMPU.

(3) $N(\mu,\sigma^2)$. $\begin{pmatrix}\Sigma x_i\\ \Sigma(x_i-\bar{x})^2\end{pmatrix}^*$. $\begin{pmatrix}N(n\mu,n\sigma^2)\\ \sigma^2\chi_{n-1}^2\end{pmatrix}$.

Point estimation of μ. $\bar{x}$: UMV, ML, AD.

Interval estimation of μ. $\left[\bar{x}\pm t_{n-1}(\alpha/2)\frac{\sqrt{\Sigma(x_i-\bar{x})^2}}{\sqrt{n(n-1)}}\right]$.

Hypothesis $[\mu\leqslant k]$. $\frac{\bar{x}-k}{\sqrt{\Sigma(x_i-\bar{x})^2}}>\frac{t_{n-1}(\alpha)}{\sqrt{n(n-1)}}$: UMPU, LR.

Hypothesis $[\mu=k]$. $\frac{|\bar{x}-k|}{\sqrt{\Sigma(x_i-\bar{x})^2}}>\frac{t_{n-1}(\alpha/2)}{\sqrt{n(n-1)}}$: UMPU, LR, UMPI(S, O) for $k=0$.

Point estimation of σ^2. $\frac{\Sigma(x_i-\bar{x})^2}{n-1}$: UMV, IAD. $\frac{\Sigma(x_i-\bar{x})^2}{n}$: ML, IAD.

Point estimation of σ. $\frac{\Gamma[(n-1)/2]}{\sqrt{2}\,\Gamma(n/2)}\sqrt{\frac{\Sigma(x_i-\bar{x})^2}{n-1}}$: UMV, IAD.

Interval estimation of σ^2. $(A\Sigma(x_i-\bar{x})^2,\ B\Sigma(x_i-\bar{x})^2)$.

Hypothesis $[\sigma^2\leqslant k]$. $\Sigma(x_i-\bar{x})^2>\chi_{n-1}^2(\alpha)k$: UMP, LR.

Hypothesis $[\sigma^2=k]$. $\Sigma(x_i-\bar{x})^2<Ak$ or $\Sigma(x_i-\bar{x})^2>Bk$: UMPU.

Hypothesis $[\sigma^2\geqslant k]$. $\Sigma(x_i-\bar{x})^2<\chi_{n-1}^2(1-\alpha)k$: UMPU, UMPI(L).

Hypothesis $\left[\frac{\mu}{\sigma}\leqslant k\right]$. $\frac{\bar{x}}{\sqrt{\Sigma(x_i-\bar{x})^2}}>E$: UMPI(S), AD.

(4) $Bin(N,\theta)$. Σx_i*. $Bin(Nn,\theta)$.

Point estimation of θ. $\frac{\bar{x}}{N}$: UMV, ML, AD.

Hypothesis $[\theta\leqslant k]$. $\bar{x}>A$: UMP.

Hypothesis $[h\leqslant\theta\leqslant l]$. $\bar{x}<B$ or $\bar{x}>C$: UMPU.

(5) $H(N,m,\theta)$ $(n=1)$. x*.

Point estimation of θ. $\frac{Nx}{m}$: UMV, AD.

Hypothesis $[\theta\leqslant k]$. $x>A$: UMP.

(6) $NB(N,\theta)$. Σx_i*. $NB(Nn,\theta)$.
Point estimation of θ. $\dfrac{Nn-1}{Nn+\Sigma x_i-1}$ (1 when the denominator is 0): UMV, AD.

$\dfrac{Nn}{Nn+\Sigma x_i}$: ML.

Hypothesis $[\theta\leqslant k]$. $\Sigma x_i<A$: UMP.
Hypothesis $[h\leqslant\theta\leqslant l]$. $\Sigma x_i<B$ or $\Sigma x_i>C$: UMPU.

(7) $P(\lambda)$. Σx_i*. $P(n\lambda)$.
Point estimation of λ. $\bar{x}$: UMV, ML, AD.
Hypothesis $[\lambda\leqslant k]$. $\bar{x}>A$: UMP.
Hypothesis $[h\leqslant\lambda\leqslant l]$. $\bar{x}<B$ or $\bar{x}>C$: UMPU.

(8) $G(\theta)$. Σx_i*. $NB(n,\theta)$.
For the point estimation of θ and hypothesis testing → (6).

(9) $U[0,\theta]$. $\max x_i$*.
Point estimation of θ. $\max x_i$: ML, IAD. $\dfrac{n+1}{n}\max x_i$: UMV, IAD.
Hypothesis $[\theta\leqslant k]$. $\max x_i>(1-\alpha)^{1/n}k$: UMP.
Hypothesis $[\theta=k]$. $\max x_i<k\alpha^{1/n}$ or $\max x_i>k$: UMP.

(10) $U[\xi,\eta]$. $(\min x_i, \max x_i)$*.
Point estimation of ξ. $\dfrac{n\min x_i-\max x_i}{n-1}$: UMV, IAD. $\min x_i$: ML, IAD.
Point estimation of $\dfrac{\xi+\eta}{2}$. $\dfrac{\min x_i+\max x_i}{2}$: UMV, AD.
Hypothesis $[\eta-\xi\leqslant k]$. $\max x_i-\min x_i>k\alpha^{1/n}$: UMP.

(11) $U\left[\theta-\dfrac{1}{2},\theta+\dfrac{1}{2}\right]$. $(\min x_i, \max x_i)^{\#}$.
Point estimation of θ. $\dfrac{\min x_i+\max x_i}{2}$: ML, AD.
Hypothesis $[\theta\leqslant k]$. $\min x_i>k+\dfrac{1}{2}-\alpha^{1/n}$ or $\max x_i>k+\dfrac{1}{2}$: UMP.

(12) $e(\mu,\sigma)$. $\begin{pmatrix}\Sigma x_i\\ \min x_i\end{pmatrix}^*$. $\begin{pmatrix}\Gamma(n,\sigma)+n\mu\\ e(\mu,\sigma/n)\end{pmatrix}$.

Point estimation of σ. $\dfrac{\Sigma x_i-n\min x_i}{n-1}$: UMV, IAD. $\bar{x}-\min x_i$: ML, IAD.
Point estimation of μ. $\dfrac{n}{n-1}\min x_i-\dfrac{1}{n-1}\bar{x}$: UMV, IAD. $\min x_i$: ML, IAD.
Hypothesis $[\sigma\leqslant k,\mu=h]$. $\Sigma x_i<h$ or $\Sigma x_i>k\log\alpha^{-1/n}+h$: UMP.
Hypothesis $[h\leqslant\sigma\leqslant l]$. $\Sigma x_i-n\min x_i<A$ or $\Sigma x_i-n\min x_i>B$: UMPU.
Hypothesis $[\mu=k]$. $\dfrac{n\min x_i-k}{\Sigma x_i-n\min x_i}<0$ or $\dfrac{n\min x_i-k}{\Sigma x_i-n\min x_i}>C$: UMPU.

(13) $\Gamma(p,\sigma)$. Σx_i*. $\Gamma(np,\sigma)$.
Point estimation of σ. $\dfrac{\bar{x}}{p}$: UMV, ML, IAD.
Interval estimation of σ. $(C\Sigma x_i, D\Sigma x_i)$.
Hypothesis $[\sigma\leqslant k]$. $\Sigma x_i>A$: UMP.
Hypothesis $[\sigma=k]$. $\Sigma x_i<Ck$ or $\Sigma x_i>Dk$: UMPU.

(14) $\begin{matrix}N(\mu_1,a^2)\\ N(\mu_2,b^2).\end{matrix}$ $\begin{pmatrix}\Sigma x_i\\ \Sigma y_i\end{pmatrix}^*$. $\begin{pmatrix}N(n_1\mu_1,n_1a^2)\\ N(n_2\mu_2,n_2b^2)\end{pmatrix}$.

Point estimation of $\mu_1-\mu_2$. $\bar{x}-\bar{y}$: UMV, ML, AD.

Interval estimation of $\mu_1-\mu_2$. $\left(\bar{x}-\bar{y}\pm u(\alpha/2)\sqrt{\dfrac{a^2}{n_1}+\dfrac{b^2}{n_2}}\right)$.

Hypothesis $[\mu_1-\mu_2\leqslant k]$. $\bar{x}-\bar{y}>k+u(\alpha)\sqrt{\dfrac{a^2}{n_1}+\dfrac{b^2}{n_2}}$: UMP, LR.

Hypothesis $[\mu_1-\mu_2=k]$. $|\bar{x}-\bar{y}-k|>u(\alpha/2)\sqrt{\dfrac{a^2}{n_1}+\dfrac{b^2}{n_2}}$: UMPU, UMPI(L), LR.

(15) $N(\mu_1,\sigma^2)$ $N(\mu_2,\sigma^2)$. $\begin{bmatrix}\Sigma x_i\\ \Sigma y_i\\ s^2=\Sigma(x_i-\bar{x})^2+\Sigma(y_i-\bar{y})^2\end{bmatrix}^*.\quad \begin{bmatrix}N(n_1\mu_1,n_1\sigma^2)\\ N(n_2\mu_2,n_2\sigma^2)\\ \sigma^2\chi^2_{n_1+n_2-2}\end{bmatrix}.$

Point estimation of $\mu_1-\mu_2$. $\bar{x}-\bar{y}$: UMV, ML, AD.

Interval estimation of $\mu_1-\mu_2$. $\left(\bar{x}-\bar{y}\pm t_{n_1+n_2-2}(\alpha/2)\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}\sqrt{\dfrac{s^2}{n_1+n_2-2}}\right)$.

Hypothesis $[\mu_1-\mu_2\leqslant k]$. $t=\dfrac{(\bar{x}-\bar{y}-k)\sqrt{n_1n_2}\sqrt{n_1+n_2-2}}{\sqrt{n_1+n_2}\sqrt{s^2}}>t_{n_1+n_2-2}(\alpha)$: UMPU, UMPI(L), LR.

Hypothesis $[\mu_1-\mu_2=k]$. $|t|>t_{n_1+n_2-2}(\alpha)$: UMPU, UMPI(L), LR.

Point estimation of σ^2. $\dfrac{s^2}{n_1+n_2-2}$: UMV, IAD. $\dfrac{s^2}{n_1+n_2}$: ML, IAD.

Interval estimation of σ^2. (As^2,Bs^2).

Hypothesis $[\sigma^2\leqslant k]$. $s^2>\chi^2_{n_1+n_2-2}(\alpha)k$: UMP, LR.

Hypothesis $[\sigma^2=k]$. $s^2<Ak$ or $s^2>Bk$: UMPU.

Hypothesis $[\sigma^2\geqslant k]$. $s^2>\chi^2_{n_1+n_2-2}(1-\alpha)k$: UMPU, UMPI(L), LR.

(16) $N(\mu_1,\sigma_1^2)$ $N(\mu_2,\sigma_2^2)$. $\begin{bmatrix}\Sigma x_i, & \Sigma(x_i-\bar{x})^2\\ \Sigma y_i, & \Sigma(y_i-\bar{y})^2\end{bmatrix}^*.$

Interval estimation of $\dfrac{\sigma_1^2}{\sigma_2^2}$. $\left[A\dfrac{\Sigma(x_i-\bar{x})^2}{\Sigma(y_i-\bar{y})^2},\ B\dfrac{\Sigma(x_i-\bar{x})^2}{\Sigma(y_i-\bar{y})^2}\right]$.

Hypothesis $\left[\dfrac{\sigma_1^2}{\sigma_2^2}\leqslant k\right]$. $\dfrac{(n_2-1)}{(n_1-1)}\dfrac{\Sigma(x_i-\bar{x})^2}{\Sigma(y_i-\bar{y})^2}>F^{n_1-1}_{n_2-1}(\alpha)k$: UMPU, UMPI(L, S), LR.

(17) $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. $\begin{bmatrix}\Sigma x_i, & \Sigma(x_i-\bar{x})^2, & \\ & & \Sigma(x_i-\bar{x})(y_i-\bar{y})\\ \Sigma y_i, & \Sigma(y_i-\bar{y})^2, & \end{bmatrix}^*.$

Point estimation of ρ. $r=\dfrac{\Sigma(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\Sigma(x_i-\bar{x})^2\Sigma(y_i-\bar{y})^2}}$: ML.

Hypothesis $[\rho=0]$. $|r|>\dfrac{t_{n-1}(\alpha/2)}{\sqrt{t_{n-1}(\alpha/2)^2+n-2}}$: UMPU, LR.

Appendix B
Numerical Tables

1. Prime Numbers and Primitive Roots (→ 297 Number Theory, Elementary

In the following table, p is a prime number and r is a corresponding primitive root.

p	r	p	r	p	r	p	r	p	r	p	r	p	r	p	r
2		79	3	191	19	311	17	439	17	577	5	709	2	857	3
3	2	83	2	193	5	313	17	443	2	587	2	719	11	859	2
5	2	89	3	197	2	317	2	449	3	593	3	727	5	863	5
7	3	97	5	199	3	331	3	457	13	599	7	733	7	877	2
11	2	101	2	211	2	337	19	461	2	601	7	739	3	881	3
13	2	103	5	223	3	347	2	463	3	607	3	743	5	883	2
17	3	107	2	227	2	349	2	467	2	613	2	751	3	887	5
19	2	109	11	229	7	353	3	479	13	617	3	757	2	907	2
23	5	113	3	233	3	359	7	487	3	619	2	761	7	911	17
29	2	127	3	239	7	367	11	491	2	631	3	769	11	919	7
31	3	131	2	241	7	373	2	499	7	641	3	773	2	929	3
37	2	137	3	251	11	379	2	503	5	643	11	787	2	937	5
41	7	139	2	257	3	383	5	509	2	647	5	797	2	941	2
43	3	149	2	263	5	389	2	521	3	653	2	809	3	947	2
47	5	151	7	269	2	397	5	523	2	659	2	811	3	953	3
53	2	157	5	271	43	401	3	541	2	661	2	821	2	967	5
59	2	163	2	277	5	409	29	547	2	673	5	823	3	971	11
61	2	167	5	281	3	419	2	557	2	677	2	827	2	977	3
67	2	173	2	283	3	421	2	563	2	683	5	829	2	983	5
71	7	179	2	293	2	431	7	569	3	691	3	839	11	991	7
73	5	181	2	307	5	433	5	571	3	701	2	853	2	997	7

†Mersenne numbers. A prime number of the form 2^p-1 is called a Mersenne number. There exist 27 such p's less than 44500: $p=2$, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213, 19937, 21701, 23209, 44497. The even perfect numbers are the numbers of the form $2^{p-1}(2^p-1)$, where 2^p-1 is a Mersenne number.

2. Indices Modulo p (→ 297 Number Theory, Elementary)

Let r be a primitive root corresponding to a prime number p. The index $l=\mathrm{Ind}_r a$ of a with respect to the basis r is the integer l in $0\leqslant l<p-1$ satisfying $r^l\equiv a(\mathrm{mod}\,p)$. $a\equiv b(\mathrm{mod}\,p)$ is equivalent to $\mathrm{Ind}_r a\equiv\mathrm{Ind}_r b(\mathrm{mod}(p-1))$. The index satisfies the following congruence relations with respect to $\mathrm{mod}(p-1)$: $\mathrm{Ind}_r ab\equiv\mathrm{Ind}_r a+\mathrm{Ind}_r b$, $\mathrm{Ind}_r a^n\equiv n\,\mathrm{Ind}_r a$, $\mathrm{Ind}_s a\equiv\mathrm{Ind}_s r\,\mathrm{Ind}_r a$.

We can solve congruence equations using these relations. The following is a table of indices.

p	$p-1$	r	2	3	5	7	11	13	17	19	23	29	31	37	41	43	47
2	1		—														
3	2	2	1	—													
5	4	2	1	3	—												
7	$2\cdot 3$	3	2	1	5	—											
11	$2\cdot 5$	2	1	8	4	7	—										
13	$2^2\cdot 3$	2	1	4	9	11	7	—									
17	2^4	3	14	1	5	11	7	4	—								
19	$2\cdot 3^2$	2	1	13	16	6	12	5	10	—							
23	$2\cdot 11$	5	2	16	1	19	9	14	7	15	—						
29	$2^2\cdot 7$	2	1	5	22	12	25	18	21	9	20	—					
31	$2\cdot 3\cdot 5$	3	24	1	20	28	23	11	7	4	27	9	—				
37	$2^2\cdot 3^2$	2	1	26	23	32	30	11	7	35	15	21	9	—			
41	$2^3\cdot 5$	7	14	25	18	1	37	9	7	31	4	33	12	8	—		
43	$2\cdot 3\cdot 7$	3	27	1	25	35	30	32	38	19	16	41	34	7	6	—	
47	$2\cdot 23$	5	18	20	1	32	7	11	16	45	5	35	3	42	15	13	—
53	$2^2\cdot 13$	2	1	17	47	14	6	24	10	37	39	46	33	30	45	22	44
59	$2\cdot 29$	2	1	50	6	18	25	45	40	38	15	28	49	55	14	33	23
61	$2^2\cdot 3\cdot 5$	2	1	6	22	49	15	40	47	26	57	35	59	39	54	43	20
67	$2\cdot 3\cdot 11$	2	1	39	15	23	59	19	64	10	28	44	47	22	53	9	50
71	$2\cdot 5\cdot 7$	7	6	26	28	1	31	39	49	16	15	68	11	20	25	48	9
73	$2^3\cdot 3^2$	5	8	6	1	33	55	59	21	62	46	35	11	64	4	51	31
79	$2\cdot 3\cdot 13$	3	4	1	62	53	68	34	21	32	26	11	56	19	75	49	59
83	$2\cdot 41$	2	1	72	27	8	24	77	56	47	60	12	38	20	40	71	23
89	$2^3\cdot 11$	3	16	1	70	81	84	23	6	35	57	59	31	11	21	29	54
97	$2^5\cdot 3$	5	34	70	1	31	86	25	89	81	77	13	46	91	85	4	84
101	$2^2\cdot 5^2$	2	1	69	24	9	13	66	30	96	86	91	84	56	45	42	58
103	$2\cdot 3\cdot 17$	5	44	39	1	4	61	72	70	80	24	86	57	93	50	77	85
107	$2\cdot 53$	2	1	70	47	43	22	14	29	78	62	32	27	38	40	59	66
109	$2^2\cdot 3^3$	11	15	80	92	20	1	101	87	105	3	98	34	43	63	42	103
113	$2^4\cdot 7$	3	12	1	83	8	74	22	5	99	41	89	50	67	94	47	31
127	$2\cdot 3^2\cdot 7$	3	72	1	87	115	68	94	38	84	121	113	46	98	80	71	60
131	$2\cdot 5\cdot 13$	2	1	72	46	96	56	18	43	35	23	51	29	41	126	124	105
137	$2^3\cdot 17$	3	10	1	75	42	122	25	38	46	125	91	73	102	119	97	19
139	$2\cdot 3\cdot 23$	2	1	41	86	50	76	64	107	61	27	94	56	80	32	115	98
149	$2^2\cdot 37$	2	1	87	104	142	109	53	124	84	95	120	132	72	41	93	138
151	$2\cdot 3\cdot 5^2$	7	10	93	136	1	82	23	124	120	145	42	34	148	3	74	128
157	$2^2\cdot 3\cdot 13$	5	141	82	1	147	28	26	40	124	135	129	62	116	21	113	92
163	$2\cdot 3^4$	2	1	101	15	73	47	51	57	125	9	107	69	33	160	38	28
167	$2\cdot 83$	5	40	94	1	118	28	103	53	58	99	150	90	61	97	87	132
173	$2^2\cdot 43$	2	1	27	39	95	23	130	73	33	20	144	102	162	138	84	64
179	$2\cdot 89$	2	1	108	138	171	15	114	166	54	135	118	62	149	155	80	36
181	$2^2\cdot 3^2\cdot 5$	2	1	56	156	15	62	164	175	135	53	48	99	26	83	20	13
191	$2\cdot 5\cdot 19$	19	44	116	50	171	85	112	98	1	134	33	175	15	165	8	123
193	$2^6\cdot 3$	5	34	84	1	104	183	141	31	145	162	123	82	5	151	24	29
197	$2^2\cdot 7^2$	2	1	181	89	146	29	25	159	154	120	36	141	192	110	78	66
199	$2\cdot 3^2\cdot 11$	3	106	1	138	142	189	172	123	55	118	70	164	11	167	88	76
211	$2\cdot 3\cdot 5\cdot 7$	2	1	43	132	139	162	144	199	154	21	179	115	118	17	80	124
223	$2\cdot 3\cdot 37$	3	180	1	89	210	107	147	144	172	163	128	82	152	204	118	50
227	$2\cdot 113$	2	1	46	11	154	28	61	99	178	34	8	197	77	131	150	218
229	$2^2\cdot 3\cdot 19$	7	111	68	214	1	42	195	24	52	131	191	175	164	73	12	193

(table continued on following page)

p	$p-1$	r	2	3	5	7	11	13	17	19	23	29	31	37	41	43	47
233	$2^3\cdot 29$	3	72	1	165	222	197	158	103	136	112	132	182	8	85	25	139
239	$2\cdot 7\cdot 17$	7	66	74	138	1	4	43	52	155	63	160	188	31	99	15	113
241	$2^4\cdot 3\cdot 5$	7	190	182	138	1	25	47	111	85	57	154	151	73	6	219	114
251	$2\cdot 5^3$	11	135	6	80	218	1	162	184	233	134	203	226	187	64	77	85
257	2^8	3	48	1	55	85	196	106	120	125	28	94	242	219	19	207	61
263	$2\cdot 131$	5	190	50	1	79	166	62	126	43	156	221	136	170	17	154	65
269	$2^2\cdot 67$	2	1	109	208	19	230	142	105	223	176	187	259	56	200	254	32
271	$2\cdot 3^3\cdot 5$	43	266	153	220	98	92	15	16	261	75	45	222	182	156	1	213
277	$2^2\cdot 3\cdot 23$	5	147	188	1	22	7	222	103	252	208	74	47	87	126	55	218
281	$2^3\cdot 5\cdot 7$	3	204	1	186	182	253	9	166	221	197	172	62	135	23	132	75

3. Bernoulli Numbers and Euler Numbers

(→ 177 Generating Functions)

B_n are Bernoulli numbers; E_n are Euler numbers.

n	Numerator of B_n	Denominator of B_n	B_n	E_n
2	1	6	0.16667	1
4	1	30	0.03333	5
6	1	42	0.02381	61
8	1	30	0.03333	1385
10	5	66	0.07576	50521
12	691	2730	0.25311	2702765
14	7	6	1.16667	199360981
16	3617	510	7.09216	19391512145
18	43867	798	54.97118	2404879675441
20	174611	330	529.12424	370371188237525
22	854513	138	6192.12319	6.934887×10^{16}
24	236364091	2730	86580.25311	1.551453×10^{19}
26	8553103	6	1425517.16667	4.087073×10^{21}
28	23749461029	870	27298231.06782	1.252260×10^{24}
30	8615841276005	14322	601580873.90064	4.415439×10^{26}

4. Class Numbers of Algebraic Number Fields

(I) Class Numbers of Real Quadratic Field

(→ 347 Quadratic Fields)

Let $k=\mathbf{Q}(\sqrt{m})$, where m is a positive integer without square factor $(1<m\leqslant 501)$. h is the class number (in the wider sense) of k. The $-$sign in the row of $N(\varepsilon)$ means that the norm $N(\varepsilon)$ of the fundamental unit is -1. When $N(\varepsilon)=+1$, the class number in the narrow sense is $2h$, and when $N(\varepsilon)=-1$, the class number in the narrow sense is also h.

m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$
2	1	–	85	2	–	170	4	–	253	1		335	2		421	1	–
3	1		86	1		173	1	–	254	3		337	1	–	422	1	
5	1	–	87	2		174	2		255	4		339	2		426	2	
6	1		89	1	–	177	1		257	3	–	341	1		427	6	
7	1		91	2		178	2		258	2		345	2		429	2	
10	2	–	93	1		179	1		259	2		346	6	–	430	2	
11	1		94	1		181	1	–	262	1		347	1		431	1	
13	1	–	95	2		182	2		263	1		349	1	–	433	1	–
14	1		97	1	–	183	2		265	2	–	353	1	–	434	4	

m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$	m	h	$N(\varepsilon)$
15	2		101	1	–	185	2	–	266	2		354	2		435	4	
17	1	–	102	2		186	2		267	2		355	2		437	1	
19	1		103	1		187	2		269	1	–	357	2		438	4	
21	1		105	2		190	2		271	1		358	1		439	5	
22	1		106	2	–	191	1		273	2		359	3		442	8	–
23	1		107	1		193	1	–	274	4	–	362	2	–	443	3	
26	2	–	109	1	–	194	2		277	1	–	365	2	–	445	4	–
29	1	–	110	2		195	4		278	1		366	2		446	1	
30	2		111	2		197	1	–	281	1	–	367	1		447	2	
31	1		113	1	–	199	1		282	2		370	4	–	449	1	–
33	1		114	2		201	1		283	1		371	2		451	2	
34	2		115	2		202	2	–	285	2		373	1	–	453	1	
35	2		118	1		203	2		286	2		374	2		454	1	
37	1	–	119	2		205	2		287	2		377	2		455	4	
38	1		122	2	–	206	1		290	4	–	379	1		457	1	–
39	2		123	2		209	1		291	4		381	1		458	2	–
41	1	–	127	1		210	4		293	1	–	382	1		461	1	–
42	2		129	1		211	1		295	2		383	1		462	4	
43	1		130	4	–	213	1		298	2	–	385	2		463	1	
46	1		131	1		214	1		299	2		386	2		465	2	
47	1		133	1		215	2		301	1		389	1	–	466	2	
51	2		134	1		217	1		302	1		390	4		467	1	
53	1	–	137	1	–	218	2	–	303	2		391	2		469	3	
55	2		138	2		219	4		305	2		393	1		470	2	
57	1		139	1		221	2		307	1		394	2	–	471	2	
58	2	–	141	1		222	2		309	1		395	2		473	3	
59	1		142	3		223	3		310	2		397	1	–	474	2	
61	1	–	143	2		226	8	–	311	1		398	1		478	1	
62	1		145	4	–	227	1		313	1	–	399	8		479	1	
65	2	–	146	2		229	3	–	314	2	–	401	5	–	481	2	–
66	2		149	1	–	230	2		317	1	–	402	2		482	2	
67	1		151	1		231	4		318	2		403	2		483	4	
69	1		154	2		233	1	–	319	2		406	2		485	2	–
70	2		155	2		235	6		321	3		407	2		487	1	
71	1		157	1	–	237	1		322	4		409	1	–	489	1	
73	1	–	158	1		238	2		323	4		410	4		491	1	
74	2	–	159	2		239	1		326	3		411	2		493	2	–
77	1		161	1		241	1	–	327	2		413	1		494	2	
78	2		163	1		246	2		329	1		415	2		497	1	
79	3		165	2		247	2		330	4		417	1		498	2	
82	4	–	166	1		249	1		331	1		418	2		499	5	
83	1		167	1		251	1		334	1		419	1		501	1	

One can find a table of fundamental units and representatives of ideal classes for $0<m<2025$ in E. L. Ince, *Cycles of reduced ideals in quadratic fields*, Royal Society, London, 1968.

(II) Class Numbers of Imaginary Quadratic Fields (→ 347 Quadratic Fields)

Let $k=\mathbf{Q}(\sqrt{-m})$, where m is a positive integer without square factor ($1 \leqslant m \leqslant 509$). h is the class number of k. In the present case, there is no distinction between the class numbers in the wider and narrow senses.

m	h	m	h	m	h	m	h	m	h	m	h	m	h	m	h
1	1	65	8	129	12	193	4	255	12	319	10	389	22	447	14
2	1	66	8	130	4	194	20	257	16	321	20	390	16	449	20
3	1	67	1	131	5	195	4	258	8	322	8	391	14	451	6
5	2	69	8	133	4	197	10	259	4	323	4	393	12	453	12
6	2	70	4	134	14	199	9	262	6	326	22	394	10	454	14
7	1	71	7	137	8	201	12	263	13	327	12	395	8	455	20
10	2	73	4	138	8	202	6	265	8	329	24	397	6	457	8
11	1	74	10	139	3	203	4	266	20	330	8	398	20	458	26

m	h	m	h	m	h	m	h	m	h	m	h	m	h	m	h
13	2	77	8	141	8	205	8	267	2	331	3	399	16	461	30
14	4	78	4	142	4	206	20	269	22	334	12	401	20	462	8
15	2	79	5	143	10	209	20	271	11	335	18	402	16	463	7
17	4	82	4	145	8	210	8	273	8	337	8	403	2	465	16
19	1	83	3	146	16	211	3	274	12	339	6	406	16	466	8
21	4	85	4	149	14	213	8	277	6	341	28	407	16	467	7
22	2	86	10	151	7	214	6	278	14	345	8	409	16	469	16
23	3	87	6	154	8	215	14	281	20	346	10	410	16	470	20
26	6	89	12	155	4	217	8	282	8	347	5	411	6	471	16
29	6	91	2	157	6	218	10	283	3	349	14	413	20	473	12
30	4	93	4	158	8	219	4	285	16	353	16	415	10	474	20
31	3	94	8	159	10	221	16	286	12	354	16	417	12	478	8
33	4	95	8	161	16	222	12	287	14	355	4	418	8	479	25
34	4	97	4	163	1	223	7	290	20	357	8	419	9	481	16
35	2	101	14	165	8	226	8	291	4	358	6	421	10	482	20
37	2	102	4	166	10	227	5	293	18	359	19	422	10	483	4
38	6	103	5	167	11	229	10	295	8	362	18	426	24	485	20
39	4	105	8	170	12	230	20	298	6	365	20	427	2	487	7
41	8	106	6	173	14	231	12	299	8	366	12	429	16	489	20
42	4	107	3	174	12	233	12	301	8	367	9	430	12	491	9
43	1	109	6	177	4	235	2	302	12	370	12	431	21	493	12
46	4	110	12	178	8	237	12	303	10	371	8	433	12	494	28
47	5	111	8	179	5	238	8	305	16	373	10	434	24	497	24
51	2	113	8	181	10	239	15	307	3	374	28	435	4	498	8
53	6	114	8	182	12	241	12	309	12	377	16	437	20	499	3
55	4	115	2	183	8	246	12	310	8	379	3	438	8	501	16
57	4	118	6	185	16	247	6	311	19	381	20	439	15	502	14
58	2	119	10	186	12	249	12	313	8	382	8	442	8	503	21
59	3	122	10	187	2	251	7	314	26	383	17	443	5	505	8
61	6	123	2	190	4	253	4	317	10	385	8	445	8	506	28
62	8	127	5	191	13	254	16	318	12	386	20	446	32	509	30

There are only 9 instances of m for which $h=1$, and only 18 instances of m for which $h=2$ (Baker, Stark). All these cases are in this table.

One can find a table of structures of the ideal class groups and representatives of ideal classes for $m<24000$ in H. Wada, *A table of ideal class groups of imaginary quadratic fields*, Proc. Japan Acad., 46 (1970), 401–403.

(III) Class Numbers of Cyclotomic Fields

Cyclotomic field $k=\mathbf{Q}(e^{2\pi i/l})$ $(1<l<100;\ l$ prime). h_1 is the first factor of the class number of k (→ 14 Algebraic Number Fields).

l	h_1	l	h_1	l	h_1	l	h_1	l	h_1	l	h_1
3	1	13	1	29	2^3	43	211	61	$41\cdot 1861$	79	$5\cdot 53\cdot 377911$
5	1	17	1	31	3^2	47	$5\cdot 139$	67	$67\cdot 12739$	83	$3\cdot 279405653$
7	1	19	1	37	37	53	4889	71	$7^2\cdot 79241$	89	$113\cdot 118401449$
11	1	23	3	41	11^2	59	$3\cdot 59\cdot 233$	73	$89\cdot 134353$	97	$577\cdot 3457\cdot 206209$

$h_1>1$ for $l>19$ (Uchida).

5. Characters of Finite Groups; Crystallographic Groups

(I) Symmetric Groups S_n, Alternating Groups A_n $(3\leqslant n\leqslant 7)$, and Mathieu Groups M_n $(n=11,12,22,23,24)$

(1) In each table, the first column gives the representation of the conjugate class as we represent a permutation by the product of cyclic permutations. For example, $(3)(2)^2$ means the conjugate class containing (123)(45)(67).

(2) The second column gives the order of the centralizer of the elements of the conjugate class.
(3) In the table of S_n, the first row gives the type of Young diagram corresponding to each irreducible character. For example, $[3,2^2,1]$ means $T(3,2,2,1)$.
(4) In the table of A_n, when we restrict the self-conjugate character of S_n (the character with *) to A_n, it is decomposed into two mutually algebraically conjugate irreducible characters, and therefore we show only one of them. The other irreducible character of A_n is given by the restriction to A_n of the character of S_n that is not self-conjugate.
(5) In the table of M_n, each character with a bar over the degree is one of the two mutually algebraically conjugate characters.

$\varepsilon_1^{\pm}=(-1\pm\sqrt{-3})/2,\quad \varepsilon_2^{\pm}=(1\pm\sqrt{5})/2,\quad \varepsilon_3^{\pm}=(-1\pm\sqrt{-7})/2$

$\varepsilon_4^{\pm}=(-1\pm\sqrt{-11})/2,\quad \varepsilon_5^{\pm}=(-1\pm\sqrt{-15})/2,\quad \varepsilon_6^{\pm}=(-1\pm\sqrt{-23})/2$

S_3		[3]	[2, 1]*	[1^3]
(1)	6	1	2	1
(2)	2	1	0	−1
(3)	3	1	−1	1

A_3		[2, 1]*
(1)	3	1
(3)	3	ε_1^+
(3)	3	ε_1^-

S_4		[4]	[3, 1]	[2^2]*	[2, 1^2]	[1^4]
(1)	24	1	3	2	3	1
(2)	4	1	1	0	−1	−1
(3)	3	1	0	−1	0	1
(4)	4	1	−1	0	1	−1
$(2)^2$	8	1	−1	2	−1	1

A_4		[2^2]*
(1)	12	1
(3)	3	ε_1^+
(3)	3	ε_1^-
$(2)^2$	4	1

S_5		[5]	[4, 1]	[3, 2]	[3, 1^2]*	[2^2, 1]	[2, 1^3]	[1^5]
(1)	120	1	4	5	6	5	4	1
(2)	12	1	2	1	0	−1	−2	−1
(3)	6	1	1	−1	0	−1	1	1
(4)	4	1	0	−1	0	1	0	−1
$(2)^2$	8	1	0	1	−2	1	0	1
(3)(2)	6	1	−1	1	0	−1	1	−1
(5)	5	1	−1	0	1	0	−1	1

A_5		[3, 1^2]*
(1)	60	3
(3)	3	0
$(2)^2$	4	−1
(5)	5	ε_2^+
(5)	5	ε_2^-

S_6		[6]	[5, 1]	[4, 2]	[4, 1^2]	[3^2]	[3, 2, 1]*	[2^3]	[3, 1^3]	[2^2, 1^2]	[2, 1^4]	[1^6]
(1)	720	1	5	9	10	5	16	5	10	9	5	1
(2)	48	1	3	3	2	1	0	−1	−2	−3	−3	−1
(3)	18	1	2	0	1	−1	−2	−1	1	0	2	1
(4)	8	1	1	−1	0	−1	0	1	0	1	−1	−1
$(2)^2$	16	1	1	1	−2	1	0	1	−2	1	1	1
(3)(2)	6	1	0	0	−1	1	0	−1	1	0	0	−1
(5)	5	1	0	−1	0	0	1	0	0	−1	0	1
(6)	6	1	−1	0	1	0	0	0	−1	0	1	−1
(4)(2)	8	1	−1	1	0	−1	0	−1	0	1	−1	1
$(2)^3$	48	1	−1	3	−2	−3	0	3	2	−3	1	−1
$(3)^2$	18	1	−1	0	1	2	−2	2	1	0	−1	1

A_6		$[3,2,1]^*$
(1)	360	8
(3)	9	-1
$(2)^2$	8	0
(5)	5	ε_2^+
(5)	5	ε_2^-
(4)(2)	4	0
$(3)^2$	9	-1

S_7		[7]	[6, 1]	[5, 2]	$[5, 1^2]$	[4, 3]	[4, 2, 1]	$[3^2, 1]$	$[4, 1^3]^*$	$[3, 2^2]$	$[3, 2, 1^2]$	$[2^3, 1]$	$[3, 1^4]$	$[2^2, 1^3]$	$[2, 1^5]$	$[1^7]$
(1)	5040	1	6	14	15	14	35	21	20	21	35	14	15	14	6	1
(2)	240	1	4	6	5	4	5	1	0	-1	-5	-4	-5	-6	-4	-1
(3)	72	1	3	2	3	-1	-1	-3	2	-3	-1	-1	3	2	3	1
(4)	24	1	2	0	1	-2	-1	-1	0	1	1	2	-1	0	-2	-1
$(2)^2$	48	1	2	2	-1	2	-1	1	-4	1	-1	2	-1	2	2	1
(3)(2)	12	1	1	0	-1	1	-1	1	0	-1	1	-1	1	0	-1	-1
(5)	10	1	1	-1	0	-1	0	1	0	1	0	-1	0	-1	1	1
(6)	6	1	0	-1	0	0	1	0	0	0	-1	0	0	1	0	-1
(4)(2)	8	1	0	0	-1	0	1	-1	0	-1	1	0	-1	0	0	1
$(2)^3$	48	1	0	2	-3	0	1	-3	0	3	-1	0	3	-2	0	-1
$(3)^2$	18	1	0	-1	0	2	-1	0	2	0	-1	2	0	-1	0	1
(5)(2)	10	1	-1	1	0	-1	0	1	0	-1	0	1	0	-1	1	-1
$(3)(2)^2$	24	1	-1	2	-1	-1	-1	1	2	1	-1	-1	-1	2	-1	1
(4)(3)	12	1	-1	0	1	1	-1	-1	0	1	1	-1	-1	0	1	-1
(7)	7	1	-1	0	1	0	0	0	-1	0	0	0	1	0	-1	1

A_7		$[4, 1^3]^*$
(1)	2520	10
(3)	36	1
$(2)^2$	24	-2
(5)	5	0
(4)(2)	4	0
$(3)^2$	9	1
$(3)(2)^2$	12	1
(7)	7	ε_3^+
(7)	7	ε_3^-

M_{11}	(1)	g	1	10	11	55	45	44	$\overline{16}$	$\overline{10}$
	$(2)^4$	48	1	2	3	-1	-3	4	0	-2
	$(4)^2$	8	1	2	-1	-1	1	0	0	0
	$(3)^3$	18	1	1	2	1	0	-1	-2	1
	$(5)^2$	5	1	0	1	0	0	-1	1	0
	(8)(2)	8	1	0	-1	1	-1	0	0	$\pm i\sqrt{2}$
	(8)(2)	8	1	0	-1	1	-1	0	0	$\mp i\sqrt{2}$
	(6)(3)(2)	6	1	-1	0	-1	0	1	0	1
	(11)	11	1	-1	0	0	1	0	ε_4^+	-1
	(11)	11	1	-1	0	0	1	0	ε_4^-	-1

$g = 11 \cdot 10 \cdot 9 \cdot 8 = 7920.$

M_{12}	(1)	g	1	11	11	55	55	55	45	54	66	99	120	144	176	$\overline{16}$
	$(2)^4$	192	1	3	3	-1	-1	7	-3	6	2	3	-8	0	0	0
	$(4)^2$	32	1	3	-1	3	-1	-1	1	2	-2	-1	0	0	0	0
	$(3)^3$	54	1	2	2	1	1	1	0	0	3	0	3	0	-4	-2
	$(5)^2$	10	1	1	1	0	0	0	0	-1	1	-1	0	-1	1	1
	(8)(2)	8	1	1	-1	-1	1	-1	-1	0	0	1	0	0	0	0
	(6)(3)(2)	6	1	0	0	-1	-1	1	0	0	-1	0	1	0	0	0
	(11)	11	1	0	0	0	0	0	1	-1	0	0	-1	1	0	ε_4^+
	(11)	11	1	0	0	0	0	0	1	-1	0	0	-1	1	0	ε_4^-
	$(2)^6$	240	1	-1	-1	-5	-5	-5	5	6	6	-1	0	4	-4	4
	(10)(2)	10	1	-1	-1	0	0	0	0	1	1	-1	0	-1	1	-1
	$(4)^2(2)^2$	32	1	-1	3	-1	3	-1	1	2	-2	-1	0	0	0	0
	$(3)^4$	36	1	-1	-1	1	1	1	3	0	0	3	0	-3	-1	1
	$(6)^2$	12	1	-1	-1	1	1	1	-1	0	0	-1	0	1	-1	1
	(8)(4)	8	1	-1	1	1	-1	-1	-1	0	0	1	0	0	0	0

$g = 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 = 95040.$

M_{22}

(1)	g	1	21	55	154	210	$\overline{280}$	231	385	99	$\overline{45}$
$(2)^8$	384	1	5	7	10	2	−8	7	1	3	−3
$(3)^6$	36	1	3	1	1	3	1	−3	−2	0	0
$(5)^4$	5	1	1	0	−1	0	0	1	0	−1	0
$(4)^4(2)^2$	16	1	1	−1	2	−2	0	−1	1	−1	1
$(4)^4(2)^2$	32	1	1	3	−2	−2	0	−1	1	3	1
$(7)^3$	7	1	0	−1	0	0	0	0	0	1	ε_3^+
$(7)^3$	7	1	0	−1	0	0	0	0	0	1	ε_3^-
$(8)^2(4)(2)$	8	1	−1	1	0	0	0	−1	1	−1	−1
$(6)^2(3)^2(2)^2$	12	1	−1	1	1	−1	1	1	−2	0	0
$(11)^2$	11	1	−1	0	0	1	ε_4^+	0	0	0	1
$(11)^2$	11	1	−1	0	0	1	ε_4^-	0	0	0	1

$g = 22 \cdot 21 \cdot 20 \cdot 48 = 443520.$

M_{23}

(1)	g	1	22	230	231	$\overline{770}$	1035	2024	$\overline{45}$	$\overline{990}$	$\overline{231}$	253	$\overline{896}$
$(2)^8$	2688	1	6	22	7	−14	27	8	−3	−18	7	13	0
$(3)^6$	180	1	4	5	6	5	0	−1	0	0	−3	1	−4
$(5)^4$	15	1	2	0	1	0	0	−1	0	0	1	−2	1
$(4)^4(2)^2$	32	1	2	2	−1	−2	−1	0	1	2	−1	1	0
$(7)^3$	14	1	1	−1	0	0	−1	1	ε_3^+	ε_3^+	0	1	0
$(7)^3$	14	1	1	−1	0	0	−1	1	ε_3^-	ε_3^-	0	1	0
$(8)^2(4)(2)$	8	1	0	0	−1	0	1	0	−1	0	−1	−1	0
$(6)^2(3)^2(2)^2$	12	1	0	1	−2	1	0	−1	0	0	1	1	0
$(11)^2$	11	1	0	−1	0	0	1	0	1	0	0	0	ε_4^+
$(11)^2$	11	1	0	−1	0	0	1	0	1	0	0	0	ε_4^-
$(15)(5)(3)$	15	1	−1	0	1	0	0	−1	0	0	ε_5^+	1	1
$(15)(5)(3)$	15	1	−1	0	1	0	0	−1	0	0	ε_5^-	1	1
$(14)(7)(2)$	14	1	−1	1	0	0	−1	1	$-\varepsilon_3^+$	ε_3^+	0	−1	0
$(14)(7)(2)$	14	1	−1	1	0	0	−1	1	$-\varepsilon_3^-$	ε_3^-	0	−1	0
(23)	23	1	−1	0	1	ε_6^+	0	0	−1	1	1	0	−1
(23)	23	1	−1	0	1	ε_6^-	0	0	−1	1	1	0	−1

$g = 23 \cdot 22 \cdot 21 \cdot 20 \cdot 48 = 10200960.$

M_{24}

$(1)^{24}$	g	1	23	$7 \cdot 36$	$23 \cdot 11$	$23 \cdot 77$	$55 \cdot 64$	$\overline{45}$	$\overline{22 \cdot 45}$	$\overline{23 \cdot 45}$	$23 \cdot 45$	$\overline{11 \cdot 21}$	$\overline{770}$
$(2)^8$	$21 \cdot 2^{10}$	1	7	28	13	−21	64	−3	−18	−21	27	7	−14
$(3)^6$	$27 \cdot 40$	1	5	9	10	16	10	0	0	0	0	−3	5
$(5)^4$	60	1	3	2	3	1	0	0	0	0	0	1	0
$(4)^4(2)^2$	128	1	3	4	1	−5	0	1	2	3	−1	−1	−2
$(7)^3$	42	1	2	0	1	0	−1	ε_3^+	ε_3^+	$2\varepsilon_3^+$	−1	0	0
$(7)^3$	42	1	2	0	1	0	−1	ε_3^-	ε_3^-	$2\varepsilon_3^-$	−1	0	0
$(8)^2(4)(2)$	16	1	1	0	−1	−1	0	−1	0	−1	1	−1	0
$(6)^2(3)^2(2)^2$	24	1	1	1	−2	0	−2	0	0	0	0	1	1
$(11)^2$	11	1	1	−1	0	0	0	1	0	1	1	0	0
$(15)(5)(3)$	15	1	0	−1	0	1	0	0	0	0	0	ε_5^+	0
$(15)(5)(3)$	15	1	0	−1	0	1	0	0	0	0	0	ε_5^-	0
$(14)(7)(2)$	14	1	0	0	−1	0	1	$-\varepsilon_3^-$	ε_3^+	0	−1	0	0
$(14)(7)(2)$	14	1	0	0	−1	0	1	$-\varepsilon_3^+$	ε_3^-	0	−1	0	0
(23)	23	1	0	−1	0	0	1	−1	1	0	0	1	ε_6^+
(23)	23	1	0	−1	0	0	1	−1	1	0	0	1	ε_6^-
$(12)^2$	12	1	−1	0	1	−1	0	1	1	−1	0	0	1
$(6)^4$	24	1	−1	0	1	−1	0	−1	−1	1	2	0	1
$(4)^6$	96	1	−1	0	1	−1	0	1	−2	−1	3	3	−2
$(3)^8$	$7 \cdot 72$	1	−1	0	1	7	−8	3	3	−3	6	0	−7
$(2)^{12}$	$15 \cdot 2^9$	1	−1	12	−11	11	0	5	−10	−5	35	−9	10
$(10)^2(2)^2$	20	1	−1	2	−1	1	0	0	0	0	0	1	0
$(21)(3)$	21	1	−1	0	1	0	−1	ε_3^-	ε_3^-	$-\varepsilon_3^+$	−1	0	0
$(21)(3)$	21	1	−1	0	1	0	−1	ε_3^+	ε_3^+	$-\varepsilon_3^-$	−1	0	0
$(4)^4(2)^4$	$3 \cdot 2^7$	1	−1	4	−3	3	0	−3	6	3	3	−1	2
$(12)(6)(4)(2)$	12	1	−1	1	0	0	0	0	0	0	0	−1	−1

$(1)^{24}$	g	$23\cdot21$	$23\cdot55$	$23\cdot88$	$23\cdot99$	$23\cdot144$	$23\cdot11\cdot21$	$23\cdot7\cdot36$	$77\cdot72$	$11\cdot35\cdot27$
$(2)^8$	$21\cdot2^{10}$	35	49	8	21	48	49	−28	−56	−21
$(3)^6$	$27\cdot40$	6	5	−1	0	0	−15	−9	9	0
$(5)^4$	60	−2	0	−1	−3	−3	3	1	−1	0
$(4)^4(2)^2$	128	3	1	0	1	0	−3	4	0	−1
$(7)^3$	42	0	−2	1	2	1	0	0	0	0
$(7)^3$	42	0	−2	1	2	1	0	0	0	0
$(8)^2(4)(2)$	16	−1	1	0	−1	0	−1	0	0	1
$(6)^2(3)^2(2)^2$	24	2	1	−1	0	0	1	−1	1	0
$(11)^2$	11	−1	0	0	0	1	0	−1	0	0
$(15)(5)(3)$	15	1	0	−1	0	0	0	1	−1	0
$(15)(5)(3)$	15	1	0	−1	0	0	0	1	−1	0
$(14)(7)(2)$	14	0	0	1	0	−1	0	0	0	0
$(14)(7)(2)$	14	0	0	1	0	−1	0	0	0	0
(23)	23	0	0	0	0	0	0	0	1	−1
(23)	23	0	0	0	0	0	0	0	1	−1
$(12)^2$	12	0	0	0	0	0	0	0	0	0
$(6)^4$	24	0	0	0	2	−2	0	0	0	0
$(4)^6$	96	3	−3	0	−3	0	−3	0	0	3
$(3)^8$	$7\cdot72$	0	8	8	6	−6	0	0	0	0
$(2)^{12}$	$15\cdot2^9$	3	−15	24	−19	16	9	36	24	−45
$(10)^2(2)^2$	20	−2	0	−1	1	1	−1	1	−1	0
$(21)(3)$	21	0	1	1	−1	1	0	0	0	0
$(21)(3)$	21	0	1	1	−1	1	0	0	0	0
$(4)^4(2)^4$	$3\cdot2^7$	3	−7	8	−3	0	1	−4	−8	3
$(12)(6)(4)(2)$	12	0	−1	−1	0	0	1	−1	1	0

$g=24\cdot23\cdot22\cdot21\cdot20\cdot48=244823040.$

(II) General Linear Groups $GL(2,q)$, Unitary Groups $U(2,q)$, and Special Linear Groups $SL(2,q)$ (q is a power of a prime) (→ 151 Finite Groups I)

(1) The notations are as follows. $\varepsilon=\exp[2\pi\sqrt{-1}/(q-1)]$, $\eta=\exp[2\pi\sqrt{-1}/(q^2-1)]$, $\sigma=\exp[2\pi\sqrt{-1}/(q+1)]$, ρ is the generator of the multiplicative group of $GF(q)-\{0\}$, ω is the generator of the multiplicative group of $GF(q^2)-\{0\}$, $\omega^{q-1}=\alpha$, B is an element of $GL(2,q)$ with order q^2-1, and $B_1=B^{q-1}$.
(2) The first column gives a representative of the conjugate class.

General Linear Group $GL(2,q)$.

	$X_n(1)$	$X_n(q)$	$Y_{m,n}$	Z_n
$\begin{pmatrix}\rho^a & \\ & \rho^a\end{pmatrix}$	ε^{2na}	$q\varepsilon^{2na}$	$(q+1)\varepsilon^{(m+n)a}$	$(q-1)\eta^{na(q+1)}$
$\begin{pmatrix}\rho^a & \\ 1 & \rho^a\end{pmatrix}$	ε^{2na}	0	$\varepsilon^{(m+n)a}$	$-\eta^{na(q+1)}$
$\begin{pmatrix}\rho^a & \\ & \rho^b\end{pmatrix}$	$\varepsilon^{n(a+b)}$	$\varepsilon^{n(a+b)}$	$\varepsilon^{ma+nb}+\varepsilon^{mb+na}$	0
B^c	ε^{nc}	$-\varepsilon^{nc}$	0	$-(\eta^{nc}+\eta^{ncq})$

(1) $1\leqslant a\leqslant q-1$, $1\leqslant b\leqslant q-1$, $a\not\equiv b\ (\mathrm{mod}\,q-1)$, $1\leqslant c<q^2-1$, $c\not\equiv 0\ (\mathrm{mod}\,q+1)$.
(2) We assume that $1\leqslant n\leqslant q-1$, for $X_n(1), X_n(q)$, $1\leqslant m<n\leqslant q-1$, for $Y_{m,n}$, $1\leqslant n<q^2-1$ for Z_n, $n\not\equiv 0(\mathrm{mod}\,q+1)$. Here, $Z_n=Z_{n'}$ when $n\equiv n'q(\mathrm{mod}\,q^2-1)$.

Unitary Group $U(2,q)$.

	$X'_n(1)$	$X'_n(q)$	$Y'_{m,n}$	Z'_n
$\begin{pmatrix}\alpha^s & \\ & \alpha^s\end{pmatrix}$	σ^{2ns}	$q\sigma^{2ns}$	$(q-1)\sigma^{(m+n)s}$	$(q+1)\sigma^{ns}$
$\begin{pmatrix}\alpha^s & \\ 1 & \alpha^s\end{pmatrix}$	σ^{2ns}	0	$-\sigma^{(m+n)s}$	σ^{ns}
$\begin{pmatrix}\alpha^s & \\ & \alpha^t\end{pmatrix}$	$\sigma^{n(s+t)}$	$-\sigma^{n(s+t)}$	$-(\sigma^{ms+nt}+\sigma^{mt+ns})$	0
$\begin{pmatrix}\omega^u & \\ & \omega^{-uq}\end{pmatrix}$	σ^{-nu}	σ^{-nu}	0	$\eta^{nu}+\eta^{-nuq}$

(1) $\begin{pmatrix}\alpha^s & \\ 1 & \alpha^s\end{pmatrix}$, $\begin{pmatrix}\omega^u & \\ & \omega^{-uq}\end{pmatrix}$ are the canonical forms of an element of $U(2,q)$ in $GL(2,q^2)$.
(2) $1 \leqslant s \leqslant q+1$, $1 \leqslant t \leqslant q+1$, $s \not\equiv t \pmod{q+1}$, $1 \leqslant u < q^2-1$, $u \not\equiv 0 \pmod{q-1}$. When $u \equiv -u'q \pmod{q^2-1}$ u, u' gives the same conjugate class.
(3) The ranges are $1 \leqslant n \leqslant q+1$ for $X_n'(1)$, $X_n'(q)$, $1 \leqslant m < n \leqslant q+1$ for $Y_{m,n}'$, $1 \leqslant n < q^2-1$ for Z_n', $n \not\equiv 0 \pmod{q-1}$. When $n' \equiv -nq \pmod{q^2-1}$, we have $Z_n' = Z_{n'}'$.

Special Linear Group $SL(2,2^n)$ (the case when $q=2^n$).

			Y_n	Z_m
$\begin{pmatrix}1 & \\ & 1\end{pmatrix}$	1	q	$q+1$	$q-1$
$\begin{pmatrix}1 & \\ 1 & 1\end{pmatrix}$	1	0	1	-1
$\begin{pmatrix}\rho^a & \\ & \rho^{-a}\end{pmatrix}$	1	1	$\varepsilon^{na}+\varepsilon^{-na}$	0
B_1^c	1	-1	0	$-(\sigma^{mc}+\sigma^{-mc})$

(1) $1 \leqslant a \leqslant (q-2)/2$, $1 \leqslant c \leqslant q/2$.
(2) $1 \leqslant n \leqslant (q-2)/2$, $1 \leqslant m \leqslant q/2$.

Special Linear Group $SL(2,q)$ ($q=$ power of an odd prime number, $e=(q-1)/2$, $e'=(q+1)/2$).

			Y_n	Z_m		
$\begin{pmatrix}1 & \\ & 1\end{pmatrix}$	1	q	$q+1$	$q-1$	$\frac{q+1}{2}$	$\frac{q-1}{2}$
$Z=\begin{pmatrix}-1 & \\ & -1\end{pmatrix}$	1	q	$(-1)^n(q+1)$	$(-1)^m(q-1)$	$(-1)^e\frac{q+1}{2}$	$(-1)^{e'}\frac{q-1}{2}$
$P_1=\begin{pmatrix}1 & \\ 1 & 1\end{pmatrix}$	1	0	1	-1	$\mu^{\pm}$	$\lambda^{\pm}$
$P_2=\begin{pmatrix}1 & \\ \beta & 1\end{pmatrix}$	1	0	1	-1	$\mu^{\mp}$	$\lambda^{\mp}$
P_1Z	1	0	$(-1)^n$	$-(-1)^m$	$(-1)^e\mu^{\pm}$	$(-1)^{e'}\lambda^{\pm}$
P_2Z	1	0	$(-1)^n$	$-(-1)^m$	$(-1)^e\mu^{\mp}$	$(-1)^{e'}\lambda^{\mp}$
$\begin{pmatrix}\rho^a & \\ & \rho^{-a}\end{pmatrix}$	1	1	$\varepsilon^{na}+\varepsilon^{-na}$	0	$(-1)^a$	0
B_1^c	1	-1	0	$-(\sigma^{mc}+\sigma^{-mc})$	0	$-(-1)^c$

(1) $1 \leqslant a \leqslant (q-3)/2$, $1 \leqslant c \leqslant (q-1)/2$, $1 \leqslant n \leqslant (q-3)/2$, $1 \leqslant m \leqslant (q-1)/2$,
$\lambda^{\pm}=\{-1\pm[(-1)^e q]^{1/2}\}/2$, $\mu^{\pm}=\{1\pm[(-1)^e q]^{1/2}\}/2$.
(2) The last two columns mean two characters (with the same signs), respectively.

(III) Ree group $Re(q)$, Suzuki Group $Sz(q)$, and Janko Group J.

Ree group $Re(q)$ ($q=3^{2n+1}=3m^2$).
The order of $Re(q)$ is $q^3(q^3+1)(q-1)$, $q_0=q^2-q+1$, $m_+=q+3m+1$, $m_-=q-3m+1$.

						A	B	C	X_μ
1	1	1	q_0	q^3	qq_0	$(q-1)mm_+/2$	$(q-1)mm_-/2$	$m(q^2-1)$	q^3+1
J	2	1	-1	q	$-q$	$-(q-1)/2$	$(q-1)/2$	0	$q+1$
X	3	1	$-(q-1)$	0	q	$-(q+m)/2$	$(q-m)/2$	$-m$	1
Y	9	1	1	0	0	m	m	$-m$	1
T	3	1	1	0	0	α	α	2α	1
T^{-1}	3	1	1	0	0	$\bar{\alpha}$	$\bar{\alpha}$	$2\bar{\alpha}$	1
YT	9	1	1	0	0	β	β	$-\beta$	1
YT^{-1}	9	1	1	0	0	$\bar{\beta}$	$\bar{\beta}$	$-\bar{\beta}$	1
JT	6	1	-1	0	0	γ	$-\gamma$	0	1
JT^{-1}	6	1	-1	0	0	$\bar{\gamma}$	$-\bar{\gamma}$	0	1
R^a		1	1	1	1	0	0	0	$\rho^{\mu a}+\rho^{-\mu a}$
S^b		1	3	-1	-3	1	-1	0	0
JR^a		1	-1	1	-1	0	0	0	$\rho^{\mu a}+\rho^{-\mu a}$
JS^b		1	-1	-1	1	1	-1	0	0
V^s		1	0	-1	0	-1	0	-1	0
W^t		1	0	-1	0	0	1	1	0

		X'_μ	Y_ν	Y'_λ	Z_κ	Z'_τ
1	1	q^3+1	$(q-1)q_0$	$(q-1)q_0$	$(q^2-1)m_+$	$(q^2-1)m_-$
J	2	$-(q+1)$	$3(q-1)$	$-(q-1)$	0	0
X	3	1	$2q-1$	$2q-1$	$-m_+$	$-m_-$
Y	9	1	-1	-1	-1	-1
T	3	1	-1	-1	$-3m-1$	$3m-1$
T^{-1}	3	1	-1	-1	$-3m-1$	$3m-1$
YT	9	1	-1	-1	-1	-1
YT^{-1}	9	1	-1	-1	-1	-1
JT	6	-1	-3	1	0	0
JT^{-1}	6	-1	-3	1	0	0
R^a		$\rho^{\mu a}+\rho^{-\mu a}$	0	0	0	0
S^b		0	$\sigma(\nu b)$	$\sigma'(\lambda b)$	0	0
JR^a		$-(\rho^{\mu a}+\rho^{-\mu a})$	0	0	0	0
JS^b		0	$\sigma(\nu b)$	$\sigma'(\lambda b)$	0	0
V^s		0	0	0	$-\sum_{i=0}^{2}(v^{\kappa s q^i}+v^{-\kappa s q^i})$	0
W^t		0	0	0	0	$-\sum_{i=0}^{2}(w^{\tau t q^i}+w^{-\tau t q^i})$

(1) The first column gives a representative of conjugate class, and the second column gives its order. The orders of R, S, V, W are $(q-1)/2$, $(q+1)/4$, m_-, m_+, respectively. R, S, T are commutative with J.

(2) $R^a \sim R^{-a}$, $V^s \sim V^{sq} \sim V^{sq^2} \sim V^{-s} \sim V^{-sq} \sim V^{-sq^2}$, $W^t \sim W^{tq} \sim W^{tq^2} \sim W^{-t} \sim W^{-tq} \sim W^{-tq^2}$. Here we fix an integer δ satisfying $\delta^3 \equiv 1$ [mod$(q+1)/4$], $(\delta-1, (q+1)/4)=1$.

$$S^b \sim S^{b\delta} \sim S^{b\delta^2} \sim S^{-b} \sim S^{-b\delta} \sim S^{-b\delta^2}, \qquad JR^a \sim JR^{-a}, \qquad JS^b \sim JS^{-b},$$

where $A \sim B$ means that A and B are mutually conjugate.

(3) $\rho=\exp[4\pi\sqrt{-1}/(q-1)]$, $v=\exp(2\pi\sqrt{-1}/m_-)$, $w=\exp(2\pi\sqrt{-1}/m_+)$,
$\sigma=\exp[8\pi\sqrt{-1}/(q+1)]$.

(4) $1 \leqslant \mu \leqslant (q-3)/4$, $1 \leqslant \lambda \leqslant (q-3)/8$.
Here ν is considered mod$(q+1)/4$ and

$$Y_\nu = Y_{\nu\delta} = Y_{\nu\delta^2} = Y_{-\nu} = Y_{-\nu\delta} = Y_{-\nu\delta^2},$$

κ is considered mod m_- and

$$Z_\kappa = Z_{\kappa q} = Z_{\kappa q^2} = Z_{-\kappa} = Z_{-\kappa q} = Z_{-\kappa q^2},$$

τ is considered mod m_+ and

$$Z'_\tau = Z'_{\tau q} = Z'_{\tau q^2} = Z'_{-\tau} = Z'_{-\tau q} = Z'_{-\tau q^2}.$$

(5) $\sigma(\nu b) = -\sum_{i=0}^{2}(\sigma^{\nu b\delta^i}+\sigma^{-\nu b\delta^i})$, $\sigma'(\lambda b)=\sum_{i=0}^{1}(\sigma^{\lambda b\delta^i}+\sigma^{-\lambda b\delta^i})-(\sigma^{\lambda b\delta^2}+\sigma^{-\lambda b\delta^2})$.

(6) $\alpha = \dfrac{-m+m\sqrt{-q}}{2}$, $\beta=\dfrac{-m-\sqrt{-q}}{2}$, $\gamma=\dfrac{1-\sqrt{-q}}{2}$. We show one of the two mutually complex conjugate characters, for the characters A, B, C.

Suzuki group $Sz(q)$. The order of $Sz(q)$ is $q^2(q^2+1)(q-1)$ $(q=2^{2n+1}, 2q=r^2)$.

			X_α	Y_β	Z_γ		
1	1	q^2	q^2+1	$(q-r+1)(q-1)$	$(q+r+1)(q-1)$	$r(q-1)/2$	$r(q-1)/2$
σ	1	0	1	$r-1$	$-r-1$	$-r/2$	$-r/2$
ρ	1	0	1	-1	-1	$r\sqrt{-1}/2$	$-r\sqrt{-1}/2$
ρ^{-1}	1	0	1	-1	-1	$-r\sqrt{-1}/2$	$r\sqrt{-1}/2$
π_0^i	1	1	$\varepsilon_0^{\alpha i}+\varepsilon_0^{-\alpha i}$	0	0	0	0
π_1^j	1	-1	0	$-(\varepsilon_1^{\beta j}+\varepsilon_1^{\beta j q}+\varepsilon_1^{-\beta j}+\varepsilon_1^{-\beta j q})$	0	1	1
π_2^k	1	-1	0	0	$-(\varepsilon_2^{\gamma k}+\varepsilon_2^{\gamma k q}+\varepsilon_2^{-\gamma k}+\varepsilon_2^{-\gamma k q})$	-1	-1

(1) The first column gives a representative of the conjugate class.

(2) π_0, π_1, π_2 are the elements of order $q-1$ $q+r+1$, $q-r+1$, respectively.

(3) ε_0, ε_1, ε_2 are the primitive $q-1$, $q+r+1$, $q-r+1$ roots of 1, respectively.

(4) π_0^i and π_0^{-i} are mutually conjugate elements, and hence X_α and $X_{-\alpha}$ give the same character. i, α run over the representatives of mod $q-1$, and $i, \alpha \not\equiv 0$ (mod $q-1$).

(5) π_1^j, π_1^{-j}, π_1^{jq}, π_1^{-jq} are mutually conjugate, and hence Y_β, $Y_{-\beta}$, $Y_{\beta q}$, $Y_{-\beta q}$ give the same character. j, β run over the representatives of $\bmod q+r+1$, and j, $\beta \not\equiv 0 \pmod{q+r+1}$.
(6) π_2^k, π_2^{-k}, π_2^{kq}, π_2^{-kq} are mutually conjugate, and hence Z_γ, $Z_{-\gamma}$, $Z_{\gamma q}$, $Z_{-\gamma q}$ give the same character. k, γ run over the representatives of mod $q-r+1$, and k, $\gamma \not\equiv 0 \pmod{q-r+1}$.

Janko Group J.

1	1	77	133	209	133	77	77	133	76	76	56	56	120	120	120
2	1	5	5	1	-3	-3	-3	-3	4	-4	0	0	0	0	0
3	1	-1	1	-1	-2	2	2	-2	1	1	2	2	0	0	0
5	1	2	-2	-1	ε^+	$-\varepsilon^+$	$-\varepsilon^-$	ε^-	1	1	$2\varepsilon^-$	$2\varepsilon^+$	0	0	0
5	1	2	-2	-1	ε^-	$-\varepsilon^-$	$-\varepsilon^+$	ε^+	1	1	$2\varepsilon^+$	$2\varepsilon^-$	0	0	0
6	1	-1	-1	1	0	0	0	0	1	-1	0	0	0	0	0
7	1	0	0	-1	0	0	0	0	-1	-1	0	0	1	1	1
10	1	0	0	1	$-\varepsilon^+$	$-\varepsilon^+$	$-\varepsilon^-$	$-\varepsilon^-$	-1	1	0	0	0	0	0
10	1	0	0	1	$-\varepsilon^-$	$-\varepsilon^-$	$-\varepsilon^+$	$-\varepsilon^+$	-1	1	0	0	0	0	0
11	1	0	1	0	1	0	0	1	-1	-1	1	1	-1	-1	-1
15	1	-1	1	-1	ε^+	$-\varepsilon^+$	$-\varepsilon^-$	ε^-	1	1	$-\varepsilon^-$	$-\varepsilon^+$	0	0	0
15	1	-1	1	-1	ε^-	$-\varepsilon^-$	$-\varepsilon^+$	ε^+	1	1	$-\varepsilon^+$	$-\varepsilon^-$	0	0	0
19	1	1	0	0	0	1	1	0	0	0	-1	-1	λ_1	λ_2	λ_3
19	1	1	0	0	0	1	1	0	0	0	-1	-1	λ_2	λ_3	λ_1
19	1	1	0	0	0	1	1	0	0	0	-1	-1	λ_3	λ_1	λ_2

(1) The order of J is $8\cdot3\cdot5\cdot7\cdot11\cdot19=175560$.
(2) The first column gives the order of the elements of each conjugate class.
(3) $\rho=\exp(2\pi\sqrt{-1}\,/19)$, $\lambda_1=\rho+\rho^7+\rho^8+\rho^{11}+\rho^{12}+\rho^{18}$, $\lambda_2=\rho^2+\rho^{14}+\rho^{16}+\rho^3+\rho^5+\rho^{17}$, $\lambda_3=\rho^4+\rho^9+\rho^{13}+\rho^6+\rho^{10}+\rho^{15}$, $\varepsilon^{\pm}=(1\pm\sqrt{5}\,)/2$.

References

For $S_n (2\leqslant n\leqslant 10)$, $A_n (3\leqslant n\leqslant 9)$:
[1] D. E. Littlewood, The theory of group characters, second edition, Oxford Univ. Press, 1950.
For $S_n (n=11,12,13)$:
[2] M. Zia-ud-Din, Proc. London Math. Soc., 39(1935), 200–204, 42 (1937), 340–355.
For S_{14}:
[3] K. Kondô, Table of characters of the symmetric group of degree 14, Proc. Phys. Math. Soc. Japan, 22 (1940), 585–593.
For M_n $(n=12,24)$:
[4] G. Frobenius, Über die Charaktere der mehrfach transitive Gruppen, S. B. Preuss. Akad. Wiss., 1904, 558–571.
For M_n $(n=11,22,23)$:
[5] N. Burgoyne and P. Fong, The Schur multipliers of the Mathieu groups, Nagoya Math. J., 27 (1966), 733–745.
For $GL(2,q)$, $SL(2,q)$:
[6] I. Schur, Untersuchungen über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen, J. Reine Angew. Math., 132 (1907), 85–137.
[7] H. E. Jordan, Group characters of various types of linear groups, Amer. J. Math., 29 (1907), 387–405.
For $GL(3,q)$, $GL(4,q)$, $GL(n,q)$:
[8] R. Steinberg, The representations of $GL(3,q)$, $GL(4,q)$, $PGL(3,q)$ and $PGL(4,q)$, Canad. J. Math. 3 (1951), 225–235.
[9] J. A. Green, The characters of the finite general linear groups, Trans. Amer. Math. Soc., 80 (1955), 402–447.
For $U(2,q)$, $U(3,q)$:
[10] V. Ennola, On the characters of the finite unitary groups, Ann. Acad. Sci. Fenn., 323 (1963), 1–34.
For $Sz(q)$, $Re(q)$, J:
[11] M. Suzuki, On a class of doubly transitive groups, Ann. of Math., (2) 75 (1962), 105–145.
[12] H. N. Ward, On Ree's series of simple groups, Trans. Amer. Math. Soc., 121 (1966), 62–89.
[13] Z. Janko, A new finite simple group with Abelian Sylow 2-subgroups and its characterization, J. Algebra, 3 (1966), 147–186.

For other Lie-type groups:
[14] B. Srinivasan, The characters of the finite symplectic group $Sp(4, q)$, Trans. Amer. Math. Soc., 131 (1968), 488–525.
[15] H. Enomoto, The characters of the finite symplectic group $Sp(4, q)$, $q = 2^f$, Osaka J. Math., 9 (1972), 75–94.
[16] G. I. Lehrer, The characters of the finite special linear groups, J. Algebra, 26 (1973), 564–583.
[17] P. Deligne and G. Lusztig, Representations of reductive groups over finite fields, Ann. of Math., (2) 103 (1976), 103–161.
For sporadic groups:
[18] M. Hall, Jr. and D. Wales, The simple groups of order 604, 800, J. Algebra, 9 (1968), 417–450.
[19] J. S. Frame, Computation of characters of the Higman-Sims group and its automorphism group, J. Algebra, 20 (1972), 320–349.
[20] D. Fendel, A characterization of Conway's group 3, J. Algebra, 24 (1973), 159–196.
[21] D. Wright, The irreducible characters of the simple group of M. Suzuki of order 448, 345, 497, 600, J. Algebra, 29 (1974), 303–323.

(IV) Three-Dimensional Crystal Classes (→ 92 Crystallographic Groups)

Crystal System Bravais Types	Geometric Crystal Classes			Arithmetic Crystal Classes	Number of Space Groups[2]
	Schoenflies Notation	International Notation[1]			
		Short	Full		
Triclinic	C_1	1	1	$(P, 1)$	1
P	$S_2(C_i)$	$\bar{1}$	$\bar{1}$	$(P, \bar{1})$	2
Monoclinic	C_2	2	2	$(P, 2)\ (C, 2)$	3–5
P, C	C_{1h}	m	m	$(P, m)\ (C, m)$	6–9
	C_{2h}	$2/m$	$\frac{2}{m}$	$(P, 2/m)\ (C, 2/m)$	10–15
Orthorhombic	$D_2(V)$	222	222	$(P, 222)\ (C, 222)\ (F, 222)\ (I, 222)$	16–24
P, C, F, I	C_{2v}	$mm2$	$mm2$	$(P, mm2)\ (C, mm2)\ (A, mm2)\ (F, mm2)\ (I, mm2)$	25–46
	$D_{2h}(V_h)$	mmm	$\frac{2}{m}\frac{2}{m}\frac{2}{m}$	$(P, mmm)\ (C, mmm)\ (F, mmm)\ (I, mmm)$	47–74
Tetragonal	C_4	4	4	$(P, 4)$[3] $(I, 4)$	75–80
P, I	S_4	$\bar{4}$	$\bar{4}$	$(P, \bar{4})\ (I, \bar{4})$	81–82
	C_{4h}	$4/m$	$\frac{4}{m}$	$(P, 4/m)\ (I, 4/m)$	83–88
	D_4	422	422	$(P, 422)$[4] $(I, 422)$	89–98
	C_{4v}	$4mm$	$4mm$	$(P, 4mm)\ (I, 4mm)$	99–110
	$D_{2d}(V_d)$	$\bar{4}2m$	$\bar{4}2m$	$(P, \bar{4}2m)\ (P, \bar{4}m2)\ (I, \bar{4}m2)\ (I, \bar{4}2m)$	111–122
	D_{4h}	$4/mmm$	$\frac{4}{m}\frac{2}{m}\frac{2}{m}$	$(P, 4/mmm)\ (I, 4/mmm)$	123–142
Trigonal	C_3	3	3	$(P, 3)$[5] $(R, 3)$	143–146
P, R	$S_6(C_{3i})$	$\bar{3}$	$\bar{3}$	$(P, \bar{3})\ (R, \bar{3})$	147–148
	D_3	32	32	$(P, 312)$[6] $(P, 321)$[7] $(R, 32)$	149–155
	C_{3v}	$3m$	$3m$	$(P, 3m1)\ (P, 31m)\ (R, 3m)$	156–161
	D_{3d}	$\bar{3}m$	$\bar{3}\frac{2}{m}$	$(P, \bar{3}1m)\ (P, \bar{3}m1)\ (R, \bar{3}m)$	162–167
Hexagonal	C_6	6	6	$(P, 6)$[8]	168–173
P	C_{3h}	$\bar{6}$	$\bar{6}$	$(P, \bar{6})$	–174
	C_{6h}	$6/m$	$\frac{6}{m}$	$(P, 6/m)$	175–176
	D_6	622	622	$(P, 622)$[9]	177–182

Crystal System Bravais Types	Geometric Crystal Classes: Schoenflies Notation	International Notation[1]: Short	International Notation[1]: Full	Arithmetic Crystal Classes	Number of Space Groups[2]
Hexagonal	C_{6v}	$6mm$	$6mm$	$(P, 6mm)$	183–186
P	D_{3h}	$\bar{6}m2$	$\bar{6}m2$	$(P, \bar{6}m2)$ $(P, \bar{6}2m)$	187–190
(cont.)	D_{6h}	$6/mmm$	$\frac{6}{m}\frac{2}{m}\frac{2}{m}$	$(P, 6/mmm)$	191–194
Cubic	T	23	23	$(P, 23)$ $(F, 23)$ $(I, 23)$	195–199
P, F, I	T_h	$m3$	$\frac{2}{m}\bar{3}$	$(P, m3)$ $(F, m3)$ $(I, m3)$	200–206
	O	432	432	$(P, 432)$[10] $(F, 432)$ $(I, 432)$	207–214
	T_d	$\bar{4}3m$	$\bar{4}3m$	$(P, \bar{4}3m)$ $(F, \bar{4}3m)$ $(I, \bar{4}3m)$	215–220
	O_h	$m3m$	$\frac{4}{m}\bar{3}\frac{2}{m}$	$(P, m3m)$ $(F, m3m)$ $(I, m3m)$	221–230

Notes

(1) The notation is based upon *International tables for X-ray crystallography* I, Kynoch, 1969. In each crystal system, the lowest class is a holohedry.

(2) These correspond to the consecutive numbers of space groups in the book cited in (1).

(3)–(10) Enantiomorphic pairs arise from these classes: two pairs for (4), (8), (9), and one pair for the others.

For the shapes of Bravais lattices → 92 Crystallographic Groups E, Fig. 3.

6. Miscellaneous Constants

$\sqrt{2}$ = 1.41421 35623 73095, $\sqrt{10}$ = 3.16227 76601 68379.

$\sqrt[3]{2}$ = 1.25992 10498 94873, $\sqrt[3]{100}$ = 4.64158 88336 12779.

$\log_{10}2$ = 0.30102 99956 63981 = 1/3.32192 80948 87364.

(I) Base of Natural Logarithm e (1000 decimals)

e = 2.71828 18284 59045 23536 02874 71352 66249 77572 47093 69995 95749 66967 62772 40766 30353 54759
45713 82178 52516 64274 27466 39193 20030 59921 81741 35966 29043 57290 03342 95260 59563 07381
32328 62794 34907 63233 82988 07531 95251 01901 15738 34187 93070 21540 89149 93488 41675 09244
76146 06680 82264 80016 84774 11853 74234 54424 37107 53907 77449 92069 55170 27618 38606 26133
13845 83000 75204 49338 26560 29760 67371 13200 70932 87091 27443 74704 72306 96977 20931 01416
92836 81902 55151 08657 46377 21112 52389 78442 50569 53696 77078 54499 69967 94686 44549 05987
93163 68892 30098 79312 77361 78215 42499 92295 76351 48220 82698 95193 66803 31825 28869 39849
64651 05820 93923 98294 88793 32036 25094 43117 30123 81970 68416 14039 70198 37679 32068 32823
76464 80429 53118 02328 78250 98194 55815 30175 67173 61332 06981 12509 96181 88159 30416 90351
59888 85193 45807 27386 67385 89422 87922 84998 92086 80582 57492 79610 48419 84443 63463 24496
84875 60233 62482 70419 78623 20900 21609 90235 30436 99418 49164 31409 34317 38143 64054 62531
52096 18369 08887 07016 76839 64243 78140 59271 45635 49061 30310 72085 10383 75051 01157 47704
17189 86106 87396 96552 12671 54688 95703 50354.

e (in octal) = 2.55760 52130 50535 5.

$1/e$ = 0.36787 94411 71442, e^2 = 7.38905 60989 30650 = 1/0.13533 52832 36613,

$\sqrt{e}$ = 1.64872 12707 00128 = 1/0.60653 06597 12633.

$\log_e 10$ = 2.30258 50929 94046 = 1/0.43429 44819 03252,

$\log_e 2$ = 0.69314 71805 59945 = 1/1.44269 50408 88964.

(II) The Number π (1000 decimals) (→ 328 Pi(π))

π=3.14159 26535 89793 23846 26433 83279 50288 41971 69399 37510 58209 74944 59230 78164 06286 20899
86280 34825 34211 70679 82148 08651 32823 06647 09384 46095 50582 23172 53594 08128 48111 74502
84102 70193 85211 05559 64462 29489 54930 38196 44288 10975 66593 34461 28475 64823 37867 83165
27120 19091 45648 56692 34603 48610 45432 66482 13393 60726 02491 41273 72458 70066 06315 58817
48815 20920 96282 92540 91715 36436 78925 90360 01133 05305 48820 46652 13841 46951 94151 16094
33057 27036 57595 91953 09218 61173 81932 61179 31051 18548 07446 23799 62749 56735 18857 52724
89122 79381 83011 94912 98336 73362 44065 66430 86021 39494 63952 24737 19070 21798 60943 70277
05392 17176 29317 67523 84674 81846 76694 05132 00056 81271 45263 56082 77857 71342 75778 96091
73637 17872 14684 40901 22495 34301 46549 58537 10507 92279 68925 89235 42019 95611 21290 21960
86403 44181 59813 62977 47713 09960 51870 72113 49999 99837 29780 49951 05973 17328 16096 31859
50244 59455 34690 83026 42522 30825 33446 85035 26193 11881 71010 00313 78387 52886 58753 32083
81420 61717 76691 47303 59825 34904 28755 46873 11595 62863 88235 37875 93751 95778 18577 80532
17122 68066 13001 92787 66111 95909 21642 01989.

π (in octal) = 3.11037 55242 10264 3.
$1/\pi$ = 0.31830 98861 83791, π^2 = 9.86960 44010 89359 = 1/0.10132 11836 42338,
$\sqrt{\pi}$ = 1.77245 38509 05516 = 1/0.56418 95835 47756,
$\sqrt{2\pi}$ = 2.50662 82746 31001 = 1/0.39894 22804 01433,
$\sqrt{\pi/2}$ = 1.25331 41373 15500 = 1/0.79788 45608 02865,
$\sqrt[3]{\pi}$ = 1.46459 18875 61523 = 1/0.68278 40632 55296.
$\log_{10}\pi$ = 0.49714 98726 94134, $\log_e \pi$ = 1.14472 98858 49400.

(III) Radian rad

1 rad = 57°.29577 95130 82321 = 3437′.74677 07849 393 = 20626 4″.80624 70964.
1° = 0.01745 32925 19943 rad, 1′ = 0.00029 08882 08666 rad, 1″ = 0.00000 48481 36811 rad.

(IV) Euler's Constant C (100 decimals) (→ 174 Gamma Function)

C = 0.57721 56649 01532 86060 65120 90082 40243 10421 59335 93992
35988 05767 23488 48677 26777 66467 09369 47063 29174 67495.

e^C = 1.78107 24179 90197 98522.

$$S_n = \sum_{i=1}^{n} \frac{1}{i}.$$

n	S_n	n	S_n	n	S_n	n	S_n
3	1.83333 333	6	2.45000 000	15	3.31822 899	100	5.18737 752
4	2.08333 333	8	2.71785 714	20	3.59773 966	500	6.79282 343
5	2.28333 333	10	2.92896 825	50	4.79920 534	1000	7.48547 086

7. Coefficients of Polynomial Approximations

In this table, we give some typical examples of approximation formulas for computation of functions on a digital computer (→ 19 Analog Computation, 336 Polynomial Approximation).

(I) Exponential Function

(1) Putting $\frac{x}{\log 2}+1=q+y+\frac{1}{2}\left(q \text{ is an integer}, -\frac{1}{2}\leqslant y<\frac{1}{2}\right)$, we have
$e^x=2^q v(y), v(y)\doteqdot \sum a_i y^i$, which gives an approximation by a polynomial of the 7th degree, where the maximal error is 3×10^{-11}.
a_0=0.70710 67811 6, a_1=0.49012 90717 2, a_2=0.16986 57957 2, a_3=0.03924 73321 5,
a_4=0.00680 09712, a_5=0.00094 28173, a_6=0.00010 93869, a_7=0.00001 0826.

(2) An approximation by a polynomial of the 11th degree: $e^x \doteqdot \sum a_i x^i$ $(-1\leqslant x\leqslant 0)$.
Maximal error 1×10^{-12}.
a_0=0.99999 99999 990, a_1=0.99999 99999 995, a_2=0.50000 00000 747,
a_3=0.16666 66666 812, a_4=0.04166 66657 960, a_5=0.00833 33332 174,
a_6=0.00138 88925 998, a_7=0.00019 84130 955, a_8=0.00002 47944 428,
a_9=0.00000 27550 711, a_{10}=0.00000 02819 019, a_{11}=0.00000 00255 791.

(3) $e^x \fallingdotseq 1 + \cfrac{x}{-\cfrac{x}{2} + \cfrac{k_0 + k_1 x^2 + k_2 x^4}{1 + k_3 x^2}}$ $(-\log\sqrt{2} \leqslant x \leqslant \log\sqrt{2})$.

Maximal error 1.4×10^{-14}.
$k_0 = 1.00000\,00000\,00327\,1$, $k_1 = 0.10713\,50664\,56464\,2$,
$k_2 = 0.00059\,45898\,69018\,8$, $k_3 = 0.02380\,17331\,57418\,6$.

(II) Logarithmic Function

(1) An approximation by a polynomial of the 11th degree: $\log(1+x) \fallingdotseq \Sigma a_i x^i$ $(0 \leqslant x \leqslant 1)$.
Maximal error 1.1×10^{-10}.
$a_0 = 0.00000\,00001\,10$, $a_1 = 0.99999\,99654\,98$, $a_2 = -0.49999\,82537\,98$,
$a_3 = 0.33329\,85059\,64$, $a_4 = -0.24963\,72428\,65$, $a_5 = 0.19773\,31015\,60$,
$a_6 = -0.15744\,88954\,13$, $a_7 = 0.11712\,91156\,18$, $a_8 = -0.07364\,03719\,14$,
$a_9 = 0.03469\,74937\,56$, $a_{10} = -0.01046\,82295\,69$, $a_{11} = 0.00148\,19917\,22$.

(2) For $1 \leqslant x \leqslant 2$, and putting $y = \frac{x - \sqrt{2}}{x + \sqrt{2}}(3 + 2\sqrt{2})$ $(-1 \leqslant y \leqslant 1)$, then $\log x \fallingdotseq \log\sqrt{2} + \sum a_i y^{2i+1}$ gives an approximation by a polynomial of the 11th degree $(0 \leqslant i \leqslant 5)$, where the maximal error is 9.2×10^{-15}.
$a_0 = 0.34314\,57505\,07610\,6$, $a_1 = 0.00336\,70892\,56222\,5$, $a_2 = 0.00005\,94707\,04347\,4$,
$a_3 = 0.00000\,12504\,99776\,2$, $a_4 = 0.00000\,00285\,68292\,8$, $a_5 = 0.00000\,00007\,43713\,9$.

(III) Trigonometric Functions

(1) We put $\frac{x}{2\pi} = p + \frac{q}{2} + \frac{r}{4} + \frac{z}{8}$ (p is an integer; $q = 0, 1$; $r = 0, 1$; $-1 \leqslant z < 1$), and $s = \sin\frac{\pi z}{4}$, $c = \cos\frac{\pi z}{4}$.
If $r = 0$, $\sin x = (-1)^q s$, $\cos x = (-1)^q c$,
If $r = 1$, $\sin x = (-1)^q c$, $\cos x = -(-1)^q s$.
Here s and c are computed by the following approximation formulas. Putting $-z^2/2 = y$, $s(y) = \sin(\pi z/4) \fallingdotseq z\Sigma a_i y^i$, $c(y) = \cos(\pi z/4) \fallingdotseq \Sigma b_i y^i$ gives an approximation by a polynomial of the 5th degree, where the maximal errors are s: 2×10^{-15}, c: 2×10^{-13}.
$a_0 = 0.78539\,81633\,97426$, $a_1 = 0.16149\,10243\,75338$, $a_2 = 0.00996\,15782\,61200$,
$a_3 = 0.00029\,26094\,99152$, $a_4 = 0.00000\,50133\,389$, $a_5 = 0.00000\,00555\,1357$,
$b_0 = 0.99999\,99999\,999$, $b_1 = 0.61685\,02750\,601$, $b_2 = 0.06341\,73767\,885$,
$b_3 = 0.00260\,79335\,007$, $b_4 = 0.00005\,74476\,09$, $b_5 = 0.00000\,07765\,93$.

(2) $\frac{\sin(\pi x/2)}{x} \fallingdotseq \Sigma(-1)^i a_i x^{2i}$ $(-1 \leqslant x \leqslant 1)$. This gives an approximation by a polynomial of 10th degree $(0 \leqslant i \leqslant 5)$, where the maximal error is 2.67×10^{-11}.
$a_0 = 1.57079\,63267\,682$, $a_1 = 0.64596\,40955\,820$, $a_2 = 0.07969\,26037\,435$,
$a_3 = 0.00468\,16578\,837$, $a_4 = 0.00016\,02547\,767$, $a_5 = 0.00000\,34318\,696$.

(3) $\tan\frac{\pi x}{4} \fallingdotseq x\left(k_0 + \frac{x^2|}{|k_1} + \ldots + \frac{x^2|}{|k_4}\right)$ (continued fraction) $(-1 \leqslant x \leqslant 1)$.
Maximal error 9.8×10^{-12}.
$k_0 = 0.78539\,81634\,9907$, $k_1 = 6.19229\,46807\,1350$, $k_2 = -0.65449\,83095\,2316$,
$k_3 = 520.24599\,06398\,9939$, $k_4 = -0.07797\,95098\,7751$.

(IV) Inverse Trigonometric Functions

(1) An approximation by a polynomial of the 21st degree $(0 \leqslant i \leqslant 10)$:
$\arcsin x \fallingdotseq \Sigma a_i x^{2i+1}$ $(|x| \leqslant 1/\sqrt{2})$.
Maximal error 10^{-10}.
$a_0 = 1.00000\,00005\,3$, $a_1 = 0.16666\,65754\,5$, $a_2 = 0.07500\,46066\,5$, $a_3 = 0.04453\,58425\,7$,
$a_4 = 0.03175\,26509\,6$, $a_5 = 0.01176\,58281\,9$, $a_6 = 0.06921\,26185\,7$, $a_7 = -0.14821\,09628\,8$,
$a_8 = 0.32889\,76635\,2$, $a_9 = -0.35020\,41201\,5$, $a_{10} = 0.19740\,50325\,0$.

(2) Putting $x = w + u$ $(w = \frac{1}{8}, \frac{3}{8}, \frac{5}{8}, \frac{7}{8};\ -\frac{1}{8} \leqslant u \leqslant \frac{1}{8})$, $v = \dfrac{x-w}{1+xw}$ $(|v| \leqslant \frac{1}{2})$

$\arctan x = \arctan w + t(v)$, $t(v) = \arctan v$.

The values of $\arctan w$:

$\arctan(1/8) = 0.12435\ 49945\ 46711$, $\arctan(3/8) = 0.35877\ 06702\ 70611$,
$\arctan(5/8) = 0.55859\ 93153\ 43560$, $\arctan(7/8) = 0.71882\ 99996\ 21623$.

$t(v)$ is computed by an approximation by a polynomial of the 9th degree $(0 \leqslant i \leqslant 4)$, where $t(v) = \arctan v \fallingdotseq \Sigma(-1)^i a_i v^{2i+1}$.

Maximal error 1.6×10^{-13}.

$a_0 = 0.99999\ 99999\ 9992$, $a_1 = 0.33333\ 33328\ 220$, $a_2 = 0.19999\ 97377\ 6$,
$a_3 = 0.14280\ 9976$, $a_4 = 0.10763\ 60$.

(3) $\arctan x \fallingdotseq x\left(k_0 + \dfrac{x^2|}{|k_1} + \ldots + \dfrac{x^2|}{|k_6}\right)$ (continued fraction) $(-1 \leqslant x \leqslant 1)$.

Maximal error 3.6×10^{-10}.

$k_0 = 0.99999\ 99936\ 2$, $k_1 = -3.00000\ 30869\ 4$, $k_2 = -0.55556\ 97728\ 4$,
$k_3 = -15.77401\ 81127\ 3$,
$k_4 = -0.16190\ 80978\ 0$, $k_5 = -44.57191\ 79508\ 8$, $k_6 = -0.10810\ 67493\ 1$.

(V) Gamma Function

An approximation by a polynomial of the 8th degree:

$\Gamma(2+x) \fallingdotseq \sum a_i x^i$ $(-1/2 \leqslant x \leqslant 1/2)$.

Maximal error 7.6×10^{-8}.

$a_0 = 0.99999\ 9926$, $a_1 = 0.42278\ 4604$, $a_2 = 0.41184\ 9671$, $a_3 = 0.08156\ 52323$,
$a_4 = 0.07406\ 48982$. $a_5 = -0.00012\ 51376\ 7$, $a_6 = 0.01229\ 95771$, $a_7 = -0.00349\ 61289$,
$a_8 = 0.00213\ 85778$.

(VI) Normal Distribution

(1) $\dfrac{2}{\sqrt{\pi}}\displaystyle\int_x^\infty e^{-t^2}\,dt \fallingdotseq \dfrac{1}{(1+\sum a_i x^{i+1})^{16}}$ $(0 \leqslant x < \infty)$. This gives an approximation by a polynomial of the 6th degree.

Maximal error 2.8×10^{-7}.

$a_0 = 0.07052\ 30784$, $a_1 = 0.04228\ 20123$, $a_2 = 0.00927\ 05272$,
$a_3 = 0.00015\ 20143$, $a_4 = 0.00027\ 65672$, $a_5 = 0.00004\ 30638$.

(2) $P(x) \equiv \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_x^\infty e^{-t^2/2}\,dt$,

$$4P(x)(1-P(x)) \fallingdotseq \left[\exp\left(-\frac{2x^2}{\pi}\right)\right]\left[1 + x^4\left(a_0 + \frac{a_1}{x^2+a_2}\right)\right] \qquad (0 \leqslant x < \infty).$$

Maximal error 2×10^{-5}.

$a_0 = 0.0055$, $a_1 = 0.0551$, $a_2 = 14.4$.

(3) The inverse function of (2)

$$x \fallingdotseq \left[y\left(a_0 + \frac{a_1}{y+a_2}\right)\right]^{1/2}, \quad y = -\log[4P(x)(1-P(x))] \qquad (0 \leqslant y < \infty).$$

Maximal error 4.9×10^{-4}.

$a_0 = 2.06117\ 86$, $a_1 = -5.72622\ 04$, $a_2 = 11.64059\ 5$.

Statistical Tables for Reference

Statistical Tables

[1] J. A. Greenwood and H. O. Hartley, Guide to tables in mathematical statistics, Princeton Univ. Press, 1962.
[2] Research Group for Statistical Sciences (T. Kitagawa and M. Masuyama, eds.) New statistical tables (Japanese), explanation p. 264, table p. 214, Kawade, 1952.
[3] R. A. Fisher and F. Yates, Statistical tables for biological, agricultural and medical research, explanation p. 30, table p. 137, Oliver & Boyd, third edition, 1948.
[4] E. S. Pearson and H. O. Hartley, Biometrika tables for statisticians, explanation p. 104, table p. 154, Cambridge Univ. Press, third edition, 1970.
[5] K. Pearson, Tables for statisticians and biometricians I, 1930, explanation p. 83, table p. 143; II, 1931, explanation p. 250, table p. 262, Cambridge Univ. Press.
[6] Statistical tables JSA-1972 (Japanese), table p. 454, explanation p. 260, Japanese Standards Association, 1972.

Tables of Special Statistical Values

[8] Harvard Univ., Tables of the cumulative binomial probability distribution, Harvard, 1955,

$\sum_{i=r}^{n}\binom{n}{i}p^iq^{n-i}$: 5 dec.,

$p=0.01(0.01)0.50$,
$n=1(1)50(2)100(10)200(20)500(50)1000$.
[9] National Bureau of Standards, NBS applied mathematical series, no. 6, Tables of the binomial probability distribution, 1950, $\binom{n}{i}p^iq^{n-i}$ and the partial sum: 7 dec., $p=0.01(0.01)0.50$, $n=2(1)49$.
[10] T. Kitagawa, Table of Poisson distribution (Japanese), Baihûkan, 1951, $e^{-m}m^i/i!$: 7-8 dec., $m=0.001(0.001)1.000(0.01)10.00$.
[11] G. J. Lieberman and D. B. Owen, Tables of the hypergeometric probability distribution, Stanford, 1961, $\binom{k}{x}\binom{N-k}{n-x}\Big/\binom{N}{n}$: 6 dec., $N=2(1)50(10)100(100)2000$, $n=1(1)(N/2)$, $k=1(1)n$.
[12] National Bureau of Standards, NBS no. 23, Tables of normal probability functions, 1942,
$\varphi(x)=(1/\sqrt{2\pi})\exp(-\frac{1}{2}x^2)$,
$\Phi(x)=\int_{-x}^{x}\varphi(x)dx$: 15 dec.,
$x=0(0.00001)1.0000(0.001)8.285$.
[13] K. Pearson, Tables of the incomplete beta-functions, Cambridge, second edition, 1968, $I_x(p,q)$: 8 dec., $p, q=0.5(0.5)\ 11(1)50$.
[14] K. Pearson, Tables of the incomplete gamma-function, Cambridge, 1922, revised edition, 1951,
$I(u,p)=\int_0^{u\sqrt{p+1}}(1/e^v)(v^p/\Gamma(p+1))dv$:
7 dec., $p=0.0(0.1)5.0(0.2)50.0$, $u=0.1(0.1)20.0$; $p=-1.0(0.05)0.0$, $u=0.1(0.1)51.3$.
[15] N. V. Smirnov, Tables for the distribution and density function of the t-distribution, Pergamon, 1961, 6 dec., $f=1(1)35$, $t=0(0.01)3.00(0.02)4.50(0.05)6.50$.
[16] G. J. Resnikoff and G. J. Lieberman, Tables of the non-central t-distribution, Stanford, 1957.
[17] F. N. David, Tables of the ordinates and probability integral of the distribution of the correlation coefficient in small samples, Cambridge, 1938.
[18] D. B. Owen, The bivariate normal probability distribution, Sandia Corp., 1957, $T(h,a)$: 6 dec., $a=0.000(0.025)1.000, \infty$, $h=0.00(0.01)3.50(0.05)4.75$, $T(h,a)$
$=\int_0^h\int_0^{ax}\frac{1}{2\pi}\exp\left(-\frac{x^2+y^2}{2}\right)dy\,dx$.
[19] National Bureau of Standards, NBS no. 50, Tables of the bivariate normal distribution function and related functions, 1959, $L(h,k,r)$
$=\int_h^\infty\int_k^\infty\frac{1}{2\pi\sqrt{1-r^2}}$
$\times\exp\left[-\frac{x^2+y^2-2rxy}{2(1-r^2)}\right]dy\,dx$: 6 dec.,
$r=\pm 0.00(0.05)0.95(0.01)0.99$,
$h, k=0.0(0.1)4.0$.

Tables of Allocation

[20] T. Kitagawa and M. Midome, Table of allocation of elements for experimental design (Japanese), Baihûkan, 1953.
[21] R. C. Bose, W. H. Clatworthy, and S. S. Shrikhande, Tables of partially balanced designs with two associate classes, North Carolina Agric. Expt. Station Tech. Bull., 1954 (table of PBIBD).

Numerical Tables for Reference

General Tables

[1] M. Boll, Tables numériques universelles, Dunod, 1947.
[2] P. Barlow, Barlow's tables, Robinson, 1814, third edition, 1930.
[3] W. Shibagaki, 0.01% table of elementary functions (Japanese), Kyôritu, 1952.
[4] K. Hayashi, Table of higher functions (Japanese), Iwanami, second edition, 1967.
[5] E. Jahnke and F. Emde, Funktionentafeln mit Formelin und Kurven, Teubner, second edition 1933 (English translation: Tables of functions with formulae and curves, Dover, fourth edition, 1945).
[6] Y. Yoshida and M. Yoshida, Mathematical tables (Japanese), Baihûkan, 1958.
[7] M. Abramowitz and I. A. Stegun (eds.), Handbook of mathematical functions with formulas, graphs and mathematical tables, National Bureau of Standards, 1964 (Dover, 1965).
[8] A. Fletcher et al. (eds.), Index of mathematical tables I, II, Scientific Computing Service, Addison-Wesley, second edition, 1962.

Multiplication Table

[9] A. L. Crelle, Rechentafeln welche alles Mutiplicieren und Dividiren mit Zahlen unter 1000 ersparen, bei grösseren Zahlen aber die Rechung erleichtern und sicherer machen, W. de Gruyter, new edition 1944.

Table of Prime Numbers

[10] D. N. Lehmer, List of prime numbers from 1 to 10,006,721, Carnegie Institution of Washington, 1914.

Series of Tables of Functions

[11] British Association for the Advancement of Science, Mathematical tables, vol. 2, Emden functions, 1932; vol. 6, Bessel functions, pt. 1, 1937; vol. 8, Number-divisor tables, 1940; vol. 9, Tables of powers giving integral powers of integrals, 1940; vol. 10, Bessel functions, pt. 2, 1952.
[12] Harvard University, Computation Laboratory, Annals, vol. 2, Tables of the modified Hankel functions of order one-third and their derivatives, 1945; vol. 3, Tables of the Bessel functions of the first kind of orders zero and one, 1947; vol. 14, Orders seventy-nine through one hundred thirty-five, 1951; vol. 18, Tables of generalized sine- and cosine-integral functions, pt. 1, 1949; vol. 19, pt. 2, 1949; vol. 20, Tables of inverse hyperbolic functions, 1949; vol. 21, Tables of the generalized exponential-integral functions, 1949.
[13] National Bureau of Standards, Applied Mathematics Series (AMS), AMS 1, Tables of the Bessel functions $Y_0(x)$, $Y_1(x)$, $K_0(x)$, $K_1(x)$, $0 \leq x \leq 1$, 1948; AMS 5, Tables of sines and cosines to fifteen decimal places at hundredths of a degree, 1949; AMS 11, Table of arctangents of rational numbers, 1951; AMS 14, Tables of the exponential function e^x (including e^{-x}), 1951; AMS 16, Tables of $n!$ and $\Gamma(n+\frac{1}{2})$ for the first thousand values of n, 1951; AMS 23, Tables of normal probability functions, 1953; AMS 25, Tables of the Bessel functions $Y_0(x)$, $Y_1(x)$, $K_0(x)$, $K_1(x)$, $0 \leq x \leq 1$, 1952; AMS 26, Tables of $\operatorname{Arctan} x$, 1953; Tables of 10^x, 1953; AMS 32, Table of sine and cosine integrals for arguments from 10 to 100, 1954; AMS 34, Table of the gamma function for complex arguments, 1954; AMS 36, Tables of circular and hyperbolic sines and cosines for radian arguments, 1953; AMS 40, Table of secants and cosecants to nine significant figures at hundredths of a degree, 1954; AMS 41, Tables of the error function and its derivative, 1954; AMS 43, Tables of sines and cosines for radian arguments, 1955; AMS 45, Table of hyperbolic sines and cosines, 1955; AMS 46, Table of the descending exponential, 1955.

Tables of Special Functions

[14] Akademiya Nauk SSSR, Tables of the exponential integral functions, 1954.
[15] J. Brownlee, Table of $\log \Gamma(x)$, Tracts for computers, no. 9, Cambridge Univ. Press, 1923.
[16] L. Dolansky and M. P. Dolansky, Table of $\log_2(1/p)$, $p \cdot \log_2(1/p)$ and $p \cdot \log_2(1/p) + (1-p)\log_2(1/(1-p))$, M.I.T. Research Lab. of Electronics tech. report 227, 1952.
[17] L. M. Milne-Thomson, Die elliptischen Funktionen von Jacobi, Springer, 1931; English translation: Jacobian elliptic function tables, Dover, 1950.
[18] H. T. Davis, Tables of the higher mathematical functions, Principia Press, vol. 1, Gamma function, 1933; vol. 2, Polygamma functions, 1935; vol. 3, Arithmetical tables, 1962.
[19] R. G. Selfridge and J. E. Maxfield, A table of the incomplete elliptic integral of the third kind, Dover, 1958.

[20] W. Shibagaki, 0.01% table of modified Bessel functions and the method of numerical computation for them (Japanese), Baihûkan, 1955.
[21] W. Shibagaki, Theory and application of gamma function, Appendix. Table of gamma function of complex variable effective up to 6 decimals (Japanese), Iwanami, 1952.
[22] L. J. Slater, A short table of the Laguerre polynomials, Proc. IEE, monograph no. 103c, 1955, 46–50.
[23] G. N. Watson, A treatise on the theory of Bessel functions, Appendix, Cambridge Univ. Press, 1922, second edition, 1958.
See also Statistical Tables for Reference.

Tables of Approximation Formulas of Functions

[24] S. Hitotumatu, Approximation formula (Japanese) Takeuti, 1963.
[25] T. Uno (ed.), Approximation formulas of functions for computers 1–3, Joint research work for mathematical sciences, pt. IV, section 5, 1961–1963 (Japanese, mimeographed).
[26] Z. Yamauti, S. Moriguti, and S. Hitotumatu (eds.), Numerical methods for electronic computers (Japanese), Baihûkan, I, 1965; II, 1967.
[27] C. Hastings (ed.), Approximations for digital computers, Princeton, 1955.
[28] A. J. W. Duijvestijn and A. J. Dekkers, Chebyshev approximations of some transcendental functions for use in digital computing, Philips Research Reports, 16 (1961), 145–174.
[29] L. A. Lyusternik, O. A. Chervonenkis, and A. P. Yanpol'skiĭ, Mathematical analysis: functions, limits, series, continued fractions, Pergamon, 1965. (Original in Russian, 1963.)
[30] J. F. Hart et al. (eds.), Computer approximations, Wiley, 1968.

Journals

Abh. Akad. Wiss. Göttingen
Abhandlungen der Akademie der Wissenschaften in Göttingen (Göttingen)

Abh. Bayer. Akad. Wiss. Math.-Nat. Kl.
Abhandlungen der Bayerschen Akademie der Wissenschaften. Mathematisch-Naturwissenschaftliche Klasse (Munich)

Abh. Math. Sem. Univ. Hamburg
Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg (Hamburg)

Acta Arith.
Acta Arithmetica, Polska Akademia Nauk, Instytut Matematyczny (Warsaw)

Acta Informat.
Acta Informatica (Berlin)

Acta Math.
Acta Mathematica (Uppsala)

Acta Math. Acad. Sci. Hungar.
Acta Mathematica Academiae Scientiarum Hungaricae (Budapest)

Acta Math. Sinica
Acta Mathematica Sinica (Peking). Chinese Math. Acta is its English translation

Acta Sci. Math. Szeged.
Acta Universitatis Szegediensis, Acta Scientiarum Mathematicarum (Szeged)

Actualités Sci. Ind.
Actualités Scientifiques et Industrielles (Paris)

Advances in Math.
Advances in Mathematics (New York)

Aequationes Math.
Aequationes Mathematicae (Basel-Waterloo)

Algebra and Logic
Algebra and Logic (New York). Translation of Algebra i Logika

Algebra i Logika
Akademiya Nauk SSSR Sibirskoe Otdelenie. Institut Matematiki. Algebra i Logika (Novosibirsk). Translated as Algebra and Logic

Algebra Universalis
Algebra Universalis (Basel)

Amer. J. Math.
American Journal of Mathematics (Baltimore)

Amer. Math. Monthly
The American Mathematical Monthly (Menasha)

Amer. Math. Soc. Colloq. Publ.
American Mathematical Society Colloquium Publications (Providence)

Amer. Math. Soc. Math. Surveys
American Mathematical Society Mathematical Surveys (Providence)

Amer. Math. Soc. Proc. Symp. Pure Math.
American Mathematical Society Proceedings of Symposia in Pure Mathematics (Providence)

Amer. Math. Soc. Transl.
American Mathematical Society Translations (Providence)

Amer. Math. Soc. Transl. Math. Monographs
American Mathematical Society Translations of Mathematical Monographs (Providence)

Ann. Acad. Sci. Fenn.
Suomalaisen Tiedeakatemian Toimituksia, Annales Academiae Scientiarum Fennicae. Series A. I. Mathematica (Helsinki)

Ann. der Phys.
Annalen der Physik (Leipzig)

Ann. Fac. Sci. Univ. Toulouse
Annales de la Faculté des Sciences de l'Université de Toulouse pour les Sciences Mathématiques et les Sciences Physiques (Toulouse)

Ann. Inst. Fourier
Annales de l'Institut Fourier, Université de Grenoble (Grenoble)

Ann. Inst. H. Poincaré
Annales de l'Institut Henri Poincaré (Paris)

Ann. Inst. Statist. Math.
Annals of the Institute of Statistical Mathematics (Tokyo)

Ann. Mat. Pura Appl.
Annali di Matematica Pura ed Applicata (Bologna)

Ann. Math. Statist.
The Annals of Mathematical Statistics (Baltimore)

Ann. Math.
Annals of Mathematics (Princeton)

Ann. Physique
Annales de Physique (Paris)

Ann. Probability
The Annals of Probability (San Francisco)

Ann. Polon. Math.
Annales Polonici Mathematici. Polska Akademia Nauk (Warsaw)

Ann. Roumaines Math.
Annales Roumaines de Mathématiques. Journal de l'Institut Mathématique Roumain (Bucharest)

Ann. Sci. Ecole Norm. Sup.
Annales Scientifiques de l'Ecole Normale Supérieure (Paris)

Ann. Scuola Norm. Sup. Pisa
Annali della Scuola Normale Superiore de Pisa, Scienze Fisiche e Matematiche (Pisa)

Ann. Statist.
The Annals of Statistics (San Francisco)

Arch. History Exact Sci.
Archive for History of Exact Sciences (Berlin-New York)

Arch. Math.
Archiv der Mathematik (Basel-Stuttgart)

Arch. Rational Mech. Anal.
Archive for Rational Mechanics and Analysis (Berlin)

Ark. Mat.
Arkiv för Matematik (Stockholm)

Ark. Mat. Astr. Fys.
Arkiv för Matematik, Astronomi och Fysik (Uppsala)

Astérique
Astérique. La Société Mathématique de France (Paris)

Atti Accad. Naz. Lincei, Mem.
Atti della Accademia Nazionale dei Lincei, Memorie (Rome)

Atti Accad. Naz. Lincei, Rend.
Atti della Accademia Nazionale dei Lincei, Rendiconti (Rome)

Automat. Control Comput. Sci.
Automatic Control and Computer Sciences (New York). Translation of Avtomatika i Vychislitel'naya Tekhnik. Akademiya Nauk Latviĭskoĭ SSR (Riga)

Bell System Tech. J.
The Bell System Technical Journal (New York)

Biometrika
Biometrika, A Journal for the Statistical Study of Biological Problems (London)

Bol. Soc. Mat. São Paulo
Boletim da Sociedade de Matemática de São Paulo (São Paulo)

Bull. Acad. Pol. Sci.
Bulletin de l'Academie Polonaise des Sciences (Warsaw)

Bull. Amer. Math. Soc.
Bulletin of the American Mathematical Society (Providence)

Bull. Calcutta Math. Soc.
Bulletin of the Calcutta Mathematical Society (Calcutta)

Bull. Math. Statist.
Bulletin of Mathematical Statistics (Fukuoka, Japan)

Bull. Nat. Res. Council
Bulletin of the National Research Council (Washington)

Bull. Sci. Math.
Bulletin des Sciences Mathématiques (Paris)

Bull. Soc. Math. Belg.
Bulletin de la Société Mathématique de Belgique (Brussels)

Bull. Soc. Math. France
Bulletin de la Société Mathématique de France (Paris)

Bull. Soc. Roy. Sci. Liège
Bulletin de la Société Royale des Sciences de Liège (Liège)

C. R. Acad. Sci. Paris
Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences (Paris)

Canad. J. Math.
Canadian Journal of Mathematics (Toronto)

Časopis Pěst. Mat.
Časopis pro Pěstováni Matematiky, Československá Akademie Věd (Prague)

Colloq. Math.
Colloquium Mathematicum (Warsaw)

Comm. ACM
Communications of the Association for Computing Machinery (New York)

Comm. Math. Phys.
Communications in Mathematical Physics (Berlin)

Comm. Pure Appl. Math.
Communications on Pure and Applied Mathematics (New York)

Comment. Math. Helv.
Commentarii Mathematici Helvetici (Zurich)

Compositio Math.
Compositio Mathematica (Groningen)

Comput. J.
The Computer Journal (London)

Crelles J.
= J. Reine Angew. Math.

CWI Newslett.
Centrum voor Wiskunde en Informatica. Newsletter (Amsterdam)

Cybernetics
Cybernetics (New York). Translation of Kibernetika (Kiev)

Czech. Math. J.
Czechoslovak Mathematical Journal (Prague)

Deutsche Math.
Deutsche Mathematik (Berlin)

Differentsial'nye Uravneniya
Differentsial'nye Uravneniya (Minsk). Translated as Differential Equations

Differential Equations
Differential Equations (New York). Translation of Differentsial'nye Uravneniya

Dokl. Akad. Nauk SSSR
Doklady Akademii Nauk SSSR (Moscow). Soviet Math. Dokl. is the English translation of its mathematics section

Duke Math. J.
Duke Mathematical Journal (Durham)

Econometrica
Econometrica, Journal of the Econometric Society (Chicago)

Edinburgh Math. Notes
The Edinburgh Mathematical Notes (Edinburgh)

Enseignement Math.
L'Enseignement Mathématique (Geneva)

Enzykl. Math.
Enzyklopädie der Mathematischen Wissenschaften mit Einschluss ihrer Anwendungen (Berlin)

Erg. Angew. Math.
Ergebnisse der Angewandte Mathematik (Berlin-New York)

Erg. Math.
Ergebnisse der Mathematik und ihrer Grenzgebiete (Berlin-New York)

Fund. Math.
Fundamenta Mathematicae (Warsaw)

Funkcial. Ekvac.
Fako de l'Funkcialaj Ekvacioj Japana Matematika Societo. Funkcialaj Ekvacioj (Serio Internacia) (Kobe, Japan)

Functional Anal. Appl.
Functional Analysis and its Applications (New York). Translation of Funktsional. Anal. Prilozhen.

Funktsional. Anal. Prilozhen.
Funktsional'nyi Analiz i ego Prilozheniya. Akademiya Nauk SSSR (Moscow). Translated as Functional Anal. Appl.

General Topology and Appl.
General Topology and its Applications (Amsterdam)

Hiroshima Math. J.
Hiroshima Mathematical Journal. Hiroshima Univ. (Hiroshima, Japan)

Hokkaido Math. J.
Hokkaido Mathematical Journal. Hokkaido Univ. (Sapporo, Japan)

IBM J. Res. Develop.
IBM Journal of Research and Development (Armonk, N.Y.)

Illinois J. Math.
Illinois Journal of Mathematics (Urbana)

Indag. Math.
Indagationes Mathematicae = Nederl. Akad. Wetensch. Proc.

Indian J. Math.
Indian Journal of Mathematics (Allahabad)

Indiana Univ. Math. J.
Indiana University Mathematics Journal (Bloomington)

Information and Control
Information and Control (New York)

Inventiones Math.
Inventiones Mathematicae (Berlin)

Izv. Akad. Nauk SSSR
Izvestiya Akademii Nauk SSSR (Moscow). Math. USSR-Izv. is the English translation of its mathematics section

J. Algebra
Journal of Algebra (New York)

J. Analyse Math.
Journal d'Analyse Mathématiques (Jerusalem)

J. Appl. Math. Mech.
Journal of Applied Mathematics and Mechanics (New York). Translation of Prikl. Mat. Mekh.

J. Approximation Theory
Journal of Approximation Theory (New York)

J. Assoc. Comput. Mach. (J. ACM)
Journal of the Association for Computing Machinery (New York)

J. Austral. Math. Soc.
The Journal of the Australian Mathematical Society (Sydney)

J. Combinatorial Theory
Journal of Combinatorial Theory. Series A and Series B (New York)

J. Comput. System Sci.
Journal of Computer and System Sciences (New York)

J. Differential Equations
Journal of Differential Equations (New York)

J. Differential Geometry
Journal of Differential Geometry (Bethlehem, Pa.)

J. Ecole Polytech.
Journal de l'Ecole Polytechnique (Paris)

J. Fac. Sci. Hokkaido Univ.
Journal of the Faculty of Science, Hokkaido University. Series I. Mathematics (Sapporo, Japan)

J. Fac. Sci. Univ. Tokyo
Journal of the Faculty of Science, University of Tokyo. Section I. (Tokyo)

J. Franklin Inst.
Journal of the Franklin Institute (Philadelphia)

J. Functional Anal.
Journal of Functional Analysis (New York)

J. Indian Math. Soc.
The Journal of the Indian Mathematical Society (Madras)

J. Inst. Elec. Engrs.
Journal of the Institution of Electrical Engineers (London)

J. Inst. Polytech. Osaka City Univ.
Journal of the Institute of Polytechnics, Osaka City University. Series A. Mathematics (Osaka)

J. London Math. Soc.
The Journal of the London Mathematical Society (London)

J. Math. Anal. Appl.
Journal of Mathematical Analysis and Applications (New York)

J. Math. and Phys.
Journal of Mathematics and Physics (Cambridge, Massachusetts, for issues prior to 1975; for 1975 and later, New York)

J. Math. Econom.
Journal of Mathematical Economics (Amsterdam)

J. Math. Kyoto Univ.
Journal of Mathematics of Kyoto University (Kyoto)

J. Math. Mech.
Journal of Mathematics and Mechanics (Bloomington)

J. Math. Pures Appl.
Journal de Mathématiques Pures et Appliquées (Paris)

J. Math. Soc. Japan
Journal of the Mathematical Society of Japan (Tokyo)

J. Mathematical Phys.
Journal of Mathematical Physics (New York)

J. Multivariate Anal.
Journal of Multivariate Analysis (New York)

J. Number Theory
Journal of Number Theory (New York)

J. Operations Res. Soc. Japan
Journal of the Operations Research Society of Japan (Tokyo)

J. Optimization Theory Appl.
Journal of Optimization Theory and Applications (New York)

J. Phys. Soc. Japan
Journal of the Physical Society of Japan (Tokyo)

J. Pure Appl. Algebra
Journal of Pure and Applied Algebra (Amsterdam)

J. Rational Mech. Anal.
Journal of Rational Mechanics and Analysis (Bloomington)

J. Reine Angew. Math.
Journal für die Reine und Angewandte Mathematik (Berlin). = Crelles J.

J. Res. Nat. Bur. Standards
Journal of Research of the National Bureau of Standards. Section B. Mathematics and Mathematical Physics (Washington)

J. Sci. Hiroshima Univ.
Journal of Science of Hiroshima University. Series A (Mathematics, Physics, Chemistry); Series A-I. (Mathematics) (Hiroshima)

J. Soviet Math.
Journal of Soviet Mathematics (New York). Translation of (1) Itogi Nauki—Seriya Matematika (Progress in Science—Mathematical Series); (2) Problemy Matematicheskogo Analiza (Problems in Mathematical Analysis); (3) Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.

J. Symbolic Logic
The Journal of Symbolic Logic (New Brunswick)

Japan. J. Math.
Japanese Journal of Mathematics (Tokyo)

Jber. Deutsch. Math. Verein. (Jber. D.M.V.)
Jahresbericht der Deutschen Mathematiker Vereinigung (Stuttgart)

Kibernetika (Kiev)
Otdelenie Matematiki, Mekhaniki i Kibernetiki Akademii Nauk Ukrainskoĭ SSR. Kibernetika (Kiev). Translated as Cybernetics

Kôdai Math. Sem. Rep.
Kôdai Mathematical Seminar Reports (Tokyo)

Linear Algebra and Appl.
Linear Algebra and Its Applications (New York)

Linear and Multilinear Algebra.
Linear and Multilinear Algebra (New York)

Mat. Sb.
Matematicheskiĭ Sbornik (Recueil Mathématique). Akademiya Nauk SSSR (Moscow). Translated as Math. USSR-Sb.

Mat. Tidsskr. A
Matematisk Tidsskrift. A (Copenhagen)

Mat. Zametki
Matematicheskiĭ Zametki.
Akademiya Nauk SSSR (Moscow). Translated as Math. Notes

Math. Ann.
Mathematische Annalen (Berlin-Göttingen-Heidelberg)

Math. Comp.
Mathematics of Computation (Providence). Formerly Math. Tables Aids Comput.

Math. J. Okayama Univ.
Mathematical Journal of Okayama University (Okayama, Japan)

Math. Japonicae
Mathematica Japonicae (Osaka)

Math. Nachr.
Mathematische Nachrichten (Berlin)

Math. Notes
Mathematical Notes of the Academy of Sciences of the USSR (New York). Translation of Mat. Zametki

Math. Rev.
Mathematical Reviews (Ann Arbor)

Math. Scand.
Mathematica Scandinavica (Copenhagen)

Math. Student
The Mathematical Student (Madras)

Math. Tables Aids Comput. (MTAC)
Mathematical Tables and Other Aids to Computation (Washington). Name changed to Mathematics of Computation in 1960 (vol. 14ff.)

Math. USSR-Izv.
Mathematics of the USSR-Izvestiya (Providence). Translation of Izv. Akad. Nauk SSSR

Math. USSR-Sb.
Mathematics of the USSR-Sbornik (Providence). Translation of Mat. Sb.

Math. Z.
Mathematische Zeitschrift (Berlin-Göttingen-Heidelberg)

Mathematika
Mathematika, A Journal of Pure and Applied Mathematics (London)

Meed. Lunds Univ. Mat. Sem.
Meddelanden från Lunds Universitets Matematiska Seminarium = Communications du Séminaire Mathématique de l'Université de Lund (Lund)

Mem. Amer. Math. Soc.
Memoirs of the American Mathematical Society (Providence)

Mem. Coll. Sci. Univ. Kyôto
Memoirs of the College of Science, University of Kyôto. Series A (Kyoto)

Mem. Fac. Sci. Kyushu Univ.
Memoirs of the Faculty of Science, Kyushu University. Series A. Mathematics (Fukuoka, Japan)

Mémor. Sci. Math.
Mémorial des Sciences Mathématiques (Paris)

Michigan Math. J.
The Michigan Mathematical Journal (Ann Arbor)

Mitt. Math. Ges. Hamburg
Mitteilungen der Mathematischen Gesellschaft in Hamburg (Hamburg)

Monatsh. Math. Phys.
Monatschefte für Mathematik und Physik (Vienna)

Monatsh. Math.
Monatshefte für Mathematik (Vienna)

Monograf. Mat.
Monografje Matematyczne (Warsaw)

Moscow Univ. Math. Bull.
Moscow University Mathematics Bulletin (New York). Translation of the mathematics section of Vestnik Moskov. Univ., Ser. I, Mat. Mekh.

Nachr. Akad. Wiss. Göttingen
Nachrichten der Akademie der Wissenschaften in Göttingen. Math.-Phys. Kl. (Göttingen)

Nagoya Math. J.
Nagoya Mathematical Journal (Nagoya)

Nederl. Akad. Wetensch. Proc.
Koninklijke Nederlandse Akademie van Wetenschappen, Proceedings. Series A. Mathematical Sciences (Amsterdam) = Indag. Math., Proc. Acad. Amsterdam

Nieuw Arch. Wisk.
Nieuw Archief voor Wiskunde (Groningen)

Numerische Math.
Numerische Mathematik (Berlin-Göttingen-Heidelberg)

Nuovo Cimento
Il Nuovo Cimento (Bologna)

Osaka J. Math.
Osaka Journal of Mathematics (Osaka)

Osaka Math. J.
Osaka Mathematical Journal (Osaka)

Pacific J. Math.
Pacific Journal of Mathematics (Berkeley)

Philos. Trans. Roy. Soc. London
Philosophical Transactions of the Royal Society of London. Series A (London)

Phys. Rev.
The Physical Review (New York)

Portugal. Math.
Portugaliae Mathematica (Lisbon)

Prikl. Mat. Mekh.
Adademiya Nauk SSSR. Otdelenie Tekhnicheskikh Nauk. Institut Mekhaniki Prikladnaya Matematika i Mekhanika (Moscow). Translated as J. Appl. Mat. Mech.

Proc. Acad. Amsterdam
= Nederl. Akad. Wetensch. Proc.

Proc. Amer. Math. Soc.
Proceedings of the American Mathematical Society (Providence)

Proc. Cambridge Philos. Soc.
Proceedings of the Cambridge Philosophical Society (Cambridge)

Proc. Imp. Acad. Tokyo
Proceedings of the Imperial Academy (Tokyo)

Proc. Japan Acad.
Proceedings of the Japan Academy (Tokyo)

Proc. London Math. Soc.
Proceedings of the London Mathematical Society (London)

Proc. Nat. Acad. Sci. US
Proceedings of the National Academy of Sciences of the United States of America (Washington)

Proc. Phys.-Math. Soc. Japan
Proceedings of the Physico-Mathematical Society of Japan (Tokyo)

Proc. Roy. Soc. London
Proceedings of the Royal Society of London. Series A (London)

Proc. Steklov Inst. Math.
Proceedings of the Steklov Institute of Mathematics (Providence). Translation of Trudy Mat. Inst. Steklov.

Prog. Theoret. Phys.
Progress of Theoretical Physics (Kyoto)

Publ. Inst. Math.
Publications de l'Institut Mathématique (Belgrade)

Publ. Inst. Math. Univ. Strasbourg
Publications de l'Institut de Mathêmatiques de l'Université de Strasbourg (Strasbourg)

Publ. Math. Inst. HES
Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques (Paris)

Publ. Res. Inst. Math. Sci.
Publications of the Research Institute for Mathematical Sciences (Kyoto)

Quart. Appl. Math.
Quarterly of Applied Mathematics (Providence)

Quart. J. Math.
The Quarterly Journal of Mathematics, Oxford. Second Series (Oxford)

Quart. J. Mech. Appl. Math.
The Quarterly Journal of Mechanics and Applied Mathematics (Oxford)

Rend. Circ. Mat. Palermo
Rendiconti del Circolo Matematico de Palermo (Palermo)

Rend. Sem. Mat. Univ. Padova
Rendiconti del Seminario Matematico dell'Universitá di Padova (Padua)

Rev. Mat. Hisp. Amer.
Revista Matemática Hispaño-Americana (Madrid)

Rev. Mod. Phys.
Reviews of Modern Physics (New York)

Rev. Un. Mat. Argentina
Revista de la Unión Matemática Argentina (Buenos Aires)

Rev. Univ. Tucumán
Revista Universidad Nacional de Tucumán, Facultad de Ciencias Exactas y Tecnologia. Serie A. Matemáticas y Fisica Teorica (Tucumán)

Roczniki Polsk. Towar. Mat.
Roczniki Polskiego Towarzystwa Matematycznego. Serja I. Prace Matematyczne (Krakow)

Rozprawy Mat.
Rozprawy Matematyczne, Polska Akademia Nauk, Instytut Matematyczny (Warsaw)

Russian Math. Surveys.
Russian Mathematical Surveys (London). Translation of Uspekhi Mat. Nauk

Sammul. Göschen
Sammulung Göschen (Leipzig)

Sankhyā
Sankhyā, The Indian Journal of Statistics. Series A and Series B (Calcutta)

S.-B. Berlin. Math. Ges.
Sitzungsberichte der Berliner Mathematischen Gesellschaft (Berlin)

S.-B. Deutsch. Akad. Wiss. Berlin
Sitzungsberichte der Deutschen Akademie der Wissenschaften zu Berlin, Mathematisch-Naturwissenschaftliche Klasse (Berlin)

S.-B. Heidelberger Akad. Wiss.
Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Klasse (Heidelberg)

S.-B. Math.-Nat. Kl. Bayer. Akad. Wiss.
Sitzungsberichte der Mathematisch-Naturwissenschaflichen Klasse der Bayerischen Akademie der Wissenschaften zu München (Munich)

S.-B. Öster. Akad. Wiss.
Sitzungsberichte der Österreichische Akademie der Wissenschaften (Vienna)

S.-B. Phys.-Med. Soz. Erlangen
Sitzungsberichte der Physikalisch-Medizinischen Sozietät zu Erlangen (Erlangen)

S.-B. Preuss. Akad. Wiss.
Sitzungsberichte der Preussischen Akademie der Wissenschaften. Physikalisch-Mathematische Klasse (Berlin)

Schr. Math. Inst. u. Inst. Angew. Math. Univ. Berlin
Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik der Universität Berlin (Berlin)

Schr. Math. Inst. Univ. Münster
Schriftenreihe des Mathematischen Instituts der Universität Münster (Münster)

Sci. Papers Coll. Gen. Ed. Univ. Tokyo
Scientific Papers of the College of General Education, University of Tokyo (Tokyo)

Sci. Rep. Tokyo Kyoiku Daigaku
Science Reports of the Tokyo Kyoiku Daigaku. Section A (Tokyo)

Scripta Math.
Scripta Mathematica. A Quarterly Journal devoted to the Philosophy, History, and Expository Treatment of Mathematics (New York)

Sém. Bourbaki
Séminaire Bourbaki (Paris)

SIAM J. Appl. Math.
SIAM Journal of Applied Mathematics. A Publication of the Society for Industrial and Applied Mathematics (Philadelphia)

SIAM J. Comput.
SIAM Journal on Computing (Philadelphia)

SIAM J. Control
SIAM Journal on Control (Philadelphia)

SIAM J. Math. Anal.
SIAM Journal on Mathematical Analysis (Philadelphia)

SIAM J. Numer. Anal.
SIAM Journal on Numerical Analysis (Philadelphia)

SIAM Rev.
SIAM Review (Philadelphia)

Siberian Math. J.
Siberian Mathematical Journal (New York). Translation of Sibirsk. Mat. Zh.

Sibirsk. Mat. Zh.
Akademiya Nauk SSSR. Sibirskoe Otdelenie. Sibirskiĭ Matematicheskiĭ Zhurnal (Moscow). Translated as Siberian Math. J.

Skr. Norske Vid. Akad. Oslo
Skrifter Utgitt av det Norske Videnskaps-Akademii Oslo. Matematisk-Naturvidenskapelig Klasse (Oslo)

Soviet Math. Dokl.
Soviet Mathematics, Doklady (Providence). Translation of mathematical section of Dokl. Akad. Nauk SSSR

SRI J.
Stanford Research Institute Journal (Menlo Park)

Studia Math.
Studia Mathematica. (Wrocław)

Sûbutu-kaisi
Nihon Sûgaku-buturi-gakkai Kaisi (Tokyo)

Sûgaku
Sûgaku, Mathematical Society of Japan (Tokyo)

Summa Brasil. Math.
Summa Brasiliensis Mathematicae (Rio de Janeiro)

Tensor
Tensor (Chigasaki, Japan)

Teor. Veroyatnost. i Primenen.
Teoriya Veroyatnosteĭ i ee Primenenie. Akademiya Nauk SSSR (Moscow). Translated as Theor. Prob. Appl.

Theor. Prob. Appl.
Theory of Probability and Its Applications. Society for Industrial and Applied Mathematics. English translation of Teor. Veroyatnost. i Primenen. (Philadelphia)

Tôhoku Math. J.
The Tôhoku Mathematical Journal (Sendai, Japan)

Tôhoku-rihô
Tôhoku Teikokudaigaku Rikahôkoku (Sendai, Japan)

Topology
Topology. An International Journal of Mathematics (Oxford)

Trans. Amer. Math. Soc.
Transactions of the American Mathematical Society (Providence)

Trans. Moscow Math. Soc.
Transactions of the Moscow Mathematical Society (Providence). Translation of Trudy Moskov. Mat. Obshch.

Trudy Mat. Inst. Steklov.
Trudy Matematicheskogo Instituta im. V. A.

Steklova. Akademiya Nauk SSSR (Moscow-Leningrad). Translated as Proc. Steklov Inst. Math.

Trudy Moskov. Obshch.
Trudy Moskovskogo Matematicheskogo Obshchestva (Moscow). Translated as Trans. Moscow Math. Soc.

Tsukuba J. Math.
Tsukuba Journal of Mathematics. Univ. Tsukuba (Ibaraki, Japan)

Ukrain. Mat. Zh.
Akademiya Nauk Ukrainskoĭ SSR. Institut Matematiki. Ukrainskiĭ Matematicheskiĭ Zhurnal (Kiev). Translated as Ukrainian Math. J.

Ukrainian Math. J.
Ukrainian Mathematical Journal (New York). Translation of Ukrain. Mat. Zh.

Uspekhi Mat. Nauk
Uspekhi Matematicheskikh Nauk (Moscow-Leningrad). Translated as Russian Math. Surveys

Vestnik Moskov. Univ.
Vestnik Moskovskogo Universiteta. I, Matematika i Mekhanika (Moscow). Mathematical section translated as Moscow Univ. Math. Bull.

Vierteljschr. Naturf. Ges. Zürich
Vierteljahrsschrifte der Naturforschenden Gesellschaft in Zürich (Zurich)

Z. Angew. Math. Mech. (Z.A.M.M.)
Zeitschrift für Angewandte Mathematik und Mechanik, Ingenieurwissenschaftliche Forschungsarbeiten (Berlin)

Z. Angew. Math. Phys. (Z.A.M.P.)
Zeitschrift für Angewandte Mathematik und Physik (Basel)

Z. Wahrscheinlichkeitstheorie
Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete (Berlin)

Zbl. Angew. Math.
Zentralblatt für Angewandte Mathematik (Berlin)

Zbl. Math.
Zentralblatt für Mathematik und ihre Grenzgebiete (Berlin-Göttingen-Heidelberg)

Zh. Èksper. Teoret. Fiz.
Zhurnal Èksperimental'noĭ i Teoreticheskoĭ Fiziki (Moscow)

Publishers

Academic Press
Academic Press Inc., New York-London

Addison-Wesley
Addison-Wesley Publishing Company, Inc., Reading (Massachusetts)-Menlo Park (California)-London-Don Mills (Ontario)

Akadémiai Kiadó
A kiadásért felös: az Adadémiai Kiadó igazatója (Publishing House of the Hungarian Academy of Sciences), Budapest

Akademie-Verlag
Berlin

Akademische Verlag.
Akademische Verlagsgesellschaft, Leipzig

Allen
W. H. Allen & Co. Ltd., London

Allen & Unwin
Allen & Unwin, Inc., Winchester (Massachusetts)

Allyn & Bacon
Allyn & Bacon, Inc., Newton (Massachusetts)

Almqvist and Wiksell
Almqvist och Wiksell Förlag, Stockholm

Asakura
Asakura-syoten, Tokyo

Aschelhoug
H. Aschelhoug and Company, Oslo

Baihûkan
Tokyo

Benjamin
W. A. Benjamin, Inc., New York-London

Birkhäuser
Birkhäuser Verlag, Basel-Stuttgart

Blackie
Blackie & Son Ltd., London-Glasgow

Cambridge Univ. Press
Cambridge University Press, London-New York

Chapman & Hall
Chapman & Hall Ltd., London

Chelsea
Chelsea Publishing Company, New York

Clarendon Press
Oxford University Press, Oxford

Cremona
Edizioni Cremonese, Rome

de Gruyter
Walter de Gruyter and Company, Berlin

Dekker
Marcel Dekker, Inc., New York

Deutscher Verlag der Wiss.
Deutscher Verlag der Wissenschaften, Berlin

Dover
Dover Publications, Inc., New York

Dunod
Dunod, Editeur, Paris

Elsevier
Elsevier Publishing Company, Amsterdam-London-New York

Fizmatgiz
Gosudarstvennoe Izdatel'stvo Fiziko-Matematicheskoĭ Literatury, Moscow

Gauthier-Villars
Gauthier-Villars & C^{ie}, Editeur, Paris

Ginn
Ginn and Company, Waltham (Massachusetts)-Toronto-London

Gordon & Breach
Gordon & Breach, Science Publishers Ltd., London

Goztekhizdat
Gosudarstvennoe Izdatel'stvo Tekhniko-Teoreticheskoĭ Literatury, Moscow

Griffin
Charles Griffin and Company Ltd., London

Hafner
Hafner Publishing Company, New York

Harper & Row
Harper & Row Publishers, New York-Evanston-London

Hermann
Hermann & C^{ie}, Paris

Hirokawa
Hirokawa-syoten, Tokyo

Hirzel
Verlag von S. Hirzel, Leipzig

Holden-Day
Holden-Day, Inc., San Francisco-London-Amsterdam

Holt, Rinehart and Winston
Holt, Rinehart and Winston, Inc., New York-Chicago-San Francisco-Toronto-London

Interscience
Interscience Publishers, Inc., New York-London

Iwanami
Iwanami Shoten, Tokyo

Kawade
Kawade-syobô, Tokyo

Kinokoniya
Kinokoniya Company, Tokyo

Kyôritu
Kyôritu-syuppan, Tokyo

Lippincott
J. B. Lippincott Company, Philadelphia

Longman
Longman Group, Ltd., Harlow (Essex)

Longmans, Green
Longmans, Green and Company, Ltd., London-New York-Toronto-Bombay-Calcutta-Madras

Macmillan
The Macmillan Company, New York-London

Maki
Maki-syoten, Tokyo

Maruzen
Maruzen Company Ltd., Tokyo

Masson
Masson et C^{ie}, Paris

Math-Sci Press
Math-Sci Press, Brookline (Massachusetts)

McGraw-Hill
McGraw-Hill Book Company, Inc., New York-London-Toronto

Methuen
Methuen and Company Ltd., London

MIT Press
The MIT Press, Cambridge (Massachusetts)-London

Nauka
Izdatel'stvo Nauka, Moscow

Noordhoff
P. Noordhoff Ltd., Groningen

North-Holland
North-Holland Publishing Company, Amsterdam

Oldenbourg
Verlag von R. Oldenbourg, Munich-Vienna

Oliver & Boyd
Oliver & Boyd Ltd., Edinburgh-London

Oxford Univ. Press
Oxford University Press, London-New York

Pergamon
Pergamon Press, Oxford-London-Edinburgh-New York-Paris-Frankfurt

Polish Scientific Publishers
Państwowe Wydawnictwo Naukowe, Warsaw

Prentice-Hall
Prentice-Hall, Inc., Englewood Cliffs (New Jersey)

Princeton Univ. Press
Princeton University Press, Princeton

Random House
Random House, Inc., New York

Sibundô
Tokyo

Springer
Springer-Verlag, Berlin (-Göttingen)-Heidelberg-New York

Teubner
B. G. Teubner Verlagsgesellschaft, Leipzig-Stuttgart

Tôkai
Tôkai-syobô, Tokyo

Tokyo-tosyo
Tokyo

Tokyo Univ. Press
Tokyo University Press, Tokyo

Ungar
Frederick Ungar Publishing Company, New York

Univ. of Tokyo Press
University of Tokyo Press, Tokyo

Utida-rôkakuho
Tokyo

Van Nostrand
D. Van Nostrand Company, Inc., Toronto-New York-London

Vandenhoeck & Ruprecht
Göttingen

Veit
Verlag von Veit & Company, Leipzig

Vieweg
Friedr Vieweg und Sohn Verlagsgesellschaft mbH, Wiesbaden

Wiley
Wiley & Sons, Inc., New York-London

Wiley-Interscience
Wiley & Sons, Inc., New York-London

Zanichelli
Nicola Zanichelli Editore, Bologna

Special Notation

This list contains the notation commonly and frequently used throughout this work. The symbol * means that the same notation is used with more than one meaning. For more detailed definitions or further properties of the notation, see the articles cited.

Notation	Example	Definition	Article and Section
I. Logic			
$\forall$	$\forall x F(x)$	Universal quantifier (for all x, $F(x)$ holds)	411B, C
$\exists$	$\exists x F(x)$	Existential quantifier (there exists an x such that $F(x)$ holds)	411B, C
$\wedge$, &	$A \wedge B$, $A \& B$	Conjunction, logical product (A and B)	411B*
$\vee$	$A \vee B$	Disjunction, logical sum (A or B)	411B*
$\neg$	$\neg A$	Negation (not A)	411B
$\rightarrow$, $\supset$, $\Rightarrow$	$A \rightarrow B$, $A \Rightarrow B$	Implication (A implies B)	411B*
$\leftrightarrow$, $\Leftrightarrow$, $\rightleftarrows$	$A \Leftrightarrow B$	Equivalence (A and B are logically equivalent)	411B
II. Sets			
$\in$	$x \in X$	Membership (element x is a member of the set X)	381A
$\notin$	$x \notin X$	Nonmembership (element x is not a member of the set X)	381A
$\subset$	$A \subset B$	Inclusion (A is a subset of B)	381A
$\not\subset$	$A \not\subset B$	Noninclusion (A is not a subset of B)	381A
$\subsetneqq$	$A \subsetneqq B$	Proper inclusion (A is a proper subset of B)	381A
$\varnothing$		Empty set	381A
$\cup$, $\bigcup$	$A \cup B$, $\bigcup A_\lambda$	Union, join	381B, D*
$\cap$, $\bigcap$	$A \cap B$, $\bigcap A_\lambda$	Intersection, meet	381B, D*
^{c}, C	A^c, $C(A)$	Complement (of a set A)	381B
$-$, $\setminus$	$A - B$, $A \setminus B$	Difference ($A - B = A \cap B^c$)	381B
$\times$	$A \times B$	Cartesian product (of A and B)	381B*
R, $\sim$	xRy, $x \sim y$	Equivalence relation (for two elements x, y)	135A*
$/$	A/R	Quotient set (set of equivalence classes of A with respect to an equivalence relation R)	135B*
Π	$\Pi_\lambda A_\lambda$	Cartesian product (of the A_λ)	381E
Σ, $\coprod$	ΣA_λ, $\coprod A_\lambda$	Direct sum (of the A_λ)	381E
$\mathfrak{P}$	$\mathfrak{P}(A)$	Power set (set of all subsets of A)	381E
	B^A	Set of all mappings from A to B	381C
$\{\mid\}$	$\{x \mid P(x)\}$	Set of all elements x with the property $P(x)$	381A

Notation	Example	Definition	Article and Section
{ }	$\{a_\lambda\}_{\lambda\in\Lambda}$	Family with index set Λ	165D
	$\{a_n\}$	Sequence (of numbers, points, functions, or sets)	165D
$\bar{\bar{\ }}$, | |, #	$\bar{\bar{X}}$, $\|X\|$, $\#X$	Cardinal number (of the set X)	49A*
$\aleph$	$\aleph_\beta$	Aleph (transfinite cardinal)	49E
$\to$	$f:X\to Y$	Mapping (f from X to Y)	381C*
$\mapsto$	$f:X\mapsto Y$	Mapping (where $f(X)=Y$, but not in the present volumes)	381C
1, id	1_A, id_A	Identity mapping (identity function)	381C
c, χ	$c_X(x)$, $\chi_X(x)$	Characteristic function (representing function)	381C
|	$f\|A$	Restriction (of a mapping f to A)	381C*
$\circ$	$g\circ f$	Composite (of mappings f and g)	381C
lim sup, $\overline{\lim}$	$\limsup A_n$	Superior limit (of the sequence of sets A_n)	270C*
lim inf, $\underline{\lim}$	$\liminf A_n$	Inferior limit (of the sequence of sets A_n)	270C*
lim	$\lim A_n$	Limit (of the sequence of sets A_n)	270C*
$\varinjlim$	$\varinjlim A_\lambda$	Inductive limit (of A_λ)	210B
$\varprojlim$	$\varprojlim A_\lambda$	Projective limit (of A_λ)	210B
III. Order			
(,)	(a,b)	Open interval $\{x\|a<x<b\}$	355C*
[,]	$[a,b]$	Closed interval $\{x\|a\leqslant x\leqslant b\}$	355C*
(,]	$(a,b]$	Half-open-interval $\{x\|a<x\leqslant b\}$	355C
[,)	$[a,b)$	Half-open interval $\{x\|a\leqslant x<b\}$	355C
max	$\max A$	Maximum (of A)	311B
min	$\min A$	Minimum (of A)	311B
sup	$\sup A$	Supremum, least upper bound (of A)	311B
inf	$\inf A$	Infimum, greatest lower bound (of A)	311B
$\ll$	$a\ll b$	Very large (b is very large compared to a)	
$\cup$, $\vee$	$a\cup b$, $a\vee b$	Join of a, b in an ordered set	243A*
$\cap$, $\wedge$	$a\cap b$, $a\wedge b$	Meet of a, b in an ordered set	243A*
IV. Algebra			
mod	$a\equiv b\ (\mathrm{mod}\ n)$	Modulo (a and b are congruent modulo n)	297G
|	$a\|b$	Divisibility (a divides b)	297A*
$\nmid$	$a\nmid b$	Nondivisibility (a does not divide b)	297A
det, | |	$\det A$, $\|A\|$	Determinant (of a square matrix A)	103A*

Notation	Example	Definition	Article and Section
tr, Sp	tr A, Sp A	Trace (of a square matrix A)	269F
^{t}, ^{T}, $'$	^tA; A^t, A^T, A'	Transpose (of a matrix A)	269B
I	I_n	Unit matrix (of degree n)	269A
E_{ij}		Matrix unit (matrix whose (i,j)-component is 1 and all others are 0)	269B
$\otimes$	$A\otimes B$	Kronecker product (of two matrices A and B)	269C*
$\cong$	$M\cong N$	Isomorphism (of two algebraic systems M and N)	256B
/	M/N	Quotient space (of an algebraic system M by N)	256F*
dim	dim M	Dimension (of a linear space, etc.)	256C
Im	Im f	Image (of a mapping f)	277E*
Ker	Ker f	Kernel (of a mapping f)	277E
Coim	Coim f	Coimage (of a mapping f)	277E
Coker	Coker f	Cokernel (of a mapping f)	277E
δ_{ij}, δ_i^j		Kronecker delta ($\delta_{ii}=1$ and $\delta_{ij}=0$ for $i\neq j$)	269A
(,), ·	$(\mathbf{a},\mathbf{b})$, $\mathbf{a}\cdot\mathbf{b}$	Inner product (of two vectors **a** and **b**)	442B*
[,], ×	$[\mathbf{a},\mathbf{b}]$, $\mathbf{a}\times\mathbf{b}$	Vector product (of two 3-dimensional vectors **a** and **b**)	442C*
$\otimes$	$M\otimes N$	Tensor product (of two modules M and N)	277J, 256I*
Hom	Hom(M,N)	Set of all homomorphisms (from M to N)	277B
Hom_A	$\mathrm{Hom}_A(M,N)$	Set of all A-homomorphisms (of an A-module M to an A-module N)	277E
Tor	$\mathrm{Tor}_n(M,N)$	Torsion product (of M, N)	200D
Ext	$\mathrm{Ext}^n(M,N)$	Extension (of M, N)	200G
$\wedge$, $\wedge^p$	$\wedge M$, $\wedge^p M$	Exterior algebra (of a linear space M), pth exterior product (of M)	256O
V. Algebraic Systems			
$\mathbf{N}$		Set of all natural numbers	294A
$\mathbf{Z}$		Set of all rational integers	294A
$\mathbf{Z}_m$		$\mathbf{Z}/m\mathbf{Z}$ (set of all residue classes modulo m)	297G*
$\mathbf{Q}$		Set of all rational numbers	294A
$\mathbf{R}$		Set of all real numbers	294A
$\mathbf{C}$		Set of all complex numbers	294A
$\mathbf{H}$		Set of all quaternions	29B
$GF(q)$, $\mathbf{F}_q$		Finite field (with q elements)	149M

Notation	Example	Definition	Article and Section
$\mathbf{Q}_p$		p-adic number field (p is a prime)	439F
$\mathbf{Z}_p$		Ring of p-adic integers	439F
[ ]	$k[x_1,\ldots,x_n]$	Polynomial ring (of variables $x_1,\ldots,x_n$ with coefficients in k)	369A
( )	$k(x_1,\ldots,x_n)$	Field extension (of k by $x_1,\ldots,x_n$)	149D
[[ ]], { }	$k[[x_1,\ldots,x_n]]$	Formal power series ring (with coefficients in k).	370A
		Note: The symbols **N**, **Z**, **Q**, **R**, **C**, and **H** stand for sets, each with its own natural mathematical structure	
VI. Groups			
GL	$GL(V)$, $GL(n,K)$	General linear group (over V, or over K of degree n)	60B
SL	$SL(n,K)$	Special linear group (over K of degree n)	60B
PSL	$PSL(n,K)$	Projective special linear group (over K of degree n)	60B
U	$U(n)$	Unitary group (of degree n)	60F
SU	$SU(n)$	Special unitary group (of degree n)	60F
O	$O(n)$	Orthogonal group (of degree n)	60I
SO	$SO(n)$	Special orthogonal group, rotation group (of degree n)	60I
$Spin$	$Spin(n)$	Spinor group (of degree n)	61D
Sp	$Sp(n)$	Symplectic group (of degree n)	60L
[For $PGL(n,K)$, $LF(n,K)$, $PU(n)$, $Sp(n)$, $PSp(n,K)$ → 60 Classical Groups]			
VII. Topology (Convergence)			
$\rightarrow$	$a_n \rightarrow a$	Convergence (sequence a_n converges to a)	87B, E*
$\downarrow$, $\searrow$	$a_n \downarrow a$, $a_n \searrow a$	Convergence monotonically decreasing	87B
$\uparrow$, $\nearrow$	$a_n \uparrow a$, $a_n \nearrow a$	Convergence monotonically increasing	87B
lim	$\lim a_n$	Limit (of a sequence a_n)	87B, E*
lim sup, $\overline{\lim}$	$\limsup a_n$, $\overline{\lim}\, a_n$	Superior limit (of a sequence a_n)	87C*
lim inf, $\underline{\lim}$	$\liminf a_n$, $\underline{\lim}\, a_n$	Inferior limit (of a sequence a_n)	87C*
^{a}, $^-$, Cl	E^a, $\bar{E}$, $\mathrm{Cl}\, E$	Closure (of a set E)	425B
^{i}, $^\circ$, Int	E^i, E°, $\mathrm{Int}\, E$	Interior (of a set E)	425B
ρ, d	$\rho(x,y)$, $d(x,y)$	Distance (between two points x and y)	273B*
$\Vert\ \Vert$	$\Vert x \Vert$	Norm (of x)	37B
l.i.m.	$\mathrm{l.i.m.}\, f_n$	Limit in the mean (of a sequence f_n)	168B

Notation	Example	Definition	Article and Section
s-lim	s-lim x_n	Strong limit (of a sequence x_n)	37B
w-lim	w-lim x_n	Weak limit (of a sequence x_n)	37E
$\simeq$	$f \simeq g$	Homotopy (of two mappings f and g)	202B
$\approx$	$X \approx Y$	Homeomorphism (of two topological spaces X and Y)	425G

VIII. Geometry and Algebraic Topology

Notation	Example	Definition	Article and Section
E^n		Euclidean space (of dimension n)	140
P^n		Projective space (of dimension n)	343B
S^n		Spherical surface (of dimension n)	140
$\mathbf{T}^n$		Torus (of dimension n)	422E
H^n	$H^n(X, A)$	n-dimensional cohomology group (of X with coefficients in A)	201H
H_n	$H_n(X, A)$	n-dimensional homology group (of X with coefficients in A)	201G
	$H_n(C)$	(of chain complex C)	201B
π_n	$\pi_n(X)$	n-dimensional homotopy group (of X)	202J, 170
∂	∂C	Boundary (of C)	201B
δ	δf	Coboundary (of f)	201H*
Sq	$Sq^i x$	Streenrod square (of x)	64B
$\mathscr{P}$	$\mathscr{P}_p^r(x)$	Steenrod pth power (of x)	64B
$\smile$	$z_1 \smile z_2$	Cup product (of z_1 and z_2)	201I
$\frown$	$z_1 \frown z_2$	Cap product (of z_1 and z_2)	201K
$\wedge$	$\omega \wedge \eta$	Exterior product (of two differential forms ω and η)	105Q*
d	$d\omega$	Exterior derivative (of a differential form ω)	105Q
grad	grad φ	Gradient (of a function φ)	442D
rot	rot $\mathbf{u}$	Rotation (of a vector $\mathbf{u}$)	442D
div	div $\mathbf{u}$	Divergence (of a vector $\mathbf{u}$)	442D
Δ	$\Delta\varphi$	Laplacian (of a function φ)	323A
$\square$	$\square\varphi$	d'Alembertian (of a function φ)	130A
D	$D\varphi$	Differential operator	112A*
$\dfrac{D(u_1,\ldots,u_n)}{D(x_1,\ldots,x_n)}, \left\lvert\dfrac{\partial u_i}{\partial x_j}\right\rvert, \det\left(\dfrac{\partial u_i}{\partial x_j}\right)$		Jacobian determinant (of $(u_1,\ldots,u_n)$ with respect to $(x_1,\ldots,x_n)$)	208B
$\dfrac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)}, \left(\dfrac{\partial u_i}{\partial x_j}\right)$		Jacobian matrix (of $(u_1,\ldots,u_n)$ with respect to $(x_1,\ldots,x_n)$)	208B

IX. Function Spaces

Notation	Example	Definition	Article and Section
C	$C(\Omega)$	Space of continuous functions (on Ω)	168B(1)

Notation	Example	Definition	Article and Section
L_p	$L_p(\Omega)$, $L_p(a,b)$	Space of functions such that $\lvert f\rvert^p$ is integrable on Ω	168B(2)
C^l	$C^l(L)(1 \leqslant l \leqslant \infty)$	Space of functions of class C^l	168B(9)
$\mathscr{D}$	$\mathscr{D}(\Omega)$	Space of C^∞ functions with compact support	168B(13)
$\mathscr{E}$	$\mathscr{E}(\Omega)$	Space of C^∞ functions	168B(13)

[For $\mathscr{A}(\Omega)$, $A(\Omega)$, $A_p(\Omega)$, $\mathscr{B}(\Omega)(=\mathrm{D}_{L^\infty}(\Omega))$, $BMO(\mathbf{R}^n)$, $BV(\Omega)$, c, C, $C_0(\Omega)$, $C_\infty(\Omega)$, $C_0^l(\Omega)$, $\mathscr{D}_{L^p}(\Omega)$, $\mathscr{D}_{\{M^p\}}(\Omega)$, $\mathscr{D}_{(M^p)}(\Omega)$, $\mathscr{E}_{\{M^p\}}(\Omega)$, $\mathscr{E}_{(M^p)}(\Omega)$, $H_p(\mathbf{R}^n)$, $H^l(\Omega)$, $H_0^l(\Omega)$, $\Lambda^S(\mathbf{R}^n)$, $\bigcap \lambda(\alpha^{(k)})$, $\sum \lambda^\times(\alpha^{(k)})$, l_p, $L_{(p,q)}(\Omega)$, m, $M(\Omega)$, $\mathscr{O}(\Omega)$, $\mathscr{O}_p(\Omega)$, $\mathscr{S}$, s, $S(\Omega)$, $W_p^l(\Omega)$ → 168 Function Spaces. For $\mathscr{B}(\Omega)$ (Space of Sato hyperfunctions), $\mathscr{D}'(\Omega)$, $\mathscr{E}'(\Omega)$, $\mathscr{O}_c$, $\mathscr{O}_M$, $\mathscr{S}'(\mathbf{R}^n)$ → 125 Distributions and Hyperfunctions]

X. Functions

$\lvert\ \rvert$	$\lvert z\rvert$	Absolute value (of a complex number z)	74B*
Re	$\mathrm{Re}\, z$	Real part (of a complex number z)	74A
Im	$\mathrm{Im}\, z$	Imaginary part (of a complex number z)	74A*
$\bar{\ }$	$\bar{z}$	Complex conjugate (of a complex number z)	74A
arg	$\arg z$	Argument (of a complex number z)	74C
[ ]	$[\alpha]$	Gauss symbol (greatest integer not exceeding a real number α)	83A
O	$f(x)=O(g(x))$	Landau's notation ($f(x)/g(x)$ is bounded for $x\rightarrow\alpha$)	87G
o	$f(x)=o(g(x))$	Landau's notation ($f(x)/g(x)$ tends to 0 for $x\rightarrow\alpha$)	87G
$\sim$	$f(x)\sim g(x)$	Infinite or infinitesimal of the same order (for $x\rightarrow\alpha$)	87G*
D	$D(T)$	Domain (of an operator T)	251A
R	$R(T)$	Range (of an operator T)	251A
supp	$\mathrm{supp}\, f$	Support (of a function f)	168B(1)
p.v.	$\mathrm{p.v.}\int_a^b f(x)dx$	Cauchy's principal value (of an integral)	216D
Pf	$\mathrm{Pf}\int f(x)dx$	Finite part (of an integral)	125C
δ	$\delta(x)$, δ_x	Dirac's delta function (measure or distribution)	125C*
exp	$\exp x$	Exponential function ($\exp x = e^x$)	113D, 269H
log, Log	$\log x$, $\mathrm{Log}\, x$	Natural logarithmic function and its principal value, respectively	131D, G
$\sin x$, $\cos x$, $\tan x$, $\sec x$, $\mathrm{cosec}\, x$, $\mathrm{cotan}\, x$		Trigonometric functions	131E
$\arcsin x$, $\arccos x$, $\arctan x$		Inverse trigonometric functions	131E
$\mathrm{Arc}\sin x$, $\mathrm{Arc}\cos x$, $\mathrm{Arc}\tan x$		Principal value of inverse trigonometric functions	131E

Notation	Example	Definition	Article and Section
$\sinh x$, $\cosh x$, $\tanh x$		Hyperbolic functions	131F
$\left(\ \right)$, C	$\binom{n}{r}$, ${}_nC_r$	Binomial coefficient, combination	330
P	${}_nP_r$	Permutation	330
!	$n!$	Factorial (of n)	330
φ	$\varphi(n)$	Euler function	295C*
μ	$\mu(n)$	Möbius function	295C
ζ	$\zeta(z)$	Riemann zeta function	450B*
J_ν	$J_\nu(z)$	Bessel function of the first kind	39B
Γ	$\Gamma(x)$	Gamma function	174A
B	$B(x,y)$	Beta function	174C
F	$F(\alpha,\beta,\gamma;z)$	Gauss's hypergeometric function	206A
P	$P\begin{Bmatrix} a & b & c \\ \lambda & \mu & \nu \\ \lambda' & \mu' & \nu' \end{Bmatrix} x$	Riemann's P function	253B
Li	$\mathrm{Li}(x)$	Logarithmic integral	167D
XI. Probability			
P, Pr	$P(E)$, $\Pr(\varepsilon)$	Probability (of an event)	342B*
E	$E(X)$	Mean or expectation (of a random variable X)	342C
V, σ^2	$V(X)$, $\sigma^2(X)$	Variance (of a random variable X)	342C
ρ	$\rho(X,Y)$	Correlation coefficient (of two random variables X and Y)	342C*
$P(\mid)$	$P(E\mid F)$	Conditional probability (of an event E under the condition F)	342E
$E(\mid)$	$E(X\mid Y)$	Conditional mean (of a random variable X under the condition Y)	342E
N	$N(m,\sigma^2)$	One-dimensional normal distribution (with mean m and variance σ^2)	Appendix A, Table 22
	$N(\boldsymbol{\mu},\Sigma)$	Multidimensional normal distribution (with mean vector $\boldsymbol{\mu}$ and variance matrix Σ)	Appendix A, Table 22
P	$P(\lambda)$	Poisson distribution (with parameter λ)	Appendix A, Table 22*

Systematic List of Articles

XI
Complex Analysis

XII
Functional Analysis

XIII
Differential, Integral, and Functional Equations

XIV
Special Functions

XV
Numerical Analysis

XVI
Computer Science and Combinatorics

XVII
Probability Theory

XVIII
Statistics

XIX
Mathematical Programming and Operations Research

Alphabetical List of Articles

Contributors to the Second Edition

Abe Eiichi
Akahira Masafumi
Akaike Hirotugu
Akaza Touru
Amari Shun-ichi
Ando Tsuyoshi
Araki Huzihiro
Araki Shôrô
Arimoto Suguru
Asai Akira
Fuji'i'e Tatuo
Fujiki Akira
Fujisaki Genjiro
Fujita Hiroshi
Fujiwara Daisuke
Fujiwara Masahiko
Fukushima Masatoshi
Furuya Shigeru
Hamada Yûsaku
Hasumi Morisuke
Hattori Akio
Hayashi Kiyoshi
Hida Takeyuki
Hijikata Hiroaki
Hirose Ken
Hitotumatu Sin
Huzii Mituaki
Ibaraki Toshihide
Ihara Shin-ichiro
Iitaka Shigeru
Ikawa Mitsuru
Ikebe Teruo
Ikebe Yasuhiko
Ikeda Nobuyuki
Inagaki Nobuo
Inoue Atsushi
Iri Masao
Itaya Nobutoshi
Itô Kiyosi
Itô Seizô
Ito Yuji
Itoh Mitsuhiro
Iwahori Nobuyoshi
Iyanaga Shôkichi
Kadoya Norihiko
Kageyama Sampei
Kamae Teturo
Kaneko Akira
Kaneyuki Soji
Kasahara Kōji
Kataoka Shinji
Kato Junji
Kato Mitsuyoshi
Kawai Takahiro
Kawasaki Tetsuro
Kimura Tosihusa
Kishi Masanori
Kobayashi Shoshichi
Kodama Yukihiro
Komatsu Hikosaburo
Kômura Yukio
Kondo Takeshi
Konishi Yoshio
Konno Hiroshi
Kotani Shinichi
Kubota Yoshihisa
Kumano-go Hajime
Kunita Hiroshi
Kuramoto Yuki
Kuroda Shige-Toshi
Kuroda Tadashi
Kusama Tokitake
Kusano Takaŝi
Kusunoki Yukio
Mabuchi Toshiki
Maeda Shūichiro
Maeda Yoshiaki
Maehara Shôji
Makabe Hajime
Maruyama Masaki
Matsumoto Kikuji
Matsumoto Yukio
Matsumura Mutsuhide
Mitsui Takayoshi
Mizohata Sigeru
Mizutani Tadayoshi
Mori Masatake
Morimoto Akihiko
Morimura Hidenori
Morita Shigeyuki
Morita Yasuo
Motohashi Nobuyoshi
Murasugi Kunio
Nagami Keio
Nagata Masayoshi
Nakagawa Hisao
Nakamura Tokushi
Nakaoka Minoru
Namba Kanji
Namba Makoto
Namikawa Yukihiko
Naruki Isao
Niiro Fumio
Nishikawa Seiki
Nishimura Junichi
Nishitani Kensaku
Nisio Makiko
Noguchi Hiroshi
Nozaki Akihiro
Ôaku Toshinori
Obata Morio
Ochiai Takushiro
Ogiue Koichi
Ohta Masami
Ohya Yujiro
Oikawa Kôtaro
Oka Mutsuo
Okabe Yasunori

Okuno Tadakazu
Omori Hideki
Oshima Toshio
Ozawa Mitsuru
Saito Kyoji
Saito Masahiko
Saito Tosiya
Sakai Fumio
Sakawa Yoshiyuki
Sato Fumitaka
Sato Ken-iti
Sawashima Ikuko
Shibata Katsuyuki
Shibata Keiichi
Shiga Koji
Shiga Tokuzo
Shikata Yoshihiro
Shimada Nobuo
Shimakura Norio
Shimizu Ryoichi
Shioda Tetsuji
Shiohama Katsuhiro
Shiraiwa Kenichi
Shiratani Katsumi
Sibuya Masaaki
Sugie Toru
Sugiura Mitsuo
Sugiura Shigeaki
Suita Nobuyuki
Sumihiro Hideyasu
Suzuki Masuo
Suzuki Mitsuo
Takahashi Tsuneo
Takakuwa Shoichiro
Takami Hideo
Takasu Toru
Takeuchi Kei
Takeuchi Masaru
Takeuti Gaishi
Tamura Itiro
Tanabe Kunio
Tanaka Hiroshi
Tanaka Masatsugu
Tanaka Shunichi
Tanno Shukichi
Tatsumi Tomomasa
Terada Toshiaki
Toda Hideo
Toda Hirosi
Toda Morikazu
Toda Nobushige
Tomiyama Jun
Totoki Haruo
Tugé Tosiyuki
Uchida Fuichi
Uchiyama Saburô
Ueno Kenji
Ukai Seiji
Ura Shoji
Ushijima Teruo
Wakabayashi Isao
Washizu Kyuichiro
Watanabe Kimio
Watanabe Shinzo
Watanabe Takesi
Yamaguchi Masaya
Yamamoto Koichi
Yamamoto Sumiyasu
Yamamoto Tetsuro
Yamazaki Masao
Yamazaki Tadashi
Yanagawa Takashi
Yanagiwara Hiroshi
Yanai Haruo
Yano Tamaki
Yokonuma Takeo
Yoshikawa Atsushi
Yoshida Junzo
Yosida Kôsaku

Contributors to the First Edition

Aizawa Sadakazu
Akao Kazuo
Akizuki Yasuo
Amemiya Ayao
Anzai Hirotada
Aoki Kiyoshi
Araki Huzihiro
Araki Shôrô
Arima Akito
Asaka Tetsuichi
Asano Keizo
Asatani Teruo
Azumaya Gorô
Fujinaka Megumu
Fujita Hiroshi
Fujiwara Daisuke
Fukuda Takeo
Fukutomi Setsuo
Furuya Shigeru
Goto Morikuni
Gotô Motinori
Gotō Yūzo
Hagihara Yusuke
Harada Manabu
Hasimoto Hidenori
Hattori Akio
Hattori Akira
Hayashi Chikio
Hayashi Keiichi
Hayashi Tsuyoshi
Hida Takeyuki
Hidaka Koji
Hirakawa Junkô
Hirano Tomoharu
Hirasawa Yoshikazu
Hirayama Akira
Hirose Ken
Hisatake Masao
Hitotumatu Sin
Hokari Shisanji
Homma Kiyomi
Homma Tatsuo
Honbu Hitoshi
Hong Imsik
Hosokawa Fujitsugu
Hukuhara Masuo
Husimi Kozi
Huzii Mituaki
Ichida Asajiro
Ihara Shin-ichiro
Ihara Yasutaka
Iitaka Shigeru
Ikawa Mitsuru
Ikeda Mineo
Ikeda Nobuyuki
Ikehara Shikao
Imai Isao
Inaba Eiji
Inagaki Takeshi
Inoue Masao
Inui Teturo
Iri Masao
Irie Shoji
Ise Mikio
Iseki Kanesiroo
Ishida Yasushi
Ishii Goro
Isizu Takehiko
Itô Juniti
Itô Kiyosi
Itô Noboru
Itô Seizô
Itô Teiiti
Ito Tsutomu
Itoh Makoto
Iwahashi Ryōsuke
Iwahori Nagayoshi
Iwamura Tsurane
Iwano Masahiro
Iwasawa Kenkichi
Iwata Giiti
Iyanaga Shôkichi
Izumi Shin-ichi
Kaburaki Masaki
Kamae Keiko
Kametani Shunji
Kanbe Tsunekazu
Kanitani Jōyō
Kasahara Koji
Katase Kiyoshi
Kato Tosio
Katuura Sutezo
Kawada Yukiyosi
Kawaguchi Akitsugu
Kawai Saburo
Kawakami Riiti
Kawata Tatuo
Kihara Taro
Kimura Motoo
Kimura Tosihusa
Kinoshita Shin'ichi
Kitagawa Toshio
Kiyasu Zen'ichi
Koba Ziro
Kobayashi Mikio
Kobori Akira
Koda Akira
Kodaira Kunihiko
Koga Yukiyoshi
Komatsu Hikosaburo
Komatu Atuo
Komatu Yûsaku
Kondo Jiro
Kondo Kazuo
Kondô Motokiti
Kondō Ryōji
Kondo Takeshi
Kondo Yōitsu

Kôta Osamu
Kotake Takesi
Kotani Masao
Kozai Yoshihide
Kubo Izumi
Kubo Ryogo
Kubota Tadahiko
Kudo Akio
Kudō Hirokichi
Kudo Tatsuji
Kuga Michio
Kunisawa Kiyonori
Kunita Hiroshi
Kunugi Kinjiro
Kuranishi Masatake
Kuroda Shige-Toshi
Kuroda Sigekatu
Kuroda Sige-Nobu
Kusama Tokitake
Kusano Takasi
Kusunoki Yukio
Maehara Shôji
Maruyama Gisiro
Mashio Shinji
Masuda Kyūya
Masuyama Motosaburo
Matsukuma Ryozai
Matsumoto Kikuji
Matsumoto Makoto
Matsumoto Toshizô
Matsumura Hideyuki
Matsumura Mutsuhide
Matsushima Yozo
Matsuyama Noboru
Matuda Tizuko
Matusita Kameo
Matuzaka Kazuo
Mikami Masao
Mimura Yukio
Minagawa Takizo
Mita Hiroo
Mitome Michio
Mitsui Takayoshi
Miyake Kazuo
Miyasawa Koichi
Miyashita Totaro
Miyazawa Hironari
Miyazima Tatuoki
Mizohata Sigeru
Mizuno Katsuhiko
Mogi Isamu
Mori Shigeo
Moriguti Sigeiti
Morimoto Haruki
Morimoto Seigo
Morimura Hidenori
Morishima Taro
Morita Kiiti
Morita Yuzo
Moriya Mikao
Moriya Tomijiro
Motohashi Nobuyoshi
Motoo Minoru
Murakami Shingo
Muramatu Tosinobu
Murata Tamotsu
Nabeya Seiji
Nagano Tadashi
Nagao Hirosi
Nagashima Takashi
Nagata Jun-iti
Nagata Masayoshi
Nagumo Mitio
Nakai Mitsuru
Nakai Yoshikazu
Nakamori Kanji
Nakamura Koshiro
Nakamura Masahiro
Nakamura Tokushi
Nakanishi Shizu
Nakano Shigeo
Nakaoka Minoru
Nakayama Tadasi
Nakayama Takashi
Namikawa Yukihiko
Nanba Kanji
Narita Masao
Narumi Seimatu
Niiro Fumio
Nikaido Hukukane
Nishi Mieo
Nishimura Toshio
Noguchi Hiroshi
Nomizu Katsumi
Nomoto Hisao
Nomura Yukichi
Nonaka Toshio
Noshiro Kiyoshi
Nozaki Akihiro
Nozaki Yasuo
Ochiai Kiichiro
Ogasawara Tôjirō
Ogawa Junjiro
Ogawara Masami
Ogino Shusaku
Ōhasi Saburo
Ohtsuka Makoto
Oikawa Kotaro
Ôishi Kiyoshi
Oka Syoten
Okamoto Kiyosato
Okamoto Masashi
Okubo Kenjiro
Okubo Tanjiro
Okugawa Kōtaro
Ōnari Setsuo
Ono Akimasa
Ono Katuzi
Ono Shigeru
Ono Takashi
Onoyama Takuji
Oshida Isao

Oshima Nobunori
Osima Masaru
Otsuki Tominosuke
Ozaki Shigeo
Ozawa Mitsuru
Ozeki Hideki
Saito Kinichiro
Saito Tosiya
Saito Yoshihiro
Sakai Eiichi
Sakai Takuzo
Sakamoto Heihachi
Sakata Shôichi
Sasaki Shigeo
Satake Ichiro
Sato Ken-iti
Sato Ryoichiro
Sato Tokui
Sato Yumiko
Sawashima Ikuko
Seimiya Toshio
Seki Setsuya
Shibagaki Wasao
Shiga Kôji
Shimada Nobuo
Shimizu Hideo
Shimizu Tatsujiro
Shimura Goro
Shintani Takuro
Shizuma Ryoji
Shoda Kenjiro
Shono Shigekata
Sibuya Masaaki
Simauti Takakazu
Sirao Tunekiti
Sueoka Seiichi
Suetuna Zyoiti
Sugawara Masahiro
Sugawara Masao
Sugiura Mitsuo
Sugiura Nariaki
Sugiyama Shohei
Suita Nobuyuki
Sumitomo Takeshi
Sunouchi Genichiro
Suzuki Michio
Suzuki Yoshindo
Takada Masaru
Takagi Teiji
Takagi Yutaka
Takahashi Hidetoshi
Takahashi Tsunero
Takenouchi Osamu
Takenouchi Tadao
Takeuchi Kei
Takeuti Gaisi
Takizawa Seizi
Tamagawa Tsuneo
Tamura Itiro
Tamura Jirô
Tanaka Chuji
Tanaka Hiroshi
Tanaka Hisao
Tanaka Minoru
Tanaka Sen-ichiro
Tannaka Tadao
Tatsumi Tomomasa
Tatuzawa Tikao
Terasaka Hidetaka
Toda Hideo
Toda Hirosi
Toda Morikazu
Tôki Yukinari
Tomotika Susumu
Totoki Haruo
Toyama Hiraku
Tsuboi Chuji
Tsuchikura Tamotsu
Tsuji Masatsugu
Tsujioka Kunio
Tsukamoto Yôtarô
Tsuzuku Toshiro
Tugué Tosiyuki
Tumura Yosiro
Uchiyama Saburo
Uchiyama Tatsuo
Udagawa Masatomo
Ueno Kenji
Ueno Tadashi
Ugaeri Tadashi
Umegaki Hisaharu
Umezawa Hiroomi
Umezawa Toshio
Uno Toshio
Ura Shoji
Ura Taro
Urabe Minoru
Wada Yasaku
Washimi Shinichi
Washio Yasutoshi
Washizu Kyuichiro
Watanabe Hiroshi
Watanabe Hisao
Watanabe Masaru
Watanabe Shinzo
Yamada Isamu
Yamaguti Masaya
Yamamoto Koichi
Yamanaka Takesi
Yamanoshita Tsuneyo
Yamanouti Takahiko
Yamashita Hideo
Yamauti Ziro
Yamazaki Keijiro
Yanagihara Kichiji
Yano Kentaro
Yano Shigeki
Yoneda Nobuo
Yoshizawa Hisaaki
Yosida Kôsaku
Yosida Setuzô
Yosida Yôiti

Translators (for the First Edition)

Adachi Masahisa
Aizawa Sadakazu
Akizuki Yasuo
Akô Kiyoshi
Araki Shôrô
Arima Akito
Asano Keizo
Azumaya Goro
Chiba Katsuhiro
Fujisawa Takehisa
Fujita Hiroshi
Fukada Ichiro
Fukuyama Masaru
Furuya Shigeru
Hashimoto Yasuko
Hasimoto Hidenori
Hasumi Morisuke
Hattori Akio
Hattori Akira
Hayashi Kazumichi
Hida Takeyuki
Hinata Shigeru
Hirasawa Yoshikazu
Hirose Ken
Hitotumatu Sin
Honda Taira
Hong Imsik
Hosoi Tsutomu
Hukuhara Masuo
Ihara Shin-ichiro
Ihara Yasutaka
Ikeda Mineo
Ikeda Nobuyuki
Imai Isao
Inoue Masao
Iri Masao
Ishii Goro
Itô Seizô
Ito Yuji
Iwahashi Ryōsuke
Iwahori Nagayoshi
Iwarmura Tsurane
Iwano Masahiro
Iwata Giichi
Iyanaga Kenichi
Iyanaga Shôkichi
Kametani Shunji
Kasahara Koji
Kato Junji
Kawada Yukiyosi
Kawano Sanehiko
Kimura Tosihusa
Kitagawa Toshio
Kobori Akira
Kodaira Kunihiko
Komatsu Hikosaburo
Komatu Atuo
Komatu Yûsaku
Kondo Ryōji
Kondo Takeshi
Kozai Yoshihide
Kubo Ryogo
Kubota Tomio
Kudo Akio
Kudo Hirokichi
Kumano-go Hitoshi
Kunisawa Kiyonori
Kunita Hiroshi
Kurita Minoru
Kuroda Shige-Toshi
Kusama Tokitake
Kusano Takaŝi
Maruyama Gisiro
Matsuda Michihiko
Matsumoto Kikuji
Matsumoto Makoto
Matsumura Hideyuki
Matsushima Yozo
Mimura Yukio
Miyazawa Hironari
Mizohata Sigeru
Mori Toshio
Morimoto Haruki
Morimura Hidenori
Morita Kiiti
Morita Yasuo
Motoo Minoru
Murakami Shingo
Muramatu Tosinobu
Murase Ichiro
Nagao Hirosi
Nagasawa Masao
Nagata Masayoshi
Nakai Mitsuru
Nakai Yoshikazu
Nakamura Tokushi
Nakanishi Shizu
Nakano Shigeo
Nakaoka Minoru
Namikawa Yukihiko
Namioka Isaac
Nanba Kanji
Niiro Fumio
Nikaido Hukukane
Nishi Mieo
Nishijima Kazuhiko
Nishimiya Han
Noguchi Hiroshi
Nonaka Toshio
Noshiro Kiyoshi
Nozaki Akihiro
Ôbayashi Tadao
Ogino Shusaku
Ohtsuka Makoto
Oikawa Kotaro
Okamoto Masashi
Okubo Kenjiro
Okugawa Kōtaro

Ono Katuzi
Ono Shigeru
Osima Masaru
Ozawa Mitsuru
Ozeki Hideki
Saito Tosiya
Sakai Shoichiro
Sakamoto Minoru
Sasaki Shigeo
Satake Ichiro
Sato Kenkichi
Sawashima Ikuko
Shibuya Masaaki
Shiga Kôji
Shimada Nobuo
Shintani Takuro
Shioda Tetsuji
Shiraiwa Kenichi
Shiratani Katsumi
Shizuma Ryoji
Simauti Takakazu
Sugawara Masahiro
Sugiura Mitsuo
Sugiyama Shohei
Suita Nobuyuki
Sumitomo Takeshi
Sunouchi Genichiro
Suzuki Michio
Takahashi Moto-o
Takahashi Reiji
Takahashi Tsunero
Takami Hideo
Takaoka Seiki
Takenouchi Osamu
Takeuchi Kei
Takeuti Gaisi
Tamura Itiro
Tanaka Hiroshi
Tanaka Hisao
Tanaka Minoru
Tatsumi Tomomasa
Tatuzawa Tikao
Terasaka Hidetaka
Toda Hideo
Toda Hirosi
Tsuka-da Nobutaka
Tsukamoto Yôtarô
Tugué Tosiyuki
Uchiyama Saburo
Uesu Tadahiro
Umegaki Hisaharu
Umezawa Toshio
Uno Toshio
Ura Shoji
Ura Taro
Urabe Minoru
Uzawa Hirobumi
Wada Junzo
Watanabe Hiroshi
Watanabe Hisao
Watanabe Shinzo
Watanabe Takesi
Yamaguti Masaya
Yamamoto Koichi
Yamanaka Takesi
Yamazaki Keijiro
Yano Kentaro
Yano Shigeki
Yoshizawa Taro
Yosida Kôsaku
Yosida Setuzô

Name Index

Note. Citation is to article and section, not to page. r = references; STR = Statistical Tables for Reference (to be found after App. B); NTR = Numerical Tables for Reference. * means that the article is not divided into sections.

A

B

C

E

F

H

K

L

N

Q

R

T

V

W

Encyclopedic Dictionary of Mathematics

Second Edition

Z

Encyclopedic Dictionary of Mathematics

Second Edition

by the
Mathematical Society
of Japan

edited by
Kiyosi Itô

Volume II
O–Z
Appendices and Indexes

The MIT Press
Cambridge, Massachusetts,
and London, England

Subject Index

Note: Citation is to article and section or to a table in an appendix, not to page.

First MIT Press paperback edition, 1993

Originally published in Japanese in 1954 by Iwanami Shoten, Publishers, Tokyo, under the title *Iwanami Sūgaku Ziten.*

This book was set in Monolaser Times Roman by Asco Trade Typesetting Ltd., Hong Kong, and printed and bound by Arcata Graphics, Kingsport in the United States of America.

Library of Congress Cataloging-in-Publication Data

Iwanami sūgaku jiten (ziten). English.
Encyclopedic dictionary of mathematics.
Translation of: Iwanami sūgaku jiten.
Includes bibliographies and indexes.
1. Mathematics—Dictionaries. I. Itô, Kiyosi, 1915– . II. Nihon Sūgakkai. III. Title.
QA5.I8313 1986 510'.3'21 86-21092
ISBN 0-262-09026-0 (HB), 0-262-59020-4 (PB)

B

C

D

E

G

H

I

K

L

N

O

P

Q

R

S

U

W