

Table 4. A Short Table of Integrals. (Continued)

$$110. \int \ln x \, dx = x \ln x - x$$

$$111. \int x^n \ln x \, dx = x^{n+1} \left[ \frac{\ln x}{n+1} - \frac{1}{(n+1)^2} \right]$$

$$112. \int \frac{1}{x \ln x} \, dx = \ln(\ln x)$$

$$113. \int e^{ax} \sin nx \, dx = \frac{e^{ax}(a \sin nx - n \cos nx)}{a^2 + n^2}$$

$$114. \int e^{ax} \cos nx \, dx = \frac{e^{ax}(a \cos nx + n \sin nx)}{a^2 + n^2}$$

Definite Integrals:

$$115. \int_0^\infty x^{s-1} e^{-x} \, dx = \int_0^1 \left( \ln \frac{1}{x} \right)^{s-1} \, dx = \Gamma(s)$$

$$116. \int_1^\infty \frac{dx}{x^m} = \frac{1}{m-1} \quad m > 1$$

$$117. \int_0^\infty \frac{a \, dx}{a^2 + x^2} = \frac{\pi}{2}, \text{ if } a > 0; 0, \text{ if } a = 0; -\frac{\pi}{2}, \text{ if } a < 0$$

$$118. \int_0^\infty \frac{\sin mx \, dx}{x} = \frac{\pi}{2}, \text{ if } m > 0; 0, \text{ if } m = 0; -\frac{\pi}{2}, \text{ if } m < 0$$

$$119. \int_0^\infty \frac{\cos x \, dx}{x} = \infty$$

$$120. \int_0^\infty \frac{\tan x \, dx}{x} = \frac{\pi}{2}$$

$$121. \int_0^\infty \frac{\sin^2 x \, dx}{x^2} = \frac{\pi}{2}$$

$$122. \int_0^\infty e^{-a^2 x^2} \, dx = \frac{1}{2a} \sqrt{\pi} = \frac{1}{2a} \Gamma\left(\frac{1}{2}\right) \quad a > 0$$

$$123. \int_0^\infty x^2 e^{-x^2} \, dx = \frac{\sqrt{\pi}}{4}$$

#### 1.4 Numerical Differentiation and Integration

Given a set of numerical values of a function, the processes of numerical differentiation and integration consist, respectively, of calculating the derivative (or derivatives) by means of these values and of computing the values of a definite integral from the set of values of the integrand. In both numerical differentiation and integration the problem is solved by representing the function by an interpolation formula and then differentiating or integrating as desired. Interpolation formulas are discussed in Art. 3.5 of Sec. 3-1.

It was noted in Art. 3.5 of Sec. 3-1 that a polynomial which agrees with  $f(x)$  at  $x_0, x_1, x_{-1}, x_2, x_{-2}, \dots$  is the central-difference formula

$$P(x) = f(x_0) + u\mu \delta f(x_0) + \frac{u^2}{2!} \delta^2 f(x_0) + \frac{u(u^2 - 1^2)}{3!} \mu \delta^3 f(x_0) \\ + \frac{u^2(u^2 - 1^2)}{4!} \delta^4 f(x_0) + \frac{u^2(u^2 - 1^2)(u^2 - 3^2)}{5!} \mu \delta^5 f(x_0) + \dots \\ = f(x_0) + u \frac{\Delta y_{-1} + \Delta y_0}{2} + \frac{u^2}{2} \Delta^2 y_{-1} + \frac{u(u^2 - 1^2)}{3!} \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \dots$$

where  $x = x_0 + uh$ . Since  $u = 0$  at  $x_0$ , the derivative at  $x_0$  is given by

$$f'(x_0) \approx P'(x_0) = \frac{1}{h} \left( \frac{\Delta y_{-1} + \Delta y_0}{2} - \frac{1}{3!} \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \dots \right)$$

Higher derivatives can be obtained in like fashion by further differentiating  $P(x)$ . Near the beginning of a set of tabular values Newton's forward-difference formula

is more convenient and near the end of the set Newton's backward-difference formula is more convenient than the central-difference formula.

There are a large number of quadrature formulas for the approximate integration of a function specified by a set of numerical values. As previously noted, any of the interpolation formulas can lead to quadrature formulas. The trapezoidal rule and Simpson's rule are the most commonly used. If  $h$  is the length of each subinterval, the *trapezoidal rule* is

$$\int_{x_0}^{x_0+nh} y \, dx = \frac{h}{2} (y_0 + 2y_1 + 2y_2 + \cdots + 2y_{n-1} + y_n)$$

and *Simpson's rule* is

$$\int_{x_0}^{x_0+nh} y \, dx = \frac{h}{3} (y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n)$$

where  $n$  in this last case must be an even number; i.e., the number of subintervals is even.

Gauss's formula (see references) is the most accurate of the formulas ordinarily used and can be used advantageously with high-speed machines.

## 2 FUNCTION THEORY

### 2.1 Real Variables

The subject matter of real variables may include the following topics: the system of real numbers, sequences, infinite series, ordinal and cardinal numbers, set theory, functions and limits, continuity and discontinuity, differentiation and integration, and measure theory. Infinite series and differentiation and integration are of such scope that they are discussed in other articles. Ordinal and cardinal numbers and measure theory are not felt to be essential here.

**2.11 The System of Real Numbers and Sequences.** The foundation of the theory of functions of a real variable depends upon the real-number system. Although the refined concept of the real number is the starting point for any discussion of the fundamental parts of higher analysis, only the more important concepts and results will be indicated.

The concept of the *natural numbers* or *positive integers*—1, 2, 3, . . . —may be taken as a starting point. The class of *rational numbers* is obtained from the positive integers by allowing the inverse operations of addition and multiplication, namely, subtraction and division. The totality of positive integers, negative integers, zero, and fractions constitute the class of *rational numbers*.

It is generally appreciated that certain numbers such as  $\sqrt{2}$  and  $\pi$  are not rational numbers and cannot therefore be represented by the ratio of two integers. Irrational numbers are generally derived from the rational numbers by either Cantor's theory or Dedekind's theory.

Cantor's theory of irrational numbers depends upon the concept of a sequence of rational numbers. If by some suitable process a first, a second, a third, . . . rational number can be formed successively, and if to every positive integer  $n$  one and only one rational number  $a_n$  corresponds, then the numbers

$$a_1, a_2, \dots, a_n, \dots$$

in this order, corresponding to the natural order of the positive integer, are said to form a *sequence* of rational numbers. The individual numbers that form the sequence are called the *elements* of the sequence. The sequence

$$a_1, a_2, \dots, a_n, \dots$$

will be denoted symbolically by  $\{a_n\}$ .

A sequence of rational numbers  $\{a_n\}$  is called convergent or regular if for an arbitrary  $\epsilon > 0$  there exists a number  $N$  such that for every  $n > N$

$$|a_n - a_{n+m}| < \epsilon \quad \text{where } m = 1, 2, 3, \dots$$

The essential feature of Cantor's theory of irrational numbers is the assumption that corresponding to every convergent sequence of rational numbers there exists a uniquely determined object called a *real number*. Any real number can, therefore, be regarded as being represented by a convergent sequence of rational numbers. Two real numbers  $A$  and  $B$  defined by the convergent sequences of rational numbers  $\{a_n\}$  and  $\{b_n\}$  are said to be the same number or are equal if there exists an integer  $N$  such that for all values of  $n > N$

$$|a_{n+m} - b_{n+m}| < \epsilon \quad m = 1, 2, 3, \dots$$

where  $\epsilon$  is an arbitrarily small positive number.

The real-number system consists of rational numbers, since any rational number  $a$  can be represented by a convergent sequence  $\{a_n\}$ , where  $a_n = a$  for all  $n$ , and the real numbers that are not rational, i.e., the irrational numbers.

In the Dedekind theory of the real-number system the real numbers correspond to partitions of the rational numbers. A *partition* is formulated in the following manner. Divide all the rational numbers into two classes  $R$  and  $S$ . In class  $R$  every number is less than any number in  $S$ , and in class  $S$  every number is greater than any number in  $R$ . For an irrational number there is no largest number in  $R$  and no smallest number in  $S$ . For a rational number there is either a largest number in  $R$  or a smallest number in  $S$ . It is possible to show that Cantor's method of convergent sequences and Dedekind's method of partitions are equivalent in that starting with the rational numbers the same system of real numbers can be developed.

If there exist a sequence  $\{a_n\}$  and a real number  $A$  such that the sequence  $\{a_n - A\}$  forms a null sequence, then the sequence  $\{a_n\}$  is said to *converge to the limit*  $A$  and is denoted by

$$\lim_{n \rightarrow \infty} a_n = A$$

This definition says that for every  $\epsilon > 0$  there exists an  $N$  such that for all  $n > N$ ,  $|a_n - A| < \epsilon$ . Every sequence that does not converge in the above sense is called *divergent*. The *Cauchy or general principle of convergence* states that the necessary and sufficient condition for the convergence of the sequence  $\{a_n\}$  is that for every  $\epsilon > 0$ , there exists an  $N$  such that for  $n > N$ ,  $|a_{n+m} - a_n| < \epsilon$ , where  $m = 1, 2, 3, \dots$

The system of real numbers can be considered sufficient for the needs of the theory of functions of real variables, since the real numbers form a closed system with respect to arithmetic operations, such as addition, subtraction, multiplication, division, extraction of roots of positive numbers, and powers, and the limiting process. A number system is called *closed* with respect to an operation or process if this process results in a number contained in the system.

It is possible to set up a one-to-one correspondence between the points of a straight line and the real-number system. Because of this possibility, the properties and definitions of the real-number system have a geometrical interpretation.

**2.12 Set Theory.** The starting point for most mathematical developments is certain objects such as numbers or letters. A *set* (or class or aggregate or collection) is defined by any property that any particular one of these objects does or does not have. The objects that have the property are called *elements* of the set. This is symbolized by

$$s \in S$$

where  $s$  is an element and  $S$  is the set. An *empty set* does not contain any elements; i.e., there are no objects having the property of the set.

Two sets  $S_1$  and  $S_2$  are called equal, and in symbols  $S_1 = S_2$  if every element of  $S_1$  is an element of  $S_2$  and conversely if every element of  $S_2$  is an element of  $S_1$ . If all

the elements of a set  $S_1$  are simultaneously elements of a set  $S_2$ , then  $S_1$  is called a *subset* of  $S_2$ , and this relationship is denoted by  $S_1 \subset S_2$ . The notation  $S_2 \supset S_1$  indicates the same relationship, and  $S_2$  is said to *include*  $S_1$ . If  $S_1 \subset S_2$  and  $S_2 \subset S_1$ , then  $S_1 = S_2$ . If  $S_1 \subset S_2$  and  $S_2 \subset S_3$ , then  $S_1 \subset S_3$ . If  $S_1 \subset S_2$  but  $S_1$  is not equal to  $S_2$ , symbolized  $S_1 \neq S_2$ , then  $S_1$  is called a *proper subset* of  $S_2$ .

The *intersection* (or logical product or meet) of two sets  $S_1$  and  $S_2$  is denoted by  $S_1 \cap S_2$  and is the set consisting of all elements common to the sets  $S_1$  and  $S_2$ . The *union* (or logical sum) of two sets  $S_1$  and  $S_2$  is denoted by  $S_1 \cup S_2$  and is the set consisting of all elements that belong to at least one of the sets  $S_1$  and  $S_2$ . The definitions of intersection and union hold for an arbitrary number of sets. If the intersection of two sets  $S_1$  and  $S_2$  is the empty set, then the two sets are called *disjoint* or *mutually exclusive*. If  $S_1$  is a subset of a set  $S$ , then the *complementary set* of  $S_1$  with respect to  $S$  is the set of elements of  $S$  obtained by omitting the elements of  $S_1$  that are elements of  $S_1$ . Generally the term complement of a set is used with respect to a fundamental, and therefore understood, set, such as the set of real numbers.

Sets may first of all be classified into finite and infinite according to whether they contain a finite or infinite number of elements. An infinite set is called *enumerable* (or denumerable or countable) if, and only if, a one-to-one correspondence can be set up between the elements of the set and the positive integers. Here the term *countable* will be used to indicate either a finite or an enumerable set. A *noncountable* or *nonenumerable* set is a set that is neither finite nor enumerable. The following results dealing with countable and nonenumerable sets are well known:

1. Any subset of a countable set is also countable.
2. The sum of a countable set of countable sets is also countable.
3. The set of rational numbers is enumerable.
4. The set of irrational numbers and the set of real numbers are nonenumerable.
5. The set of all algebraic numbers is enumerable. An *algebraic number* is the root of the polynomial equation

$$\sum_{i=0}^n a_i x^i = 0$$

where  $a_n \neq 0$  and all the  $a_i$ 's are integers.

6. The set of transcendental numbers is nonenumerable. The real numbers that are not algebraic are called *transcendental*.

In the discussion of the real-number system the set of points on a line was noted to correspond to the set of real numbers. This set of points is called a *linear point set* or, briefly, a *linear set*. A linear set is *bounded* if all its points lie in a finite interval. An *open interval*, symbolized by  $(a, b)$ , consists of all points  $x$  such that  $a < x < b$ . A *closed interval*, symbolized by  $[a, b]$ , consists of all points  $x$  such that  $a \leq x \leq b$ .

A point  $a$  is called a *limit point* (or limiting point or accumulation point) of a set  $S$  if there exists a point of the set  $S$ , different from  $a$ , in every neighborhood of the point  $a$ . For a linear set a *neighborhood* of a point  $a$  means the open interval  $(a - \epsilon, a + \epsilon)$ , where  $\epsilon > 0$ . It can be shown from the definition that every neighborhood of a limit point contains infinitely many elements of the set. If every point of an interval is a limit point of a set  $S$ , then the set  $S$  is said to be *everywhere dense*. The set of all limit points of a set  $S$  is called the *derived set* (or derivative) of the set  $S$  and is denoted by  $S'$ . The *closure* of  $S$ , denoted by  $\bar{S}$ , consists of all the points of  $S$  and the limit points of  $S$ , that is,  $\bar{S} = S \cup S'$ . Any point of a set which is not a limit point is called an *isolated point*.

A point  $a$  is called an *interior point* (or inner point) of a set  $S$  if there exists a neighborhood of  $a$  containing only points of  $S$ . A point  $a$  is called an *exterior point* of a set  $S$  if there exists a neighborhood of  $a$  containing no points of  $S$ . If a point is neither an interior nor an exterior point of a set  $S$ , then it is called a *boundary point* of the set. If a set contains all its boundary points and therefore its limit points, it is called *closed*. If every point of a set is an interior point, the set is called an *open set*. A *perfect set* is a closed set where every point of the set is a limit point of the set. Finally, a

*continuum* is a set that is perfect and everywhere dense. It may be noted that the complement of an open set is a closed set and conversely.

For a linear set  $S$  a point  $a$  is called an *upper bound* if  $s \leq a$  for every point  $s$  of the set  $S$ . The point  $a$  is a lower bound if  $s \geq a$  for every point  $s$  of the set. The point  $a$  is called the *least upper bound* for the set  $S$  if it is an upper bound and if for any  $\epsilon > 0$  there exists a point of  $S$  greater than  $a - \epsilon$ . The *greatest lower bound*  $a$  for a set is similarly a lower bound such that there is a point of the set less than  $a + \epsilon$ , where  $\epsilon > 0$ .

Two of the important theorems of set theory follow.

*Heine-Borel Theorem.* Let the closed set of points  $S$  be covered by a set of intervals; then there exists a finite number of intervals that likewise cover  $S$ . A set  $S$  is said to be covered by a set of intervals  $I$  if every point of the set  $S$  is interior to at least one of the intervals of the set  $I$ .

*Weierstrass-Bolzano Theorem.* If  $S$  is an infinite bounded set there exists at least one limit point.

**2.13 Functions and Their Limits.** If in the course of a discussion a symbol may be assigned various numerical values, the symbol is called a *variable*. A *constant* assumes but one numerical value during a discussion. A *real variable* has values in the set of real numbers. Given two variables  $x$  and  $y$ ,  $y$  is called a *function* of  $x$  if to every value of  $x$  in the domain of  $x$  there is determined a definite value or values of  $y$ . This functional relationship is denoted symbolically by  $y = f(x)$ .  $x$  is called the *independent variable* and  $y$  is called the *dependent variable*. The vital aspect of the definition of a function is that for every value the independent variable takes on, the corresponding value or values of the dependent variable are uniquely determined. The set of values that the independent variable assumes is prescribed and is called the *domain*. The set of values taken on by the *dependent variable* is called the *range*. A function is called *single-valued* if the dependent variable takes on but one value for each value of the independent variable. If for any value of the independent variable the dependent variable takes on more than one value, the function is called *multivalued* (or multiple-valued).

A *polynomial function* has the form

$$a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \equiv \sum_{i=0}^n a_i x^{n-i}$$

where the  $a_i$  are constants and  $n$  is a positive integer. A *rational function* is the ratio of two polynomial functions. An *algebraic function* is defined by means of the equation

$$\sum_{i=0}^n f_i(x)a_{n-1} = 0$$

where the  $f_i(x)$  are rational functions of  $x$ . *Transcendental functions* are functions that are not algebraic.

The theory of functions of a real variable deals with correspondences between two sets of real numbers, designated the independent and dependent variables. The terminology of set theory therefore applies to the set  $y$ . The function  $f(x)$  is said to be bounded, have a least upper bound, etc., if the set  $y$  does. If the least upper bound of a function is a point taken on by the function, then it is called the *maximum* (or maximum value) of the function. The *minimum* is associated similarly with the greatest lower bound.

The quantity  $f(x)$  is said to have a *limit*  $b$  as  $x$  tends to  $x_0$  if for any  $\epsilon > 0$  there exists a  $\delta > 0$  such that  $|f(x) - b| < \epsilon$  for all  $x$  for which  $0 < |x - x_0| \leq \delta$ . Symbolically this relation is written

$$\lim_{x \rightarrow x_0} f(x) = b$$

From the definition of the limit of a function, it follows that if the limit exists, the value approached by  $f(x)$  as  $x$  approaches  $x_0$  does not depend upon the value of  $f(x)$  at  $x_0$  and also is independent of the particular set of values that  $x$  takes on in approaching  $x_0$ . If the independent variable  $x$  is allowed to take on only values larger than  $x_0$  or less than  $x_0$ , then the respective limits are called *right-hand* and *left-hand limits*. These are symbolized, respectively, by

$$\begin{aligned} & \lim_{x \rightarrow x_0^+} f(x) = f(x_0^+) \quad \text{or} \quad \lim_{x > x_0} f(x) \\ \text{and} \quad & \lim_{x \rightarrow x_0^-} f(x) = f(x_0^-) \quad \text{or} \quad \lim_{x < x_0} f(x) \end{aligned}$$

The Cauchy or general principle of convergence states that a necessary and sufficient condition for the existence of a limit to  $f(x)$  as  $x$  tends to  $x_0$  is that for  $\epsilon > 0$  there exists a  $\delta > 0$  such that  $|f(x'') - f(x')| < \epsilon$  for all values of  $x', x''$  for which  $0 < |x'' - x_0| < |x' - x_0| \leq \delta$ .

A function may depend upon the values taken on by two or more independent variables. Again the vital aspect of the functional relationship is that whenever each of the independent variables assumes a value, a corresponding value or set of values of the dependent variable is uniquely determined. Given a function of two or more variables, there exist two types of limits: iterated limits and simultaneous limits. Let  $f(x, y)$  be the function of two independent variables  $x$  and  $y$ , and let  $(x_0, y_0)$  be the limit point; then

$$\lim_{x \rightarrow x_0} \left[ \lim_{y \rightarrow y_0} f(x, y) \right] \quad \text{and} \quad \lim_{y \rightarrow y_0} \left[ \lim_{x \rightarrow x_0} f(x, y) \right]$$

are called iterated (or repeated) limits. An *iterated limit* indicates that first an ordinary limit is taken for one variable holding the other variable (or variables) fixed and then a limit is taken for the other variable (or other variable with the remaining fixed). The *simultaneous* limit

$$\lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} f(x, y)$$

has the value  $A$  if for  $\epsilon > 0$  there exists a positive number  $\delta$  such that  $|f(x, y) - A| < \epsilon$  for all  $x$  and  $y$  such that  $0 < |x - x_0| < \delta$  and  $0 < |y - y_0| < \delta$ . If the simultaneous limit exists, then the two iterated limits exist and are equal. The converse does not hold, since the simultaneous limit can be nonexistent and yet the two iterated limits may exist and even be equal.

**2.14 Continuous and Discontinuous Functions.** A function  $f(x)$  is said to be *continuous at a point*  $x_0$  if

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

that is, for  $\epsilon > 0$ , there exists a  $\delta > 0$  such that  $|f(x) - f(x_0)| < \epsilon$  for all  $x$  such that  $|x - x_0| < \delta$ . In words this definition states that the limit shall exist at  $x_0$ , that the function is defined at  $x_0$ , and that these two values are equal.  $f(x)$  is *continuous in the interval*  $[a, b]$  if it is continuous at every point  $x$ , where  $a < x < b$ , and if at the end points

$$\lim_{x \rightarrow a^+} f(x) = f(a) \quad \text{and} \quad \lim_{x \rightarrow b^-} f(x) = f(b)$$

The quantity  $f(x)$  is said to have an ordinary discontinuity (or jump discontinuity or simple discontinuity or discontinuity of the first kind) at the point  $x_0$  if the right-hand and left-hand limits at a point exist but are not equal; i.e.,

$$\lim_{x \rightarrow x_0^+} f(x) \neq \lim_{x \rightarrow x_0^-} f(x)$$

If the right-hand and left-hand limits exist and are equal but the function has a different value, i.e.,

$$\lim_{x \rightarrow x_0^+} f(x) = \lim_{x \rightarrow x_0^-} f(x) \neq f(x_0)$$

then the function has a *removable discontinuity* at the point  $x_0$ . When the right-hand limit or the left-hand limit or both these limits fail to exist at point  $x_0$ , then the function has a *discontinuity of the second kind* at the point  $x_0$ .

A function is *continuous on the right at a point* if the right-hand limit has the same value as the function at the point. Continuity on the left at a point and continuity on the right and left in an interval are defined in a corresponding fashion.  $f(x)$  is called *uniformly continuous* in the interval  $[a, b]$  if for  $\epsilon > 0$  there exists a  $\delta > 0$  independent of the  $x_0$  in the interval  $[a, b]$  such that  $|f(x) - f(x_0)| < \epsilon$  for all  $x$  satisfying  $|x - x_0| < \delta$ .

Let  $f(x)$  be continuous in the interval  $[a, b]$ , then the following results hold:

1.  $f(x)$  is uniformly continuous in the interval.
2.  $f(a)$  and  $f(b)$  have opposite signs; then there exists at least one value of  $x$  in the interval for which  $f(x)$  vanishes.
3.  $f(a) \neq f(b)$ ; then as  $x$  takes on all values between  $a$  and  $b$ ,  $f(x)$  takes on at least once all values between  $f(a)$  and  $f(b)$ .
4. If  $f(x)$  is single-valued in  $[a, b]$ , then there exists at least one point of  $[a, b]$  at which  $f(x)$  takes on a maximum value. Likewise there exists a value of  $x$  where the minimum is attained.
5. The function is bounded in that interval.
6. The function is uniquely determined at every point of the interval by prescribing the function at a set of points everywhere dense in the interval  $[a, b]$ .

## 2.2 Complex Variables

**2.21 Complex Plane and Sphere.** Complex numbers have been discussed in Art. 2 of Sec. 3-1. The set of complex numbers can be put into one-to-one correspondence with the points of a plane. This correspondence associates the complex number  $z = x + iy$  with the point in the plane whose rectangular or Cartesian coordinates are  $(x, y)$ . Because of this association this plane is called the complex or  $z$  plane. This geometric association for complex numbers not only gives a geometrical interpretation for operations involving complex numbers but also allows the use of geometric terminology such as points and distances when discussing complex numbers. When the improper point  $z = \infty$  is added, the complex plane is closed.

The number or point  $z$  can also be thought of as a vector that originates at the origin of the coordinate system and ends at the coordinates  $(x, y)$ .

The points of the closed complex plane can be mapped by stereographic projection one-to-one onto the points of a sphere called the Riemann sphere (or sphere of complex numbers). In this mapping, the south pole is placed at the origin and corresponding points for the sphere and plane lie on a ray that originates at the north pole.

The following relations involve the absolute value  $|z| = \sqrt{x^2 + y^2}$  and the conjugate  $\bar{z} = x - iy$ :

1.  $z_1 \pm z_2 = \bar{z}_1 \pm \bar{z}_2$ .
2.  $z_1 z_2 = \bar{z}_1 \bar{z}_2$ ;  $(z_1 / z_2) = \bar{z}_1 / \bar{z}_2$ .
3.  $|\bar{z}| = |z|$ .  $z\bar{z} = |z|^2$ .
4.  $|z_1 - z_2|$  is the distance between the points  $z_1$  and  $z_2$ .
5.  $|z_2| > |z_1|$  says that the point  $z_2$  is farther from the origin than  $z_1$ .
6.  $|z_1 + z_2| \leq |z_1| + |z_2|$  corresponds to the geometric statement that no side of a triangle is greater than the sum of the other two sides.
7.  $|z_1 - z_2| \geq ||z_1| - |z_2||$  states that no side of a triangle is less than the difference of the other two sides.

**2.22 Functions of a Complex Variable.** The concepts and definitions for real variables generally have significance for complex variables.  $z$  is called a *complex variable* if in the course of a discussion it assumes various complex values. Given two complex variables  $z$  and  $w$ ,  $w$  is called a *complex function of the complex variable  $z$*  if to every value of  $z$  in the domain of  $z$  there is determined a value or values of  $w$ . Again  $w$  is called *single-valued* if it takes on only one value for each value of  $z$  and  $w$  is *multiple-valued* if it takes on two or more values for any value of  $z$ . Polynomial, rational, algebraic, and transcendental functions are defined for complex variables in the same

fashion as for real variables. The distinguishing feature is the use of complex constants and the complex variable  $z$  in place of the real variable  $x$ .

A neighborhood of a point  $z_1$  is the circular region

$$|z - z_1| < \epsilon$$

where  $\epsilon > 0$ . The definitions of limit point, interior point, open set, and closed set as given in the article on set theory are applicable to a set of complex numbers. An open region, or just region, denotes an open set that is connected; that is, any two points of the set may be joined by a continuous curve all of whose points belong to the set. A closed region is a region plus all its limit points. A region is called simply connected if every closed curve within the region encloses only points of the region. A region that is not simply connected is called multiply connected.

If an integral is taken along a closed curve that is the boundary of a region, the integral is commonly called a contour integral. A closed curve is called positively oriented if the interior of the curve lies to the left of the curve as the curve is traversed, i.e., the direction is counterclockwise. If only the initial and terminal points of a curve coincide, then the closed curve is called simple.

The quantity  $f(z)$  is said to have the limit  $w_0$  as  $z$  approaches  $z_0$  if for any  $\epsilon > 0$  there exists a  $\delta > 0$  such that  $|f(z) - w_0| < \epsilon$  for all  $z$  for which  $0 < |z - z_0| \leq \delta$ . Let  $w = f(z) = u + iy = u(x, y) + iv(x, y)$ , where  $z = x + iy$  and  $u$  and  $v$  are real functions of  $x$  and  $y$ ; then the limit can be expressed by

$$\lim_{z \rightarrow z_0} f(z) = \lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} u(x, y) + i \lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} v(x, y) = u_0 + iv_0 = w_0$$

The results for limits with a complex variable follow almost directly, therefore, from corresponding results for real functions of two real variables.

A single-valued function  $f(z)$  of a complex variable is called continuous at a point  $z_0$  if

$$\lim_{z \rightarrow z_0} f(z) = f(z_0)$$

The three requirements of the definition are, again, that first the function be defined at  $z_0$ , second the limit must exist as  $z_0$  is approached, and finally the limit value must equal the value of the function at  $z_0$ . The results of real variables concerned with continuity lead to analogous results for complex variables. For example, if a function  $f(z)$  is continuous in a bounded closed region, then it is uniformly continuous in that region; it is bounded in that region, that is,  $|f(z)| < M$  where  $M$  is a finite positive number;  $|f(z)|$  has a finite upper limit in the region; and if  $z_0$  is an inner point of the region such that  $f(z_0) \neq 0$ , then there exists a neighborhood of  $z_0$  for which  $f(z) \neq 0$ .

The derivative of  $f(z)$  at the point  $z_0$  is defined by

$$\lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0} = \lim_{z \rightarrow z_0} \frac{\Delta f}{\Delta z} \equiv f'(z_0)$$

where  $f'(z_0)$  denotes the complex number that the limit, if it exists, assumes. The differentiation rules for real variables can be extended essentially without change to complex variables. A function that has a derivative at every point of a region is called differentiable in the region.

Let  $w = f(z)$  be a continuous single-valued function of  $z$  in a region  $R$ , let  $\alpha$  and  $\beta$  be two points of  $R$ , and let  $C$  be a curve of finite length connecting the two points and lying in  $R$ . Furthermore, let  $z_0 = \alpha, z_1, \dots, z_n = \beta$  be a sequence of points on  $C$ , and let  $\eta_i$  be any point on the curve between  $z_{i-1}$  and  $z_i$ . The integral of  $f(z)$  along  $C$  between the limits  $\alpha$  and  $\beta$ , which is symbolized by  $\int_C f(z) dz$ , is defined by

$$\int_{\alpha}^{\beta} f(z) dz \equiv \int_C f(z) dz = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\eta_i)(z_i - z_{i-1})$$

where the limit indicates that  $n$  tends to infinity and that the absolute value of every subdivision  $|z_i - z_{i-1}|$  tends to zero.

Since  $f(z) = u(x, y) + iv(x, y)$  and  $dz = dx + i dy$ , the complex integral can formally be written as follows in terms of real integrals:

$$\int_C f(z) dz = \int_C [u(x, y) dx - v(x, y) dy] + i \int_C [v(x, y) dx + u(x, y) dy]$$

The curve  $C$  may be represented by the *real parametric equations*

$$x = g(t) \quad \text{and} \quad y = h(t)$$

where  $0 \leq t \leq 1$ ,  $\alpha = g(0) + ih(0)$ , and  $\beta = g(1) + ih(1)$ . Let the functions  $g(t)$  and  $h(t)$  be single-valued and have continuous first-order derivatives; then

$$\int_C f(z) dz = \int_0^1 [ug'(t) - vh'(t)] dt + i \int_0^1 [vg'(t) + uh'(t)] dt$$

The following elementary properties for complex integrals may be noted:

1.  $\int_{\alpha C_1}^{\beta} f(z) dz + \int_{\beta C_2}^{\gamma} f(z) dz = \int_{\alpha C_1 + C_2}^{\gamma} f(z) dz$ ; that is, the sum of two integrals taken along two successive curves is equal to the integral taken over the entire curve.
2.  $\int_{\alpha C}^{\beta} f(z) dz = - \int_{\beta C}^{\alpha} f(z) dz$ ; that is, if the direction of integration is reversed, the value of the integral remains the same except for sign.
3.  $\int_C [k_1 f_1(z) + k_2 f_2(z)] dz = k_1 \int_C f_1(z) dz + k_2 \int_C f_2(z) dz$ .
4.  $\left| \int_C f(z) dz \right| \leq ML$ , where  $|f(z)| \leq M$  for any  $z$  on  $C$  and  $L$  is the length of the curve  $C$ .

**2.23 Analytic Functions.** A single-valued function  $f(z)$  which is differentiable, that is, has a first derivative at every point, in a region is called *analytic* (or regular or holomorphic) *in the region*. A function  $f(z)$  is called *analytic at a point*  $z_0$  if its derivative exists at every point of some neighborhood of  $z_0$ .

The concept of analytic functions, or analyticity, is particularly unifying and important for mathematical physics. Two necessary and sufficient conditions for the function  $f(z) = u(x, y) + iv(x, y)$  to be analytic in the region  $D$  follow:

1. *Cauchy-Riemann Equations.* The four first-order partial derivatives of  $u$  and  $v$  with respect to  $x$  and  $y$  exist and are continuous in the region  $D$ , and they satisfy the *Cauchy-Riemann differential equations*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \frac{\partial u}{\partial y} = - \frac{\partial v}{\partial x}$$

2. *Cauchy-Goursat Theorem and Morera's Theorem.* The integral  $\int_C f(z) dz$  of the continuous function  $f(z)$ , when taken along the entire boundary curve  $C$  of any sub-region of the region  $D$ , is zero.

a. **CAUCHY-GOURSAT THEOREM (NECESSITY).** Let  $f(z)$  be single-valued and analytic within and on a closed curve  $C$ ; then

$$\int_C f(z) dz = 0$$

b. **MORERA'S THEOREM (SUFFICIENCY).** Let  $f(z)$  be continuous in a simply connected region; then if

$$\int_C f(z) dz = 0$$

for every closed curve lying within the region,  $f(z)$  is analytic in that region.

Although it is possible to formulate other necessary and sufficient conditions for analyticity, it is customary to consider the consequences of analyticity whether they, in turn, imply analyticity or not. Some important consequences of analyticity follow.

1. *Cauchy's Integral Formula.* If  $f(z)$  is analytic in a region  $D$ , then *Cauchy's integral formula*

$$f(z_0) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - z_0} dz$$

is valid for every simple, closed, positively oriented curve  $C$  and for every point  $z_0$  interior to the curve  $C$ .

2. *Higher Derivatives.* If a single-valued function  $f(z)$  is analytic in a region, then not only does the function by definition have a first derivative in the region but it also has all higher derivatives. The formulas

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_C \frac{f(z)}{(z - z_0)^{n+1}} dz \quad n = 1, 2, \dots$$

are valid with the same conditions used for Cauchy's integral formula.

3. *Laplace's Equation.* If  $f(z) = u + iv$  is analytic in a region, then the functions  $u(x, y)$  and  $v(x, y)$  satisfy in that region the partial differential equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

This equation is called *Laplace's differential equation* and is of great importance in mathematical physics. A function that has continuous second partial derivatives and satisfies Laplace's equation is called a *harmonic function*. If  $f(z) = u + iv$  is analytic, then  $u$  and  $v$  are called *conjugate harmonic functions*.

4. *Taylor's Series.* If  $f(z)$  is analytic in a region  $D$  with a boundary  $C$ , then at each interior point  $z_0$ ,  $f(z)$  can be represented uniquely by a power series of the form

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n$$

where

$$a_n = \frac{1}{n!} f^{(n)}(z_0)$$

This series, called a *Taylor's series*, converges and represents  $f(z)$  in the largest circle with center  $z_0$  that encloses only points of  $D$ . If  $z_0 = 0$ , a Taylor's series is called a *Maclaurin's series*.

5. *Laurent's Series.* Let  $f(z)$  be analytic in the annular region  $D$  bounded by two concentric circles with center  $z_0$ ; then  $f(z)$  can be represented by the *Laurent's series*

$$\sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$$

where

$$a_n = \frac{1}{2\pi i} \int_C (\zeta - z_0)^{-n-1} f(\zeta) d\zeta$$

and  $C$  is a simple closed curve lying in  $D$  and enclosing the inner circle.

6. *Identity Theorem.* If two functions are analytic in a region, and if they coincide, in any neighborhood of any point  $z_0$  of the region or any curve terminating at  $z_0$  or even for an infinite number of distinct points with the limit point  $z_0$ , then the two functions are equal throughout the region.

7. *Principle of the Maximum Modulus.* The maximum modulus of a function analytic in a closed region always lies on the boundary of the region.

8. *Liouville's Theorem.* If a function  $f(z)$  is analytic and its modulus  $|f(z)|$  is bounded for all values of  $z$  in the complex plane, then  $f(z)$  is a constant.

**2.24 Singularities and the Classification of Functions.** If a function can be made analytic at a point  $z_0$  by merely assigning the function a new value at the point  $z_0$ , then the function is said to have a *removable singularity* at  $z_0$ . An *isolated singular point* of a function is a singular point that can be enclosed by a circle containing no other singular point of the function. An isolated singular point  $z_0$  of a function  $f(z)$  is called a *pole of order  $n$*  if a positive integer  $n$  exists such that

$$(z - z_0)^n f(z)$$

is analytic at  $z = z_0$  and is different from zero when  $z = z_0$ . In case  $n = 1$ , the pole is called a *simple pole*. An isolated singular point  $z_0$  of a function  $f(z)$  is called an *essential singularity* of  $f(z)$  if

$$\lim_{z \rightarrow z_0} (z - z_0)^n f(z)$$

tends to infinity for all finite values of  $n$ .  $z = z_0$  is called a *branch point* of the function  $f(z)$  if  $f(z_0 + \rho e^{i\varphi})$  is not periodic in  $\varphi$  with period  $2\pi$ , where  $\rho$  is chosen so that  $z_0 + \rho e^{i\varphi}$  is in the region of analyticity of  $f$  for all  $\varphi$ .

The nature and location of the singularities of a function  $f(z)$  lead to the following classification:

1. By Liouville's theorem if  $f(z)$  has no singularities, it is a constant.
2. If  $f(z)$  has only an isolated singularity at infinity, it is called an *entire* (or integral) function. If the singularity is a pole of order  $n$ , then the entire function becomes a polynomial of order  $n$ .
3. If  $f(z)$  has only poles in a finite region, then it is called a *meromorphic function* of  $z$  in that region. If  $f(z)$  has only poles in the finite  $z$  plane and either is analytic or has an isolated singularity at infinity, then it is called *meromorphic*.
4. If  $f(z)$  has a branch point it is a *multivalued function*.

**2.25 Residues.** At an isolated singular point a function  $f(z)$  may be represented by the Laurent's series

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} b_n (z - z_0)^{-n}$$

The expression  $\sum_{n=1}^{\infty} b_n (z - z_0)^{-n}$  is called the *principal part* of  $f(z)$ . The coefficient  $b_1$

is defined as the *residue* of the function  $f(z)$  at the point  $z_0$ . The formula for  $b_1$  is given from the Laurent's expansion by

$$b_1 = \frac{1}{2\pi i} \int_C f(z) dz$$

where  $C$  is a simple closed curve enclosing  $z_0$ .

The fundamental theorem of the calculus of residues follows:

*Residue Theorem.* Let  $f(z)$  be analytic except for a finite number  $n$  of isolated singular points within and on the closed curve  $C$ ; then

$$\int_C f(z) dz = 2\pi i \sum_{j=1}^n R_j$$

where  $R_1, \dots, R_n$  are the residues at the  $n$  singular points.

When  $f(z)$  has a simple pole at  $z_0$  and has the form

$$f(z) = \frac{g(z)}{h(z)}$$

where  $g(z_0) \neq 0$ ,  $h(z_0) = 0$ , and  $h'(z_0) \neq 0$ , the residue of  $f(z)$  at  $z_0$  is

$$\frac{g(z_0)}{h'(z_0)}$$

**2.26 Conformal Mapping.** Let  $w = f(z)$  denote a mapping or transformation from the  $z$  plane to the  $w$  plane. If the transformation preserves the magnitude of angles but not necessarily the sense, it is called *isogonal*. If in a mapping both the magnitude and sense are preserved between every pair of curves through a point  $z_0$ , then the mapping is called *conformal*.

If a function is analytic at a point  $z_0$ , then either  $f'(z) = 0$  or the mapping  $w = f(z)$  is conformal at  $z_0$ . A *critical point*  $z_0$  of a mapping is a point at which  $f'(z_0) = 0$ . One of the most important results of a conformal mapping is that a harmonic function, that is, a function  $K(x, y)$  which satisfies Laplace's equation  $(\partial^2 K / \partial x^2) + (\partial^2 K / \partial y^2) = 0$ , remains harmonic under the change of variables that arises from the conformal mapping  $w = u + iv = f(z)$ ; that is,  $(\partial^2 K / \partial u^2) + (\partial^2 K / \partial v^2) = 0$ . Furthermore, a boundary condition of the type  $K(x, y) = C$  or of the type  $dK/dn = 0$ , where  $dK/dn$  is the normal derivative, transforms into a boundary condition of the same type. Therefore by using analytic functions it is possible to find in many cases a function that is harmonic in a given region and satisfies boundary conditions of the above type.

The transformations  $w = (\alpha z + \beta) / (\alpha z + \delta)$ , where  $\alpha\delta - \beta\gamma \neq 0$ , called linear fractional transformations, are conformal. In particular they map circles, which include straight lines, since the lines are circles with infinite radius, into circles. Other important properties of linear fractional transformations may be found in the references.

The transformation

$$w = C_1 \int (z - x_1)^{-k_1} (z - x_2)^{-k_2} \cdots (z - x_{n-1})^{-k_{n-1}} dz + C_2 = f(z)$$

where  $C_1$  and  $C_2$  are arbitrary constants and the integral is an indefinite integral, is called a *Schwarz-Christoffel transformation*. This mapping takes the  $x$  or real axis in the  $z$  plane into a polygon of  $n$  sides in the  $w$  plane. The points  $w_i = f(x_i)$ , where  $i = 1, 2, \dots, n$  and  $x_n = \infty$ , are the vertices of the polygon. The exterior angles at the vertices  $w_i$  ( $i = 1, 2, \dots, n - 1$ ) are given by  $k_i\pi$ . The exterior angle at  $w_n$  is given by

$$k_n\pi = 2\pi - (k_1 + k_2 + \cdots + k_{n-1})\pi$$

**2.27 Analytic Continuation. Riemann Surfaces.** The identity theorem listed under analytic functions leads to the important concept of analytic continuation. Let  $f_1(z)$  be an analytic function in a region  $D_1$ , and let  $f_2(z)$  be an analytic function in a region  $D_2$ . Furthermore, let the regions  $D_1$  and  $D_2$  have a subregion in common in which the functions  $f_1(z)$  and  $f_2(z)$  coincide completely (or even on a curve in the subregion). The functions  $f_1(z)$  and  $f_2(z)$  then define the same analytic function  $F(z)$ .  $f_1(z)$  and  $f_2(z)$  are called *analytic continuations* of each other. Furthermore,  $f_1(z)$  and  $f_2(z)$  are called *elements* of (or partial representatives of) the function  $F(z)$ , which is analytic over the composite region formed by  $D_1$  and  $D_2$ .

It is often possible to start with a given element, such as a Taylor's series at a given point and by using power (or Taylor's) series at other points extend the domain of existence of the function by analytic continuation. If the process of analytic continuation is carried out as far as possible, the resulting function is called the *complete analytic function*. A complete analytic function is called *single-valued* when its value and behavior at every point  $z_0$  are always the same, independent of any path by which it may be reached by analytic continuation. A multiple-valued function is a function that is not single-valued.

In considering multiple-value functions it is convenient to introduce the geometric concept of a Riemann surface. A *Riemann surface* is a generalization of the  $z$  plane consisting of a surface of more than one sheet arranged vertically. On each point

of this surface the multiple-valued function has one and only one value, and the function is therefore single-valued on the surface.

### 2.3 Special Functions

Since over a thousand special functions have been investigated, it is possible to consider here only a few of the more important functions of mathematical and reactor physics.

**2.31 Gamma Functions. Factorial, Polygamma, Beta.** *Definitions.* 1. THE GAMMA FUNCTION  $\Gamma(z)$  is an analytic meromorphic function of  $z$  with simple poles at  $z = -n$ , where  $n = 0, 1, 2, \dots$  and with corresponding residues  $(-1)^n/n!$ . The following conditions then determine  $\Gamma(z)$  uniquely:

- $\Gamma(z+1) = z\Gamma(z)$ .
- If  $\Gamma(z)$  is real and positive, then  $(z)$  is real and positive.
- $\Gamma(1) = 1$ .
- $[(d^2/dz^2)\Gamma(z)]\Gamma(z) - (d\Gamma/dz)^2 > 0$  when  $z$  is real and positive.

#### 2. WEIERSTRASS DEFINITION:

$$\frac{1}{\Gamma(z)} = ze^{\gamma z} \prod_{n=1}^{\infty} \left[ \left(1 + \frac{z}{n}\right) e^{-z/n} \right]$$

where  $\gamma$ , the *Euler-Mascheroni constant*, is defined by

$$\gamma = \lim_{J \rightarrow \infty} \left( \sum_{j=1}^J \frac{1}{j} - \ln J \right) = 0.577215665 \dots$$

#### 3. EULER'S FORMULA:

$$\Gamma(z) = \frac{1}{z} \prod_{n=1}^{\infty} \left[ \left(1 + \frac{1}{n}\right)^z \left(1 + \frac{z}{n}\right)^{-1} \right]$$

exists except for  $z = -n$ ,  $n = 0, 1, 2, \dots$

#### 4. EULER'S INTEGRAL:

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt$$

holds for the real part of  $z$  greater than zero ( $\text{Re } z > 0$ ).

#### Functional Equations:

$$\Gamma(z+1) = z\Gamma(z)$$

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$$

$$\Gamma(z)\Gamma\left(z + \frac{1}{n}\right)\Gamma\left(z + \frac{2}{n}\right) \cdots \Gamma\left(z + \frac{n-1}{n}\right) = (2\pi)^{(n-1)/2} n^{1/2-n} \Gamma(nz)$$

For  $n = 2$  this becomes  $\Gamma(z)\Gamma(z + 1/2) = (2\pi)^{1/2} 2^{1/2-2z} \Gamma(2z)$ .

#### Special Values:

$$\Gamma(n+1) = n! = n(n-1)(n-2) \cdots 2 \cdot 1 \quad (n = 0, 1, 2, \dots)$$

$$\Gamma(1) = 0! = 1 \quad \Gamma(2) = 1$$

$$\Gamma(1/2) = \sqrt{\pi}$$

$$\Gamma'(1) = -\gamma, \text{ the Euler constant}$$

Derivatives:

$$\psi_1(z) \equiv \frac{\Gamma'(z)}{\Gamma(z)} \equiv \frac{d}{dz} [\ln \Gamma(z)]$$

$$\lambda_n(z) \equiv \frac{d}{dz} [\psi_{n-1}(z)], \quad n = 2, 3, \dots \text{ is called a } \textit{polygamma function}.$$

$$\psi_n(z) = (-1)^n n! \sum_{j=0}^{\infty} \frac{1}{(j+z)^{n+1}}$$

$$\psi_n(z+1) = \psi_n(z) - [(-1)^n n!] z^{-n}$$

*Sterling's Formula or Asymptotic Formula for Large |z|:*

$$\ln \Gamma(z) = (z - \frac{1}{2}) \ln z - z + \frac{1}{2} \ln 2\pi + R(z)$$

where

$$R(z) = \sum_{n=1}^{N-1} (-1)^{n-1} \frac{B_{2n} z^{1-2n}}{2n(2n-1)} + R_N(z)$$

where

$$|\arg z| < \pi$$

and where

$$|R_N(z)| \leq \frac{B_{2N}}{2N(2N-1)|z|^{2N-1} \left(\cos \frac{\varphi}{2}\right)^{2N-1}} \quad (z = re^{i\varphi})$$

The numbers  $B_{2n}$  are the *Bernoulli numbers* and are defined by  $B_0 = 1, B_1 = \frac{1}{2}, B_{2n+1} = 0$ , and

$$B_{2n} = \frac{2(2n)!}{(2\pi)^{2n}} \sum_{k=1}^{\infty} \frac{1}{k^{2n}}$$

In particular,  $n! = n^n e^{-n} \sqrt{2\pi n} [1 + r(x)]$ , where

$$0 < r(x) < \frac{1}{12x} + \frac{1}{288x^2}$$

*The Beta Function.* This is defined by

$$B(m, n) = \int_0^1 t^{m-1} (1-t)^{n-1} dt$$

where  $\text{Re } m > 0$  and  $\text{Re } n > 0$ .

Alternative definitions are

$$B(m, n) = \int_0^{\infty} \frac{t^{m-1}}{(1+t)^{m+n}} dt \quad B(m, n) = \int_0^{\pi/2} \sin^{2m-1} \varphi \cos^{2n-1} \varphi d\varphi$$

$$B(m, n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} = B(n, m)$$

**2.32 Hypergeometric Functions. Ordinary, General, Confluent.** *Ordinary Hypergeometric Function.* The *hypergeometric differential equation*

$$z(1-z) \frac{d^2w}{dz^2} + [c - z(a+b+1)] \frac{dw}{dz} - abw = 0$$

has at the origin the general solution

$$w(z) = AF(a, b; c; z) + Bz^{1-c}F(a+1-c, b+1-c; 2-c; z)$$

\*The notation "Re" means "the real part of."

where  $A$  and  $B$  are arbitrary constants and  $a$ ,  $b$ , and  $c$  are parameters. The function

$$F(a, b; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}$$

where  $(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)} = (a+n-1)(a+n-2) \cdots (a+1)a$   $(a)_0 = 0$

is called the *hypergeometric function or series*. The circle of convergence for this series is the unit circle  $|z| = 1$ .

The hypergeometric equation can also be written in the form

$$[\theta(\theta + c - 1) - z(\theta + a)(\theta + b)]w = 0$$

where the operator  $\theta \equiv z(d/dz)$ . The singularities (regular) of the equation occur at 0, 1, and  $\infty$ . Because of these singularities solutions of the hypergeometric equations are also often written in terms of the arguments  $1/z$ ,  $1 - z$ ,  $1/(1 - z)$ ,  $z/(1 - z)$ , and  $(1 - z)/z$ . Kummer's 24 solutions and various relationships among them involve these arguments and may be found in the references.

The six functions  $F(\alpha \pm 1, \beta; \gamma; z)$ ,  $F(\alpha, \beta \pm 1; \gamma; z)$ , and  $F(\alpha, \beta; \gamma \pm 1; z)$  are called *contiguous functions* to the functions  $F(\alpha, \beta; \gamma; z)$ . Gauss's 15 recursion formulas relate the contiguous functions by expressing one function in terms of two others. These formulas may also be found in the references.

In the references are found various integral representations, both real and contour, for the hypergeometric function. The best known integral representation follows:

$$F(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a} dt$$

Re  $c > 0$     Re  $b > 0$      $|z| < 1$

When  $z = 1$  and Re  $(a + b - c) < 0$ ,

$$F(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$$

*Generalized Hypergeometric Functions.* The *generalized hypergeometric differential equation*

$$\left[ \theta \prod_{j=1}^q (\theta + b_j - 1) - z \prod_{i=1}^p (\theta + a_i) \right] w = 0$$

is of the order  $\max(p, q + 1)$ . It has singularities at 0 and  $\infty$  if  $p \neq q + 1$  and at 0, 1, and  $\infty$  if  $p = q + 1$ . The solution regular at  $z = 0$  is

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; z) = {}_pF_q(a; b; z) = \sum_{n=0}^{\infty} \frac{(a_1)_n \cdots (a_p)_n}{(b_1)_n \cdots (b_q)_n} \frac{z^n}{n!}$$

and is called the *generalized hypergeometric series*. It is assumed that the  $b$ s are not negative integers. In general this series converges for all finite  $z$  if  $p \leq q$ , converges for  $|z| < 1$  if  $p = q + 1$ , and diverges for all  $z \neq 0$  if  $p > q + 1$ .

Contiguous relations, integral representations, and relations among various arguments for the generalized function may be found in the references.

*Confluent Hypergeometric Functions.* The *confluent hypergeometric function*, or *Kummer function*,

$${}_1F_1(a; b; z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!}$$

satisfies the *Kummer's differential equation*

$$z \frac{d^2w}{dz^2} + (b - z) \frac{dw}{dz} - aw = 0$$

*Kummer's first formula* is

$${}_1F_1(a; b; z) = e^z {}_1F_1(b - a; b; -z)$$

*Kummer's second formula* is

$${}_1F_1(a; 2a; 2z) = e^{2z} {}_1F_1(a + \frac{1}{2}; \frac{3}{2}z^2)$$

The three independent recursion formulas or contiguous relations are

$$F(a; b; z) = F(a - 1; b; z) + \frac{z}{b} F(a; b + 1; z)$$

$$(a + z)F(a; b; z) = aF(a + 1; b; z) + z \frac{b - a}{b} F(a; b + 1; z)$$

$$(a - b + 1)F(a; b; z) = aF(a + 1; b; z) + (1 - b)F(a; b - 1; z)$$

An integral representation when  $\text{Re } b > \text{Re } a > 0$  is given by

$$F(a; b; z) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b - z)} \int_0^1 t^{a-1} (1 - t)^{b-a-1} e^{tz} dt$$

Most of the functions of mathematical physics can be expressed in terms of generalized hypergeometric series. Examples of functions that can be expressed as special cases of a  ${}_2F_1$  or a  ${}_1F_1$  appear in some of the following articles.

**2.33 The Cylindrical or Bessel Functions.** The cylindrical, or Bessel, functions are solutions of the Bessel differential equation

$$z^2 \frac{d^2w}{dz^2} + z \frac{dw}{dz} + (z^2 - \nu^2)w = 0$$

This equation has a regular singularity at  $z = 0$  and an irregular singularity at  $z = \infty$ . The functions

$$J_\nu(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(k + \nu + 1)} \left(\frac{z}{2}\right)^{2k+\nu}$$

$$N_\nu(z) \equiv Y_\nu(z) = \frac{1}{\sin \nu\pi} [J_\nu(z) \cos \nu\pi - J_{-\nu}(z)]$$

$$H_\nu^{(1)}(z) = J_\nu(z) + iY_\nu(z) \quad H_\nu^{(2)}(z) = J_\nu(z) - iY_\nu(z)$$

are all solutions of the Bessel differential equation.  $J_\nu(z)$  is called a *Bessel function of the first kind*. The subscript  $\nu$  is the order of the function.  $Y_\nu$  (or  $N_\nu$ ) is called a *Neumann function* or a *Bessel function of the second kind*.  $H_\nu^{(1)}$  and  $H_\nu^{(2)}$  are called *first and second Hankel functions* or *Bessel functions of the third kind*. It may be noted that  $J_\nu(z)$  and  $Y_\nu(z)$  are real if  $\nu$  is real and  $z$  is positive.

If  $\nu$  is not an integer, it is customary to choose  $J_\nu(z)$  and  $J_{-\nu}(z)$  as the two linearly independent solutions of Bessel's equation. When  $\nu = n$ , where  $n$  is a positive integer, it is necessary to use  $J_n(z)$  and  $Y_n(z) = \lim_{\nu \rightarrow n} Y_\nu(z)$  for the two independent solutions,

since

$$J_{-n}(z) = (-1)^n J_n(z)$$

The *Bessel coefficients*, that is, the Bessel function of the first kind of integer order, occur in the following expansion:

$$\exp[\frac{1}{2}z(t - t^{-1})] = \sum_{n=-\infty}^{\infty} J_n(z)t^n$$

The Bessel differential equation has the form, when  $z$  is replaced by  $iz$ ,

$$z^2 \frac{d^2 w}{dz^2} + z \frac{dw}{dz} - (z^2 + \nu^2)w = 0$$

The functions

$$I_\nu(z) = e^{-i\frac{1}{2}\nu\pi} J_\nu(ze^{i\frac{1}{2}\pi}) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(k + \nu + 1)} \left(\frac{z}{2}\right)^{2k+\nu}$$

and

$$K_\nu(z) = \frac{\pi i}{2} e^{\nu\pi i/2} H_\nu^{(1)}(iz)$$

which are solutions of this equation, are called, respectively, modified Bessel functions of the first and second kind. Again if  $\nu$  is not an integer,  $I_\nu(z)$  and  $I_{-\nu}(z)$  are taken as the two independent solutions, and if  $\nu$  is an integer,  $I_\nu(z)$  and  $K_\nu(z)$  are taken as the two solutions.  $I_\nu(z)$  and  $K_\nu(z)$  are real when  $\nu$  is real and  $z$  is positive.

Some of the more useful relations involving Bessel functions follow.  $Z_\nu$  is used as an abbreviation for  $c_1 J_\nu(z) + c_2 Y_\nu(z)$ , where  $c_1$  and  $c_2$  denote arbitrary constants.

*Functional Equations:*

$$\begin{aligned} Z_{\nu-1} + Z_{\nu+1} &= \frac{2\nu}{z} Z_\nu \\ J_\nu J_{-\nu+1} + J_{\nu-1} J_{-\nu} &= \frac{2 \sin \nu\pi}{\pi z} \\ J_\nu Y_{\nu-1} - Y_\nu J_{\nu-1} &= \frac{2}{\pi z} \\ I_\nu I_{-\nu+1} - I_{-\nu} I_{\nu-1} &= -\frac{2 \sin \nu\pi}{\pi z} \\ I_\nu K_{\nu+1} + I_{\nu+1} K_\nu &= \frac{1}{z} \end{aligned}$$

*Derivatives:*

$$\begin{aligned} Z'_\nu &= \frac{1}{2} [Z_{\nu-1} - Z_{\nu+1}] = \frac{-\nu}{z} Z_\nu + Z_{\nu-1} = \frac{\nu}{z} Z_\nu - Z_{\nu+1} \\ [z^\nu Z_\nu(az)]' &= az^\nu Z_{\nu-1}(az) \\ Z'_0 &= -Z_1 \quad Z'_1 = Z_0 - \frac{1}{z} Z_1 \\ J'_\nu J_{-\nu} - J'_{-\nu} J_\nu &= -\frac{2 \sin \nu\pi}{\pi z} \\ J'_\nu Y_\nu - J_\nu Y'_\nu &= \frac{2}{\pi z} \\ I'_\nu I_{-\nu} - I'_{-\nu} I_\nu &= -\frac{2 \sin \nu\pi}{\pi z} \\ I'_\nu K_\nu - K'_\nu I_\nu &= -\frac{1}{z} \end{aligned}$$

*Integral Representations:*

$$\begin{aligned} J_n(z) &= \frac{1}{\pi} \int_0^\pi \cos(z \sin \varphi - n\varphi) d\varphi \\ J_\nu(z) &= \frac{2\pi^{-\frac{1}{2}}}{\Gamma(\nu + \frac{1}{2})} \left(\frac{z}{2}\right)^\nu \int_0^1 (1-t^2)^{\nu-\frac{1}{2}} \cos zt dt \end{aligned}$$

*Representation as a Hypergeometric Function:*

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu \frac{e^{-i\pi\nu}}{\Gamma(\nu+1)} {}_1F_1\left(\nu + \frac{1}{2}; 2\nu + 1; 2iz\right) = \frac{(z/2)^\nu}{\Gamma(\nu+1)} {}_0F_1\left(\nu + 1; -\frac{1}{4}z^2\right)$$

General Differential Equation:

if  $m \neq 0$  and  $b \neq 0$ , 
$$z^2 w'' + azw' + (bz^m + c)z = 0$$

$$w = z \frac{1-a}{2} Z_\nu \left( \frac{2}{m} \sqrt{b} z^{m/2} \right)$$

where 
$$\nu = \frac{1}{m} \sqrt{(1-a)^2 - 4c}$$

Special Cases:

$$J_{\frac{1}{2}}(z) = Y_{-\frac{1}{2}}(z) = \frac{\sin z}{(\frac{1}{2}\pi z)^{\frac{1}{2}}}$$

$$J_{-\frac{1}{2}}(z) = -Y_{\frac{1}{2}}(z) = \frac{\cos z}{(\frac{1}{2}\pi z)^{\frac{1}{2}}}$$

$$H_{\frac{1}{2}}^{(1)}(z) = \frac{-i}{(\frac{1}{2}\pi z)^{\frac{1}{2}}} e^{iz} \quad H_{\frac{1}{2}}^{(2)}(z) = \frac{i}{(\frac{1}{2}\pi z)} e^{-iz}$$

**2.34 Legendre Functions.** The Legendre functions are solutions of Legendre's differential equation

$$(1 - z^2) \frac{d^2 w}{dz^2} - 2z \frac{dw}{dz} + \left[ \nu(\nu + 1) - \frac{\mu^2}{(1 - z^2)} \right] w = 0$$

where  $\nu$  and  $\mu$  are arbitrary parameters. The functions

$$P_\nu^\mu(z) = \frac{1}{\Gamma(1 - \mu)} \left( \frac{z + 1}{z - 1} \right)^{\frac{1}{2}\mu} F \left( -\nu, \nu + 1; 1 - \mu; \frac{1}{2} - \frac{1}{2} z \right)$$

and

$$Q_\nu^\mu(z) = \frac{e^{i\mu\pi} \pi^{\frac{1}{2}} \Gamma(\nu + \mu + 1)}{2^{\nu+1} \Gamma(\nu + \frac{3}{2})} \frac{(z^2 - 1)^{\frac{1}{2}\mu}}{z^{\nu+\mu+1}} F \left( \frac{1}{2} \nu + \frac{1}{2} \mu + 1, \frac{1}{2} \nu + \frac{1}{2} \mu + \frac{1}{2}; \nu + \frac{3}{2}, \frac{1}{2} z \right)$$

are linearly independent solutions of Legendre's equation.  $P_\nu^\mu(z)$  and  $Q_\nu^\mu(z)$  are known respectively as Legendre functions of the first and second kind. Since Legendre's equation is not changed when  $z$  is replaced by  $-z$ ,  $\mu$  by  $-\mu$ , and  $\nu$  by  $-(\nu + 1)$ , it follows that

$$P_{\nu, \pm\mu}(\pm z) \quad P_{-\nu-1, \pm\mu}(\pm z) \quad Q_{\nu, \pm\mu}(\pm z) \quad Q_{-\nu-1, \pm\mu}(\pm z)$$

are also solutions of the equation.

The most common Legendre functions occur when  $\mu = 0$ . The superscript 0 is dropped in this case, and the Legendre functions are written as  $P_\nu(z)$  and  $Q_\nu(z)$ . In case  $\mu \neq 0$ , the functions are often called associated Legendre functions. The following simple relationships exist:

$$P_\nu^k(z) = (z^2 - 1)^{k/2} \frac{d^k P_\nu(z)}{dz^k}$$

and 
$$Q_\nu^k(z) = (z^2 - 1)^{\frac{1}{2}k} \frac{d^k P_\nu(z)}{dz^k} \quad k = 1, 2, \dots$$

When  $\nu$  is a nonnegative integer  $n$ , the functions  $P_n(z)$  become polynomials called the Legendre polynomials. In hypergeometric function notation

$$P_{2n}(z) = \frac{(-1)^n (2n)!}{2^{2n} (n!)^2} F \left( -n, n + \frac{1}{2}; \frac{1}{2}; z^2 \right)$$

and 
$$P_{2n+1}(z) = \frac{(-1)^n (2n + 1)!}{2^{2n} (n!)^2} z F \left( -n, n + \frac{3}{2}; \frac{3}{2}; z^2 \right)$$

A convenient formula for the Legendre polynomial is

$$P_n(z) = \frac{1}{2^n n!} \frac{d^n}{dz^n} (z^2 - 1)^n$$

known as *Rodrigues' formula*. These polynomials form an orthogonal system on the interval  $[-1, 1]$  and have all their roots real and simple and between  $-1$  and  $1$ . The Legendre polynomials occur in the following expansion:

$$(1 - 2zt + t^2)^{-1/2} = \sum_{n=0}^{\infty} P_n(z)t^n$$

Some useful relations follow:

*Functional Equation:*

$$(n + 1)P_{n+1} = z(2n + 1)P_n - nP_{n-1}$$

*Derivatives:*

$$(2n + 1)P_n = P'_{n+1} - P'_{n-1}$$

$$nP_n = zP'_n - P'_{n-1}$$

$$(n + 1)P_n = P'_{n+1} - zP'_n$$

*Laplace's Integral:*

$$P_n(z) = \frac{1}{\pi} \int_0^\pi [(z^2 - 1)^{1/2} \cos \varphi + z]^n d\varphi$$

*Special Cases and Values:*

$$P_0(x) = 1 \quad P_1(x) = x = \cos \varphi \quad P_2(x) = \frac{1}{2}(3x^2 - 1) = \frac{1}{4}(3 \cos 2\theta + 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x) = \frac{1}{8}(5 \cos 3\varphi + 3 \cos \varphi)$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3) = \frac{1}{64}(35 \cos 4\varphi + 20 \cos 2\varphi + 9)$$

$$P_n(1) = 1 \quad P_n(-1) = (-1)^n \quad P_{2n+1}(0) = 0 \quad P_{2n}(0) = (-1)^n \frac{(2n)!}{2^{2n}(n!)^2}$$

*A Sum Formula:*

$$\sum_{k=0}^n (2k + 1)P_k(x)P_k(y) = (n + 1) \frac{P_n(x)P_{n+1}(y) - P_n(y)P_{n+1}(x)}{y - x}$$

*Orthogonality Relations:*

$$\int_{-1}^1 P_n(z)P_m(z) dz = 0 \quad n \neq m \quad \int_{-1}^1 P_n(z)^2 dz = \frac{2}{2n + 1}$$

### 2.35 Orthogonal Polynomials. *Tschebyscheff Polynomials:*

$$T_n(z) = \cos(n \arccos z) \quad U_n(z) = \sin(n \arccos z)$$

$$T_n(z) = F\left(n, -n; \frac{1}{2}; \frac{1-z}{2}\right)$$

$$\int_{-1}^1 \frac{T_m(z)T_n(z)}{\sqrt{1-z^2}} dz = 0 \quad \text{if } m \neq n$$

$$= \frac{\pi}{2} \quad \text{if } m = n \neq 0$$

$$= \pi \quad \text{if } m = n = 0$$

The *Tschebyscheff* polynomials of the first and second kind,  $T_n(z)$  and  $U_n(z)$ , are linearly independent solutions of

$$(1 - z^2)w'' - zw' + n^2w = 0$$

*Hermite Polynomials:*

$$H_n(z) = (-1)^n e^{z^2/2} \frac{d^n}{dz^n} (e^{-z^2/2})$$

$$H_{2n}(z) = \frac{(-1)^n (2n)!}{2^n n!} {}_1F_1 \left( -n; \frac{1}{2}; \frac{z^2}{2} \right)$$

$$H_{2n+1}(z) = \frac{(-1)^n (2n+1)!}{2^n n!} z {}_1F_1 \left( -n; \frac{3}{2}; \frac{z^2}{2} \right)$$

$$\int_{-\infty}^{\infty} e^{-z^2/2} H_n(z) H_m(z) dz = 0 \quad \text{if } m \neq n$$

$$= (2\pi)^{1/2} n! \quad \text{if } m = n$$

$$H''_n - zH'_n + nH_n = 0$$

*Laguerre Polynomials:*

$$L_n^{(\alpha)}(z) = \frac{e^{z^2-\alpha} d^n}{n! dz^n} (e^{-z^2+\alpha}) = \frac{(\alpha+1)_n}{n!} {}_1F_1(-n; \alpha+1; z)$$

$$\int_0^{\infty} e^{-z^2\alpha} L_m^{(\alpha)} L_n^{(\alpha)}(z) dz = 0 \quad \text{if } m \neq n$$

$$= \frac{(\alpha+1)_n}{n!} \Gamma(1+\alpha) \quad \text{if } m = n$$

$$L_n^{(\alpha)}(z) \text{ satisfies } zw'' + (\alpha+1-z)w' + nw = 0.$$

*Jacobi Polynomials:*

$$P_n^{(\alpha,\beta)}(z) = \frac{(-1)^n (1-z)^{-\alpha} (1+z)^{-\beta} d^n}{2^n n! dz^n} [(1-z)^{\alpha+n} (1+z)^{\beta+n}]$$

$$P_n^{(\alpha,\beta)}(z) = \frac{(\alpha+1)_n}{n!} F \left( -n, n+\alpha+\beta+1; \alpha+1; \frac{1}{2} - \frac{1}{2}z \right)$$

$P_n^{(\alpha,\beta)}(z)$  satisfies

$$(1-z^2)y'' + [\beta-\alpha-z(\alpha+\beta+2)]y' + n(n+\alpha+\beta+1)y = 0$$

*Gegenbauer Polynomials:*

$$C_n^\alpha(z) = \frac{(2\alpha)_n}{n!} F \left( -n, n+2\alpha; \alpha+\frac{1}{2}; \frac{1-z}{2} \right)$$

**2.36 Dirac Delta Function.** In many problems of physics and engineering it is expedient to introduce a quantity  $\delta(x)$  called the *Dirac delta function* (or just  $\delta$  function).  $\delta(x)$  is not a proper function in the sense of having a definite value for each value of the independent variable. Formally a Dirac delta function  $\delta(x)$  is defined to be zero except at  $x = 0$ , that is,

$$\delta(x) = 0 \quad x \neq 0$$

but is such that

$$\int_{-\infty}^{\infty} \delta(x) dx = 1$$

This formal definition does not give a clear picture. It may, however, be inferred that the  $\delta$  function is very large near  $x = 0$  but is zero outside a very small interval about  $x = 0$ . It is not important in applications to know the precise variation of  $\delta(x)$  with  $x$ . For example, the  $\delta$  function may be defined by

$$\delta(x) = \lim_{\Delta \rightarrow 0} \delta_\Delta(x)$$

where

$$\delta_{\Delta}(x) \begin{cases} = 0 & x < -\frac{\Delta}{2} \\ = \frac{1}{\Delta} & -\frac{\Delta}{2} < x < \frac{\Delta}{2} \\ = 0 & x > \frac{\Delta}{2} \end{cases}$$

The following formal formulas may be used with caution:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\delta(x) dx &= f(0) & \int_{-\infty}^{\infty} f(x)\delta(x-a) dx &= f(a) \\ f(x)\delta(x-a) &= f(a)\delta(x-a) \\ x\delta(x) &= 0 & \delta(-x) &= \delta(x) \\ \int_{-\infty}^{\infty} \delta(a-x)\delta(x-b) dx &= \delta(a-b) \\ \delta(x^2-a^2) &= \frac{1}{2a} [\delta(x-a) + \delta(x+a)] & a > 0 \end{aligned}$$

*Laplace Transform:*

$$\begin{aligned} \int_0^{\infty} \delta(x)e^{-sx} dx &= 1 \\ \delta(x) &= \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} e^{sx} dx \end{aligned}$$

*Fourier Transform:*

$$\begin{aligned} \int_{-\infty}^{\infty} \delta(x)e^{isx} dx &= 1 \\ \delta(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-isx} dx \end{aligned}$$

*Derivative of Step Function.* Let

$$\begin{aligned} S(x) &= 0 & -\infty < x < 0 \\ &= 1 & 0 < x < \infty \\ \delta(x) &= S'(x) \end{aligned}$$

**2.37 Other Functions.** *Incomplete Gamma Functions:*

$$\begin{aligned} \gamma(a,x) &= \int_0^x e^{-t}t^{a-1} dt = \frac{x^a}{a} {}_1F_1(a; 1+a; -x) \\ \Gamma(a,x) &= \int_x^{\infty} e^{-t}t^{a-1} dt = \Gamma(a) - \gamma(a,x) \end{aligned}$$

*Error Functions:*

$$\begin{aligned} \operatorname{Erf} x &= \int_0^x e^{-t^2} dt = \frac{1}{2}\gamma(\frac{1}{2}, x^2) = x {}_1F_1(\frac{1}{2}; \frac{3}{2}; -x^2) \\ \operatorname{Erfc} x &= \int_x^{\infty} e^{-t^2} dt = \frac{1}{2}\Gamma(\frac{1}{2}, x^2) = \frac{1}{2}\pi^{1/2} - \operatorname{Erf} x \end{aligned}$$

*Exponential and Logarithmic Integrals:*

$$\begin{aligned} E_1(x) &= -E_i(-x) = \int_x^{\infty} e^{-t}t^{-1} dt = \Gamma(0,x) \\ \operatorname{li}(x) &= \int_0^x \frac{dt}{\log t} = E_i(\log x). \end{aligned}$$

*Sine and Cosine Integrals:*

$$\text{si}(x) = \int_{\infty}^x \frac{\sin t}{t} dt = \frac{1}{2i} [E_i(ix) - E_i(-ix)]$$

$$\text{Si}(x) = \int_0^x \frac{\sin t}{t} dt = \frac{\pi}{2} + \text{si}(x)$$

$$\text{Ci}(x) = \int_{\infty}^x \frac{\cos t}{t} dt = \frac{1}{2} [E_i(ix) + E_i(-ix)]$$

*Fresnel Integrals:*

$$C(x) = (2\pi)^{-1/2} \int_0^x t^{-1/2} \cos t dt$$

$$S(x) = (2\pi)^{-1/2} \int_0^x t^{-1/2} \sin t dt$$

*Elliptic Integrals:*

First kind: 
$$F(\lambda, k) = \int_0^{\lambda} \frac{dx}{[(1-x^2)(1-k^2x^2)]^{1/2}}$$

Second kind: 
$$E(\lambda, k) = \int_0^{\lambda} \left( \frac{1-k^2x^2}{1-x^2} \right)^{1/2} dx$$

Third kind: 
$$\pi(\lambda, \nu, k) = \int_0^{\lambda} \frac{dx}{(1+\nu x^2)[(1-x^2)(1-k^2x^2)]^{1/2}}$$

*Complete Elliptic Integrals:*

$$\begin{aligned} F(1, k) &= K(k) = \frac{1}{2}\pi F\left(\frac{1}{2}, \frac{1}{2}; 1; k^2\right) \\ E(1, k) &= E(k) = \frac{1}{2}\pi F\left(-\frac{1}{2}, \frac{1}{2}; 1; k^2\right) \\ \pi(1, \nu, k) &= \pi(\nu, k) \end{aligned}$$

### 3 SERIES AND EXPANSIONS OF FUNCTIONS

#### 3.1 Infinite Series

Let  $s_0, s_1, s_2, \dots$  be an infinite sequence of numbers, where  $s_0 = a_0, s_1 = a_0 + a_1$ , and generally

$$s_n = a_0 + a_1 + \dots + a_n$$

then if the sequence  $\{s_n\}$  has some convergence property, the *infinite series*

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + \dots$$

is said to have that convergence property. The numbers  $s_n$  are called the *partial sums*.

If the infinite series  $\sum_{n=0}^{\infty} a_n$  is convergent, the limit of the sequence  $\{s_n\}$  is called the *sum* or *value* of the series. If all the *terms*  $a_n$  are such that  $a_n \geq 0$ , then

$\sum_{n=0}^{\infty} a_n$  is called a *series with positive terms*. A series whose terms are alternately positive and negative is called an *alternating series*.

An infinite series  $\sum_{n=0}^{\infty} a_n$  is called *absolutely convergent* if the series of absolute values

$\sum_{n=0}^{\infty} |a_n|$  is convergent. If an infinite series is convergent but not absolutely con-

vergent, it is called *conditionally convergent*. The following results deal primarily with convergence properties. The summation symbol  $\Sigma$  by itself will imply  $\sum_{n=0}^{\infty}$ .

**3.11 Fundamental Theorem.** A necessary and sufficient condition that the series  $\Sigma a_n$  is convergent is that for any  $\epsilon > 0$  there exists a number  $N = N(\epsilon)$  such that for every  $n > N$  and every integer  $m \geq 1$ ,

$$|s_{n+m} - s_n| \equiv |a_{n+1} + a_{n+2} + \cdots + a_{n+k}| < \epsilon$$

**3.12 Comparison Test.** Let  $\Sigma c_n$  and  $\Sigma d_n$  be two series with positive terms, and furthermore, let  $\Sigma c_n$  converge and  $\Sigma d_n$  diverge. If  $a_n \leq c_n$  for all  $n$  greater than some  $N$ , then  $\Sigma a_n$  converges. If  $a_n \geq d_n$  for all  $n$  greater than some  $N$ , then  $\Sigma a_n$  diverges.

*Root Test.* If the series with positive terms  $\Sigma a_n$  is such that for all  $n > N$ ,

$$\sqrt[n]{a_n} \leq a < 1$$

then the series is convergent. However, if for all  $n > N$ ,

$$\sqrt[n]{a_n} \geq 1$$

then the series diverges.

**3.13 Ratio Test.** If for all  $n > N$ ,  $a_n > 0$ , and

$$\frac{a_{n+1}}{a_n} \leq a < 1$$

the series  $\Sigma a_n$  is convergent. However, if for all  $n > N$ ,

$$\frac{a_{n+1}}{a_n} \geq 1$$

the series  $\Sigma a_n$  is divergent.

**3.14 Alternating Series.** An alternating series  $\Sigma a_n$  is convergent if  $|a_n| < |a_{n-1}|$  and if  $\lim_{n \rightarrow \infty} a_n = 0$ . The error made in taking the sum of the first  $n$  terms of the

alternating series  $\Sigma a_n$  as an approximation to the sum of the series is less than the absolute value of the  $(n + 1)$ st term.

**3.15 Rearrangement Theorem.** The sum of an absolutely convergent series remains the same whatever change is made in the order of the terms of the series. In a conditionally convergent series it is possible to rearrange the order of the terms so that the new series converges to any desired value.

**3.16 Cauchy's Products.** If two series  $\Sigma c_j$  and  $\Sigma d_k$  are absolutely convergent

with respective sums  $C$  and  $D$ , then the *Cauchy's product*  $\sum_{j=0}^{\infty} \sum_{k=0}^j c_k d_{j-k} = \Sigma c_j \Sigma d_k$  is

also absolutely convergent and has the sum  $CD$ .

**3.17 Geometric Series.** The geometric series  $\sum_{n=0}^{\infty} r^n$  is convergent if  $|r| < 1$ ,

with sum  $1/(1 - r)$ , and is divergent if  $|r| \geq 1$ .

**3.18 Harmonic Series.** The harmonic series  $\sum_{n=1}^{\infty} (1/n)$  is divergent.

**3.19 Hyperharmonic Series.** The hyperharmonic series  $\sum_{n=1}^{\infty} (1/n^k)$  is convergent when  $k > 1$  and is divergent when  $k \leq 1$ .



and in general, it can be shown inductively that

$$\begin{aligned} A_i &= a_i A_{i-1} + b_i A_{i-2} \\ B_i &= a_i B_{i-1} + b_i B_{i-2} \end{aligned}$$

As functions of the variables  $a$  and  $b$ , the  $A_i$  and  $B_i$  are called continuants. If one represents

$$A_i = K \left( \begin{matrix} b_1, \dots, b_i \\ a_0, a_1, \dots, a_i \end{matrix} \right)$$

thus exhibiting the variables, then

$$B_i = K \left( \begin{matrix} b_2, \dots, b_i \\ a_1, a_2, \dots, a_i \end{matrix} \right)$$

Two important properties of continuants are expressed by the identities

$$K \left( \begin{matrix} b_1, \dots, b_i \\ a_0, a_1, \dots, a_i \end{matrix} \right) = K \left( \begin{matrix} b_i, \dots, b_1 \\ a_i, a_{i-1}, \dots, a_0 \end{matrix} \right)$$

$$\text{and} \quad K \left( \begin{matrix} c_1 b_1, c_1 c_2 b_2, \dots, c_{n-1} c_n b_n \\ a_0, c_1 a_1, c_2 a_2, \dots, c_n a_n \end{matrix} \right) = c_1 \cdots c_n K \left( \begin{matrix} b_1, \dots, b_n \\ a_0, a_1, \dots, a_n \end{matrix} \right)$$

for any  $c_1, \dots, c_n$ . In particular it is always possible to choose the  $c$ s so that every  $a_1 = a_2 = \dots = 1$  or so that every  $b_1 = b_2 = \dots = 1$ , provided only that no  $a_i = 0$  in the one case and no  $b_i = 0$  in the other.

If every  $b_i = 1$ , then  $F$  is said to be *periodic* or *recurring* in case for some positive integers  $n$  and  $\nu$  it is true that  $a_{i+\nu} = a_i$  whenever  $i \geq n$ . The value of any periodic continued fraction can be expressed in the form  $F = (A + B^2)/C$ , where  $A$ ,  $B$ , and  $C$  are polynomials in the variables  $a_0, \dots, a_n$ , and conversely.

If the  $a_i$  or the  $b_i$  or both are functions of a variable  $x$ , then the continued fraction defines a function  $F(x)$  wherever it converges. For questions of convergence the following identity is often useful:

$$F_n - F_{n-1} = (-1)^{n-1} b_1 b_2 \cdots b_n / (B_n B_{n-1})$$

The following are some useful continued fractions:

$$\tan x = x/\sqrt{1} - x^2/\sqrt{3} - x^2/\sqrt{5} - x^2/\sqrt{7} - \cdots \quad (1)$$

$$\tanh x = x/\sqrt{1} + x^2/\sqrt{3} + x^2/\sqrt{5} + x^2/\sqrt{7} + \cdots \quad (2)$$

$$\tan^{-1} x = x/\sqrt{1} + 1^2 x^2/\sqrt{3} + 2^2 x^2/\sqrt{5} + 3^2 x^2/\sqrt{7} + \cdots \quad (3)$$

$$\tanh^{-1} x = x/\sqrt{1} - 1^2 x^2/\sqrt{3} - 2^2 x^2/\sqrt{5} - 3^2 x^2/\sqrt{7} - \cdots \quad (4)$$

$$\exp x = 1/\sqrt{1} - x/\sqrt{1} + x/\sqrt{2} - x/\sqrt{3} + x/\sqrt{4} - x/\sqrt{5} + \cdots \quad (5)$$

### 3.3 Power Series

The series

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

where  $x$  is a variable and the numbers  $a_n$ , called *coefficients*, do not depend upon  $x$ , is called a power series in  $x$ . For power series the primary question is for what values of  $x$  does the resulting series of constants have various convergence properties. The following theorem applies to power series that converge for some value of  $x$  and diverge for other values.

**3.31 Fundamental Power Series Theorem.** There exists a positive number  $r$  such that  $\sum a_n x^n$  converges absolutely for every  $|x| < r$  but diverges for every  $|x| > r$ . If  $x$  is a real variable, the interval  $(-r, r)$  is called the interval of convergence and  $r$  is called the radius of convergence. If the real variable  $x$  is replaced by a complex

variable  $z$ , the number  $r$  is again called the *radius of convergence* and the circle  $|z| = r$  is called the *circle of convergence*.

The series  $\sum_{n=0}^{\infty} a_n(x - x_0)^n$  is called a power series in  $x - x_0$ . Letting  $x' = x - x_0$ , this series converges by the above theorem when  $|x - x_0| = |x'| < r$ . The quantity  $x_0$  is called the *center of the series*.

Some of the properties of power series in a complex variable are discussed in the article on complex variables. Some operations for power series follow:

In the common interval of convergence for two power series  $\sum a_n x^n$  and  $\sum b_n x^n$ ,

$$\sum a_n x^n \pm \sum b_n x^n = \sum (a_n \pm b_n) x^n \tag{6}$$

$$\left( \sum_{j=0}^{\infty} a_j x^j \right) \left( \sum_{n=0}^{\infty} b_n x^n \right) = \sum_{n=0}^{\infty} \left( \sum_{j=0}^n a_{n-j} b_j \right) x^n \tag{7}$$

Let  $x$ ,  $\alpha$ , and  $\beta$  be in the interval of convergence; then

$$\frac{d}{dx} \left( \sum a_n x^n \right) = \sum n a_n x^{n-1} \tag{8}$$

$$\int_{\alpha}^{\beta} \left( \sum a_n x^n \right) = \sum \frac{a_n}{n+1} x^{n+1} \tag{9}$$

Table 5 presents some of the common power series.

Table 5. Power Series

*Maclaurin Series for f(x):*

$$1. \sum_{n=0}^{\infty} a_n x^n$$

$$\text{where } a_n = \frac{f^{(n)}(0)}{n!}$$

*Taylor Series for f(x):*

$$2. \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

$$\text{where } a_n = \frac{f^{(n)}(x_0)}{n!}$$

*Binomial:*

$$3. (x + y)^n = x^n + n x^{n-1} y + \frac{n(n-1)}{2!} x^{n-2} y^2 + \frac{n(n-1)(n-2)}{3!} x^{n-3} y^3 + \dots \quad y^2 < x^2$$

When  $n$  is a positive integer, the series terminates at  $y^n$

$$4. (1 \pm x)^n = 1 \pm n x + \frac{n(n-1)}{2!} x^2 \pm \frac{n(n-1)(n-2)}{3!} x^3 + \dots \quad x < 1$$

When  $n$  is a positive integer, the series terminates at  $x^n$

*Exponential and Logarithmic:*

$$5. e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots = \lim_{n \rightarrow \infty} \left( 1 + \frac{1}{n} \right)^n = 2.718281828$$

$$6. e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

$$7. a^x = 1 + x \ln a + \frac{(x \ln a)^2}{2!} + \frac{(x \ln a)^3}{3!} + \dots$$

$$8. \ln x = \frac{x-1}{x} + \frac{1}{2} \left( \frac{x-1}{x} \right)^2 + \frac{1}{3} \left( \frac{x-1}{x} \right)^3 + \dots \quad x > \frac{1}{2}$$

$$9. \ln x = 2 \left[ \frac{x-1}{x+1} + \frac{1}{3} \left( \frac{x-1}{x+1} \right)^3 + \frac{1}{5} \left( \frac{x-1}{x+1} \right)^5 + \dots \right] \quad x > 0$$

Table 5. Power Series. (Continued)

$$10. \ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots \quad x^2 < 1$$

$$11. \ln(a+x) = \ln a + 2 \left[ \frac{x}{2a+x} + \frac{1}{3} \left( \frac{x}{2a+x} \right)^3 + \frac{1}{5} \left( \frac{x}{2a+x} \right)^5 + \cdots \right] \\ a > 0, -a < x < a$$

$$12. \ln \left( \frac{1+x}{1-x} \right) = 2 \left( x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots \right) \quad x^2 < 1$$

*Trigonometric and Inverse Trigonometric:*

$$13. \sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$14. \cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

$$15. \tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \frac{62x^9}{2,835} + \cdots \quad x^2 < \frac{\pi^2}{4}$$

$$16. \csc x = \frac{1}{x} + \frac{x}{3!} + \frac{7x^3}{3 \cdot 5!} + \frac{31x^5}{3 \cdot 7!} + \cdots \quad x^2 < \pi^2$$

$$17. \sec x = 1 + \frac{x^2}{2!} + \frac{5x^4}{4!} + \frac{61x^6}{6!} + \cdots \quad x^2 < \frac{\pi^2}{4}$$

$$18. \cot x = \frac{1}{x} - \frac{x}{3} - \frac{x^3}{45} - \frac{2x^5}{945} - \frac{x^7}{4,725} - \cdots \quad x^2 < \pi^2$$

$$19. \sin^{-1} x = x + \frac{x^3}{2 \cdot 3} + \frac{3x^5}{2 \cdot 4 \cdot 5} + \frac{3 \cdot 5x^7}{2 \cdot 4 \cdot 6 \cdot 7} + \cdots \quad x^2 < 1$$

$$20. \tan^{-1} x = \frac{\pi}{2} - \frac{1}{x} + \frac{1}{3x^3} - \frac{1}{5x^5} + \cdots \quad x^2 > 1$$

$$21. \tan^{-1} x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{1}{7}x^7 + \cdots \quad x^2 < 1$$

$$22. \sec^{-1} x = \frac{\pi}{2} - \frac{1}{x} - \frac{1}{6x^3} - \frac{3}{2 \cdot 4 \cdot 5x^5} - \frac{3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 7x^7} - \cdots \quad x^2 < 1$$

$$23. \cos^{-1} x = \frac{\pi}{2} - \sin^{-1} x, \cot^{-1} x = \frac{\pi}{2} - \tan^{-1} x, \csc^{-1} x = \frac{\pi}{2} - \sec^{-1} x$$

*Hyperbolic and Inverse Hyperbolic:*

$$24. \sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots$$

$$25. \cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots$$

$$26. \tanh x = x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315} + \cdots \quad x^2 < \frac{\pi^2}{4}$$

$$27. \sinh^{-1} x = x - \frac{x^3}{2 \cdot 3} + \frac{3x^5}{2 \cdot 4 \cdot 5} - \frac{3 \cdot 5x^7}{2 \cdot 4 \cdot 6 \cdot 7} + \cdots \quad x^2 < 1$$

$$28. \sinh^{-1} x = \log 2x + \frac{1}{2 \cdot 2x^2} - \frac{3}{2 \cdot 4 \cdot 4x^4} + \frac{3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 6x^6} + \cdots \quad x > 1$$

$$29. \cosh^{-1} x = \pm \log 2x - \frac{1}{2 \cdot 2x^2} - \frac{3}{2 \cdot 4 \cdot 4x^4} - \frac{3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 6x^6} - \cdots \quad x > 1$$

$$30. \tanh^{-1} x = x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \cdots \quad x^2 < 1$$

$$31. \ln \sin x = \ln x - \frac{x^2}{6} - \frac{x^4}{180} - \frac{x^6}{2,835} - \cdots \quad x^2 < \pi^2$$

$$32. \ln \cos x = -\frac{x^2}{2} - \frac{x^4}{12} - \frac{x^6}{45} - \frac{17x^8}{2,520} - \cdots \quad x^2 < \frac{\pi^2}{4}$$

$$33. \ln \tan x = \ln x + \frac{x^2}{3} + \frac{7x^4}{90} + \frac{62x^6}{2,835} + \cdots \quad x^2 < \frac{\pi^2}{4}$$

$$34. e^{\sin x} = 1 + x + \frac{x^2}{2!} - \frac{3x^4}{4!} - \frac{8x^5}{5!} - \frac{3x^6}{6!} - \frac{56x^7}{7!} + \cdots$$

$$35. e^{\cos x} = e \left( 1 - \frac{x^2}{2!} + \frac{4x^4}{4!} - \frac{31x^6}{6!} + \cdots \right)$$

$$36. e^{\tan x} = 1 + x + \frac{x^2}{2!} + \frac{3x^3}{3!} + \frac{9x^4}{4!} + \frac{37x^5}{5!} + \cdots \quad x^2 < \frac{\pi^2}{4}$$

3.4 Orthogonal Functions

3.41 Sturm-Liouville Series and Other Expansions. Two vectors  $\mathbf{a}$  and  $\mathbf{b}$  in  $n$ -dimensional space may be represented by rectangular coordinates; that is,

$$\mathbf{a} = (a_1, a_2, \dots, a_n)$$

and  $\mathbf{b} = (b_1, b_2, \dots, b_n)$ . The square of the length of the vector  $\mathbf{a}$  is given by  $\sum_{j=1}^n a_j^2$ . If

$\sum_{j=1}^n a_j^2 = 1$ , the vector  $\mathbf{a}$  is a unit vector and is said to be *normed or normalized*. The cosine of the angle between the two vectors  $\mathbf{a}$  and  $\mathbf{b}$  is given by

$$\cos \theta = \frac{\sum_{j=1}^n a_j b_j}{\left[ \left( \sum_{j=1}^n a_j^2 \right) \left( \sum_{j=1}^n b_j^2 \right) \right]^{1/2}} \tag{10}$$

The quantity  $\sum_{j=1}^n a_j b_j$ , denoted by  $(\mathbf{a}, \mathbf{b})$ , is called the *inner product* (or scalar product)

of the two vectors. The quantity  $\sum_{j=1}^n a_j^2 = (\mathbf{a}, \mathbf{a})$  is called the *norm* of  $\mathbf{a}$ . The vectors

$\mathbf{a}$  and  $\mathbf{b}$  are *orthogonal* (or perpendicular) if  $(\mathbf{a}, \mathbf{b}) = 0$ . A set of vectors  $\{a_j\}$  is called *orthonormal*, that is, orthogonal and normalized, if  $(\mathbf{a}_j, \mathbf{a}_k) = \delta_{jk}$ , where  $\delta_{jk}$  is Kronecker's  $\delta$ , which is 0 or 1 according to whether  $j = k$  or  $j \neq k$ . An arbitrary vector  $\mathbf{g}$  in  $n$ -dimensional space can be expressed as a linear combination of a set of  $n$  orthonormal vectors  $\{\mathbf{a}_j\}$ ; that is,

$$\mathbf{g} = \sum_{j=1}^n (\mathbf{g}, \mathbf{a}_j) \mathbf{a}_j = \sum_{j=1}^n c_j \mathbf{a}_j \tag{11}$$

Most of the notations used for  $n$ -dimensional space can be extended not only to countably infinite spaces but also to certain spaces of functions. An example of a countably infinite space is the *special Hilbert space* consisting of all points

$$\mathbf{a} = (a_1, a_2, \dots, a_j, \dots)$$

such that  $\sum_{j=1}^{\infty} a_j^2 = (\mathbf{a}, \mathbf{a}) < \infty$ .

In dealing with a function  $f(x)$  it is customary to speak of  $f(x)$  as a point or a vector in the function space. Let the real functions  $f(x)$  and  $g(x)$  be defined on the finite interval  $a \leq x \leq b$ ; then the *inner product* of the two functions is given by

$$\int_a^b f(x)g(x) dx = (f, g) \tag{12}$$

If the inner product is zero, that is,  $(f, g) = 0$ , the two functions are said to be *orthogonal*. The *norm* of the function  $f(x)$  is given by

$$\int_a^b [f(x)]^2 dx = (f, f) \tag{13}$$

A set of functions  $\{f_j(x)\}$  is said to be orthonormal on the interval  $a \leq x \leq b$  if

$$\int_a^b f_j(x)f_k(x) dx = (f_j, f_k) = \delta_{jk} \quad (14)$$

The concept of orthogonality for functions of a single variable can be generalized in two particular important ways. If  $f(x)$  and  $g(x)$  are complex functions of a real variable  $x$  on the interval  $a \leq x \leq b$ , then  $f$  and  $g$  are said to be *orthogonal in the Hermitian sense* if

$$\int_a^b f(x)\overline{g(x)} dx = (f, g) = 0 \quad (15)$$

where  $\overline{g(x)}$  is the conjugate of  $g(x)$ . Often there is associated with a set of real functions  $f_j(x)$  a function  $w(x) \geq 0$ , on the interval  $a \leq x \leq b$ , such that

$$\int_a^b w(x)f_j(x)f_k(x) dx = \delta_{ij} \quad (16)$$

In this case the set is called *orthonormal on  $[a, b]$  with respect to the weight function  $w(x)$* . The orthogonality concept can also be extended to infinite intervals and to functions of more than one variable.

The functions  $f_1(x), f_2(x), \dots, f_n(x)$  are called *linearly independent* on the interval  $a \leq x \leq b$  if the only constants  $c_j$  that satisfy the equation  $\sum_{j=1}^n c_j f_j(x) = 0$  for all  $x$

of the interval are the constants  $c_1 = c_2 = \dots = c_n = 0$ . A set of orthogonal functions is likewise a set of linearly independent functions. An orthogonal set  $\{f_j(x)\}$  is called *complete* in a given function space if the only element of the space that is orthogonal to every  $f_j(x)$  is the zero function; that is,  $(f, f_j) = 0$ , for all  $j$ , implies  $f = 0$ .

Let  $\{\varphi_j(x)\}$ , ( $j = 1, 2, \dots$ ), be a countable orthonormal set of real functions on the interval  $a \leq x \leq b$ , and consider the possibility of representing a function  $f(x)$  on that interval by

$$f(x) = \sum_{j=1}^{\infty} c_j \varphi_j(x)$$

Formally multiply both sides of this equation by  $\varphi_k(x)$ , and integrate both sides over the interval, the integration being carried out term by term on the right-hand side. This leads to the formal result

$$(f, \varphi_k) = \sum_{j=1}^{\infty} c_j (\varphi_j, \varphi_k) = c_k$$

The numbers

$$c_j = (f, \varphi_j) = \int_a^b f(x)\varphi_j(x) dx \quad j = 1, 2, \dots \quad (17)$$

are called the *Fourier constants or coefficients* of  $f(x)$  corresponding to the orthonormal system  $\{\varphi_j(x)\}$ . The series

$$\sum_{j=1}^{\infty} c_j \varphi_j(x) \quad \text{where } c_j = (f, \varphi_j) \quad (18)$$

is called a *generalized Fourier series or expansion* corresponding to  $f(x)$ .

Given a finite orthonormal set  $\{\varphi_j(x)\}$  ( $j = 1, 2, \dots, n$ ), the best approximation in the mean, or in the sense of least squares, to the function  $f(x)$  on the interval  $a \leq x \leq b$  for the quantity  $\sum_{j=1}^n a_j \varphi_j(x)$  is given by letting  $a_j$  be the Fourier coefficients

$$c_j = (f, \varphi_j) = \int_a^b f(x) \varphi_j(x) dx; \text{ that is,}$$

$$\int_a^b \left[ f(x) - \sum_{j=1}^n a_j \varphi_j(x) \right]^2 dx$$

is minimized when  $a_j = c_j$ .

The relation

$$\sum_{j=1}^n (f, \varphi_j)^2 = \sum_{j=1}^n c_j^2 \leq \int_a^b [f(x)]^2 dx = (f, f) \tag{19}$$

is true for any  $n$  and is known as *Bessel's inequality*. If  $\{\varphi_j(x)\}$  is a complete orthonormal set, then *Parseval's theorem*

$$\sum_{j=1}^{\infty} (f, \varphi_j)^2 = \sum_{j=1}^{\infty} c_j^2 = \int_a^b [f(x)]^2 dx \tag{20}$$

holds.

**3.42 Sturm-Liouville Series.** Systems of the following types:

$$\begin{aligned} [ry'(x)]' + (\lambda p + g)y &= 0 \\ \alpha_1 y(a) + \alpha_2 y'(a) + \alpha_3 y(b) + \alpha_4 y'(b) &= 0 \\ \beta_1 y(a) + \beta_2 y'(a) + \beta_3 y(b) + \beta_4 y'(b) &= 0 \end{aligned} \tag{21}$$

where the prime (') denotes differentiation with respect to  $x$ , are known as Sturm-Liouville systems. The coefficients  $r(x) > 0$ ,  $g(x)$ , and  $p(x)$  are taken to be continuous functions of  $x$  in the interval  $a \leq x \leq b$ , and  $\lambda$  is an arbitrary parameter. The great interest in S-L (Sturm-Liouville) theory arises largely from the fact that many boundary-value problems in physics and engineering lead to questions that can be answered by using S-L theory.

Consider the general second-order linear differential equation

$$f_2(x)y'' + f_1(x)y' + [f_0(x) + \lambda g_0(x)]y = 0 \tag{22}$$

where  $\lambda$  is a parameter and where  $f_2 > 0$ ,  $f_1$ ,  $f_0$ , and  $g_0$  are continuous functions of  $x$ . The general homogeneous boundary conditions

$$\begin{aligned} \alpha_1 y(a) + \alpha_2 y'(a) + \alpha_3 y(b) + \alpha_4 y'(b) &= 0 \\ \beta_1 y(a) + \beta_2 y'(a) + \beta_3 y(b) + \beta_4 y'(b) &= 0 \end{aligned}$$

are taken to be linearly independent. Upon dividing the general differential equation by  $f_2(x)$  and then multiplying the result by  $r = \exp \int (f_1/f_2) dx$ , the S-L equation

$$[ry'(x)]' + [\lambda p + g]y = 0 \tag{23}$$

where  $p = (g_0/f_2)r$  and  $g = (f_0/f_2)r$ , is obtained. It follows that the S-L system is quite general.

The differential equation of an S-L is self-adjoint. An S-L system is called self-adjoint when

$$(\alpha_1 \beta_2 - \beta_1 \alpha_2)r(b) = (\alpha_4 \beta_3 - \alpha_3 \beta_4)r(a)$$

The most common type of self-adjoint S-L system has two additional conditions. First  $p(x)$  is assumed not to vanish for  $a \leq x \leq b$ . In this case both  $r(x)$  and  $p(x)$

can be assumed greater than zero. The second condition is that the boundary conditions are of *Sturmian type*; that is,

$$\text{where } \begin{array}{l} \alpha_1 y(a) + \alpha_2 y'(a) = 0 \\ |\alpha_1| + |\alpha_2| > 0 \end{array} \quad \text{and} \quad \begin{array}{l} \beta_1 y(b) + \beta_2 y'(b) = 0 \\ |\beta_1| + |\beta_2| > 0 \end{array} \quad (24)$$

The following three theorems summarize the primary results of S-L theory:  
*Oscillation Theorem.* The system of the Sturm-Liouville differential equation with Sturmian boundary conditions and positive  $p(x)$  and  $r(x)$  has an infinite number of real characteristic numbers that may be arranged in a monotone increasing sequence  $\lambda_0 < \lambda_1 < \lambda_2 < \dots$  that tends to  $+\infty$ . Corresponding to each simple eigenvalue  $\lambda_n$ , there exists an eigenfunction  $\varphi_n(x)$ , unique except for a multiplicative constant. Each  $\varphi_n(x)$  has exactly  $n$  zeros in the interval  $a < x < b$ .

*Expansion Theorem.* Given an arbitrary continuous and piece-wise differentiable function  $f(x)$  that vanishes at the end points of the interval when  $\varphi_0(x)$  vanishes; then the series

$$\sum_{n=0}^{\infty} c_n \varphi_n(x) \quad \text{where } c_n = \int_a^b p(x) f(x) \varphi_n(x) dx = (f, \varphi_n) \quad (25)$$

converges uniformly and absolutely and has the sum  $f(x)$ .

The eigenfunctions fulfill the orthonormality conditions

$$(\varphi_i, \varphi_k) = \int_a^b p(x) \varphi_i(x) \varphi_k(x) dx = \delta_{ik} \quad (26)$$

where  $p(x)$  is the weight function. The normality can easily be attained by letting

$$\tilde{\varphi}_i(x) = \varphi_i (\varphi_i, \varphi_i)^{-1/2} = \varphi_i \left[ \int_a^b p(x) \varphi_i^2(x) dx \right]^{-1/2}$$

*Equi-convergence Theorem.* Given that  $f(x)$  is integrable over the interval  $(a, b)$ ; then the S-L expansion behaves as regards convergence in the same way as an ordinary Fourier series.

The transformation

$$\text{where } \begin{array}{l} Y(z) = (pr)^{1/4} y \\ z = \frac{1}{J} \int_a^x \left( \frac{p}{r} \right)^{1/2} dx \quad \text{and} \quad J = \frac{1}{\pi} \int_a^b \left( \frac{p}{r} \right)^{1/2} dx \end{array}$$

takes the S-L equation into the *Liouville normal form*

$$\frac{d^2 Y}{dz^2} + [k^2 - q(z)] Y(z) = 0 \quad (27)$$

where  $k^2 = J^2 \lambda$  and  $q(z) = (pr)^{-1/4} (d^2/dz^2)(pr)^{1/4} - J^2(q/p)$ .

If the interval under consideration is infinite, or if  $r$  vanishes at either or both ends of the interval, the S-L system is called *singular*. Results for singular S-L systems may be found in the references. In certain problems, such as those occurring in reactor theory, it is necessary to consider generalized S-L systems that have discontinuous coefficients or solutions. This arises because the problem is multiregion and has interface conditions. The results of ordinary S-L theory can often be extended to these cases.

*The trigonometric series*

$$\frac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \quad (28)$$

is called a *Fourier series* corresponding to  $f(x)$  in the interval  $(-\pi, \pi)$  if its coefficients are given by

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, dt \quad \text{and} \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, dt$$

When  $f(x)$  is an *odd function*, that is,  $f(-x) = -f(x)$ ,  $a_n = 0$  and the corresponding series is known as the *Fourier sine series*. If  $f(x)$  is an *even function*, that is

$$f(-x) = f(x)$$

$b_n = 0$  and the corresponding series is called a *Fourier cosine series*. More generally the Fourier series corresponding to  $f(x)$  in the interval  $a \leq x \leq b$  is

$$\frac{1}{b-a} \int_a^b f(t) \, dt + \frac{2}{b-a} \sum_{n=1}^{\infty} \int_a^b f(t) \cos \frac{2n\pi(t-x)}{b-a} \, dt \quad (29)$$

The exponential form for a Fourier series is

$$\sum_{k=-\infty}^{\infty} c_k \exp\left(\frac{2k\pi ix}{b-a}\right) \quad \text{where} \quad c_k = \frac{1}{b-a} \int_a^b f(t) \exp\left(-\frac{2k\pi it}{b-a}\right) \, dt \quad (30)$$

A number of theorems giving conditions under which a Fourier series corresponding to a function converges in some manner to that function are given in the references. The following is an example of a Fourier theorem.

*Fourier Theorem.* If  $f(x)$  is sectionally continuous on the interval  $(-\pi, \pi)$  and periodic with period  $2\pi$ , then the Fourier series corresponding to  $f(x)$  in the interval  $(-\pi, \pi)$  converges to the value  $\frac{1}{2}[f(x^+) + f(x^-)]$  at every point where the right-hand and left-hand derivatives exist.

The set  $\{1/\sqrt{2\pi}, \cos nx/\sqrt{\pi}, \sin nx/\sqrt{\pi}\}$  forms a complete orthonormal set for the space of all functions that are sectionally continuous, are assigned the value  $\frac{1}{2}[f(x^+) + f(x^-)]$  at the points of discontinuity and the value  $\frac{1}{2}[f(-\pi^+) + f(\pi^-)]$  at the end of the intervals, and possess right-hand and left-hand derivatives at every point of the interval  $(-\pi, \pi)$ .

Fourier series not only are special cases of S-L series but indeed are the simplest cases of S-L series, being associated with the differential equation  $y'' + \lambda y = 0$ .

*Legendre Series.* The Legendre expansion or series for a function  $f(x)$  in the interval  $(-1, 1)$  is given by

$$\sum_{n=0}^{\infty} a_n P_n(x) \quad \text{where} \quad a_n = \frac{2n+1}{2} \int_{-1}^1 f(t) P_n(t) \, dt \quad (31)$$

and where  $P_n(x)$  is the Legendre polynomial of degree  $n$ . The Legendre series is a special case of an S-L series. Legendre's equation

$$\frac{d}{dx} \left[ (1-x^2) \frac{dy}{dx} \right] + n(n+1)y(x) = 0 \quad (32)$$

is seen to be an S-L equation where  $r(x) = 1-x^2$ ,  $p(x) = 1$ ,  $g(x) = 0$ , and

$$\lambda = n(n+1)$$

The set  $\{(n + \frac{1}{2})^{1/2} P_n(x), (n = 0, 1, 2, \dots)\}$  forms an orthonormal set; i.e.,

$$\int_{-1}^1 P_n(t) P_m(t) \, dt = (n + \frac{1}{2})^{-1} \delta_{mn}$$

*Fourier-Bessel Expansion.* The Bessel function  $J_n(\lambda x)$  satisfies the Bessel equation

$$\frac{d}{dx} \left( x \frac{dy}{dx} \right) + \left( \lambda x - \frac{n^2}{x} \right) y = 0 \quad (33)$$

For each  $n$  this is seen to be an S-L equation where  $r = x$ ,  $p = x$ , and  $g = -n^2/x$ . Let  $\lambda_j$  ( $j = 1, 2, \dots$ ) be the roots of the equation  $J_n(\lambda a) = 0$  and let

$$N_{nj} = \int_0^a t [J_n(\lambda_j t)]^2 dt = \frac{a^2}{2} [J_{n+1}(\lambda_j a)]^2$$

The set of functions

$$\left\{ \varphi_{nj} = \frac{J_n(\lambda_j x)}{(N_{nj})^{1/2}} \right\}$$

forms an orthonormal set in the interval  $(0, a)$  with respect to the weight function  $x$ ; that is,

$$\int_0^a t \varphi_{nj}(t) \varphi_{nk}(t) dt = \delta_{ij}$$

A Bessel expansion or series for a function  $f(x)$  in the interval  $(0, a)$  is given by

$$\sum_{j=1}^{\infty} c_{nj} \varphi_{nj} = \sum_{j=1}^{\infty} a_j J_n(\lambda_j x) \quad (34)$$

where  $c_{nj} = \int_0^a t \varphi_{nj}(t) f(t) dt$  and  $a_j = \frac{1}{N_{nj}} \int_0^a t J_n(\lambda_j t) f(t) dt$

## 4 DIFFERENTIAL EQUATIONS

### 4.1 Introduction

A *differential equation* is an equation that involves a differential or a derivative. The equation may contain algebraic and transcendental functions of a differential or a derivative. It is assumed that a differential equation is not an identity.

It is difficult to make a simple yet complete classification of differential equations because not only is the number of classes and subclasses needed immense but also any particular differential equation is likely to appear in many classes. Nevertheless certain broad classifications are commonly adhered to and have proved useful. The primary division of differential equations results from the number of independent variables that are present. An *ordinary* differential equation is an equation containing one independent variable, one or more dependent variables, and at least one derivative of a dependent variable with respect to the independent variable. By contrast, a *partial* differential equation contains two or more independent variables, one or more dependent variables, and at least one partial derivative of a dependent variable with respect to an independent variable.

A differential equation is also categorized by giving its order and degree. The *order* of an equation is the order of the highest derivative found in the equation. (The word differential is omitted in front of equation if no confusion results.) The *degree* of an equation is the power to which the highest derivative is raised. To this definition should be added the requirement that in stating the degree of an equation it is implied that the equation is a polynomial in all the derivatives.

If the dependent variables and their derivative occur only to the first power or degree, and not as products, an equation is called *linear*. For a linear equation, the coefficients of the dependent variables and their derivatives are therefore functions only of the independent variables. An equation that is not linear is called *nonlinear*.

By a *solution* of a differential equation is meant a set of functions of the independent

variables, one function identified with each dependent variable, that when substituted in the equation results in an identity in the independent variables. It is implied that the solution functions have at least as high derivatives as are needed in the differential equation.

The *general* solution of an ordinary differential equation of  $n$ th order and one dependent variable contains  $n$ , and only  $n$ , arbitrary constants. A *particular* solution of an ordinary differential equation of  $n$ th order and one dependent variable is obtained when the  $n$  arbitrary constants are given special values. Not all solutions of an ordinary differential equation need to be particular cases of the general solution. A solution that is not such a particular case is generally called a *singular* solution.

The *general* solution of a partial differential equation of  $n$ th order,  $m$  independent variables, and one dependent variable is a solution containing  $n$  arbitrary functions of  $m - 1$  variables. The question of the existence and the uniqueness of the solutions is not considered here, and the references should be consulted.

## 4.2 Ordinary Differential Equations

**4.21 Differential Equations of the First Order.** The general differential equation of the first order has the form

$$F(y', y, x) = 0$$

where the prime denotes differentiation with respect to  $x$ . This is the implicit form for the first-order equation. The equation

$$p = y' = f(y, x)$$

is the explicit form for the first-order equation. There exist a number of special types of implicit and explicit equations whose solutions can be obtained by elementary methods. The following list presents the more common types and their solutions.

*Variables Separable:*

$$p = f(x)g(y) \quad \int_a^y \frac{dy}{g(y)} = \int_a^x f(x) dx \quad (35)$$

*Homogeneous Equation:*

$$p = f\left(\frac{y}{x}\right) \quad y = xv \quad \log \frac{x}{a} = \int_{\alpha/a}^{y/x} \frac{dv}{v - f(v)} \quad (36)$$

1. Now the equation  $P(x, y) dx + Q(x, y) dy = 0$ , where  $P(tx, ty) = t^n P(x, y)$  and  $Q(tx, ty) = t^n Q(x, y)$ , has this homogeneous form.

2. An equation  $p = f[(ax + by + c)/(\alpha x + \beta y + \gamma)]$  may be brought into the above homogeneous form, when  $a\beta - \alpha b \neq 0$ , by the linear transformation  $x = u + d$ ,  $y = v + e$ , where  $u$  and  $v$  are new variables and  $d$  and  $e$  are constants. If  $a\beta - \alpha b = 0$ , the equation can be put in the variables separable form by either substitution

$$v = ax + by + c$$

or substitution  $v = \alpha x + \beta y + \gamma$ .

*Linear Equation:*

$$p + f(x)y = g(x) \quad y = e^{-h(x)} \left( \alpha - \int_a^x e^{hg} dx \right) \quad h = \int_a^x f dx \quad (37)$$

*Bernoulli's Equation:*

$$p + f(x)y = g(x)y^n \quad (38)$$

becomes the linear equation  $p + (1 - n)f(x)y = (1 - n)g(x)$  under the substitution  $v = y^{1-n}$ .

*Dependent or Independent Variable Missing:*

1.  $p = f(y)$  and  $p = g(x)$  are equations of the variables separable type.
2.  $y = f(p)$ ;  $x - c = \int [f'(p)/p] dp$ .
3.  $x = f(p)$ ;  $y - c = \int pf'(p) dp$ .

*Clairaut Equation:*

$$y = px + f(p) \quad \text{either } y = cx + f(c) \quad \text{or} \quad x + f'(p) = 0 \quad (39)$$

a singular solution.

*Exact Equation:*

$$P(x,y) dx + Q(x,y) dy = 0 \text{ is an exact equation if } \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \quad (40)$$

$$\int P dx + \int \left( Q - \frac{\partial}{\partial y} \int P dx \right) dy = c$$

is the solution.

*Integrating Factor:*

A function  $H(x,y)$  is called an integrating factor or Euler multiplier if

$$HP(x,y) dx + HQ(x,y) dy = 0$$

is an exact equation.

There exist certain other well-known first-order equations whose solution cannot be obtained by elementary methods. The following types may be listed, and for further information concerning them the references should be consulted.

*Generalized Riccati Equation:*

$$\frac{dy}{dx} = f(x)y^2 + g(x)y + h(x)$$

*Special Riccati Equation:*

$$\frac{dy}{dx} + ay^2 = bx^n$$

*Abel's Equation of the First Kind:*

$$\frac{dy}{dx} = \sum_{i=0}^3 f_i(x)y^i$$

*Abel's Equations of the Second Kind:*

$$[g_1(x)y + g_0(x)]y = \sum_{i=0}^3 f_i(x)y^i$$

*Binomial Equation:*

$$\left( \frac{dy}{dx} \right)^m = f(x,y)$$

**4.22 Linear Differential Equations.** The general linear differential equation of the  $n$ th order has the form

$$\sum_{i=0}^n f_{n-i}(x) \frac{d^i y}{dx^i} = g(x) \quad (41)$$

It is generally assumed that the coefficients  $f_i(x)$  and  $g(x)$  either are continuous and one-valued or are meromorphic functions of  $x$  throughout some region. Furthermore,

it is usually assumed that  $f_0(x)$  has at most isolated zeros in this region. If  $g(x) \equiv 0$ , Eq. (41) is called homogeneous.

The expression

$$L \equiv \sum_{i=0}^n f_{(n-i)} \frac{d^i}{dx^i} \equiv f_0 \frac{d^n}{dx^n} + f_1 \frac{d^{n-1}}{dx^{n-1}} + \cdots + f_{n-1} \frac{d}{dx} + f_n$$

is called the linear differential operator of order  $n$ . The following properties are common to all  $n$ th-order linear differential equations:

1. If  $y = y_1$  is a solution of the homogeneous equation  $L(y) = 0$ , then  $y = cy_1$  is also a solution of the equation, where  $c$  is any arbitrary constant.
2. If  $y_1, y_2, \dots, y_p$  are  $p$  solutions of  $L(y) = 0$ , then

$$y = c_1y_1 + c_2y_2 + \cdots + c_p y_p$$

is also a solution of the equation, where  $c_1, c_2, \dots, c_p$  are arbitrary constants.

3. If  $y_0$  is any solution of  $L(y) = g$  and  $y_1$  is any solution of  $L(y) = 0$ , then

$$y = y_0 + y_1$$

is a solution of  $L(y) = g$ .

4. The *complete primitive* of  $L(y) = 0$  has the form

$$Y(x) = c_1y_1 + c_2y_2 + \cdots + c_n y_n$$

where  $c_1, c_2, \dots, c_n$  are  $n$  arbitrary constants and  $y_1, y_2, \dots, y_n$  are  $n$  linearly independent functions that are solutions of  $L(y) = 0$ . The general solution of  $L(y) = g(x)$  has the form  $y = y_0 + Y$ , where  $y_0$  is a *particular* integral, that is, any solution of  $L(y) = g(x)$  containing no arbitrary constant, and where  $Y$  is the *complementary function* or complete primitive of  $L(y) = 0$ .

Certain other concepts are useful in the discussion of linear differential equations. The determinant

$$\Delta(y_1, \dots, y_n) \equiv \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y'_1 & y'_2 & \cdots & y'_n \\ \cdots & \cdots & \cdots & \cdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{vmatrix}$$

is known as the *Wronskian* of the functions  $y_1, y_2, \dots, y_n$ . The vanishing of the Wronskian of  $y_1, \dots, y_n$  at any point of the region being considered means that the functions are linearly dependent. Conversely if the Wronskian does not vanish, the functions are linearly independent. A set of  $n$  linearly independent solutions of the  $n$ th-order equation  $L(y) = 0$  is called a *fundamental set* or *fundamental system*.

The *adjoint equation* to the equation

$$L(y) \equiv \sum_{i=0}^n f_{(n-i)} \frac{d^i y}{dx^i} = 0$$

is the equation

$$L^*(v) \equiv \sum_{i=0}^n (-1)^i \frac{d^i [f_{(n-i)} v]}{dx^i} = 0$$

The operator  $L^*$  is called the formal adjoint operator to the operator  $L$ . If  $L \equiv L^*$ , then the operator  $L$  and the equation  $L(y) = 0$  are called *self-adjoint*. The relation

$$vL(y) - yL^*(v) = \frac{d}{dx} [P(y, v)] \tag{42}$$

where  $P$  is a linear function of  $y$  and  $v$  and their first  $n - 1$  derivatives, can be verified easily, and is known as the *Lagrange identity*.

The equation

$$A(y) \equiv \sum_{i=0}^n a_{n-i} \frac{d^i y}{dx^i} = 0 \quad (43)$$

where the  $a_{n-i}$  are constants, is a *homogeneous linear equation with constant coefficients*. This special linear equation not only is one of the most important and common in applications but historically is the first equation of a general type to be completely solved. The substitution  $y = e^{mx}$  in this equation leads to the characteristic or auxiliary equation

$$\sum_{i=0}^n a_{n-i} m^i = 0$$

If this auxiliary equation has  $n$  distinct roots,  $r_1, r_2, \dots, r_n$ , then the general solution of  $A(y) = 0$  is

$$y = c_1 e^{r_1 x} + c_2 e^{r_2 x} + \dots + c_n e^{r_n x}$$

where the  $c$ s are arbitrary constants. If  $m$  of the roots of the auxiliary equation are equal to some value  $s$ , then the portion of the general solution corresponding to  $s$  is

$$(c_1 + c_2 x + c_3 x^2 + \dots + c_m x^{m-1}) e^{sx}$$

The equation

$$\sum_{i=0}^n a_{n-i} x^i \frac{d^i y}{dx^i} = f(x) \quad (44)$$

where the  $a_{n-i}$  are constants, is known as *Euler's differential equation*. The substitution  $x = e^z$  transforms Euler's equation into a linear equation with constant coefficients. This follows from the relation

$$x^i \frac{d^i y}{dx^i} = \frac{d}{dz} \left( \frac{d}{dz} - 1 \right) \dots \left( \frac{d}{dz} - i + 1 \right) y$$

**4.23 Linear Second-order Differential Equations.** The second-order linear equation is a particularly important and common equation in physics and engineering. The general equation of this type has the form

$$f(x)y'' + g(x)y' + h(x)y = r(x) \quad (45)$$

The differential expression

$$L(y) \equiv fy'' + gy' + hy$$

has the adjoint expression

$$L^*(v) \equiv fv'' + (2f' - g)v' + (f'' - g' + h)v$$

if  $f' = g$ , then  $L(w) = L^*(w)$  and  $L(w)$  is self-adjoint. Although  $L(y)$  is not self-adjoint, the expression

$$\frac{1}{f} e^{F(x)} L(y) \quad \text{where } F(x) = \int_a^x \frac{g}{f} dx$$

is self-adjoint. Since any second-order linear homogeneous equation can therefore be made self-adjoint by the multiplication of a suitable factor, there is no loss of generality in considering only the self-adjoint case.

The *normal form* for the homogeneous equation  $L(y) = 0$  is  $u'' + Iu = 0$ , where  $I(x) = (h/f) - \frac{1}{2}(g/f)^2 - \frac{1}{2}(g/f)'$ . The left-hand side of the normal equation is

obtained by letting  $y = u \exp\left(-\frac{1}{2} \int_a^x \frac{g}{f} dx\right)$  in the expression  $L(y)$ .  $I(x)$  is called the invariant function of the equation.

The method of variation of constants (or parameters) and the solution of equations by the use of series are both applicable to  $n$ th-order linear differential equations. For illustrative purposes these techniques are considered for the second-order case.

If the complete primitive  $c_1y_1(x) + c_2y_2(x)$ , that is, the general solution of the equation  $y'' + gy' + hy = 0$ , is known, then the solution of the equation

$$y'' + gy' + hy = r$$

can be obtained by the method of variation of constants. Assume a solution of the nonhomogeneous equation of the form  $y = c_1(x)y_1(x) + c_2(x)y_2(x)$ , where  $c_1$  and  $c_2$  are functions of  $x$  that are to be determined. If  $c'_1y_1 + c'_2y_2$  is set equal to zero, then  $y' = c_1y'_1 + c_2y'_2$ . Furthermore,  $y'' = c_1y''_1 + c_2y''_2 + c'_1y'_1 + c'_2y'_2$ , and therefore substitution in the nonhomogeneous equation gives  $c'_1y'_1 + c'_2y'_2 = r$ . This last equation and the equation  $c'_1y_1 + c'_2y_2 = 0$  can be solved for  $c'_1$  and  $c'_2$ . Direct integration then gives

$$c_1(x) = A - \int \frac{r(t)y_2(t) dt}{W} \quad c_2(x) = B + \int \frac{r(t)y_1(t) dt}{W}$$

where  $W = y_1y'_2 - y_2y'_1$  is the Wronskian function and  $A$  and  $B$  are constants. The solution of  $y'' + gy' + hy = r$  is given by

$$y = y_2(x) \int \frac{y_1 r}{W} dt - y_1(x) \int \frac{y_2 r}{W} dt + Ay_1(x) + By_2(x)$$

The method of solving a differential equation by the use of power series consists of assuming a solution of the form  $\sum_{n=0}^{\infty} a_n x^{n+\rho}$ , substituting it formally into the differential equation, collecting terms in like powers, and setting the resulting coefficients of each power equal to zero. The objective of this process is to determine the  $\rho$  and the  $a_n$ 's and thus a solution. An elementary illustration follows. Consider  $y'' + y = 0$ ,

and assume  $y = \sum_{n=0}^{\infty} a_n x^{n+\rho}$ ,  $a_0 \neq 0$ . Direct substitution gives

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

The shifting of indices gives  $\sum_{n=0}^{\infty} a_n x^{n+\rho} = \sum_{n=2}^{\infty} a_{n-2} x^{n+\rho-2}$ . The collecting of like terms results in

$$\rho(\rho - 1)a_0 x^{\rho-2} + (\rho + 1)\rho a_1 x^{\rho-1} + \sum_{n=2}^{\infty} [(n + \rho)(n + \rho - 1)a_n + a_{n-2}]x^{n+\rho-2} = 0$$

The coefficients of the powers are zero when  $\rho(\rho - 1) = 0$ ,  $(\rho + 1)\rho a_1 = 0$ ,

$$(n + \rho)(n + \rho - 1)a_n + a_{n-2} = 0$$

( $n \geq 2$ ). The indicial equation  $\rho(\rho - 1) = 0$  is satisfied if  $\rho = 0$  or  $\rho = 1$ . First let  $\rho = 0$ . In this case  $(\rho + 1)\rho a_1 = 0$  is also satisfied for arbitrary  $a_1$ , for example,  $a_1 = 0$ . The remaining equation becomes

$$a_n = \frac{-a_{n-2}}{nn-1} = \frac{+a_{n-4}}{n(n-1)(n-2)(n-3)} = \dots = (-1)^{n/2} \frac{a_0}{n!}$$

for  $n$  even and  $a_n = 0$  for  $n$  odd. Consequently

$$y = a_0 \sum_{m=0}^{\infty} (-1)^m \frac{x^{2m}}{(2m)!} = a_0 \cos x$$

This solution is easily obtained or verified by using the properties of the  $\cos x$ . If  $\rho = 1$ , it is easy to obtain a series solution, which is  $a_0 \sin x$ .

### 4.3 Partial Differential Equations

**4.31 First-order Equations.** Let  $x_1, x_2, \dots, x_n$  be  $n$  independent variables, let  $v = v(x_1, x_2, \dots, x_n)$  be the dependent variables, and let

$$p_1 = \frac{\partial v}{\partial x_1} \quad p_2 = \frac{\partial v}{\partial x_2} \quad \dots \quad p_n = \frac{\partial v}{\partial x_n}$$

The general partial differential equation of the first order has the form

$$F(p_1, p_2, \dots, p_n, v, x_1, \dots, x_n) = 0 \quad (46)$$

When  $n = 2$ , it is customary to put  $x_1 = x, x_2 = y, p = \partial v / \partial x$ , and  $q = \partial v / \partial y$ , so that the equation has the form

$$F(p, q, v, x, y) = 0 \quad (47)$$

First-order equations are called "linear" if the equation is of the first degree in the partial derivatives. Furthermore, they are called nonlinear if at least one partial derivative is present to some degree other than one.

The linear equation for two independent variables has the form

$$P(x, y, v)p + Q(x, y, v)q = R(x, y, v) \quad (48)$$

and is called *Lagrange's linear equation*. The simultaneous ordinary differential equations

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{dv}{R} \quad (49)$$

are called the *subsidiary equations for Lagrange's equation*. If  $f(x, y, v) = c_1$  and  $g(x, y, v) = c_2$ , where  $c_1$  and  $c_2$  are arbitrary constants, are two independent solutions of the subsidiary equations, then any arbitrary functional relation  $\varphi(f, g) = 0$  satisfies Lagrange's equation.  $\varphi(f, g) = 0$  is called the *general solution or integral of the equation*. For example, the equation  $ap + bq = 1$  has the subsidiary equation  $dx/a = dy/b = dv/1$ . The subsidiary equation has the two integrals

$$x - av \equiv f = c_1 \quad \text{and} \quad y - bv \equiv g = c_2$$

and the general integral is

$$\varphi(x - av, y - bv) = 0$$

Subsidiary equations and the general integral are similarly obtained for the linear equation with  $n$  independent variables.

Consider the nonlinear equation  $F(p, q, v, x, y) = 0$ . A *complete integral* of this equation is any solution that contains two arbitrary constants or parameters  $\alpha$  and  $\beta$ . The complete integral may be denoted by  $f(x, y, v, \alpha, \beta) = 0$  and may be interpreted geometrically as a two-parameter family of surfaces. A *particular integral* is obtained

from a complete integral by assigning  $\alpha$  and  $\beta$  particular or definite values. If  $\alpha$  and  $\beta$  are eliminated from the three relations

$$f = 0 \quad \frac{\partial f}{\partial \alpha} = 0 \quad \frac{\partial f}{\partial \beta} = 0$$

then the eliminant  $\varphi(x, y, v) = 0$ , if it satisfies the original equation, is called a *singular integral* of the equation. A singular integral, if one exists, may also be found by eliminating  $p$  and  $q$  from the relations

$$F = 0 \quad \frac{\partial F}{\partial p} = 0 \quad \frac{\partial F}{\partial q} = 0$$

Next let  $\beta = \varphi(\alpha)$  so that  $f[x, y, v, \alpha, \varphi(\alpha)] = 0$  represents a one-parameter family of surfaces. The totality of solutions of the equation  $F = 0$ , which are derived from the equations

$$f[x, y, v, \alpha, \varphi(\alpha)] = 0 \quad \frac{\partial f}{\partial \alpha} = 0$$

upon eliminating  $\alpha$  for all possible choices of  $\varphi$ , is known as the *general integral*. For each choice of  $\varphi$  and  $\alpha$  the equations  $f = 0$  and  $\partial f / \partial \alpha = 0$  represent a space curve called a *characteristic curve*.

*Charpit's method* may be used for solving nonlinear equations of first order. This method is based upon determining a second equation of the type  $G(p, q, v, x, y) = 0$  that, along with  $F = 0$ , can be solved for  $p$  and  $q$  in terms of  $x, y$ , and  $v$ . The  $p$  and  $q$  thus obtained can be inserted in

$$p \, dx + q \, dy = dv$$

and an integrable expression for  $dv$  may result. It can be shown that any solution  $w(p, q, v, x, y) = \alpha$  containing  $p$  or  $q$  or both, of the system

$$\frac{dp}{F_x + pF_v} = \frac{dq}{F_y + qF_v} = \frac{dv}{-pF_p - qF_q} = \frac{dx}{-F_p} = \frac{dy}{-F_q} \tag{50}$$

where  $F_x$  signifies  $\partial F / \partial x$ , can be used for  $G$ ; that is,  $G \equiv w - \alpha = 0$ . This method can be generalized to problems involving more than two independent variables.

**4.32 Second-order Equations.** The general second-order equation has the form

$$F(v_{x_1x_1}, v_{x_1x_2}, \dots, v_{x_ix_j}, \dots, v_{x_nx_n}, v_{x_1}, \dots, v_{x_n}, v, x_1, \dots, x_n) = 0 \tag{51}$$

where  $v_{x_ix_j}$  signifies  $\partial^2 v / (\partial x_i \partial x_j)$ . Almost all second-order equations of interest are in the class of quasilinear equations. The equation

$$\sum_{i,j=1}^n A_{ij} v_{x_ix_j} = B \tag{52}$$

where the  $A_{ij}$  and  $B$  are functions of the  $x_1, \dots, x_n, v, v_{x_1}, \dots, v_{x_n}$ , is called *quasilinear*. Although most of the following concepts apply to quasilinear equations, it is easier to consider the linear equation

$$\sum_{i,j=1}^n A_{ij} v_{x_ix_j} + \sum_{i=1}^n B_i v_{x_i} + cv + D = 0 \tag{53}$$

where the  $A_{ij}, B_i, C$ , and  $D$  are functions of  $x_1, x_2, \dots, x_n$ .

In order to classify second-order linear equations it is customary to deal with the coordinate transformations

$$z_k = \sum_{i=1}^n a_{ki} x_i$$

where the matrix  $a = (a_{ki})$  is nonsingular with real elements. It is possible to find such transformations so that the above linear equation takes the canonical form

$$\sum_{i,j=1}^n A_{ij} v_{x_i x_j} + \dots = 0 \quad (54)$$

where  $A_{ij}^* = \pm 1$  if  $i = j \leq m \leq n$  and  $A_{ij}^* = 0$  if  $i \neq j$  or if  $i = j > m$  at a given point  $(x_1^0, \dots, x_n^0)$ .

With the use of this canonical form it is possible to classify a linear equation at a given point  $(x_1^0, \dots, x_n^0)$  as follows: If at the point

1. All the  $A_{ii}^* \neq 0$  and have the same sign, then the equation is called *elliptic at the point*.
2. All the  $A_{ii}^* \neq 0$  and if all but one  $A_{ii}^*$  have the same sign, the equation is called *hyperbolic at the point*.
3. All the  $A_{ii}^* \neq 0$  and if there exist more than one positive  $A_{ii}^*$  and more than one negative  $A_{ii}^*$ , the equation is called *ultrahyperbolic at the point*.
4. Some  $A_{ii}^* = 0$ , the equation is called *parabolic in the broad sense at the point*.
5. Only one  $A_{ii}^* = 0$ , the other  $A_{ii}^*$  all have the same sign, and the coefficient of  $\partial v / \partial z_i$  corresponding to the zero  $A_{ii}^*$  is not zero, then the equation is called *parabolic at the point*.

A linear equation is called elliptic, hyperbolic, etc., in a region if it has that character at every point in the region.

Next consider the equation

$$Ar + 2Bs + Ct + D(x, y, v, p, q) = 0 \quad (55)$$

where  $r = \partial^2 v / \partial x^2$ ,  $s = \partial^2 v / (\partial s \partial t)$ , and  $t = \partial^2 v / \partial t^2$  and where the coefficients  $A$ ,  $B$ , and  $C$  are functions of  $x$  and  $y$  that are twice continuously differentiable. This equation is called elliptic, parabolic, or hyperbolic according to whether the determinant

$\Delta = \begin{vmatrix} AB \\ BC \end{vmatrix} = AC - B^2$  is greater than zero, equal to zero, or less than zero. The associated first-order ordinary differential equation

$$A \left( \frac{dy}{dx} \right)^2 - 2B \frac{dy}{dx} + C = 0$$

has for solutions two one-parameter families of curves in the  $xy$  plane,

$$f_1(x, y) = \alpha \quad \text{and} \quad f_2(x, y) = \beta$$

In the hyperbolic case when  $\Delta < 0$ , these two curves, called *characteristics*, are real. In the elliptic case,  $\Delta > 0$ , the characteristics are complex. Finally for the parabolic case,  $\Delta = 0$ , the curves coincide.

The characteristics can be used to obtain the following normal forms:

*Hyperbolic:*

$$v_{\alpha\beta} = F(\alpha, \beta, v, v_\alpha, v_\beta)$$

or  $v_{\zeta\eta} - v_{\eta\eta} = F(\zeta, \eta, v, v_\zeta, v_\eta) \quad \alpha = \zeta + \eta \quad \beta = \zeta - \eta$

*Elliptic:*

$$v_{\zeta\eta} + v_{\eta\eta} = F(\zeta, \eta, v, v_\zeta, v_\eta) \quad \alpha = \zeta + i\eta \quad \beta = \zeta - i\eta$$

*Parabolic:*

$$v_{\eta\eta} = F(\zeta, \eta, v, v_\zeta, v_\eta) \quad \alpha = \beta = \zeta$$

$\eta$  arbitrary function of  $x$  and  $y$ .

*Cauchy Problem.* The Cauchy problem consists of finding a solution of the system

$$\frac{\partial^n u_i}{\partial t^n} = G_i \left( t, x_1, \dots, x_N, v_1, \dots, v_M, \dots, \frac{\partial^k v_j}{\partial t^{k_0} \partial x_1^{k_1} \dots \partial x_n^{k_n}}, \dots \right)$$

( $i, j = 1, 2, \dots, M; k_0 < n; k_0 + k_1 + \dots + k_n \leq n_i$ )

which satisfies the

$$\frac{\partial^k v_i}{\partial t^k} = H_{ik}(x_1, \dots, x_n) \quad (k = 0, 1, \dots, n_i - 1)$$

It should be noted that the number of equations is equal to the number of unknowns; the independent variable  $t$ , time in physical problems, is singled out; if  $n_i$  is the order of the highest derivatives of  $u_i$  that are present, then  $\partial^n u_i / \partial t^n$  must also be present; and the functions  $H_{ik}$  are all given in the same region in the  $(x_1, x_2, \dots, x_n)$  space.

The simplest example of a Cauchy problem is that of finding a solution of

$$\frac{dv}{dt} = F(t, v) \quad \text{where } v(t_0) = v_0$$

For partial differential equations a simple example is given by the vibrating-string equation

$$\frac{\partial^2 v}{\partial t^2} = \frac{\partial^2 v}{\partial x^2}$$

where  $v(t_0, x) = H_1(x)$  is the initial displacement and where  $v_t(t_0, x) = H_2(x)$  is the initial velocity.

The *Cauchy-Kowalewski theorem* states that if all the  $G_i$  are analytic in a neighborhood for all their arguments, and if the  $H_{ik}$  are analytic in the corresponding neighborhood of the point  $(x_1^0, \dots, x_n^0)$ , then the Cauchy problem has a unique analytic solution in the neighborhood of  $(t^0, x_1^0, \dots, x_n^0)$ .

**4.33 Elliptic Equations.** The simplest representative of the class of elliptic equations is Laplace's equation in two dimensions and rectangular coordinates, namely,

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0 \quad \text{or} \quad v_{xx} + v_{yy} = 0$$

This equation often describes steady states, such as the steady temperatures in a homogeneous body or the equilibrium form of a membrane stretched over a curve. A function  $v$  is called *harmonic* or a *potential function* in a region  $D$  if it has continuous derivatives of the first two orders and satisfies  $v_{xx} + v_{yy} = 0$  at every point of  $D$ .

Consider an elliptic partial differential equation  $L(v) = 0$  in a finite region  $D$  with a boundary  $\Gamma$ . The boundary-value problems for elliptic equations are classified as follows:

*Dirichlet's Problem, or the First Boundary-value Problem.* Find a function  $v$  that satisfies  $L(v) = 0$  in the region  $D$  and is equal to a prescribed continuous function  $f$  on the boundary  $\Gamma$ . The term  $v$  equal to a function  $f$  on  $\Gamma$  means that the limit approached by  $v$  as the boundary  $\Gamma$  is approached by points interior to  $D$  is  $f$ .

*Neumann's Problem, or the Second Boundary-value Problem.* Find a function  $v$  that satisfies  $L(v) = 0$  in the region  $D$  and whose outward normal derivative  $\partial v / \partial n$  at every point on the boundary  $\Gamma$  is equal to a prescribed function  $f$ .

*Mixed Problem, or the Third Boundary-value Problem.* Find a function  $v$  that satisfies  $L(v) = 0$  in the region  $D$  and, if  $f$  is a prescribed function on the boundary  $\Gamma$ ,  $av + b(\partial v / \partial n) = f$  on the boundary  $\Gamma$ . If  $b \equiv 0$  or  $a \equiv 0$ , then the mixed problem reduces to Dirichlet's problem or Neumann's problem.

A few of the important results associated with elliptic equations follow:

*The Minimum-Maximum Property.* If  $v$  is a harmonic function in a bounded region  $D$  and is continuous on the boundary  $\Gamma$ , then the values of  $v$  in  $D$  cannot exceed its maximum on  $\Gamma$  or be less than its minimum on  $\Gamma$ .

*Poisson's Integral.* The solution of Dirichlet's problem for the circle of radius  $R$ , where

$$\nabla^2 v \equiv \frac{\partial^2 v}{\partial r^2} + \frac{1}{r} \frac{\partial v}{\partial r} + \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} = 0$$

and  $v = f(\theta)$  when  $r = R$ , is given by

$$v(r, \theta) = \frac{1}{2\pi R} \int_0^{2\pi R} f(s) \frac{R^2 - r^2}{R^2 + r^2 - 2Rr \cos [(s/r) - \theta]} ds \tag{56}$$

*Mean Value Property.* If  $v$  is a harmonic function in the region  $D$ , then  $v$  has the mean value property in  $D$ . The function  $v$  is said to have the mean value property in  $D$  if at every point  $P$  in  $D$   $v(P)$  equals the average of  $v$  either over the circumference or over the area of every circle in  $D$  with center at  $P$ .

*Harnack's Theorem.* If  $v_k(x, y)$  ( $k = 1, 2, \dots$ ) are a sequence of functions harmonic in a finite region  $D$ , and if this sequence converges uniformly in  $D$ , then the limit function  $v(x, y)$  is harmonic in  $D$ .

*Requirement for Neumann's Problem.* In order that there exist a solution of the second boundary-value problem it is necessary that the integral of  $f$  over the boundary  $\Gamma$  vanish; i.e.,

$$\int_{\Gamma} \left( \frac{\partial v}{\partial n} \right) ds \equiv \int_{\Gamma} f(s) ds = 0$$

**4.34 Hyperbolic Equations.** The equation of a vibrating string

$$\frac{\partial^2 v}{\partial t^2} = \frac{\partial^2 v}{\partial x^2}$$

is the simplest example of hyperbolic equations. The Cauchy problem for this equation with the initial conditions

$$v(0, x) = f(x) \quad \text{and} \quad v_t(0, x) = g(x)$$

has the solution, called d'Alembert's formula,

$$v(t, x) = \frac{1}{2} [f(x+t) + f(x-t)] + \frac{1}{2} \int_{x-t}^{x+t} g(s) ds \tag{57}$$

A formal solution of the wave equation

$$v_{tt} = \sum_{i=1}^n v_{x_i x_i}$$

with the initial conditions

$$v(0, x_1, \dots, x_n) = f(x_1, \dots, x_n) \quad \text{and} \quad v_t(0, x_1, \dots, x_n) = g(x_1, \dots, x_n)$$

is given by

$$v = \frac{1}{(n-2)!} \frac{\partial^{n-2}}{\partial t^{n-2}} \int_0^t (t^2 - r^2)^{(n-3)/2} r Q_g(x, r) dr + \frac{1}{(n-2)!} \frac{\partial^{n-1}}{\partial t^{n-1}} \int_0^t (t^2 - r^2)^{(n-3)/2} r Q_f(x, r) dr \tag{58}$$

where

$$Q_h(x, r) = \frac{1}{V_n} \int \dots \int h(x_1 + \beta_1 r, \dots, x_n + \beta_n r) dV_n = \frac{2}{V_n r^{n-1}} \int_{\rho \leq r} \frac{\varphi(x_1 + \alpha_1, \dots, x_n + \alpha_n)}{(r^2 - \rho^2)^{1/2}} d\alpha_1 \dots d\alpha_n$$

where

$$V_n = \frac{2\pi^{n/2}}{\Gamma(n/2)}$$

and  $\rho^2 = \alpha_1^2 + \cdots + \alpha_n^2$  is the mean value of the function  $h$  in the  $n + 1$  dimensional space. In three dimensions ( $n = 3$ ) this formula is called Kirchoff's formula, and in two dimensions the formula is often called the Poisson formula.

Methods for solving hyperbolic equations are discussed in connection with boundary-value problems.

**4.35 Parabolic Equations.** The simplest parabolic equation is the one-dimensional heat-conduction or diffusion equation

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2}$$

This equation with the initial condition

$$v(0, x) = f(x)$$

and the boundary conditions

$$v(t, 0) = g_1(t) \quad \text{and} \quad v(t, L) = g_2(t)$$

is the first boundary-value problem for the heat equation.

#### 4.4 Differential Equations of Mathematical Physics

$$\nabla^2 V \equiv \Delta V \equiv \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}$$

in Cartesian coordinates.

##### 4.41 Laplace Equation:

$$\nabla^2 \varphi = 0$$

1. Gravitational potential in a region free of mass
2. Steady temperature, i.e., temperature in a body that depends on position but not on time
3. Velocity potential of the irrotational flow of an incompressible fluid
4. Magnetic potential
5. Electrostatic potential in a region free of charges

##### 4.42 Poisson Equation:

$$\nabla^2 \varphi = -\rho(x, y, z)$$

1. Electrostatic potential in a region containing charges
2. Steady temperature with internal sources

##### 4.43 Helmholtz Equation:

$$\nabla^2 \varphi + k^2 \varphi = 0$$

$k^2$  may be a constant (or parameter) or a function of position.

##### 4.44 Heat-conduction or Diffusion Equation:

$$\nabla^2 \varphi = \frac{1}{K} \frac{\partial \varphi}{\partial t}$$

##### 4.45 Wave Equation:

$$\nabla^2 \varphi = \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2}$$

##### 4.46 Telegraph Equation:

$$\nabla^2 \varphi = a \frac{\partial^2 \varphi}{\partial t^2} + b \frac{\partial \varphi}{\partial t} + c \varphi$$

**4.47 Maxwell's Equations:**

$$\begin{aligned} \nabla \cdot \mathbf{B} &= 0 & \text{curl } \mathbf{H} &= \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} + \frac{1}{c} 4\pi \mathbf{J} & \mathbf{D} &= \epsilon \mathbf{E} \\ \nabla \cdot \mathbf{D} &= 4\pi \rho & \text{curl } \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} & \mathbf{B} &= \mu \mathbf{H} \end{aligned}$$

where  $\mathbf{E}$  is the electric field vector,  $\mathbf{H}$  is the magnetic field vector,  $c$  is the velocity of light,  $\mathbf{J}$  is current density,  $\rho$  is the density of the charge,  $\mu$  is permeability, and  $\epsilon$  is the dielectric constant.

**4.48 Equation for the Transverse Motion of a Plate or Bar:**

$$a^2 \Delta \Delta \varphi + \frac{\partial^2 \varphi}{\partial t^2} = 0$$

**4.49 Equation of Continuity:**

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

where  $\rho$  is the density and  $\mathbf{v}$  is the velocity vector.

**4.410 Navier-Stokes Equation:**

$$\rho \frac{d\mathbf{v}}{dt} = \mathbf{F} - \nabla p + \mu \nabla^2 \mathbf{v} + \frac{\mu}{3} \nabla (\nabla \cdot \mathbf{v})$$

where  $p$  is the pressure,  $\mu$  is the coefficient of viscosity, and  $\mathbf{F}$  is the external force.

**4.5 Boundary-value Problems**

Although the term boundary-value problem can be used in connection with ordinary differential equations, difference equations, integral equations, and variational problems, the term generally denotes a partial differential equation and a set of auxiliary conditions. Some authors separate such problems into initial-condition or boundary-value problems according to whether the conditions are specified at some initial time or all the conditions depend upon the space coordinates but not upon an initial state.

The following two comments should be noted when dealing with boundary-value problems. First, the number of boundary conditions is usually equal to the sum of the orders of the highest derivative with respect to each independent variable. For example, in Problem 1 of Art. 4.51, the number of boundary conditions is four while in Problem 2 the number of boundary conditions is three.

The second and most important comment is that, whenever possible, the problem should be broken into simpler problems that can be dealt with separately. In particular, if nonhomogeneities occur in more than one of the equations, it is often possible to consider a set of simpler problems each with one nonhomogeneity. For example, consider the problem

$$V_{xx} + V_{yy} = A(x, y)$$

with the boundary conditions

$$\begin{aligned} V(0, y) &= f_1(y) & V(a, y) &= f_2(y) & 0 < y < b \\ V(x, 0) &= f_3(x) & V(x, b) &= f_4(x) \end{aligned}$$

This problem can be broken into five simpler problems.

Let  $V = U_1 + U_2 + U_3 + U_4 + W$ , where  $U_1, U_2, U_3$ , and  $U_4$  satisfy the equation  $U_{xx} + U_{yy} = 0$  and where  $W_{xx} + W_{yy} = A$ . Furthermore,

$$\begin{aligned} W(0, y) &= W(a, y) = W(x, 0) = W(x, b) = 0 \\ U_1(0, y) &= f_1(y) & U_1(a, y) &= U_1(x, 0) = U_1(x, b) = 0 \\ U_2(a, y) &= f_2(y) & U_2(0, y) &= U_2(x, 0) = U_2(x, b) = 0 \\ U_3(x, 0) &= f_3(x) & U_3(0, y) &= U_3(a, y) = U_3(x, b) = 0 \\ U_4(x, b) &= f_4(x) & U_4(0, y) &= U_4(a, y) = U_4(x, 0) = 0 \end{aligned}$$

The two most common methods of solving boundary-value problems are the method of separation of variables and the method of transforms, particularly the Laplace transform. The techniques of conformal mapping, integral equations, and calculus of variations can also be used for solving boundary-value problems. Two examples of both the method of separation of variables and the use of the Laplace transform follow.

**4.51 Separation of Variables.** *Problem 1.* Solve the two-dimensional Laplace equation

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0$$

with the rectangular boundary conditions

$$\begin{array}{lll} V(0,y) = 0 & V(a,y) = 0 & 0 < y < b \\ V(x,0) = f(x) & V(x,b) = 0 & 0 < x < a \end{array}$$

*Formal Solution.* Let  $V(x,y) = X(x)Y(y)$ , and substitute into the Laplace equation. This gives

$$\frac{d^2 X}{dx^2} Y + X \frac{d^2 Y}{dy^2} = 0$$

or upon dividing by  $XY$ ,

$$\frac{d^2 X/dx^2}{X} = \frac{d^2 Y/dy^2}{Y}$$

Since the left-hand side of this equation is a function only of  $x$  and the right-hand side is a function only of  $y$ , both sides must be independent of  $x$  and  $y$  and therefore equal to some constant  $\lambda$ . The term *separation of variables* is seen to be derived from having separated all the functions in an equation that depend only upon a certain variable from the other functions that depend upon the other variables. The constant  $\lambda$  is called the *separation constant or parameter*.

The use of the separation constant results in the two equations

$$\frac{d^2 X}{dx^2} = \lambda X \quad \text{and} \quad \frac{d^2 Y}{dy^2} = -\lambda Y$$

If  $\lambda = -\alpha^2$ , the general solutions of these two equations can be written, when  $\alpha \neq 0$ , as

$$X = A \sin \alpha x + B \cos \alpha x \quad \text{and} \quad Y = E \sinh \alpha y + F \cosh \alpha y$$

and when  $\alpha = 0$ , as  $X = Ax + B$  and  $Y = Ey + F$ .

At this point it is necessary to apply the boundary conditions. First consider  $V(0,y) = X(0)Y(y) = 0$ . This implies that  $B = 0$  for all  $\alpha$ . Next consider

$$V(a,y) = X(a)Y(y) = 0$$

In order that  $A \neq 0$  and therefore  $X \neq 0$ , it follows that  $\sin \alpha a = 0$  or  $\alpha = n\pi/a$ , where  $n$  is an integer. The conditions  $V(x,b) = X(x)Y(b) = 0$  imply that

$$E \sinh \alpha b + F \cosh \alpha b = 0$$

This will be satisfied if  $E = -D \cosh \alpha b$  and  $F = D \sinh \alpha b$ . Hence

$$Y = D \sinh \alpha(b - y)$$

The quantity

$$C_n \sin \frac{n\pi}{a} x \sinh \frac{n\pi}{a} (b - y)$$

for any integer  $n$ , is seen to satisfy the Laplace equation and the boundary conditions along the sides  $x = 0$ ,  $x = a$ , and  $y = b$ .

In order to satisfy the remaining nonhomogeneous conditions  $V(x,0) = f(x)$ , it is usually necessary to consider the expression

$$\sum_{n=1}^{\infty} C_n \sin \frac{n\pi}{a} x \sinh \frac{n\pi}{a} (b-y)$$

where the  $C_n$  are constants. It was noted that each term and indeed the sum of any finite number of terms of this series satisfy all the conditions, except perhaps  $V(x,0) = f(x)$ . The question now is if, when  $y = 0$ ,  $C_n$  can be chosen so that

$$\sum_{n=1}^{\infty} C_n \sinh \frac{n\pi}{a} b \sin \frac{n\pi}{a} x = f(x)$$

This is seen to be a Fourier series, and it is known that if  $C_n \sinh (n\pi/a)b \equiv c_n$  are the Fourier coefficients,

$$c_n = \frac{2}{a} \int_0^a f(t) \sin \frac{n\pi t}{a} dt$$

then indeed the series converges to  $f(x)$ , where  $f(x)$  is nearly an arbitrary function.

The solution of the boundary-value problem is therefore given by

$$V(x,y) = \sum_{n=1}^{\infty} \left[ \frac{2}{a} \int_0^a f(t) \sin \frac{n\pi t}{a} dt \right] \sin \frac{n\pi}{a} x \frac{\sinh (n\pi/a)(b-y)}{\sinh n\pi b/a}$$

*Problem 2.* Solve the heat equation for one-dimensional flow

$$\frac{\partial v}{\partial t} = k \frac{\partial^2 v}{\partial x^2}$$

with the boundary conditions

$$\begin{aligned} v(0,t) &= 0 & v(\pi,t) &= 0 & t > 0 \\ v(x,0) &= f(x) & 0 < x < \pi & \end{aligned}$$

*Formal Solution.* Let  $v(x,t) = X(x)T(t)$ , and substitute into the heat equation. This gives

$$X \frac{dT}{dt} = k \frac{d^2 X}{dx^2} T \quad \text{or} \quad \frac{T'}{kT} = \frac{X''}{X}$$

Therefore, since in this last equation the function on the left is a function only of  $t$  and the function on the right is a function only of  $x$ , the equations

$$X'' = \lambda X \quad \text{and} \quad T' = \lambda k T$$

are obtained, where  $\lambda$  is the separation constant. It is easily seen that the quantities

$$C_n e^{-n^2 kt} \sin (nx)$$

where  $n$  is an integer, satisfy these equations and the homogeneous boundary conditions. Again consider a series of such terms

$$\sum_{n=1}^{\infty} C_n e^{-n^2 kt} \sin (nx)$$

When  $t = 0$ , this reduces to the Fourier series

$$\sum_{n=1}^{\infty} c_n \sin (nx)$$

and to  $f(x)$ , provided the  $c_n$  are chosen as the Fourier coefficients,

$$c_n = \frac{2}{\pi} \int_0^{\pi} f(s) \sin (ns) ds$$

The solution of the problem is given by

$$v(x, t) = \sum_{n=1}^{\infty} \left[ \frac{2}{\pi} \int_0^{\pi} f(s) \sin (ns) ds \right] \sin (nx) e^{-n^2 kt}$$

**4.52 Laplace Transform.** *Problem 3.* Solve the heat equation

$$\frac{\partial U}{\partial t} = k \frac{\partial^2 U}{\partial x^2}$$

with the boundary conditions

$$\begin{aligned} U(0, t) &= F(t) & \lim_{x \rightarrow \infty} U(x, t) &= 0 & t > 0 \\ U(x, 0) &= 0 & 0 < x < 1 \end{aligned}$$

*Formal Solution.* Let  $u(x, s) = L\{U(x, t)\} = \int_0^{\infty} e^{-st} U(x, t) dt$ . The application of the Laplace transform to the above problem results in the equations

$$su(x, s) = k \frac{d^2 u}{dx^2}$$

$u(0, s) = f(s)$ ,  $\lim_{x \rightarrow \infty} u(x, s) = 0$ , where  $f(s)$  is the transform of  $F(t)$ . The solution of this transformed problem is

$$u(x, s) = f(s) \exp \left[ -x \left( \frac{s}{k} \right)^{1/2} \right]$$

Since

$$\exp \left[ -x \left( \frac{s}{k} \right)^{1/2} \right] = L \left\{ \frac{x}{2} (\pi k t^3)^{-1/2} \exp \left( -\frac{x^2}{4kt} \right) \right\}$$

the use of the convolution gives

$$U(x, t) = \frac{x}{2} (\pi k)^{-1/2} \int_0^t \tau^{-3/2} F(t - \tau) \exp \left( -\frac{x^2}{4k\tau} \right) d\tau$$

If  $F(t) = A$ , a constant, then

$$U(x, t) = A \operatorname{erfc} \left[ \frac{x}{2} (kt)^{-1/2} \right]$$

*Problem 4.* Solve the vibrating-string equation

$$Y_{it} = a^2 Y_{zz}$$

with the boundary conditions

$$\begin{aligned} Y(x, 0) &= f(x) & Y_t(x, 0) &= 0 & 0 < x < 1 \\ Y(0, t) &= 0 & Y(1, t) &= 0 & t > 0 \end{aligned}$$

The transformed problem becomes

$$\begin{aligned} s^2 y(x,s) - sf(x) &= a^2 y_{xx} \\ y(0,s) = 0 & \quad y(1,s) = 0 \end{aligned}$$

where  $y(x,s) = L[Y(x,t)]$ . The solution of this transformed problem can be obtained either by the use of variation of parameters or by transforming with respect to  $x$ . In either case the solution becomes

$$y(x,s) = \frac{\varphi(x,s)}{a \sinh s/a}$$

where

$$\varphi(x,s) = \sinh \frac{(1-x)s}{a} \int_0^x f(z) \sinh \frac{zs}{a} dz + \sinh \frac{xs}{a} \int_x^1 f(z) \sinh \frac{(1-z)s}{a} dz$$

In order to invert the transform and obtain  $Y(x,t)$ , it is necessary to use an inversion series. The terms in the inversion series correspond to zeros of  $\sinh s/a$ , that is, the poles of  $y(x,s)$ . The sum of the two residues of  $e^{st}y(x,s)$  at  $s = \pm s_n$ , where  $s_n = in\pi a$  ( $n = 1, 2, \dots$ ), is given by

$$\frac{1}{2} c_n \sin(n\pi x) [\exp(in\pi at) + \exp(-in\pi at)]$$

where

$$c_n = 2 \int_0^1 f(z) \sin(n\pi z) dz$$

The formal solution therefore becomes

$$Y(x,t) = \sum_{n=1}^{\infty} c_n \sin(n\pi x) \cos(n\pi at)$$

When  $t = 0$ , this is seen to reduce to the Fourier series for  $f(x)$ .

#### 4.6 Numerical Solution of Differential Equations

Since relatively few differential equations can be solved (or integrated) in finite terms, it is generally necessary to consider numerical methods in order to obtain approximate solutions. Even when the solution of a differential equation can be obtained in infinite terms, the solution is usually difficult to evaluate and hence often of limited practical value. The advent of high-speed digital computers has made it relatively easy to obtain the numerical solution of difficult equations, such as multi-group reactor equations. A large number of useful techniques for solving differential equations may be found in the references. Some simple techniques follow.

**4.61 Ordinary Differential Equation. Modified Euler Method.** Although this method as well as the following Runge-Kutta method can be used for a system of equations, it is easier to consider a single first-order equation in order to indicate the method of solution. Consider the differential equation

$$y' = f(x,y)$$

with the initial condition  $y(x_0) = y_0$ . An approximation for  $y(x_0 + \Delta x) \equiv y(x_1) = y_1$  is easily seen to be given by

$$y_1 \approx y_0 + y'_0 \Delta x$$

where  $y'_0$  is shorthand notation for  $f(x_0, y_0)$ . In like fashion

$$y_{n+1} \approx y_n + y'_n \Delta x \quad n = 1, 2, \dots$$

This, the oldest, simplest, and crudest technique, was devised by Euler. A modified Euler method lets

$$y_1^* \approx y_0 + y'_0 \Delta x$$

and

$$y_1 \approx y_0 + \frac{y'_0 + (y_1^*)'}{2} \Delta x$$

where  $(y_1^*)' = f(x_1, y_1^*)$ . Although this modified Euler method is slow and of limited accuracy, it is very simple and easily used.

*Runge-Kutta Method.* Again consider  $y' = f(x, y)$  subject to  $y(x_0) = y_0$ . This method consists of computing in sequence  $k_1, k_2, k_3, k_4$  and then the desired  $y_1$ , where

$$\begin{aligned} k_1 &= f(x_0, y_0) \Delta x \\ k_2 &= f\left(x_0 + \frac{\Delta x}{2}, y_0 + \frac{k_1}{2}\right) \Delta x \\ k_3 &= f\left(x_0 + \frac{\Delta x}{2}, y_0 + \frac{k_2}{2}\right) \Delta x \\ k_4 &= f(x_0 + \Delta x, y_0 + k_3) \Delta x \\ y_1 &\approx y_0 + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned}$$

In order to obtain  $y_{n+1}$  ( $n = 1, 2, \dots$ ), it is necessary only to replace the  $x_0$  and  $y_0$  on the right-hand side of the above relations by the previously obtained  $x_n$  and  $y_n$ . This technique is one of the most commonly used and is of sufficient accuracy for most applications.

**4.62 Partial Differential Equations. Elliptic Equations.** For illustrative purposes consider the Laplace equation  $v_{xx} + v_{yy} = 0$  with Dirichlet conditions on a square boundary; that is,  $v(0, y) = f_1(y), v(a, y) = f_2(y), v(x, 0) = f_3(x), v(x, a) = f_4(x)$ . Let the square of side  $a$  be divided into a network or lattice of squares of side  $h$ , where  $a$  is an integer multiple of  $h$ . This can be accomplished by drawing two families of lines parallel to the sides of the original square and spaced at integer multiples of  $h$ . The intersections of these families of lines are called *lattice points*. The numerical technique consists of solving a set of associated difference equations for a function that takes on values only at the lattice points and using it as an approximation to the function  $\varphi$  at the lattice points.

The differential equation  $v_{xx} + v_{yy} = 0$  can be associated with the difference equation

$$\frac{v(x+h, y) - 2v(x, y) + v(x-h, y)}{h^2} + \frac{v(x, y+h) - 2v(x, y) + v(x, y-h)}{h^2} = 0$$

or  $v(x, y) = \frac{1}{4}[v(x+h, y) + v(x-h, y) + v(x, y+h) + v(x, y-h)]$ . This association results from approximating a first derivative  $dw/dx$  by  $[w(x+h, y) - w(x, y)]/h$  and a second derivative  $d^2w/dx^2$  by  $[w(x+h, y) - 2w(x, y) + w(x-h, y)]/h^2$ . The last equation for  $v(x, y)$  indicates that the value at an interior lattice point is the arithmetic mean of the values at the four nearest points. With the notation  $v(ih, jh) = v_{ij}$ , where  $i, j = 0, 1, 2, \dots, N$ , and  $nh = a$ , the difference equation becomes

$$v_{ij} = \frac{1}{4}(v_{i+1, j} + v_{i-1, j} + v_{i, j+1} + v_{i, j-1}) \quad i, j = 1, 2, \dots, N-1$$

Whenever the subscripts  $i$  or  $j$  equal 0 or  $N$ , the corresponding  $v_{ij}$  is set equal to the appropriate boundary value; for example,  $v_{0j} = f_1(jh)$ . The problem now consists of solving a system of linear equations.

There are two general methods, relaxation and iteration, that can be used to obtain an approximate solution to this linear system. The relaxation method deals with the quantities

$$R_{ij} = v_{i+1, j} + v_{i-1, j} + v_{i, j+1} + v_{i, j-1} - 4v_{ij}$$

In this method a set of  $v_{ij}^0$  ( $i, j = 1, 2, \dots, N-1$ ) is chosen and the  $R_{ij}$  are computed. The  $R_{ij}$  are called the *residuals*, and the object of the relaxation process is to try to alter (or relax) the  $v_{ij}^0$  so that the residuals all reduce, as nearly as possible, to zero. Since the relaxation process is not systematic and may proceed in any fashion in order to reduce the residuals, it is a hand method and not a machine method.

Iterative methods of solving linear equations are discussed in Art. 4.1 of Sec. 3-1. The Gauss-Seidel method is a commonly used iterative technique. Consider the formula

$$v^{(n+1)}_{ij} = \frac{1}{4}[v^{(n)}_{i+1, j} + v^{(n+1)}_{i-1, j} + v^{(n)}_{i, j+1} + v^{(n+1)}_{i, j-1}]$$

where the superscript  $n$  denotes the  $n$ th iteration. Again a set  $v_{ij}^{(0)}$  is chosen for the interior points. New values  $v_{ij}^{(1)}$  are obtained from the formula by starting at the lower left-hand corner and proceeding to the right and then upward.  $v_{ij}^{(n+1)}$  are similarly obtained from the  $v_{ij}^{(n)}$ . The process is continued until the  $v_{ij}^{(n+1)}$  are sufficiently close to the  $v_{ij}^{(n)}$ .

More generally, when difference equations replace any elliptic equation and its boundary conditions, a system of linear equations results. The numerical method in such cases is called *implicit*.

*Parabolic Equations.* The numerical solution of parabolic and hyperbolic equations can be calculated with step-by-step processes. In contrast to implicit methods, these step-by-step processes are called *explicit methods*.

Consider the heat equation  $U_t = aU_{xx}$  with the boundary conditions  $U(x,0) = f(x)$ ,  $U(0,t) = g_1(t)$ ,  $U(L,t) = g_2(t)$ . Let the strip  $t > 0$ ,  $0 < x < L$  be covered with a net of equal rectangles with sides  $\Delta x = h$  and  $\Delta t = k$ , where  $L$  is an integer multiple of  $h$ . If  $U_t$  is replaced by  $[U(x, t+k) - U(x,t)]/k$  and  $U_{xx}$  is replaced by  $[U(x+h, t) - 2U(x,t) + U(x-h, t)]/h^2$  the heat equation is expressed by

$$U(x, t+k) = rU(x+h, t) + (1-2r)U(x,t) + rU(x-h, t)$$

where  $r = ka/h^2$ . With this formula it is possible to calculate the values of  $U$  on the  $t+k$  line if the values of  $U$  are known on the  $t$  line. Now the values of  $U$  on the line  $t=0$  are given by  $U(x,0) = f(x)$ , and therefore it is possible to calculate the values  $U(x,k)$ . It is, of course, not necessary to calculate the values  $U(0,k)$  and  $U(L,k)$  or, indeed, any  $U(0,t)$  and  $U(L,t)$ , since they are specified by the boundary conditions. Having obtained  $U(x,k)$ , the values  $U(x,2k)$  can be obtained, and so on, step by step to  $U(x,nk)$ , where  $t = nk$  is the time interval to be covered.

One precaution must be followed in using this explicit method. The value of  $r$  must satisfy  $0 < r \leq \frac{1}{2}$ ; otherwise the numerical solution obtained may have little connection with the actual solution. This fact, noted by Courant, Friedrichs, and Lewy, states, roughly, that if the space mesh length  $h$  is fixed, the time mesh length  $k$  cannot be too large.

*Hyperbolic Equations.* Consider the wave equation  $U_{tt} = a^2U_{xx}$  with the boundary conditions  $U(x,0) = f_1(x)$ ,  $U_t(x,0) = f_2(x)$ ,  $U(0,t) = g_1(t)$ ,  $U(L,t) = g_2(t)$ . Again cover the strip  $t > 0$ ,  $0 < x < L$  with rectangles having sides  $\Delta x = h$  and  $\Delta t = k$ , where  $L$  is an integer multiple of  $H$ . Let the quantity  $U(x,t)$  be denoted by  $U_{ij}$  when  $x = ih$  and  $t = jk$ , where  $i = 0, 1, 2, \dots, N$  and  $j = 0, 1, 2, \dots$ . If the second derivatives are replaced by second differences, the wave equation becomes

$$\frac{U_{i,j+1} - 2U_{ij} + U_{i,j-1}}{k^2} = a^2 \frac{U_{i+1,j} - 2U_{ij} + U_{i-1,j}}{h^2}$$

or  $U_{i,j+1} = rU_{i+1,j} - 2(r-1)U_{ij} + rU_{i-1,j} - U_{i,j-1}$ , where  $r = k^2a^2/h^2$ . The initial conditions  $U(x,0) = f_1(x)$  and  $U_t(x,0) = f_2(x)$  or, with forward differences,

$$U(x,k) - U(x,0) = kf_2(x)$$

and therefore  $U(x,k) = kf_2(x) + f_1(x)$  supply the information needed in order to start the step-by-step process.

Note that when  $ka = h$ , the equation reduces to  $U_{i,j+1} = U_{i+1,j} + U_{i-1,j} - U_{i,j-1}$ . It is easily verified that  $U = f(x-at) + g(x+at)$ , where  $f(z)$  and  $g(z)$  are any two functions of  $z$  with second derivatives, is a solution of both the wave equation and this last difference equation. Therefore any solution of either the wave equation or the difference equation is a solution of the other, since  $U = f(x-at) + g(x+at)$  is a general solution.

More generally it is necessary when dealing with hyperbolic equations to make sure that the numerical process makes sense; for example, high-frequency component of a wave motion will always be distorted. Here  $r = k^2a^2/h^2$  must be restricted by  $r \leq 1$  in order to ensure convergence of the numerical method.

5 OTHER TOPICS

5.1 Vector Analysis

A discussion of vectors appears in Art. 1 of Sec. 3-1. The primary objective here is to present some of the more common and useful formulas and results of vector analysis.

**5.11 Scalar and Vector Products.** The scalar (or inner or dot) product of two vectors  $\mathbf{a}$  and  $\mathbf{b}$  is denoted by  $\mathbf{a} \cdot \mathbf{b}$  or  $(\mathbf{a}, \mathbf{b})$ . The magnitude or length of a vector  $\mathbf{a}$  is denoted by  $|\mathbf{a}|$ . The vector (or cross) product of two vectors is denoted by  $\mathbf{a} \times \mathbf{b}$ . It is customary to express a vector in 3-space in terms of three unit vectors  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  along the positive  $x, y,$  and  $z$  axes. In Sec. 3-1 the unit vectors are denoted by  $\mathbf{e}_1, \mathbf{e}_2,$  and  $\mathbf{e}_3$ .  $\mathbf{e}_\alpha$  denotes a unit vector in the positive  $\alpha$  direction. The two vectors  $\mathbf{a}$  and  $\mathbf{b}$ , therefore, can be written as  $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$  and  $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ ; the following elementary relations may be noted:

$$\begin{aligned} |\mathbf{a}|^2 &= \mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + a_3^2 \\ \mathbf{a} \cdot \mathbf{b} &= \mathbf{b} \cdot \mathbf{a} = a_1b_1 + a_2b_2 + a_3b_3 = |\mathbf{a}||\mathbf{b}| \cos(\mathbf{a}, \mathbf{b}) & (59) \\ \mathbf{a} \cdot \mathbf{b} = 0 &\text{ means } \mathbf{a} = 0 \text{ or } \mathbf{b} = 0 \text{ or } \mathbf{a} \text{ is perpendicular to } \mathbf{b} & (60) \\ \mathbf{i} \cdot \mathbf{i} = \mathbf{j} \cdot \mathbf{j} = \mathbf{k} \cdot \mathbf{k} &= 1 \quad \mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{k} \cdot \mathbf{i} = 0 & (61) \\ \mathbf{a} \times \mathbf{b} &= -\mathbf{b} \times \mathbf{a} = \mathbf{i}(a_2b_3 - a_3b_2) + \mathbf{j}(a_3b_1 - b_3a_1) + \mathbf{k}(a_1b_2 - b_1a_2) & (62) \\ \mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} &= \mathbf{k} \times \mathbf{k} = 0 & (63) \\ \mathbf{i} \times \mathbf{j} = \mathbf{k} \quad \mathbf{j} \times \mathbf{k} &= \mathbf{i} \quad \mathbf{k} \times \mathbf{i} = \mathbf{j} \end{aligned}$$

*Triple Scalar Product.* Volume of parallelepiped with edges  $\mathbf{a}, \mathbf{b}$ , and  $\mathbf{c}$ :

$$(\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \quad (64)$$

*Triple Vector Products:*

$$\begin{aligned} \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) &= \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}) & (65) \\ (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) &= (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}) & (66) \\ (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) &= [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{d})]\mathbf{c} - [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})]\mathbf{d} = [\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})]\mathbf{b} \\ &\quad - [\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})]\mathbf{a} & (67) \end{aligned}$$

*The Differential Operator  $\nabla$ . Rectangular Coordinates:*

$$\begin{aligned} \nabla &= \text{del} = \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z} \\ \nabla v &= \text{grad } v = \mathbf{i} \frac{\partial v}{\partial x} + \mathbf{j} \frac{\partial v}{\partial y} + \mathbf{k} \frac{\partial v}{\partial z} & (68) \end{aligned}$$

The gradient is a measure of the rate of change of the scalar field  $v$  at the point  $(x, y, z)$ .

$$\nabla \cdot \mathbf{F} = \text{div } \mathbf{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \quad \mathbf{F} = F_x\mathbf{i} + F_y\mathbf{j} + F_z\mathbf{k} \quad (69)$$

$$\nabla \times \mathbf{F} = \text{curl } \mathbf{F} = \text{rot } \mathbf{F} = \mathbf{i} \left( \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) + \mathbf{j} \left( \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) + \mathbf{k} \left( \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) \quad (70)$$

$$\nabla^2 v = \nabla \cdot (\nabla v) = \text{del squared} = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} \quad (71)$$

*Differentiation Formulas:*

$$\begin{aligned} \nabla(uv) &= u\nabla v + v\nabla u & (72) \\ \nabla \cdot (v\mathbf{F}) &= (\nabla v) \cdot \mathbf{F} + v\nabla \cdot \mathbf{F} & (73) \\ \nabla \times (v\mathbf{F}) &= (\nabla v) \times \mathbf{F} + v\nabla \times \mathbf{F} & (74) \\ \nabla \cdot (\mathbf{F} \times \mathbf{G}) &= \mathbf{G} \cdot (\nabla \times \mathbf{F}) - \mathbf{F} \cdot (\nabla \times \mathbf{G}) & (75) \\ \nabla(\mathbf{F} \cdot \mathbf{G}) &= \mathbf{F} \cdot \nabla \mathbf{G} + \mathbf{G} \cdot \nabla \mathbf{F} + \mathbf{F} \times (\nabla \times \mathbf{G}) + \mathbf{G} \times (\nabla \times \mathbf{F}) & (76) \\ \nabla \times (\mathbf{F} \times \mathbf{G}) &= \mathbf{G} \cdot \nabla \mathbf{F} - \mathbf{F} \cdot \nabla \mathbf{G} + \mathbf{F}(\nabla \cdot \mathbf{G}) - \mathbf{G}(\nabla \cdot \mathbf{F}) & (77) \\ \nabla \cdot (\nabla \times \mathbf{F}) &= 0 \quad \nabla \times (\nabla v) = 0 & (78) \\ \nabla \times (\nabla \times \mathbf{F}) &= \nabla(\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F} & (79) \end{aligned}$$

*Cylindrical Coordinates:*

$$x = r \cos \theta \quad y = r \sin \theta \quad z = z$$

$$ds^2 = dr^2 + r^2 d\theta^2 + dz^2 \quad (80)$$

$$\text{grad } v = \mathbf{e}_r \frac{\partial v}{\partial r} + \mathbf{e}_\theta \frac{1}{r} \frac{\partial v}{\partial \theta} + \mathbf{e}_z \frac{\partial v}{\partial z} \quad (81)$$

$$\text{div } \mathbf{F} = \frac{1}{r} \frac{\partial}{\partial r} (rF_r) + \frac{1}{r} \frac{\partial F_\theta}{\partial \theta} + \frac{\partial F_z}{\partial z} \quad (82)$$

$$\text{curl } \mathbf{F} = \mathbf{e}_r \left( \frac{1}{r} \frac{\partial A_z}{\partial \theta} - \frac{\partial A_\theta}{\partial z} \right) + \mathbf{e}_\theta \left( \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right) + \mathbf{e}_z \left[ \frac{1}{r} \frac{\partial}{\partial r} (rA_\theta) - \frac{1}{r} \frac{\partial A_r}{\partial \theta} \right] \quad (83)$$

$$\nabla^2 v = \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial v}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 v}{\partial \theta^2} + \frac{\partial^2 v}{\partial z^2} \quad (84)$$

*Spherical Coordinates:*

$$x = r \cos \varphi \sin \theta \quad y = r \sin \varphi \sin \theta \quad z = r \cos \theta$$

$$ds^2 = dr^2 + r^2 \sin^2 \theta d\varphi^2 + r^2 d\theta^2 \quad (85)$$

$$\text{grad } v = \mathbf{e}_r \frac{\partial v}{\partial r} + \mathbf{e}_\varphi \frac{1}{r \sin \theta} \frac{\partial v}{\partial \varphi} + \mathbf{e}_\theta \frac{1}{r} \frac{\partial v}{\partial \theta} \quad (86)$$

$$\text{div } \mathbf{F} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 F_r) + \frac{1}{r \sin \theta} \frac{\partial F_\varphi}{\partial \varphi} + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} (\sin \theta F_\theta) \quad (87)$$

$$\begin{aligned} \text{curl } \mathbf{F} = \mathbf{e}_r \frac{1}{r \sin \theta} \left[ \frac{\partial}{\partial \theta} (\sin \theta F_\varphi) - \frac{\partial F_\theta}{\partial \varphi} \right] + \mathbf{e}_\varphi \frac{1}{r} \left[ \frac{\partial}{\partial r} (rF_\theta) - \frac{\partial F_r}{\partial \theta} \right] \\ + \mathbf{e}_\theta \frac{1}{r} \left[ \frac{1}{\sin \theta} \frac{\partial A_r}{\partial \varphi} - \frac{\partial}{\partial r} (rA_\varphi) \right] \end{aligned} \quad (88)$$

$$\nabla^2 v = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial v}{\partial r} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 v}{\partial \varphi^2} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial v}{\partial \theta} \right) \quad (89)$$

*Integral Theorems.*  $\mathbf{n}$  is a unit outward normal vector to a surface  $A$ ,  $dA$  is an element of surface, and  $dV$  is an element of volume.

$$\mathbf{r} = xi + yj + zk$$

*Stokes' Theorem:*

$$\iint_A (\nabla \times \mathbf{F}) \cdot \mathbf{n} dA = \int_C \mathbf{F} \cdot d\mathbf{r} \quad (90)$$

The double integral is taken over the area bounded by the closed curve  $C$ .

*Divergence or Gauss' Theorem:*

$$\iiint_V \nabla \cdot \mathbf{F} dV = \iint_A \mathbf{F} \cdot \mathbf{n} dA \quad (91)$$

The triple integral is taken over the volume inside the closed surface  $A$ .

*Green's Theorem:*

$$\iiint_V (\nabla v \cdot \Delta w) dV = \iint_S v(\nabla w) \cdot \mathbf{n} dA - \iiint_V v \nabla^2 w dV \quad (92)$$

$$\iiint_V (v \nabla^2 w - w \nabla^2 v) dV = \iint_S (v \nabla w - w \nabla v) \cdot \mathbf{n} dA \quad (93)$$

The triple integrals are taken over the volume inside the closed surface  $A$ .

### 5.2 Integral Transforms—Operational Mathematics

The *integral transform*  $f(s)$  of a function  $F(t)$  is defined by the integral equation

$$f(s) = \int_a^b K(s,t)F(t) dt = T_s\{F\} \quad (94)$$

The function  $K(s,t)$ , called the *kernel of the transform*, is taken to be a known function. It is apparent from the definition that

$$cf(s) = \int_a^b K(s,t)cF(t) dt = T_s\{cF\}$$

and

$$f_1(s) + f_2(s) = \int_a^b K(s,t)[F_1(t) + F_2(t)] dt = T_s\{F_1 + F_2\} = T_s\{F_1\} + T_s\{F_2\}$$

These two relations indicate that the transformation between the functions  $F(t)$  and  $f(s)$  is linear. The defining transformation above is, then, a linear integral transform, although it is customary to leave out the word linear.

The use of integral transforms is a particularly useful technique for solving boundary-value problems. In essence, this technique can be described as follows: First, apply an integral transform to a differential equation and its boundary conditions. This, in effect, reduces the number of independent variables by one and introduces a parameter. The resulting, or transform, problem is then to be solved for the transform of the dependent, or wanted, variable. It is then necessary to invert this solution, that is, undo the result of the transform, in order to obtain the solution of the given problem. An example of this process appears in the article concerned with boundary-value problems.

Each type of integral transform is particularly appropriate to the solution of certain kinds of linear boundary-value problems. The Laplace transform has proved to be especially useful for solving transient problems of the type arising in the conduction of heat in solids and in vibration theory. In contrast to the Laplace transform, where the transformed variable is usually the time, the Fourier transforms are generally taken with respect to a space variable over an infinite or semi-infinite interval. The Hankel transform is applicable when the problem has symmetry about an axis and a radial variable from 0 to  $\infty$  is present. The Mellin transform is similar to the Fourier transform in its application.

It should be noted that the transform of the product of two arbitrary or unknown functions cannot be obtained usefully in terms of the transforms of the individual functions. Consequently, the integral transform method is successful only when the coefficients involved in a boundary-value problem are either independent of or elementary functions of the transform variable.

**5.21 Laplace Transform.** The *Laplace transform*  $f(s)$  of a function  $F(t)$  is defined by

$$f(s) = \int_0^{\infty} e^{-st}F(t) dt = J\{F\} \quad (95)$$

The Laplace transform of  $F(t)$  exists if  $F(t)$  is sectionally continuous in every finite interval in  $t > 0$  and if it is of exponential order, that is, there exist constants  $C$  and  $M$  such that

$$\lim_{t \rightarrow \infty} e^{-ct}|F(t)| < M$$

The inverse Laplace transform is denoted by  $L^{-1}\{f(s)\}$ . This indicates that

$$L\{F(t)\} = f(s) \quad \text{and} \quad F(t) = L^{-1}\{f(s)\}$$

The inverse transform is not always easily obtained. The use of tables of transforms is the most convenient method of obtaining the inverse transform. The general

Table 6. Laplace Transforms

No.	$F(t) = L^{-1}\{f(s)\}$	$f(s) = L\{F(t)\}$
1	$F'(t)$	$sf(s) - F(0^+)$
2	$F^{(n)}(t)$	$s^n f(s) - s^{n-1}F(0^+) - s^{n-2}F'(0^+) - \dots - F^{(n-1)}(0^+)$
3	$\int_0^t F(x) dx$	$\frac{1}{s} f(s)$
4	$\int_0^t \int_0^{x_1} F(x) dx dx_1$	$\frac{1}{s^2} f(s)$
5	$\int_0^t F_1(t-\tau)F_2(\tau) d\tau$ = $F_1 * F_2$ , convolution	$f_1(s)f_2(s)$
6	$t^n F(t)$	$(-1)^n f^{(n)}(s)$
7	$\frac{1}{t} F(t)$	$\int_s^\infty f(x) dx$
8	$e^{at}F(t)$	$f(s-a)$
9	$F(t-b)$ where $F(t) = 0$ if $t < 0$	$e^{-bs}f(s)$
10	$\frac{1}{c} F\left(\frac{t}{c}\right)$	$f(cs)$
11	$\sum_{n=1}^N \frac{p(a_n)}{q'(a_n)} e^{a_n t}$	$\frac{p(s)}{q(s)}$ , $q(s) = (s-a_1)(s-a_2)\dots(s-a_N)$
12	$t^k$	$\frac{\Gamma(k+1)}{s^{k+1}}$ ( $k+1 > 0$ )
13	$\frac{2^n n! n^{-1/2} \pi^{-1/2}}{1 \cdot 3 \cdot 5 \dots (2n-1)}$	$s^{-(n+1/2)}$ ( $n = 1, 2, \dots$ )
14	$t^{k-1} e^{at}$	$\frac{\Gamma(k)}{(s-a)^k}$ ( $k > 0$ )
15	$\cos at$	$\frac{s}{s^2 + a^2}$
16	$\sin at$	$\frac{a}{s^2 + a^2}$
17	$\sinh at$	$\frac{a}{s^2 - a^2}$
18	$\cosh at$	$\frac{s}{s^2 - a^2}$
19	$t \cos at$	$\frac{(s^2 + a^2)^2}{2as}$
20	$t \sin at$	$\frac{(s^2 + a^2)^2}{2as}$
21	$t^n \sin at$	$\frac{2^n a^n n! s^n}{(s^2 + a^2)^{n+1}}$
22	$(\pi t)^{-1/2} \cos 2\sqrt{kt}$	$s^{-1/2} e^{-k/s}$
23	$(\pi k)^{-1/2} \sin 2\sqrt{kt}$	$s^{-3/2} e^{-k/s}$
24	$J_0(2\sqrt{kt})$	$\frac{1}{s} e^{-k/s}$
25	$J_0(at)$	$(s^2 + a^2)^{-1/2}$
26	$J_\nu(at)$ ( $\text{Re } \nu > -1$ )	$a^{-\nu} (s^2 + a^2)^{1/2} - s J^\nu (s^2 + a^2)^{-1/2}$
27	$t^\nu J_\nu(at)$ ( $\text{Re } \nu > -1/2$ )	$\pi^{-1/2} \Gamma(\nu + 1/2) (2a)^\nu (s^2 + a^2)^{-\nu-1/2}$
28	$\frac{k}{2} (\pi t^3)^{-1/2} \exp\left(-\frac{k^2}{4t}\right)$	$e^{-k\sqrt{s}}$ ( $k > 0$ )
29	$\text{erfc}\left(\frac{k}{2} t^{-1/2}\right)$	$\frac{1}{s} e^{-k\sqrt{s}}$ ( $k \geq 0$ )
30	$(\pi t)^{-1/2} \exp\left(-\frac{k^2}{4t}\right)$	$s^{-1/2} e^{-k\sqrt{s}}$ ( $k \geq 0$ )

inversion procedures, however, involve inversion theorems and most often inversion integrals or series. The references contain tables of transforms as well as more information concerning the following transforms.

The *inversion integral* for a function  $f(s)$  is given by

$$L_i^{-1}\{f(s)\} = \frac{1}{2\pi i} \lim_{\beta \rightarrow \infty} \int_{\gamma - i\beta}^{\gamma + i\beta} e^{st} f(z) dz \tag{96}$$

where the integration is along a line parallel to the imaginary axis in the complex  $z$  plane. The following theorem, essentially the same as the one found in Churchill's "Modern Operational Mathematics in Engineering," relates the inversion integral and inverse transform:

*Inversion Theorem.* If  $f(s)$  is an analytic function of  $s$  and is of order  $O(s^{-k})$  in some half plane  $\text{Re } s \geq c$ , where  $k$  and  $c$  are real constants and  $k > 1$ , then the inversion integral  $L_i^{-1}\{f(s)\}$  along any line  $\text{Re } z = \gamma$ , where  $\gamma \geq c$ , converges to a function  $F(t)$  that is independent of  $\gamma$  and whose Laplace transform is  $f(s)$ ; that is,

$$F(t) = L_i^{-1}\{f(s)\} \quad \text{and} \quad L\{F(t)\} = f(s)$$

$F(t)$  is continuous for each  $t \geq 0$ , is of order  $O(e^{\gamma t})$  for all  $t \geq 0$ , and is such that  $F(0) = 0$ .

It is often possible (see references) to express the inversion integral in terms of an infinite series or an equivalent infinite integral.

**5.22 Fourier Transforms.** The *Fourier transform*  $f(y)$  of a function  $F(x)$  is defined by

$$f(y) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} e^{iyt} F(t) dt \tag{97}$$

The function  $f(y)$  is also called the *spectral function* for  $F(x)$ . The *Fourier integral theorem* formally states that

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} F(t) e^{iy(t-x)} dt \tag{98}$$

This leads to Fourier's inversion formula

$$F(x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} f(y) e^{-iyx} dy \tag{99}$$

For real  $F(x)$  the right-hand side of the Fourier integral theorem has the form

$$\frac{1}{\pi} \int_0^{\infty} dy \int_{-\infty}^{\infty} F(t) \cos 2\pi y(t-x) dt \tag{100}$$

If  $f(x)$  is sectionally continuous in every finite interval, and if  $\int_{-\infty}^{\infty} |f(x)| dx$  converges, then at every point where  $f(x)$  has right- and left-hand derivatives

$$\frac{1}{2} [F(x^+) + F(x^-)] = \frac{1}{\pi} \int_0^{\infty} dy \int_{-\infty}^{\infty} F(t) \cos 2\pi y(t-x) dt \tag{101}$$

If  $f(x)$  is an odd function, the *Fourier sine transform*

$$f(y) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} F(t) \sin yt dt \tag{102}$$

with the inversion formula

$$F(x) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} f(y) \sin xy dy \tag{103}$$

Table 7. Fourier Transforms

No.	$F(x) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} f(y)e^{-iyx} dy$	$f(y) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} F(x)e^{iyx} dx$
1	$F(ax)$	$\frac{1}{a} f\left(\frac{y}{a}\right)$
2	$ixF(x)$	$\frac{d}{dy} f(y)$
3	$\frac{d}{dx} F(x)$	$-iyf(y)$
4	$e^{ixa}F(x)$	$f(y+a)$
5	$F(x+a)$	$e^{-ixy}f(y)$
6	1	$\delta(y)$
7	$(2\pi)^{-1/2} \int_{-\infty}^{\infty} F(t)G(x-t) dt$	$f(y)g(y)$
8	$(2\pi)^{-1/2}  x ^{-1/2}$	$ y ^{-1/2}$
9	$\frac{\sin ax}{x}$	$\left(\frac{\pi}{2}\right)^{1/2}  y  < a$ 0 $ y  > a$
10	$\sin ax^2$	$(2a)^{-1/2} \sin\left(\frac{y^2}{4a} + \frac{\pi}{4}\right)$
11	$\cos ax^2$	$(2a)^{-1/2} \cos\left(\frac{y^2}{4a} - \frac{\pi}{4}\right)$
12	$e^{ixp}$ $p < x < q$ 0 $x < p, x > q$ where $p < q$	$i(2\pi)^{-1/2} \frac{e^{ip(w+y)} - e^{iq(w+y)}}{y}$
13	$e^{-ax+ixw}$ $x > 0$ 0 $x < 0$	$i(2\pi)^{-1/2} (w+y+ic)^{-1}$
14	$\frac{1}{ x }$	$\frac{1}{ y }$
15	$ x ^{-s}$ $0 < \text{Re } s < 1$	$\left(\frac{2}{\pi}\right)^{1/2}  y ^{s-1} \Gamma(1-s) \sin \frac{\pi}{2} s$
16	$e^{-sx^2}$ $\text{Re } s > 0$	$(2s)^{-1/2} e^{-y^2/4s}$
17	$e^{-a x }  x ^{-1/2}$	$[(a^2+y^2)^{1/2} + a]^{1/2} (a^2+y^2)^{-1/2}$
18	$\frac{1}{x^2+a^2}$	$\left(\frac{\pi}{2a}\right)^{1/2} e^{-a y }$
19	$(a^2-x^2)^{-1/2}$ $ x  <  a $ 0 $ x  > 0$	$\left(\frac{\pi}{2}\right)^{1/2} J_0(ay)$
20	$(x^2+a^2)^{-1/2}$	$\left(\frac{2}{\pi}\right)^{1/2} K_0(a y )$
21	$P_n(x)$ $ x  < 1$ 0 $ x  > 1$	$i^n \pi^{-1/2} J_{n+1/2}(y)$

is obtained. If  $f(x)$  is an even function, the *Fourier cosine transform*

$$f(y) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} F(t) \cos yt dt \quad (104)$$

with the inversion formula

$$F(x) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\infty} f(y) \cos xy dy \quad (105)$$

is obtained.

**5.23 Hankel Transforms.** The *Hankel transform*  $f(y)$  of order  $\gamma$  of a function  $F(x)$  is defined by

$$f(y) = \int_0^{\infty} x J_{\gamma}(xy) F(x) dx \quad (106)$$

The inversion formula is given by

$$F(x) = \int_0^{\infty} y J_{\gamma}(xy) f(y) dy \quad (107)$$

**5.24 Mellin Transform.** The *Mellin transform*  $f(y)$  of a function  $F(x)$  is defined by

$$f(y) = \int_0^{\infty} x^{y-1} F(x) dx \quad (108)$$

The inversion formula is given by

$$F(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f(y)x^{-y} dy \quad (109)$$

**5.25 Finite Fourier Transforms.** The *finite sine transform*  $f_s(n)$  and *finite cosine transform*  $f_c(n)$  of a function  $F(x)$  are defined, respectively, by

$$f_s(n) = \int_0^{\pi} F(x) \sin nx dx \quad n = 1, 2, \dots \quad (110)$$

and

$$f_c(n) = \int_0^{\pi} F(x) \cos nx dx \quad n = 0, 1, 2, \dots \quad (111)$$

The inversion formulas for these transformations are given, respectively, by

$$F(x) = \frac{2}{\pi} \sum_{n=1}^{\infty} f_s(n) \sin nx \quad 0 < x < \pi \quad (112)$$

and

$$F(x) = \frac{1}{\pi} f_c(0) + \frac{2}{\pi} \sum_{n=1}^{\infty} f_c(n) \cos nx \quad 0 < x < \pi \quad (113)$$

### 5.3 Linear Integral Equations

The following four types of linear integral equations are commonly noted:  
Volterra equation of the first kind:

$$\int_a^x K(x,s)y(s) ds = f(x) \quad (114)$$

Volterra equation of the second kind:

$$y(x) - \int_a^x K(x,s)y(s) ds = f(x) \quad (115)$$

Fredholm equation of the first kind:

$$\int_a^b K(x,s)y(s) ds = f(x) \quad (116)$$

Fredholm equation of the second kind:

$$y(x) - \int_a^b K(x,s)y(s) ds = f(x) \quad (117)$$

The problem for these linear integral equations is to determine the unknown function  $y(x)$  that satisfies the equation in the desired interval. It is assumed that the functions  $K(x,s)$ ,  $f(x)$  and the limits  $a$  and  $b$  are known.  $K(x,s)$  is called the *kernel* of the equation. These equations are called integral equations, since the unknown function appears in the integrand. Similarly the equations are called linear, since the unknown function  $y(x)$  occurs linearly.

Although some theory is related to nonlinear integral equations, most of the existing theory deals with linear equations.

$$y(x) + \int_a^b K(x,s) \sin [y(s)] ds = f(x)$$

is an example of a nonlinear equation, since the unknown function appears nonlinearly through a sine function. Some theory has also been developed for systems of linear integral equations and linear integral equations in more than one independent variable. This article, however, deals only with linear integral equations in one independent variable.

An integral equation is called *singular* if either one or both of the limits of integration become infinite or if the kernel becomes infinite at one or more points of its domain.

$$\int_0^\infty e^{-zs} y(s) ds = f(x)$$

and

$$\int_0^x \frac{1}{\sqrt{x-s}} y(s) ds = f(x)$$

are examples of singular equations.

There are two common *methods of relating linear differential equations and Volterra integral equations*. Consider the linear differential equation

$$y^{(n)}(x) + \sum_{i=1}^n A_i(x) y^{(n-i)} = \varphi(x)$$

If one lets  $z(x) = y^{(n)}(x)$ , then

$$y^{(n-1)}(x) = \int_0^x z(s) ds + a_{n-1} \quad y^{(n-2)}(x) = \int_0^x (x-s)z(s) ds + a_{n-1}x + a_{n-2}$$

and finally

$$y(x) = \int_0^x \frac{(x-s)^{n-1}}{(n-1)!} z(s) ds + a_{n-1} \frac{x^{n-1}}{(n-1)!} + \cdots + a_0$$

Direct substitution in the differential equation leads then to the integral equation

$$z(x) + \int_0^x \left[ A_1 + (x-s)A_2 + \cdots + \frac{(x-s)^{n-1}}{(n-1)!} A_n \right] z(s) ds = f(x)$$

The other method consists of repeated integration of the differential equation. It is not difficult to verify using the formula

$$\int_0^x \int_0^{x_1} \cdots \int_0^{x_n} f(x_{n+1}) dx_{n+1} \cdots dx_1 = \int_0^x \frac{(x-s)^n}{n!} f(s) ds$$

that the differential equation will then become

$$y(x) + \int_0^x K(x,s)y(s) ds = f(x) \tag{118}$$

where  $K(x,s) = \sum_{i=0}^{n-1} k_i(s)(x-s)^i$ , the  $k_i(s)$  being functions of  $A_i(s)$ .

Most developments of linear integral equations deal with Fredholm equations, since a Volterra equation can be considered as a special case of a Fredholm equation with a kernel  $K(x,s)$  which is defined for  $a \leq s \leq x \leq b$  and is zero for  $a \leq x \leq s \leq b$ .

Often either a parameter  $\lambda$  appears naturally or it is convenient to introduce such a parameter in the discussion of an integral equation. The equation

$$y(x) - \lambda \int_a^b K(x,s)y(s) ds = f(x) \tag{119}$$

exhibits the customary location of a parameter.

There are three general methods leading to the solution of Fredholm equations of the second kind with a parameter. The first method is called the *method of successive substitution*. In this method the unknown function is obtained as a power series in  $\lambda$  with the coefficients being functions of the independent variable  $x$ . This series is usually called a *Neumann series* and converges for certain values of  $\lambda$ . Method two is due to Fredholm and will be called the *Fredholm method*. The unknown function in this technique is obtained first as the ratio of two power series in  $\lambda$ . The power series in the numerator has coefficients that are functions of  $x$ , while the denominator power series has coefficients that are independent of  $x$ . This method considers the integral equation as the limit of a set of  $n$  linear algebraic equations in  $n$  unknowns when  $n$  tends to infinity. The *Hilbert-Schmidt theory* provides the third method for obtaining a solution. In this theory the unknown function is obtained in the form of a series of *fundamental or characteristic functions*. These characteristic functions, or *eigenfunctions*, are solutions of the homogeneous integral equation and are associated with particular values of the parameters called characteristic numbers or eigenvalues.

Before indicating some of the results obtained by these three methods, it is worthwhile to consider the special case where the kernel has the form

$$K(x,s) = \sum_{i=1}^n A_i(x)B_i(s)$$

where  $A_i(x)$  and  $B_i(y)$  are continuous and linearly independent. For a Fredholm equation of the first kind it is necessary that  $f(x) = \sum_{i=1}^n C_i A_i(x)$  in order that the equation have a solution. The general solution of this equation can then be written in the form

$$y(x) = \sum_{i=1}^n a_i B_i(x) + h(x)$$

where  $h(x)$  is any function orthogonal to all the  $B_i(x)$ ; that is,  $\int_a^b h(s)B_i(s) ds = 0$ .

For Fredholm equations of the second kind with the above type of kernel a solution of the form

$$y(x) = f(x) + \sum_{i=1}^n c_i A_i(x)$$

is assumed. A necessary and sufficient condition that there exist a unique solution and set of  $c_i$  is that a certain determinant  $\Delta$  is not equal to zero. This determinant  $\Delta$  is defined by

$$\Delta \equiv \det \left[ \delta_{ij} - \int_a^b B_i(s)A_j(s) ds \right]$$

where  $\delta_{ij}$  is the Kronecker delta and  $i, j = 1, 2, \dots, n$ . This solution can be written as

$$y = f(x) + \int_a^b L(x,s)f(s) ds$$

$$L(x,s) = \frac{\begin{vmatrix} 0 & A_1(x) & \cdots & A_n(x) \\ B_1(s) & & & \\ \cdot & & \Delta & \\ \cdot & & & \\ B_n(s) & & & \end{vmatrix}}{\Delta}$$

where

If  $\Delta = 0$ , solutions may also exist. This situation parallels the linear algebraic case and corresponding results hold.

**5.31 Successive Substitution.** The method of successive substitution describes in its name the process used to obtain a solution. Successive substitution for  $y(x)$  gives

$$\begin{aligned} y(x) &= f(x) + \lambda \int_a^b K(x,s)y(s) ds = f(x) \\ &\quad + \lambda \int_a^b K(x,s) \left[ f(s) + \lambda \int_a^b K(s,s_1)y(s_1) ds_1 \right] \\ &= \dots = f(x) + \lambda \int Kf + \lambda^2 \int K \int Kf + \dots \\ &\quad + \lambda^n \left[ \underbrace{\int K \dots \int K}_n \right] f + \lambda^{n+1} \left[ \underbrace{\int K \dots \int K}_{n+1} \right] y \end{aligned}$$

where the integration is indicated symbolically. If, then,

$$\lim_{n \rightarrow \infty} \lambda^n \underbrace{\int K \dots \int K}_n y = \lim_{n \rightarrow \infty} (\lambda \int K)^n y = 0$$

then formally

$$y(x) = f(x) + \sum_{n=1}^{\infty} \left( \lambda \int K \right)^n f$$

Theorem 1 gives the essential result for this method.

*Theorem 1.* If the kernel  $K(x,s)$  is real and continuous in its domain ( $a \leq x \leq b$ ,  $a \leq s \leq b$ ), if  $|K(x,s)| \leq M$  in this domain, if  $f(x)$  is continuous, and if the parameter satisfies  $|\lambda| < 1/[M(b-a)]$ , then the equation

$$y(x) - \lambda \int_a^b K(x,s)y(s) ds = f(x)$$

has a unique continuous solution that is given by

$$y(x) = f(x) + \sum_{n=1}^{\infty} \left( \lambda \int K \right)^n f \quad (120)$$

The series involved in this solution converges absolutely and uniformly.

In the case of the Volterra equation

$$y(x) - \lambda \int_a^x K(x,s)y(s) ds = f(x)$$

the theorem holds without the restriction on the size of  $|\lambda|$ .

**5.32 Fredholm Method.** The Fredholm method evolved from considering a system of linear equation that replaces the integral equation. Divide the range of integration into  $n$  equal parts of length  $h$ , and consider the unknown function and the kernel at  $n$  corresponding values of the independent variable. The system of  $n$  equations in the  $n$  unknowns  $y(x_i)$

$$y(x_i) - \lambda \sum_{j=1}^n hK(x_i,s_j)y(s_j) = f(x_i) \quad i = 1, 2, \dots, n$$

corresponds to the integral equation

$$y(x) - \lambda \int_a^b K(x,s)y(s) ds = f(x)$$

The solution of these linear equations, providing  $\Delta \neq 0$ , is given by

$$y(x_i) = \frac{\sum_{i=1}^n \Delta_{ij} f(x_i)}{\Delta}$$

where  $\Delta \equiv \det (\delta_{ij} - \lambda h K_{ij})$ ,  $K_{ij} = K(x_i, s_j)$ , and  $\Delta_{ij}$  is the first minor of the element in the  $i$ th row and  $j$ th column. The determinant  $\Delta$  can be expanded in the form

$$\Delta = 1 - \lambda h \sum_{i=1}^n K_{ii} + \frac{\lambda^2 h^2}{2!} \sum_{i,j=1}^n \begin{vmatrix} K_{ii} & K_{ij} \\ K_{ji} & K_{jj} \end{vmatrix} + \dots + \frac{(-\lambda h)^n}{n!} \begin{vmatrix} K_{11} & \dots & K_{1n} \\ \vdots & & \vdots \\ K_{n1} & & K_{nn} \end{vmatrix}$$

The limit as  $n$  tends to infinity leads formally to

$$\lim_{n \rightarrow \infty} \Delta \equiv D(\lambda) = 1 + \sum_{j=1}^{\infty} \frac{(-\lambda)^j}{j!} \underbrace{\int \dots \int}_j K \begin{pmatrix} x_1 & x_j \\ x_1 & x_j \end{pmatrix} dx_1 \dots dx_j$$

where

$$K \begin{pmatrix} x_1 & x_j \\ s_1 & s_j \end{pmatrix} \equiv \begin{vmatrix} K(x_1, s_1) & \dots & K(x_j, s_1) \\ \vdots & & \vdots \\ K(x_1, s_j) & & K(x_j, s_j) \end{vmatrix}$$

In a corresponding fashion

$$\lim_{n \rightarrow \infty} \frac{\Delta_{ij}}{h} \equiv D(x,s;\lambda) = \lambda \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{k!} \underbrace{\int \dots \int}_k K \begin{pmatrix} x_1 & \dots & x_k \\ s_{x_1} & \dots & s_{x_k} \end{pmatrix} dx_1 \dots dx_k$$

The solution given above for  $y(x_i)$  can be rewritten in the form

$$y(x_i) = f_i \frac{\Delta_{ij}}{\Delta} + \sum_{i=1}^{n'} f_i \frac{\Delta_{ij}}{\Delta}$$

where the prime indicates that  $i \neq j$ . By starting with this solution, by taking the limit as  $n$  tends to infinity, and by noting that  $x_i$  can be any point of  $a \leq x \leq b$ , it is possible to obtain the formal result

$$y(x) = f(x) + \int_a^b L(x,s,\lambda) f(s) ds$$

where

$$L(x,s,\lambda) = \frac{D(x,s,\lambda)}{D(\lambda)}$$

The two primary results of the Fredholm theory can be expressed in Theorems 2 and 3.

*Theorem 2.* If the kernel  $K(x,s)$  is continuous in its domain, if  $f(x)$  is continuous, and if  $D(\lambda) \neq 0$ , then the equation

$$y(x) - \lambda \int_a^b K(x,s)y(s) ds = f(x)$$

has the unique continuous solution

$$y(x) = f(x) + \int_a^b L(x,s;\lambda)f(s) ds \quad (121)$$

where, as obtained formally before,

$$L(x,s,\lambda) = \frac{D(x,s,\lambda)}{D(\lambda)}$$

The power series  $D(\lambda)$  converges absolutely for all  $\lambda$ , and the power series  $D(x,s,\lambda)$  converges uniformly with respect to  $x$  and  $s$  in its domain and converges absolutely for all  $\lambda$ .

*Theorem 3.* There exists at most a denumerable number of values of  $\lambda$ , called characteristic values or eigenvalues, for which  $D(\lambda) = 0$ . In general no solution will exist for the nonhomogeneous integral equation, that is,  $f(x) \neq 0$ , for  $\lambda_0$  such that  $D(\lambda_0) = 0$ . On the other hand the homogeneous integral equation, that is,  $f(x) = 0$ , has no non-trivial solution except for those  $\lambda_0$  where  $D(\lambda_0) = 0$ . These theorems and other existing results or theorems are analogous to theorems for systems of linear algebraic equation.

**5.33 Hilbert-Schmidt Theory.** The Hilbert-Schmidt theory is developed with the same conditions on the kernel  $K(x,s)$  and the function  $f(x)$  that are used for the Fredholm theory. In addition it is assumed the kernel is symmetric. A *symmetric kernel* satisfies the condition

$$K(x,s) = K(s,x)$$

Before some of the results of Hilbert-Schmidt theory are indicated, the types of common kernel may be noted. A Hermitian kernel satisfies the condition

$$K(x,s) = -\overline{K(s,x)}$$

where the bar indicates the complex conjugate. A *skew-Hermitian kernel* satisfies

$$K(x,s) = -\overline{K(s,x)}$$

A *polar kernel* has the form

$$p(s)K(x,s)$$

where  $K(x,s)$  is a symmetric kernel. Any equation that has a polar kernel may be transformed into an equation with a symmetric kernel. A kernel  $K(x,s)$  is called *positive definite* if

$$\int_a^b \int_a^b f(t)K(t,s)f(s) ds dt > 0$$

for any continuous bounded  $f(x)$ . If the inequality is reversed, the kernel is called *negative definite*. If the equality is also permitted, the kernel is called *semidefinite*.

The following concepts appear in the results of Hilbert-Schmidt theory. A function  $z(x)$  is normalized if

$$\int_a^b z(x)\overline{z(x)} dx = \int_a^b z(x)^2 dx = 1$$

As usual the bar indicates the complex conjugate. Two functions  $z_1(x)$  and  $z_2(x)$  are said to be orthogonal if

$$\int_a^b z_1(x)\overline{z_2(x)} dx = 0$$

A set of functions  $\{z_i(x)\}$  is said to be orthonormal if

$$\int_a^b z_j(x)\overline{z_k(x)} dx = \delta_{jk}$$

where  $\delta_{jk}$  is the Kronecker delta. A set of functions  $z_i(x)$  defined on the interval  $a \leq x \leq b$  is called complete if the only function of the type being considered that is orthogonal to every  $z_i(x)$  is the zero function. A value  $\lambda_0$  such that the homogeneous equation

$$y(x) = \lambda_0 \int_a^b K(x,s)y(s) ds$$

has a nontrivial solution is called a characteristic value or eigenvalue. A corresponding function is called a characteristic function or eigenfunction.

Some of the more important results of Hilbert-Schmidt can be summarized as follows. Given a real, symmetric, continuous, nonzero kernel, the following properties hold:

1. There exists at least one eigenvalue  $\lambda_0$ .
2. All the eigenvalues are real, and the eigenfunction can be assumed real.
3. There exists a complete orthonormal set of eigenfunctions  $\{y_i(x)\}$ .
4. A continuous function  $f(x)$  can be expressed in the form

$$f(x) = \sum_i c_i y_i(x)$$

where

$$c_i = \int_b^a f(s)y_i(s) ds$$

The series if infinite is uniformly convergent on  $a \leq x \leq b$ .

#### 5.4 Calculus of Variation

The methods of calculus of variation are generally separated into direct and indirect methods. In the direct method a suitable approximate variational problem is formulated in which a finite set of  $n$  constants or parameters is to be determined. It then may be possible to let  $n$  tend to infinity in the solution of the approximate problem and thus lead to the solution of the original problem. In the indirect method a related differential equation, often called the Euler equation, is usually obtained. Some solution of the differential equation can then be shown to solve the variational problem. Only some elementary results of the indirect method follow.

**5.41 Euler Equations.** The simplest general problem of the calculus of variation consists of obtaining a function  $y = y(x)$  that gives a minimum (or maximum) value to the integral

$$I = \int_{x_1}^{x_2} f\left(x, y, \frac{dy}{dx}\right) dx = \int_{x_1}^{x_2} f(x, y, y') dx$$

where the values  $x_1$ ,  $x_2$ ,  $y(x_1)$ , and  $y(x_2)$  are given. The given function  $f$  is taken to be twice continuously differentiable with respect to its arguments. Furthermore, the sought function  $y(x)$  is assumed to be twice differentiable.

Let  $y(x)$  denote the minimizing function for the integral  $I$ , and let  $\eta(x)$  be a function defined for  $x_1 \leq x \leq x_2$  that possesses a continuous second derivative and is such that  $\eta(x_1) = 0 = \eta(x_2)$  but is otherwise arbitrary. Consider then the set of *comparison functions*  $Y(x)$  defined by

$$Y(x) = y(x) + \epsilon \eta(x) = y(x) + \delta y(x)$$

where  $\epsilon$  is a parameter. Since  $\eta(x)$  vanishes at the end points  $x_1$  and  $x_2$ ,  $Y(x_1) = y(x_1)$  and  $Y(x_2) = y(x_2)$ . The set of functions  $Y(x)$  have two important properties. First, by proper choice of  $\epsilon$  and  $\eta(x)$  it is possible to represent any suitably differentiable function satisfying the end conditions. Second, no matter what  $\eta(x)$  is chosen,  $y(x)$  is a member of the set for  $\epsilon = 0$ . The quantity  $\delta y = \epsilon \eta(x)$  is called the *variation of the function*  $y(x)$ .

Consider next the integral

$$I(\epsilon) = \int_{x_1}^{x_2} f(x, Y, Y') dx$$

Since letting  $\epsilon = 0$  is seen to be equivalent to replacing  $Y$  and  $dY/dX$  and  $y$  and  $dy/dx$ , the integral  $I(\epsilon)$  is a minimum with respect to  $\epsilon$  when  $\epsilon = 0$ . It follows that a necessary condition for a minimum is given by

$$\left. \frac{d}{d\epsilon} I(\epsilon) \right|_{\epsilon=0} = I'(0) = 0$$

Now

$$I'(\epsilon) = \int_{x_1}^{x_2} \left( \frac{\partial f}{\partial Y} \frac{\partial Y}{\partial \epsilon} + \frac{\partial f}{\partial Y'} \frac{\partial Y'}{\partial \epsilon} \right) dx = \int_{x_1}^{x_2} \left( \frac{\partial f}{\partial Y} \eta + \frac{\partial f}{\partial Y'} \eta' \right) dx$$

and therefore

$$I'(0) = \int_{x_1}^{x_2} \left( \frac{\partial f}{\partial y} \eta + \frac{\partial f}{\partial y'} \eta' \right) dx = 0$$

The integration by parts of this last integral gives

$$I'(0) = \frac{\partial f}{\partial y'} \eta \Big|_{x_1}^{x_2} + \int_{x_1}^{x_2} \eta \left[ \frac{\partial f}{\partial y} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) \right] dx = \int_{x_1}^{x_2} \eta \left[ f_y - \frac{d}{dx} f_{y'} \right] dx = 0$$

which is valid for any of the chosen  $\eta(x)$ .

The *basic or fundamental lemma of the calculus of variation* states the following: If the equation

$$\int_{x_1}^{x_2} \eta(x) H(x) dx = 0 \quad (122)$$

with  $H(x)$  a continuous function, is valid for all functions  $\eta(x)$  that vanish at the end point and are twice continuously differentiable, then  $H(x) = 0$  identically in  $x_1 \leq x \leq x_2$ . Proof of this lemma may be found in the references.

The use of this lemma with the above equation  $I'(0) = 0$  leads to the *Euler (or Euler-Lagrange) differential equation*

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \left( \frac{\partial f}{\partial y'} \right) \equiv f_y - \frac{d}{dx} f_{y'} = 0 \quad (123)$$

or in detail,

$$y'' f_{y'y'} + y' f_{y'y} + f_{y'x} - f_y = 0$$

This is seen to be a second-order differential equation for the unknown function  $y(x)$ . The two constants of integration can be evaluated from the end-point conditions.

The condition  $I'(0) = 0$ , which leads to the Euler equation, is not a sufficient condition for a minimum of  $I(\epsilon)$ . This condition  $I'(0) = 0$  indeed may imply a minimum, a maximum, or a stationary value (or inflection point) for  $I(\epsilon)$ .

The term *extremum* to the value of  $I(\epsilon)$  is applied to all three cases. The Euler differential equation is a necessary condition for the existence of an extremum. Every solution of the Euler equation is called an *extremal* (or *extremizing function*). The sufficient conditions for the minimum (or maximum) of  $I(\epsilon)$  are quite involved and may be found in the references. Often the physical situation indicates the type of extremum obtained.

The expression

$$\delta I = \epsilon I'(0) = \int_{x_1}^{x_2} \left[ f_y - \frac{d}{dx} f_{y'} \right] \delta y dx + f_{y'} \delta y \Big|_{x_1}^{x_2}$$

even if  $\epsilon \eta = \delta y$  does not vanish at the end points, is called the *variation (or first variation)* of the integral  $I$ .

The problem discussed above may be generalized in a number of ways. Consider first the integral

$$I = \int_{t_1}^{t_2} f(t, x_1, \dots, x_n, \dot{x}_1, \dots, \dot{x}_n) dt$$

where the superior dot denotes differentiation with respect to the independent variable  $t$ . Proceeding in a fashion similar to the above problem, the Euler equations become

$$\frac{\partial f}{\partial x_i} - \frac{d}{dt} \left( \frac{\partial f}{\partial \dot{x}_i} \right) \equiv f_{x_i} - \frac{d}{dt} f_{\dot{x}_i} = 0 \quad i = 1, \dots, n$$

For the integral

$$I = \int_{x_1}^{x_2} f[x, y, y', y'', \dots, y^{(n)}] dx$$

the Euler equation becomes

$$f_y - \frac{d}{dx} f_{y'} + \frac{d^2}{dx^2} f_{y''} - \dots + (-1)^n \frac{d^n}{dx^n} f_{y^{(n)}} = 0$$

Finally consider the double integral

$$I = \iint_D f(x, y, v, v_x, v_y) dx dy$$

where  $D$  is the region of integration and  $v$  takes on prescribed values on the boundary of  $D$ . The Euler differential equation becomes

$$\frac{\partial}{\partial x} f_{v_x} + \frac{\partial}{\partial y} f_{v_y} - f_v = 0$$

or in detail

$$f_{v_x v_x} v_{xx} + 2f_{v_x v_y} v_{xy} + f_{v_y v_y} v_{yy} + f_{v_x v_x} v_x + f_{v_y v_y} v_y + f_{v_x x} + f_{v_y y} - f_v = 0$$

**5.42 Lagrange's and Hamilton's Equations.** In classical dynamics, when a system is conservative, a variational principle, called Hamilton's principle, can be used to determine the equations of motion. For a system with  $n$  degrees of freedom it is possible to choose  $n$  independent quantities  $q_1, \dots, q_n$  that specify the configuration of the system. The corresponding velocities are given by  $\dot{q}_1, \dots, \dot{q}_n$  where  $dq/dt = \dot{q}$ . The kinetic energy  $T$  is given by a quadratic function in the  $\dot{q}$ s; that is,

$$T = \frac{1}{2} \sum_{i,j=1}^n a_{ij} \dot{q}_i \dot{q}_j$$

where the  $a_{ij}$ s may be functions of the  $q$ s. For conservative systems the external force is given by the gradient of a scalar potential function or energy  $V$ . The *kinetic potential or Lagrange function*  $L$  is defined by  $L = T - V$ , the difference between the kinetic and potential energy. *Hamilton's principle states that*

$$\delta \int_{t_1}^{t_2} L dt = 0 \tag{124}$$

The Euler equations that result from this principle are

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0 \quad i = 1, 2, \dots, n \tag{125}$$

These equations are customarily called *Lagrange's equations of motion*.

For a conservative system, the *total energy*  $E$ , which is equal to the sum of the kinetic and potential energy, is a constant throughout the motion of the system. When the total energy is expressed in terms of the coordinates  $q$  and the momenta  $p$ , it is called the *Hamiltonian function*  $H$ . The momentum  $p_i$  for the  $i$ th coordinate is given by

$$p_i = \frac{\partial T}{\partial \dot{q}_i} = \sum_{j=1}^n a_{ij} \dot{q}_j$$

The variational principle now gives

$$\delta \int_{t_1}^{t_2} L dt = \delta \int_{t_1}^{t_2} (T - V) dt = \delta \int_{t_1}^{t_2} (2T - H)$$

Without giving the details, the resulting Euler equations, called Hamilton's canonical equations, become

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad i = 1, 2, \dots, n$$

*Auxiliary Conditions.* In many problems a function is sought that renders one integral an extremum and that causes one or more other integrals to take on prescribed values. Variational problems that involve such auxiliary integral conditions are often called *isoperimetric problems*. Consider, for example, minimizing the integral

$$I = \int_{x_1}^{x_2} f(x, y, y') dx$$

subject to the integral

$$K = \int_{x_1}^{x_2} g(x, y, y') dx$$

having a given prescribed value. If a function  $f^*$  is defined by

$$f^* = f + \lambda g$$

where the constant  $\lambda$  is called an *undetermined or Lagrange multiplier*, the Euler equation becomes

$$\frac{\partial f^*}{\partial y} - \frac{d}{dx} \left( \frac{\partial f^*}{\partial y'} \right) = 0$$

The three constants, the two integration constants, and  $\lambda$  can be evaluated from the end conditions and by giving  $K$  its prescribed value.

#### BIBLIOGRAPHY

Some references for various sections follow:

- Art. 1. Refs. 11, 16
  - Art. 2. Refs. 2, 5, 10, 15, 17, 22, 25, 27, 36, 37
  - Art. 3. Refs. 2, 4, 21, 25, 34
  - Art. 4. Refs. 1, 3, 4, 7, 8, 9, 12, 13, 19, 28, 30, 32, 38
  - Art. 5. Refs. 6, 12, 14, 15, 18, 23, 24, 31, 33, 35, 39
  - General: 12, 20, 26, 29, 40
1. Bateman, H.: "Partial Differential Equations of Mathematical Physics," Cambridge University Press, New York, 1932.
  2. Carslaw, H. S.: "Theory of Fourier Series and Integrals," The Macmillan Company, New York, 1930.
  3. Carslaw, H. S., and J. C. Jaeger: "Conduction of Heat in Solids," Oxford University Press, New York, 1947.
  4. Churchill, R. V.: "Fourier Series and Boundary Value Problems," McGraw-Hill Book Company, Inc., New York, 1941.
  5. Churchill, R. V.: "Introduction to Complex Variables and Applications," McGraw-Hill Book Company, Inc., New York, 1948.
  6. Churchill, R. V.: "Modern Operational Mathematics in Engineering," McGraw-Hill Book Company, Inc., New York, 1944.
  7. Coddington, E. A., and N. Levinson: "Theory of Ordinary Differential Equations," McGraw-Hill Book Company, Inc., New York, 1955.