A course in mathematical analysis

Edouard Goursat

March 3, 2025

Contents

Chapter 1

Derivatives and differentials

1.1 Functions of a single variable

1. Limits. When the successive values of a variable x approach nearer and nearer a constant quantity a, in such a way that the absolute value of the difference $x - a$ finally becomes and remains less than any preassigned number, the constant a is called the limit of the variable x. This definition furnishes a criterion for determining whether a is the limit of the variable x. The necessary and sufficient condition that it should be, is that, given positive number ε, no matter how small, the absolute value of $x - a$ should remain less than ε for all values which the variable x can assume, after a certain instant.

Numerous examples of limits are to be found in Geometry and Algebra. For example, the limit of the variable quantity $x = (a^2 - m^2)/(a - m)$, as m approaches a, is $2a$; for $x - 2a$ will be less than ε whenever $m - a$ is taken less than ε. Likewise, the variable $x = a - 1/n$, where n is a positive integer, approaches the limit a when n increases indefinitely; for $a - x$ is less than ε whenever n is greater than $1/\varepsilon$. It is apparent from these examples that the successive values of the variable x, as it approaches its limit, may form a continuous or a discontinuous sequence.

It is in general very difficult to determine the limit of a variable quantity. The following proposition, which we will assume as self-evident, enables us, in many cases, to establish the existence of a limit

Any variable quantity which never decreases, and which always remains less than a constant quantity L, approaches a limit l, which is less than or at most equal to L.

Similarly, *any variable quantity which never increases, and which always remains greater than a constant quantity L', approaches a limit l', which is greater than or else equal to L'.*

For example, if each of an infinite series of positive terms is less, respectively, than the corresponding term of another infinite series of positive terms which is known to converge, then the first series converges also, for the sum $\sum_n$ of the first n terms evidently increases with n, and this sum is constantly less than the total sum of the second series.

2. Functions. When two variable quantities are so related that the value of one of them depends upon the value of the other, they are said to be functions of each other. If one of them be supposed to vary arbitrarily, it is called the *independent variable*. Let this variable be denoted by x, and let us suppose, for example, that it can assume all values between two given numbers a and b $(a < b)$. Let y be another variable, such that to each value of x between a and b, and also for the values a and b themselves, there corresponds one definitely determined value of y. Then y is called a function of of x, defined in the interval (a, b); and this dependence is indicated by writing the equation $y = f(x)$. For instance, it may happen that y is the result of certain arithmetical operations performed upon x. Such is the case for the very simplest functions studied in elementary mathematics, e.g. polynomials, rational functions, radicals, etc.

A function may also be defined graphically. Let two coordinate axes Ox, Oy be taken in a plane; let us join any two points A and B of this plane by a curvilinear arc ACB, of any shape, which is not cut in more than one point by any parallel to the axis Oy. Then the ordinate of a point of this curve will be a function of the abscissa. The arc ACB may be composed of several distinct portions which belong to different curves, such as segments of straight lines, arcs of circles, etc.

In short, any absolutely arbitrary law may be assumed for finding the value of y from that of x. The word *function*, in its most general sense, means nothing more nor less than this: to every value of x corresponds a value of y.

3. Continuity. The definition of functions to which the infinitesimal calculus applies does not admit of such broad generality. Let $y = f(x)$ be a function defined in a certain interval (a, b) and let x_0 and $x_0 + h$ be two values of x in that interval. If the difference $f(x_0 + h) - f(x_0)$ approaches zero as the absolute value of h approaches zero, the function $f(x)$ is said to be continuous for the value x_0. From the very definition of a limit we may also say that *a function $f(x)$ is continuous for $x = x_0$, if, corresponding to every positive number ε, no matter how small, we can find a positive number η, such that*

$$|f(x_0) + h) - f(x_0)| < \varepsilon$$

for every value of h less than η in absolute value.[1] We shall say that a function $f(x)$ is continuous in an interval (a, b) if it is continuous for every value of x lying in that interval, and if the differences

$$f(a + h) - f(a), \qquad f(b - h) - f(b)$$

[1]The notation $|a|$ denotes the absolute value of a

each approaches zero when h which is now to be taken only positive, approaches zero.

In elementary textbooks it is usually shown that polynomials, rational functions, the exponential and the logarithmic function, the trigonometric functions, and the inverse trigonometric functions are continuous functions, except for certain particular values of the variable. It follows directly from the definition of continuity that the sum or the product of any number of continuous functions is itself a continuous function; and this holds for the quotient of two continuous functions also, except for the values of the variable which the denominator vanishes.

It seems superfluous to explain here the reasons which lead us to assume that functions which are defined by physical conditions are, at least in general, continuous.

Among the properties of continuous functions we shall now state only the two following, which one might be tempted to think were self-evident, but which really amount to actual theorems, of which rigorous demonstrations will be given later.[2]

1. *If the function $y = f(x)$ is continuous in the interval (a, b), and if N is a number between $f(a)$ and $f(b)$, then the equation $f(x) = N$ has at least one root between a and b.*

2. *There exists at least one value of x belonging to the interval (a, b), inclusive of its end points, for which y takes on a value M which is greater than, or at least equal to, the value of the function at any other point in the interval. Likewise, there exists a value of x for which y takes on a value m, than which the function assumes no smaller value in the interval.*

The numbers M and m are called the maximum and the minimum values of $f(x)$, respectively, in the interval (a, b). It is clear that the value of x for which $f(x)$ assumes its maximum value M, or value of x corresponding to the minimum m, may be at one end points, a or b. It follows at once from the two theorems above, that if N is a number between M and m, the equation $f(x) = N$ has at least one root which lies between a and b.

4. Examples of discontinuities. The functions which we shall study will be in general continuous, but they may cease to be so for certain exceptional values of the variable. We proceed to give several examples of the kinds of discontinuity which occur most frequently.

The function $y = 1/(x - a)$ is continuous for every value x_0 of x except a. The operation necessary to determine the value of y from that of x ceases to have a meaning when x is assigned the value a; but we note that when x is very near to a the absolute value of y is very large, and y is positive or

[2]See Chapter **??**

negative with $x - a$. As the difference $x - a$ diminishes, the absolute value of y increases indefinitely, so as eventually to become and remain greater than any preassigned number. This phenomenon is described by saying that y becomes infinite when $x = a$. Discontinuity of this kind is of great importance in Analysis.

Let us consider next the function $y = \sin 1/x$. As x approaches zero, $1/x$ increases indefinitely, and y does not approach any limit whatever, although it remains between $+1$ and -1. The equation $\sin 1/x = A$. where $|A| < 1$, has an infinite number of solutions which lie between 0 and ε no matter how small ε be taken. Whatever value be assigned to y when $x = 0$, the function under consideration cannot be made continuous for $x = 0$.

An example of a still different kind of discontinuity is given by the convergent infinite series

$$S(x) = x^2 + \frac{x^2}{1+x^2} + \cdots + \frac{x^2}{(1+x^2)^n} + \ldots$$

When x approaches zero, $S(x)$ approaches the limit 1, although $S(0) = 0$. For, when $x = 0$ every term of the series is zero, and hence $S(0) = 0$. But if x be given a value different from zero, a geometric progression is obtained, of which the ratio is $1/(1 - x^2)$. Hence

$$S(x) = \frac{x^2}{1 - \frac{1}{1+x^2}} = \frac{x^2(1+x^2)}{x^2} = 1 + x^2;$$

and the limit of $S(x)$ is seen to be 1. Thus, in this example, the function approaches a definite limit as x approaches zero, but that limit is different from the value of the function for $x = 0$.

5. Derivatives. Let $f(x)$ be a continuous function. Then the two terms of the quotient

$$\frac{f(x+h) - f(x)}{h}$$

approach zero simultaneously, as the absolute value of h approaches zero, while x remains fixed. If this quotient approaches a limit, this limit is called the derivative of the function $f(x)$, and is denoted by y', or by $f'(x)$, in the notation due to Lagrange.

An important geometrical concept is associated with this analytic notion of derivative. Let us consider, in a plane XOY, the curve AMB, which represents the function $y = f(x)$, which we shall assume to be continuous in the interval (a, b). Let M and M' be two points on this curve, in the interval (a, b), and let their abscissae be x and and $x + h$, respectively. The slope of the straight line MM' is then precisely the quotient above. Now as h approaches zero the point M' approaches the point$M$; and, if the function has a derivative, the slope of the line MM' approaches the limit y'. The straight line MM', therefore, approaches a limiting position, which is called the *tangent to the curve.* It follows that the equation of the tangent is

$$Y - y = y'(X - x),$$

where X and Y are the running coordinates.

To generalize, let us consider any curve in space, and let

$$x = f(t), \quad y = \phi(t), \quad z = \psi(t)$$

be the coordinates of a point on the curve, expressed as functions of a variable parameter t. Let M and M' be two points of the curve corresponding to two values, t and $t + h$, of the parameter. The equations of the chord MM' are then

$$\frac{X - f(t)}{f(t+h) - f(t)} = \frac{Y - \phi(t)}{\phi(t+h) - \phi(t)} = \frac{Z - \psi(t)}{\psi(t+h) - \psi(t)}$$

If we divide each denominator by h and then let h approach zero, the chord MM' evidently approaches a limiting position, which given by the equations

$$\frac{X - f(t)}{f'(t)} = \frac{Y - \phi(t)}{\phi'(t)} = \frac{Z - \psi(t)}{\psi'(t)}$$

provided, of course, that each of the three functions $f(x)$, $\phi(x)$, $\psi(x)$ possesses a derivative. The determination of the tangent to a curve thus reduces, analytically, to the calculation of derivatives.

Every function which possesses a derivative is necessarily continuous, but the converse is not true. It is easy to give examples of continuous functions which do not possess derivatives for particular values of the variable. The function $y = x \sin 1/x$, for example, is a perfectly continuous function of x, for $x = 0$,[3] and y approaches zero as x approaches zero. But the ratio $y/x = \sin 1/x$ does not approach any limit whatever, as we have already seen.

Let us next consider the function $y = x^{\frac{2}{3}}$. Here y is continuous for every value of x; and $y =$ when $x = 0$. But the ratio $y/x - x^{-\frac{1}{3}}$ increases indefinitely as x approaches zero. For abbreviation the derivative is said to be infinite for $x = 0$; the curve which represents the function is tangent to the axis of y at the origin.

Finally, the function

$$y = \frac{xe^{\frac{1}{x}}}{1 + e^{\frac{1}{x}}}$$

is continuous at $x = 0$,but the ratio y/x approaches two different limits according as x is always positive or always negative while it is approaching

[3] After the value zero has been assigned to y for $x = 0$.

zero. When x is positive and small, $e^{1/x}$ is positive and very large, and the ratio y/x approaches 1. But if x is negative and very small in absolute value, $e^{1/x}$ is very small, and the ratio y/x approaches zero. There exist then two values of the derivative according to the manner in which x approaches zero: the curve which represents this function has a *corner* at the origin.

It is clear from these examples that there exist continuous functions which do not possess derivatives for particular values of the variable. But the discoverers of the infinitesimal calculus confidently believed that a continuous function had a derivative in general. Attempts at proof were even made, but these were, of course, fallacious. Finally, Weierstrass succeeded in settling the question conclusively by giving examples of continuous functions which do not possess derivatives for any values of the variable whatever.[4] But as these functions have not as yet been employed in any applications, we them shall not consider them here. In the future, when we say that a function $f(x)$ has a derivative in the interval (a, b), we shall mean that it has an *unique finite derivative* for every value of x between a and b and also for $x = a$ (h being positive) and for $x = b$ (h being negative), unless an explicit statement made to the contrary.

6. Successive derivatives. The derivative of a function $f(x)$ is in general another function of x, $f'(x)$. If $f'(x)$ in turn has a derivative the new function is called the *second derivative* of $f(x)$, and is represented by y'' or by $f''(x)$. In the same way the third derivative y''', of $f'''(x)$, is defined to be the derivative of the second, and so on. In general, the nth derivative $y^{(n)}$, or $f^{(n)}(x)$, is the derivative of the derivative of order $(n-1)$. If, in thus forming the successive derivatives, we never obtain a function which has no derivative, we may imagine the process carried on indefinitely. In this way we obtain an unlimited sequence of derivatives of the function $f(x)$ with which we started. Such is the case for all functions which have found any considerable application up to the present time.

The above notation is due to Lagrange. The notation $D_n y$ or $D_n f(x)$, due to Cauchy, is also used occasionally to represent the nth derivative. Leibniz' notation will be given presently.

7 Rolle's theorem. The use of derivatives in the study of equations depends upon the following proposition, which is known a *Rolle's Theorem*:

Let a and b be two roots of the equation $f(x) = 0$. If the function $f(x)$ is continuous and possesses a derivative in the interval (a, b), the equation $f'(x) = 0$ has at least one root which lies between a and b.

For the function $f(x)$ vanishes, by hypothesis, for $x = a$ and $x = b$. If it vanishes at every point of the interval (a, b), its derivative also vanishes at every point of the interval, and the theorem is evidently fulfilled. If the

[4]Note read at the Academy of Sciences of Berlin, July 18, 1872. Other examples are to be found in the memoir by Darboux on discontinuous functions (*Annales de l'Ecole Normale Superieure*, Vol. IV, 2d series). One of Weierstrass's examples is given later (Chapter **??**).

function $f(x)$ does not vanish throughout the interval, it will assume either positive or negative values at some points. Suppose, for instance, that it has positive values. Then it will have a maximum value M for some value of x, say x_1, which lies between a and b (§ 3, Theorem II). The ratio

$$\frac{f(x_1+h)-f(x_1)}{h}$$

where h is taken positive, is necessarily negative or else zero. Hence the limit of this ratio, i.e. $f'(x_1)$, cannot be positive; i.e. $f'(x_1) \leq 0$. But if we consider $f'(x_1)$ as the limit of the ratio

$$\frac{f(x_1-h)-f(x_1)}{-h}$$

where h is positive, it follows in the same manner that $f'(x_1) \geq 0$. From these two results it is evident that $f'(x_1) = 0$.

8. Law of the mean. It is now easy to deduce from the above theorem the important law of the mean:[5]

Let $f(x)$ be a continuous function which has a derivative in the interval (a, b). Then

$$f(b) - f(a) = (b-a)f'(c), \tag{1.1}$$

where c is a number between a and b.

In order to prove this formula, let $\phi(x)$ be another function which has the same properties as $f(x)$, i.e. it is continuous and possesses a derivative in the interval (a, b). Let us determine three constants, A, B, C such that the auxiliary function

$$\psi(x) = Af(x) + B\phi(x) + C$$

vanishes for $x = a$ and for $x = b$. The necessary and sufficient conditions for this are

$$Af(a) + B\phi(a) + C = 0, \quad Af(b) + B\phi(b) + C = 0;$$

and these are satisfied if set

$$A = \phi(a) - \phi(b), \quad B = f(b) - f(a), \quad C = f(a)\phi(b) - f(b)\phi(a).$$

The new function $\psi(x)$ thus defined is continuous and has a derivative In the interval (a, b). The derivative $\psi'(x) = Af'(x) + B\phi'(x)$ therefore vanishes

[5]"Formule de accrissemets finis." The French also use "Formule de la moyenne" as a synonym. Other English synonyms are "Average value theorem" and "Mean value theorem".

for some value c which lies between a and b, whence, replacing A and B by their values, we find a relation of the form

$$[\phi(a) - \phi(b)]f'(c) = [f(b) - f(a)]\phi'(c). \tag{1.1'}$$

It is merely necessary to take $\phi(x) = x$ in order to obtain the equality which was to be proved. It is to be noticed that this demonstration does not presuppose the continuity of the derivative $f'(x)$.

From the theorem just proven it follows that if the derivative $f'(x)$ zero at each point of the interval (a, b), the function $f(x)$ has the same value at every point of the interval; for the application of the formula to two values x_1, x_2 belonging to the interval (a, b), gives $f(x_1) = f(x_2)$. Hence, if two functions have the same derivative, their difference is a constant; and the converse is evidently true also. *If a function $F(x)$ be given whose derivative is $f(x)$, all other functions which have the same derivative are found by adding to $F(x)$ an arbitrary constant.*[6]

The geometrical interpretation of the equation (1.1) is very simple. Let us draw the curve AMB which represents the function $y = f(x)$ in the interval (a, b). Then the ratio $[f(b) - f(x)]/(b - a)$ is the slope of the chord AB, while $f'(c)$ is the slope of the tangent at a point C of the curve whose abscissa is c. Hence the equation (1.1) expresses the fact that *there exists a point C on the curve AMB, between A and B, where the tangent is parallel to the chord AB.*

If the derivative $f(x)$ is continuous, and if we let a and b approach the same limit x_0 according to any law whatever, the number c, which lies between a and b, also approaches x_0, and the equation (1.1) shows that the limit of the ratio

$$\frac{f(b) - f(x)}{b - a}$$

is $f'(x_0)$. The geometrical interpretation is as follows. Let us consider upon the curve $y = f(x)$ a point M whose abscissa is x_0, and two points A and B whose abscissae are a and b, respectively. The ratio $[f(b) - f(a)]/(b - a)$ is

[6]This theorem is sometimes applied without due regard to the conditions imposed in its statement. Let $f(x)$ and $\phi(x)$, for example, be two continuous functions which have derivatives $f'(x)$, $\phi'(x)$ in an interval (a, b). If the relation $f'(x)\phi(x) - f(x)\phi'(x) = 0$ is satisfied by these four functions, it is sometimes accepted as proved that the derivative of the function f/ϕ, or

$$\frac{f'(x)\phi(x) - f(x)\phi'(x)}{\phi(^2}$$

is zero, and that accordingly f/ϕ is constant in the interval (a, b) But this conclusion is not absolutely rigorous unless the function $\phi(x)$ does not vanish in the interval (a, b). Suppose, for instance, that $\phi(x)$ and $\phi'(x)$ vanish for a value c between a and b. A function $f(x)$ equal to $C_1\phi(x)$ between a and c, and to $C_2\phi(x)$ between c and b, where C_1 and C_2 are different constants, is continuous and has a derivative in the interval (a, b), and we have $f'(x)\phi(x) - f(x)\phi'(x) = 0$ for every value of x in the interval. The geometrical interpretation is apparent.

equal to the slope of the chord AB, while $f'(x_0)$ is the slope of the tangent at M. Hence, when the two points A and B approach the point M according to any law whatever, the secant AB approaches, as its limiting position, the tangent at the point M.

This does not hold in general however, if the derivative is not continuous. For instance, if two points be taken on the curve $y = x^{\frac{2}{3}}$, on opposite sides of the y axis, it is evident from a figure that the direction of the secant joining them can be made to approach any arbitrarily assigned limiting value by causing the two points to approach the origin according to a suitably chosen law.

The equation $(1.1')$ is sometimes called the *generalized law of the mean.* From it de l'Hospital's theorem on indeterminate forms follows at once. For, suppose $f(a) = 0$ and $\phi(x) = 0$. Replacing b by x in $(1.1')$, we find

$$\frac{f(x)}{\phi(x)} = \frac{f'(x_1)}{\phi'(x_1)}$$

where x_1 lies between a and x. This equation shows that *if the ratio* $f'(x)/\phi'(x)$ *approaches a limit as* x *approaches* a*, the ratio* $f(x)/\phi(x)$ *approaches the same limit, if* $f(a) = 0$ *and* $\phi(a) = 0$.

9. Generalizations of the law of the mean. Various generalizations of the law of the mean have been suggested. The following one is due to Stieltjes (*Bulletin de la Societe Mathematique*, Vol. XVI, p. 100). For the sake of definiteness consider three functions, $f(x)$, $g(x)$, $h(x)$, each of which has derivatives of the first and second orders. Let a, b, c be three particular values of the variable $(a < b < c)$. Let A be a number defined by the equation

$$\begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f(c) & g(c) & h(c) \end{vmatrix} - A \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} = 0,$$

and let

$$\phi(x) = \begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f(x) & g(x) & h(x) \end{vmatrix} - A \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & x & x^2 \end{vmatrix}$$

be an auxiliary function. Since this function vanishes when $x = b$ and when $x = c$, its derivative must vanish for some value ζ between b and c. Hence

$$\begin{vmatrix} f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \\ f'(\zeta) & g'(\zeta) & h'(\zeta) \end{vmatrix} - A \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 0 & 1 & 2\zeta \end{vmatrix} = 0$$

If b be replaced by x in the left-hand side of this equation, we obtain a function which vanishes when $x = a$ and when $x = b$. Its derivative therefore vanishes for some value of x between a and b which we shall call ξ. The new equation thus obtained is

$$\begin{vmatrix} f(a) & g(a) & h(a) \\ f'(\xi) & g'(\xi) & h'(\xi) \\ f'(\zeta) & g'(\zeta) & h'(\zeta) \end{vmatrix} - A \begin{vmatrix} 1 & a & a^2 \\ 0 & 1 & 2\xi \\ 0 & 1 & 2\zeta \end{vmatrix} = 0.$$

Finally, replacing ζ by x in the left-hand side of this equation, we obtain a function of x which vanishes when $x = \xi$ and when $x = \zeta$. Its derivative vanishes for some value η, which lies between ξ and ζ and therefore between a and c. Hence A must have the value

$$A = \frac{1}{1 \cdot 2} \begin{vmatrix} f(a) & g(a) & h(a) \\ f'(\xi) & g'(\xi) & h'(\xi) \\ f''(\eta) & g''(\eta) & h''(\eta) \end{vmatrix}$$

where ξ lies between a and b, and η lies between a and c.

This proof does not presuppose the continuity of the second derivatives $f''(x), g''(x), h''(x)$. If these derivatives are continuous, and if the values a, b, c approach the same limit x_0, we have, in the limit,

$$\lim A = \frac{1}{1 \cdot 2} \begin{vmatrix} f(x_0) & g(x_0) & h(x_0) \\ f'(x_0) & g'(x_0) & h'(x_0) \\ f''(x_0) & g''(x_0) & h''(x_0) \end{vmatrix}.$$

Analogous expressions exist for n functions and the proof follows the same lines. If only two functions $f(x)$ and $g(x)$ are taken, the formulae reduce to the law of the mean if we set $g(x) = 1$.

An analogous generalization has been given by Schwarz (*Annali di Mathematica*, 2d series, Vol. X).

1.2 Functions of several variables

10. Introduction. A variable quantity ω whose value depends on the values of several other variables, $x, y, z, \ldots, t$, which are independent of each other, is called *a function of the independent variables, $x, y, z, \ldots, t$;* and this relation is denoted by writing $\omega = f(x, y, z, \ldots, t)$. For definiteness, let us suppose that $\omega = f(x, y)$ is a function of the two independent variables x and y. If we think of x and y as the Cartesian coordinates of a point in the plane, each pair of values (x, y) determines a point of the plane, and conversely. If to each point of a certain region A in the xy plane, bounded by one or more contours of any form whatever, there corresponds a value of ω, the function $f(x, y)$ is said to be defined in the region A.

Let (x_0, y_0) be the coordinates of a point M_0 lying in this region.

The function $f(x,y)$ *is said to be continuous for the pair of values* (x_0, y_0) *if, corresponding to any preassigned positive number* ε*, another positive number* η *exists such that*

$$|f(x_0+h, y_0+k) - f(x_0, y_0)| < \varepsilon$$

whenever $|h| < \eta$ *and* $|k| < \eta$.

This definition of continuity may be interpreted as follows. Let us suppose constructed in the xy plane a square of side 2η about M_0 as center, with its sides parallel to the axes. The point M' whose coordinates are $x_0 + h$, $y_0 + k$ will lie inside this square, if $|h| < \eta$ and $|k| < \eta$. To say that the function is continuous for the pair of values (x_0, y_0) amounts to saying that by taking this square sufficiently small we can make the difference between the value of the function at M_0 and its value at any other point of the square less than ε in absolute value.

It is evident that we may replace the square by a circle about (x_0, y_0) as center. For, if the above condition is satisfied for all points inside a square, it will evidently be satisfied for all points inside the inscribed circle. And conversely, if the condition is satisfied for all points inside a circle, it will also be satisfied for all points inside the square inscribed in that circle. We might then define continuity by saying that an η exists for every ε, such that whenever $\sqrt{h^2+k^2} < \eta$ we also have

$$|f(x_0+h, y_0+k) - f(x_0, y_0)| < \varepsilon$$

The definition of continuity for a function of $3, 4, \ldots, n$ independent variables is similar to the above.

It is clear that any continuous function of the two independent variables x and y is a continuous function of each of the variables taken separately. However, the converse does not always hold.[7]

11. Partial derivatives. If any constant value whatever be substituted for y, for example, in a continuous function $f(x,y)$, there results a continuous function of the single variable x. The derivative of this function of x, if it exists, is denoted by $f_x(x,y)$ or by ω_x. Likewise the symbol ω_y, or $f_y(x,y)$, is used to denote the derivative of the function $f(x,y)$ when x is regarded as constant and y as the independent variable. The functions $f_x(x,y)$ and $f_y(x,y)$ are called *the partial derivatives* of the function $f(x,y)$. They are themselves, in general, functions of the two variables x and y. If we form

[7] Consider, for instance, the function $f(z,y)$, which is equal to $2xy/(x^2+y^2)$ when two two variables x and y are not both zero, and which is zero when $x = 0$ and $y = 0$. It is evident that this is a continuous function of x when y is constant, and *vice versa*. Nevertheless it not a continuous function of the two independent variables x and y for the pair of values $x = 0$, $y = 0$, For, if the point (x,y) approaches the origin upon line $x = y$, the function $f(x,y)$ approached the limit 1 and not zero. Such functions have been studied by Baire in his thesis.

their partial derivatives in turn, we get the partial derivatives of the second order of the given function $f(x, y)$. Thus there are four partial derivatives of the second order, $f_{x^2}(x, y)$, $f_{xy}(x, y)$, $f_{yx}(x, y)$, $f_{y^2}(x, y)$. The partial derivatives of the third, fourth, and higher orders are defined similarly. In general, given a function $\omega(x, y, z, \ldots, t)$ of any number of independent variables, a partial derivative of the nth order is the result of n successive differentiations of the function f, in a certain order, with respect to any of the variables which occur in f. We will now show that the result does not depend upon the order in which the differentiations are carried out.

Let us first prove the following lemma:

Let $\omega = f(x, y)$ be a function of the two variables x and y. Then $f_{xy} = f_{yx}$ provided that these two derivatives are continuous.

To prove this let us first write the expression

$$U = f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) - f(x + \Delta x, y) + f(x, y)$$

in two different forms, where we suppose that $x, y, \Delta x, \Delta y$ have definite values. Let us introduce the auxiliary function

$$\phi(v) = f(x + \Delta x, v) - f(x, v)$$

where v is an auxiliary variable. Then we may write

$$U = \phi(y + \Delta y) - \phi(y)$$

Applying the law of the mean to the function $\phi(v)$, we have

$$U = \Delta y \phi_y(y + \theta \Delta y), \quad \text{where } 0 < \theta < 1;$$

or, replacing ϕ_y by its value,

$$U = \Delta y[f_y(x + \Delta x, y + \theta \Delta y) - f_y(x, y + \theta \Delta y)].$$

If we now apply the law of the mean to the function $f_y(u, y+\theta\Delta y)$ regarding u as the independent variable, we find

$$U = \Delta x \Delta y f_{yx}(x + \theta' \Delta x, y + \theta \Delta y), \quad 0 < \theta' < 1.$$

From the symmetry of the expression U in $x, y, \Delta x, \Delta y$, we see that we would also have, interchanging x and y,

$$U = \Delta y \Delta x f_{xy}(x + \theta_1' \Delta x, y + \theta_1 \Delta y),$$

where θ_1, and θ_1' are again positive constants less than unity. Equating these two values of U and and dividing by $\Delta x \Delta y$, we have

$$f_{xy}(x + \theta_1' \Delta x, y + \theta_1 \Delta y) = f_{yx}(x + \theta' \Delta x, y + \theta \Delta y).$$

Since the derivatives $f_{xy}(x,y)$ and $f_{yx}(x,y)$ are supposed continuous, the two members of the above equation approach $f_{xy}(x,y)$ and $f_{yx}(x,y)$ respectively, as Δx and Δy approach zero, and we obtain the theorem which we wished to prove.

It is to be noticed in the above demonstration that no hypothesis whatever is made concerning the other derivatives of the second order, f_{x^2} and f^{y^2}. The proof applies also to the case where the function $f(x,y)$ depends upon any number of other independent variables besides x and y, since these other variables would merely have to be regarded as constants in the preceding developments.

Let us now consider a function of any number of independent variables,

$$\omega = f(x, y, z, \dots, t),$$

and let Ω be a partial derivative of order n of this function. Any permutation in the order of the differentiations which leads to Ω can be effected by a series of interchanges between two successive differentiations; and, since these interchanges do not alter the result as we have just seen, the same will be true of the permutation considered. It follows that in order to have a notation which is not ambiguous for the partial derivatives of the nth order, it is sufficient to indicate the number of differentiations performed with respect to each of the independent variables. For instance, any nth derivative of a function of three variables, $\omega(x,y,z)$, will be represented by one or the other of the notations

$$f_{x^p y^q z^r}(x,y,z), \quad D^n_{x^p y^q z^r} f(x,y,z),$$

where $p+q+r=n$.[8] Either of these notations represents the result of differentiating f successively p times with respect to x, q times with respect to y and r times with respect to z, these operations being carried out in any order whatever. There are three distinct derivatives of the first order, f_x, f_y, f_z; six of the second order, f_{x^2}, f_{y^2}, f_{z^2}, f_{xy}, f_{yz}, f_{xz}; and so on.

In general, a function of p independent variables has just as many distinct derivatives of order n as there are distinct terms in a homogeneous polynomial of order n in p independent variables; that is,

$$\frac{(n+1)(n+2)\dots(n+p-1)}{1\cdot 2\cdot \dots (p-2)(p-1)}$$

as is shown in the theory of combinations.

Practical rules. A certain number of practical rules for the calculation of derivatives are usually derived in elementary books on the Calculus. A

[8]The notation $f_{x^p y^q z_r}(x,y,z)$ in used instead of the notation $f^{(n)}_{x^p, y^q z_r}(x,y,z)$ for simplicity. Thus the notation $f_{xy}(x,y)$, used in place of $f''_{xy}(x,y)$, is simpler and equally clear.

table of such rules is appended, the function and its derivative being placed on the same line:

$$y = x^a, \quad y' = ax^{a-1};$$

$$y = a^x, \quad y' = a^x \log a,$$

where the symbol log denotes the natural logarithm;

$$y = \log x, \quad y' = \frac{1}{x};$$

$$y = \sin x, \quad y' = \cos x;$$

$$y = \cos x, \quad y' = -\sin x,$$

$$y = \arcsin x, \quad y' = \frac{1}{\pm\sqrt{1-x^2}};$$

$$y = \arctan x, \quad y' = \frac{1}{1+x^2};$$

$$y = uv, \quad y' = u'v + uv';$$

$$y = \frac{u}{v}, \quad y' = \frac{u'v - uv'}{v^2};$$

$$y = f(u), \quad y_x = f'(u)u_x;$$

$$y = f(u, v, w), \quad y_x = u_x f_u + v_x f_v + w_x f_w.$$

The last two rules enable us to find the derivative of a function of a function and that of a composite function if f_u, f_v, f_w are continuous. Hence we can find the successive derivatives of the functions studied in elementary mathematics, – polynomials, rational and irrational functions, exponential and logarithmic functions, trigonometric functions and their inverses, and the functions derivable from all of these by combination.

For functions of several variables there exist certain formulae analogous to the law of the mean. Let us consider, for definiteness, a function $f(x, y)$ of the two independent variables x and y. The difference $f(x+h, y+k) - f(x, y)$ may be written in the form

$$\begin{aligned} f(x+h, y+k) - f(x, y) = {} & [f(x+h, y+k) - f(x, y+k)] \\ & + [f(x, y+k) - f(x, y)] \end{aligned}$$

to each part of which we may apply the law of the mean. We thus find

$$f(x+h, y+k) - f(x, y) = hf_x(x+\theta h, y+k) + kf_y(x, y+\theta' k),$$

where θ and θ' each lie between zero and unity.

This formula holds whether the derivatives f_x and f_y are continuous or not. If these derivatives are continuous, another formula, similar to the above, but involving only one undetermined number θ, may be employed.[9] In order to derive this second formula consider the auxiliary function $\phi(t) = f(x+ht, y+kt)$, where x, y, h and k have determinate values and t denotes an auxiliary variable. Applying the law of the mean to this function, we find

$$\phi(1) - \phi(0) = \phi'(\theta), \quad 0 < \theta < 1.$$

Now $\phi(t)$ is a composite function of t, and its derivative $\phi'(t)$ is equal to $hf_x(x+ht, y+kt) + kf_y(x+ht, y+kt)$; hence the preceding formula may be written in the form

$$f(x+h, y+k) - f(x, y) = hf_x(x+\theta h, y+\theta k) + kf_y(x+\theta h, y+\theta k).$$

12. Tangent plane to a surface. We have seen that the derivative of a function of a single variable gives the tangent to a plane curve. Similarly, the partial derivatives of a function of two variables occur in the determination of the tangent plane to a surface. Let

$$z = F(x, y) \tag{1.2}$$

be the equation of a surface S, and suppose that the function $F(x, y)$, together with its first partial derivatives, is continuous at a point (x_0, y_0) of the xy plane. Let z_0 be the corresponding value of z, and $M_0(x_0, y_0, z_0)$ the corresponding point on the surface S. If the equations

$$x = f(t), \quad y = \phi(t), \quad z = \psi(t) \tag{1.3}$$

represent a curve C on the surface S through the point M_0, the three functions $f(t)$, $\phi(t)$, $\psi(t)$, which we shall suppose continuous and differentiable, must reduce to x_0, y_0, z_0, respectively, for some value t_0 of the parameter t. The tangent to this curve at the point M_0 is given by the equations (§ 5)

$$\frac{X - x_0}{f'(t_0)} = \frac{Y - y_0}{\phi'(t_0)} = \frac{Z - z_0}{\psi'(t_0)}. \tag{1.4}$$

Since the curve C lies on the surface S, the equation $\psi(t) = F[f(t), \phi(t)]$ must hold for all values of t; that is, this relation must be an identity in t.

[9]Another formula may be obtained which involves only one undetermined number θ, and which holds even when the derivatives f_x and f_y are discontinuous. For the application of the mean to the auxiliary function $\phi(t) = f(x+ht, y+k) + f(x, y+kt)$ gives

$$\phi(1) - \phi(0) = \phi'(\theta), \quad 0 < \theta < 1,$$

or

$$f(x+h, y+k) - f(x, y) = hf_x(x+\theta h, y+k) + kf_y(x, y+\theta k)$$

The operations performed, and hence the final formula, all hold provided the derivatives f_x and f_y merely exist at the points $(x+ht, y+k)$, $(x, y+kt)$, $0 \le t \le 1$.

Taking the derivative of the second member by the rule for the derivative of a composite function, and setting $t = t_0$, we have

$$\psi'(t_0) = f'(t_0)F_{x_0} + \phi'(t_0)F_{y_0}. \tag{1.5}$$

We can now eliminate $f'(t_0)$, $\phi'(t_0)$, $\psi'(t_0)$ between the equations (1.4) and (1.5), and the result of this elimination is

$$Z - z_0 = (X - x_0)F_{x_0} + (Y - y_0)F_{y_0}. \tag{1.6}$$

This is the equation of a plane which is the locus of the tangents to all curves on the surface through the point M_0. It is called *the tangent plane to the surface.*

13. Passage from increments to derivatives. We have defined the successive derivatives in terms of each other, the derivatives of order n being derived from those of order $(n-1)$ and so forth. It is natural to inquire whether we may not define a derivative of any order as the limit of a certain ratio directly, without the intervention of derivatives of lower order. We have already done something of this kind for f_{xy}; (§ 11) for the demonstration given above shows that f_{xy} is the limit of the ratio

$$\frac{f(x+\Delta x, y+\Delta y) - f(x+\Delta x, y) - f(x, y+\Delta y) + f(x,y)}{\Delta x \Delta y}$$

as Δx and Δy both approach zero. It can be shown in like manner that the second derivative f'' of a function $f(x)$ of a single variable is the limit of the ratio

$$\frac{f(x+h_1+h_2) - f(x+h_1) - f(x+h_2) + f(x)}{h_1 h_2}$$

as h_1 and h_2 both approach zero.

For, let us set

$$f_1(x) = f(x+h_1) - f(x)$$

and then write the above ratio in the form

$$\frac{f_1(x+h_2) - f_1(x)}{h_1 h_2} = \frac{f_1'(x+\theta h_2)}{h_1}, \quad 0 < \theta < 1;$$

or

$$\frac{f'(x+h_1+\theta h_2) - f'(x\theta h_2)}{h_1} = f''(x+\theta' h_1 + \theta h_2), \quad 0 < \theta' < 1.$$

The limit of this ratio is therefore the second derivative f'', provided that derivative is continuous.

Passing now to the general case, let us consider, for definiteness, a function of three independent variables, $\omega = f(x, y, z)$. Let us set

$$\begin{aligned}\Delta_x^h \omega =& f(x+h, y, x) - f(x, y, z),\\ \Delta_y^k \omega =& f(x, y+k, z) - f(x, y, z),\\ \Delta_z^l \omega =& f(x, y, z+l) - f(x, y, z),\end{aligned}$$

where $\Delta_x^h\omega$, $\Delta_y^k\omega$, $\Delta_z^l\omega$ are the *first increments* of ω. If we consider h, k, l as given constants, then these three first increments are themselves functions of x, y, z, and we may form the relative increments of these functions corresponding to increments h_1, k_1, l_1 of the variables. This gives us the second increments, $\Delta_x^{h_1}\Delta_x^h\omega, \Delta_x^{h_1}\Delta_y^k\omega, \ldots$. This process can be continued indefinitely; an increment of order n would be defined as a first increment of an increment of order $(n-1)$. Since we may invert the order of any two of these operations, it will be sufficient to indicate the successive increments given to each of the variables. An increment of order n would be indicated by some such notation as the following:

$$\Delta^{(n)}\omega = \Delta_x^{h_1}\Delta_x^{h_2}\ldots\Delta_x^{h_p}\Delta_y^{k_1}\ldots\Delta_y^{k_q}\Delta_z^{l_1}\ldots\Delta_x^{l_r}f(x,y,z)$$

where $p+q+r=n$ and where the increments h, k, l may be either equal or unequal. This increment may be expressed in terms of a partial derivative of order n, being equal to the product

$$h_1h_2\ldots h_pk_1\ldots k_ql_1\ldots l_r \\ \times f_{x^py^qz^r}(x+\theta_1h_1+\ldots\theta_ph_p, y+\theta_1'k_1+\ldots\theta_q'k_q, z+\theta_1''l_1+\ldots\theta_q''l_r),$$

where every θ lies between 0 and 1. This formula has already been proved for first and for second increments. In order to prove it in general, let us assume that it holds for an increment of order $(n-1)$, and let

$$\phi(x,y,z) = \Delta_x^{h_2}\ldots\Delta_x^{h_p}\Delta_y^{k_1}\ldots\Delta_y^{k_q}\Delta_z^{l_1}\ldots\Delta_x^{l_r}f(x,y,z).$$

Then, by hypothesis,

$$\phi(x,y,z) = h_2\ldots h_pk_1\ldots k_ql_1\ldots l_rf_{x^{p-1}y^qz^r}(x+\theta_1h_1+\ldots\theta_ph_p, y+\ldots, z+\ldots).$$

But the nth increment considered is equal to $\phi(x+h_1,y,z)-\phi(x,y,z)$; and if we apply the law of the mean to this increment, we finally obtain the formula sought.

Conversely, the partial derivative $f_{x^py^qz^r}$ is the limit the limit of the ratio

$$\frac{\Delta_x^{h_1}\Delta_x^{h_2}\ldots\Delta_x^{h_p}\Delta_y^{k_1}\ldots\Delta_y^{k_q}\Delta_z^{l_1}\ldots\Delta_x^{l_r}f(x,y,z)}{h_1h_2\ldots h_pk_1k_2\ldots k_ql_1\ldots l_r}$$

as all the increments approach zero.

It is interesting to notice that this definition is sometimes more general than the usual definition. Suppose, for example, that $\omega = f(x,y) = \phi(x) + \psi(y)$ is a function of x and y, where neither ϕ nor ψ has a derivative. Then ω also has no first derivative, and consequently second derivatives are out of the question, in ordinary sense. Nevertheless, if we adopt the new definition, the derivative f_{xy} is the limit of the fraction

$$\frac{f(x+h,y+k)-f(x+h,y)-f(x,y+k)+f(x,y)}{hk},$$

which is equal to

$$\frac{\phi(x+h)+\psi(y+k)-\phi(x+k)-\psi(y)-\phi(x)-\psi(y-k)+\phi(x)+\psi(y)}{hk}.$$

But the numerator of this ratio is identically zero. Hence the ratio approaches zero as a limit, and we find $f_{xy} = 0$.[10]

1.3 The differential notation

The differential notation, which has been in use longer than any other,[11] is due to Leibniz. Although it is by no means indispensable, it possesses certain advantages of symmetry and of generality which are convenient, especially in the study of functions of several variables. This notation is founded upon the use of infinitesimals.

14. Differentials. Any *variable* quantity which approaches zero as a limit is called an *infinitely small quantity*, or simply *infinitesimal.* The condition that the quantity be variable is essential, for a constant, however small, is not an infinitesimal unless it is zero.

Ordinarily several quantities are considered which approach zero simultaneously. One of them is chosen as the standard of comparison, and is called the *principal infinitesimal.* Let α be the principal infinitesimal, and β another infinitesimal. Then β is said to be an infinitesimal of higher order *with respect to* α, if the ratio β/α approaches zero with α. On the other hand, β is called an infinitesimal of the first order with respect to α, if the ratio β/α approaches a limit K different from zero as α approaches zero. In this case

$$\frac{\beta}{\alpha} = K + \varepsilon,$$

where ε is another infinitesimal with respect to α. Hence

$$\beta = \alpha(K+\varepsilon) = K\alpha + \alpha\varepsilon,$$

and $K\alpha$ is called the *principal part* of β. The complementary term $\alpha\varepsilon$ is an infinitesimal of higher order with respect to α. In general, if we can find

[10] A similar remark may be made regarding functions of a single variable. For example, the function $f(z) = x^3 cos 1/x$ has the derivative

$$f'(x) = 2x^2 \cos\frac{1}{x} + x\sin\frac{1}{x}$$

and $f'(x)$ has no derivative for $x = 0$. But the ratio

$$\frac{f(2\alpha) - 2f(\alpha) - f/(0)}{\alpha^2}$$

or $8\alpha\cos(1/2\alpha) - 2\alpha\cos(1/\alpha)$, has the limit zero when α approaches zero.

[11] With the possible exception of Newton's notation.

a positive power of α, say α^n, such that β/α^n approaches a finite limit K different from zero as α approaches zero, β is called an infinitesimal of order n with respect to α. Then we have

$$\frac{\beta}{\alpha^n} = K + \varepsilon,$$

or

$$\beta = \alpha^n(K + \varepsilon) = K\alpha^n + \alpha^n\varepsilon.$$

The term $K\alpha^n$ is again called the principal part of β.

Having given these definitions, let us consider a continuous function $y = f(x)$, which possesses a derivative $f'(x)$. Let Δx be an increment of x, and let Δy denote the corresponding increment of y. From the very definition of a derivative, we have

$$\frac{\Delta y}{\Delta x} = f'(x) + \varepsilon,$$

where ε approaches zero with Δx. If Δx be taken as the principal infinitesimal, Δy is itself an infinitesimal whose principal part is $f'(x)\Delta x$.[12] This principal part is called the *differential* of y and is denoted by dy.

$$dy = f'(x)\Delta x.$$

When $f(x)$ reduces to x itself, the above formula becomes $dx = \Delta x$; and hence we shall write, for symmetry,

$$dy = f'(x)dx,$$

where the increment dx of the independent variable x to be given the same fixed value, is which otherwise arbitrary and of course variable, for all of the several dependent functions of x which may be under consideration at the same time.

Let us take a curve C whose equation is $y = f(x)$ and consider two points on it, M and M', whose abscissae are x and $x + dx$, respectively. In the triangle MTN we have

$$NT = MN \tan \angle TMN = dx f'(x).$$

Hence NT represents the differential dy, while Δy is equal to NM'. It is evident from the figure that $M'T$ is an infinitesimal of higher order, in general, with respect to NT, as M' approaches M, unless MT is parallel to the x axis.

Successive differentials may be defined, as were successive derivatives, each in terms of the preceding. Thus we call the differential of the differential

[12] Strictly speaking, we should here exclude the case where $f'(x) = 0$. It is, however convenient to retain the same definition of $du = f'(x)\Delta x$ in this case also, even though it is not the principal part of Δy.

of the first order the *differential of the second order*, where dx is given the same value in both cases, as above. It is denoted by d^2y.

$$d^2y = d(dy) = [f''(x)dx]dx = f''(x)(dx)^2.$$

Similarly, the third differential is

$$d^3y = d(d^2y) = [f'''(x)dx^2]dx = f'''(x)(dx)^3,$$

and so on. In general, the differential of the differential of order $(n-1)$ is

$$d^ny = f^{(n})(x)dx^n.$$

The derivatives $f'(x), f''(x), \ldots f^{(n)}(x), \ldots$ can be expressed, on the other hand, in terms of differentials, and we have a new notation for the derivatives:

$$y' = \frac{dy}{dx}, \quad y'' = \frac{d^2y}{dx}, \quad \ldots, f^{(n})(x) = \frac{d^ny}{dx^n}, \quad \ldots.$$

To each of the rules for the calculation of a derivative corresponds a rule for the calculation of a differential. For example, we have

$$\begin{aligned} dx^m &= mx^{m-1}dx, & da^x &= a^x \log a dx; \\ d\log x &= \frac{dx}{x}, & d\sin x &= \cos x dx; \quad \ldots; \\ d\text{arc}\sin &= \frac{dx}{\pm\sqrt{1-x^2}}, & d\text{arc}\tan x &= \frac{dx}{1a+x^2}. \end{aligned}$$

Let us consider for a moment the case of a function of a function. Let $y = f(u)$, where u is a function of the independent variable x. Then

$$y_x = f'(u)u_x.$$

whence, multiplying both sides by dx, we get

$$y_x dx = f'(u) \times u_x dx;$$

that is,

$$dy = f'(u)du.$$

The formula for dy is therefore the same as if u were the independent variable. This one of the advantages of the differential notation. In the derivative notation there are two distinct formulae,

$$y_x = f'(x), \quad y_x = f'(u)u_x,$$

to represent the derivative of y with respect to x, according as y is given directly as a function of x or is given as a function of x by means of an auxiliary function u. In the differential notation the same formula applies in each case.[13]

If $y = f(u, v, w)$ is a composite function, we have

$$y_x = u_x f_u + v_x f_v + w_x f_w,$$

at least if f_u, f_v, f_w are continuous, or, multiplying by dx,

$$y_x dx = u_x dx f_u + v_x dx f_v + w_x dx f_w$$

that is

$$dy = f_u du + f_v dv + f_w dw.$$

Thus we have, for example,

$$d(uv) = udv + vdu, \quad d(\frac{u}{v}) = \frac{vdu - udv}{v^2}.$$

The same rules enable us to calculate the successive differentials. Let us seek to calculate the successive differentials of a function $y = f(x)$, for instance. We have already

$$dy = f'(u)du.$$

In order to calculate d^2y, it must be noted that du cannot be regarded as fixed, since u is not the independent variable. We must then calculate the differential of the *composite* function $f'(u)du$, where u and du are the auxiliary functions. We thus find

$$d^2y = f''(u)du^2 + f'(u)d^2u.$$

To calculate d^3y, we must consider d^2y as a composite function, with u, du, d^2u as auxiliary functions, which leads to the expression

$$d^3y = f'''(u)du^3 + 3f''(u)dud^2u + f'(u)d^3u;$$

and so on. It should be noticed that these formulae for d^2y, d^3y, etc., are not the same as if u were the independent variable, on account of the terms d^2u, d^3u etc.[14]

A similar notation is used for the partial derivatives of a function of several variables. Thus the partial derivative of order n of $f(x, y, z)$, which is represented by $f_{x^p y^q z^r}$ in our previous notation, is represented by

$$\frac{\partial^n f}{\partial x^p \partial y^q \partial z^r}, \quad p + q + r = n,$$

[13] This particular advantage is slight, however; for the last formula above is equally well a general one and covers both the cases mentioned.

[14] This disadvantage would seem completely to offset the advantage mentioned above. Strictly speaking, we should distinguish between $d_x^2 y$ and $d_u^2 y$, etc.

in the differential notation.[15] This notation is purely symbolic, and no sense represents a quotient, as it does in the case of functions of a single variable.

15. Total differentials. Let $\omega = f(x, y, z)$, be a function of the three independent variables x, y, z. The expression

$$d\omega = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

is called the total differential of ω, where dx, dy, dz are three fixed increments, which are otherwise arbitrary, assigned to the three independent variables x, y, z. The three products

$$\frac{\partial f}{\partial x}dx, \quad \frac{\partial f}{\partial y}dy, \quad \frac{\partial f}{\partial z}dz$$

are called partial differentials.

The total differential of the second order $d^2\omega$ is the total differential of the total differential of the first order, the increments dx, dy, dz remaining the same as we pass from one differential to next higher. Hence

$$d^2\omega = d(d\omega) = \frac{\partial d\omega}{\partial x}dx + \frac{\partial d\omega}{\partial y}dy + \frac{\partial d\omega}{\partial z}dz;$$

or expanding,

$$\begin{aligned} d^2\omega &= \left(\frac{\partial^2 f}{\partial x^2}dx + \frac{\partial^2 f}{\partial x \partial y}dy + \frac{\partial^2 f}{\partial x \partial z}dz\right)dx \\ &\quad + \left(\frac{\partial^2 f}{\partial x \partial y}dx + \frac{\partial^2 f}{\partial y^2}dy + \frac{\partial^2 f}{\partial y \partial z}dz\right)dy \\ &\quad + \left(\frac{\partial^2 f}{\partial x \partial z}dx + \frac{\partial^2 f}{\partial y \partial z}dy + \frac{\partial^2 f}{\partial z^2}dz\right)dz \\ &= \frac{\partial^2 f}{\partial x^2}dx^2 + \frac{\partial^2 f}{\partial y^2}dy^2 + \frac{\partial^2 f}{\partial z^2}dz^2 \\ &\quad + 2\frac{\partial^2 f}{\partial x \partial y}dxy + 2\frac{\partial^2 f}{\partial x \partial z}dxdz + 2\frac{\partial^2 f}{\partial y \partial z}dydz. \end{aligned}$$

If $\partial^2 f$ be replaced by ∂f^2, the right-hand side of this equation becomes the square of

$$\frac{\partial f}{\partial x}dx, \frac{\partial f}{\partial y}dy, \frac{\partial f}{\partial z}dz.$$

We may then write, symbolically,

$$d^2\omega = \left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right)^{(2)},$$

[15] This use of the letter ∂ to denote the partial derivatives of a function of several variables is due to Jacobi. Before his time the same letter d was used as is used for the derivatives of a function of a single variable.

it being agreed that ∂f^2 is to be replaced by $\partial^2 f$ after expansion.

In general, if we call the total differential of the total differential $(n-1)$ the *total differential of order* n, and denote it by $d^n\omega$ we may write, in the same symbolism,

$$d^n\omega = \left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right)^{(n)},$$

where ∂f^n is to be replaced by $\partial^n f$ after expansion; that is, in our ordinary notation,

$$d^n\omega = \sum A_{pqr}\frac{\partial^n f}{\partial x^p}\partial y^q \partial z^r dx^p dy^q dz^r, \quad p+q+r=n,$$

where

$$A_{pqr} = \frac{n!}{p!q!r!}$$

is the coefficient of the term $a^p b^q c^r$ in the development of $(a+b+c)^n$. For, suppose this formula holds for $d^n\omega$. We will show that then it holds for $d^{n+1}\omega$; and this will prove it in general, since we have already proved it for $n=2$. From the definition, we find

$$\begin{aligned} d^{n+1}\omega &= d(d^n\omega) \\ &= \sum A_{pqr}\left[\frac{\partial^{n+1} f}{\partial x^{p+1}\partial y^q \partial z^r}dx^{p+1}dy^q dz^r + \frac{\partial^{n+1} f}{\partial x^p \partial y^{q+1}\partial z^r}dx^p dy^{q+1}dz^r \right. \\ &\quad \left. + \frac{\partial^{n+1} f}{\partial x^p \partial y^q \partial z^{r+1}}dx^p dy^q dz^{r+1}\right]; \end{aligned}$$

whence, replacing $\partial^{n+1} f$ by ∂f^{n+1}, the right-hand side becomes

$$\sum A_{pqr}\frac{\partial f^n}{\partial x^p \partial y^q \partial z^r}dx^p dy^q dz^r \left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right),$$

or

$$\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right)^{(n)} \left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right).$$

Hence, using the same symbolism, we may write

$$d^{n+1}\omega = \left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right)^{(n+1)}.$$

Note. Let us suppose that the expression for $d\omega$, obtained in any way whatever, is

$$d\omega = Pdx + Qdy + Rdz, \tag{1.7}$$

where P, Q, R are any functions x, y, z. Since by definition

$$d\omega = \frac{\partial\omega}{\partial x}dx + \frac{\partial\omega}{\partial y}dy + \frac{\partial\omega}{\partial z}dz,$$

we must have

$$\left(\frac{\partial\omega}{\partial x} - P\right) dx + \left(\frac{\partial\omega}{\partial y} - Q\right) dy + \left(\frac{\partial\omega}{\partial z} - R\right) dz = 0$$

where dx, dy, dz are any constants. Hence

$$\frac{\partial\omega}{\partial x} = P, \frac{\partial\omega}{\partial y} = Q, \frac{\partial\omega}{\partial z} = R. \tag{1.8}$$

The single equation (1.7) is therefore equivalent to the three separate equations (1.8); and determines all three partial derivatives at once.

In general, if the nth total differential be obtained in any way whatever,

$$d^n\omega = \sum C_{pqr} dx^p dy^q dz^r;$$

then the coefficients C_{pqr} are respectively equal to the corresponding nth derivatives multiplied by certain numerical factors. Thus all these derivatives are determined at once. We shall have occasion to use these facts presently.

16. Successive differentials of composite functions. Let $\omega = F(u, v, w)$ be a composite function, u, v, w being themselves functions of the independent variables x, y, z, t. The partial derivatives may then be written down as follows

$$\begin{aligned}
\frac{\partial\omega}{\partial x} &= \frac{\partial F}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial F}{\partial w}\frac{\partial w}{\partial x},\\
\frac{\partial\omega}{\partial y} &= \frac{\partial F}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial F}{\partial w}\frac{\partial w}{\partial y},\\
\frac{\partial\omega}{\partial z} &= \frac{\partial F}{\partial u}\frac{\partial u}{\partial z} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial z} + \frac{\partial F}{\partial w}\frac{\partial w}{\partial z},\\
\frac{\partial\omega}{\partial t} &= \frac{\partial F}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial t} + \frac{\partial F}{\partial w}\frac{\partial w}{\partial t}.
\end{aligned}$$

If these four equations be multiplied by dx, dy, dz, dt, respectively, and added, the left-hand side becomes

$$\frac{\partial\omega}{\partial x}dx + \frac{\partial\omega}{\partial y}dy + \frac{\partial\omega}{\partial z}dz + \frac{\partial\omega}{\partial t}dt$$

that is, $d\omega$; and the coefficients of

$$\frac{\partial F}{\partial u}, \quad \frac{\partial F}{\partial v}, \quad \frac{\partial F}{\partial w}$$

on the right-hand side are du, dv, dw, respectively. Hence

$$d\omega = \frac{\partial F}{\partial u}du + \frac{\partial F}{\partial v}dv + \frac{\partial F}{\partial w}dw, \tag{1.9}$$

and we see that the expression of the total differential of the first order of a composite function is the same as if the auxiliary functions were the independent variables. This is one of the main advantages of the differential notation. The equation (1.9) does not depend, in form, either upon the number or upon the choice of the independent variables; and it is equivalent to as many separate equations as there are independent variables.

To calculate $d^2\omega$, let us apply the rule just found for $d\omega$, noting that the second member of (1.9) involves the six auxiliary functions u, v, w, du, dr, dw. We thus find

$$\begin{aligned} d^2\omega = &\frac{\partial^2 F}{\partial u^2}du^2 + \frac{\partial^2 F}{\partial u\partial v}dudv + \frac{\partial^2 F}{\partial u\partial w}dudw + \frac{\partial F}{\partial u}d^2u \\ &+ \frac{\partial^2 F}{\partial u\partial v}dudv + \frac{\partial^2 F}{\partial v^2}dv^2 + \frac{\partial^2 F}{\partial v\partial w}dvdw + \frac{\partial F}{\partial v}d^2v \\ &+ \frac{\partial^2 F}{\partial u\partial w}dudw + \frac{\partial^2 F}{\partial v\partial w}dvdw + \frac{\partial^2 F}{\partial w^2}dw^2 + \frac{\partial F}{\partial w}d^2w, \end{aligned}$$

or, simplifying and using the same symbolism as above,

$$d^2\omega = \left(\frac{\partial F}{\partial u}du + \frac{\partial F}{\partial v}dv + \frac{\partial F}{\partial w}dw\right)^{(2)} + \frac{\partial F}{\partial u}d^2u + \frac{\partial F}{\partial v}d^2v + \frac{\partial F}{\partial w}d^2w.$$

This formula is somewhat complicated on account of the terms in d^2u, d^2v, d^2w which drop out when u, v, w are the independent variables. This limitation of the differential notation should be borne in mind, and the distinction between $d^2\omega$ in the two cases carefully noted. To determine $d^3\omega$ we would apply the same rule to $d^2\omega$, noting that $d^2\omega$ depends upon the nine auxiliary functions $u, v, w, du, du, dw, d^2u, d^2v, d^2w$; and so forth. The general expressions for these differentials become more and more complicated; $d^n\omega$ is an integral function of $du, dv, dw, d^2u, \ldots, d^nu, d^nv, d^nw$, and the terms containing d^nu, d^nv, d^nw are

$$\frac{\partial F}{\partial u}d^nu + \frac{\partial F}{\partial v}d^nv + \frac{\partial F}{\partial w}d^nw.$$

If, in the expression for $d^n\omega$, $u, v, w, du, du, dw, \ldots$ be replaced by their values in terms of the independent variables, $d^n\omega$ becomes an integral polynomial in $dx, dy, dz, \ldots$ whose coefficients are equal (cf. *Note*, § 15) to the partial derivatives of ω of order n, multiplied by certain numerical factors. We thus obtain these derivatives at once.

Suppose, for example, that we wished to calculate the first and second derivatives of a composite function $\omega = f(u)$, where u is a function of two independent variables $u = \phi(x, y)$. If we calculate these derivatives separately, we find for the two partial derivatives of the first order

$$\frac{\partial\omega}{\partial x} = \frac{\partial\omega}{\partial u}\frac{\partial u}{\partial x}, \quad \frac{\partial\omega}{\partial y} = \frac{\partial\omega}{\partial u}\frac{\partial u}{\partial y}. \tag{1.10}$$

Again, taking the derivatives of these two equations with respect to x, and then with respect to y, we find only the three following distinct equations, which give the second derivatives:

$$\begin{aligned} \frac{\partial^2\omega}{\partial x^2} &= \frac{\partial^2\omega}{\partial u^2}\left(\frac{\partial u}{\partial x}\right)^2 + \frac{\partial\omega}{\partial u}\frac{\partial^2 u}{\partial x^2}, \\ \frac{\partial^2\omega}{\partial x\partial y} &= \frac{\partial^2\omega}{\partial u^2}\frac{\partial u}{\partial x}\frac{\partial u}{\partial y} + \frac{\partial\omega}{\partial u}\frac{\partial^2 u}{\partial x\partial y}, \\ \frac{\partial^2\omega}{\partial y^2} &= \frac{\partial^2\omega}{\partial u^2}\left(\frac{\partial u}{\partial y}\right)^2 + \frac{\partial\omega}{\partial u}\frac{\partial^2 u}{\partial y^2}. \end{aligned} \tag{1.11}$$

The second of these equations is obtained by differentiating the first of equations (1.10) with respect to y, or the second of them with respect to x. In the differential notation these five relations (1.10) and (1.11) may be written in the form

$$\begin{aligned} d\omega &= \frac{\partial\omega}{\partial u}du, \\ d^2\omega &= \frac{\partial^2\omega}{\partial u^2}du^2 + \frac{\partial\omega}{\partial u}du^2. \end{aligned} \tag{1.12}$$

If du and d^2u in these formulae be replaced by

$$\frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial u}dy \quad \text{and} \quad \frac{\partial^2 u}{\partial x^2}dx^2 + \frac{\partial^2 u}{\partial x\partial u}dxdy + \frac{\partial^2 u}{\partial y^2}dy^2,$$

respectively, the coefficients of dx and dy in the first give the first partial derivatives of ω, while the coefficients of dx^2, $2dxdy$, and dy^2 in the second give the second partial derivatives of ω.

17. Differentials of a product. The formula for the total differential of order n of a composite function becomes considerably simpler in certain special cases which often arise in practical applications. Thus, let us seek the differential of order n of the product of two functions $\omega = uv$. For the first values of n we have

$$d\omega = vdu + udv, \quad d^2\omega = vd^2u + 2dudv + ud^v; \quad \ldots;$$

and, in general, it is evident from the law of formation that

$$d^n\omega = vd^nu + C_1dvd^{n-1}u + C_2d^2vd^{n-2}u + \cdots + ud^nv,$$

where $C_1, C_2, \ldots$ are positive integers. It might be shown by algebraic induction that these coefficients are equal to those of the expansion of $(a+b)^n$; but the same end may be reached by the following method, which is much more elegant, and which applies to many similar problems. Observing that $C_1, C_2, \ldots$ do not depend upon the particular functions u and v employed, let us take the special functions $u = e^x$, $v = e^y$, where x and y are the two independent variables, and determine the coefficients for this case. We thus find

$$\omega = e^{x+y}, \quad d\omega = e^{x+y}(dx + dy), \quad \ldots, \quad d\omega = e^{x+y}(dx + dy)^n,$$

$$du = e^x dx, \quad d^2u = e^x dx^2, \quad \ldots, dv = e^y dy, \quad d^2v = e^y dy^2, \quad \ldots;$$

and the general formula, after division by e^{x+y}, becomes

$$(dx + dy)^n = dx^n + C_1 dydn^{n-1} + C_2 dy^2 dn^{n-2} + \cdots + dy^n.$$

Since dx and dy are arbitrary, it follows that

$$C_1 = \frac{n}{1}, \quad C_2 = \frac{n(n-1)}{1 \cdot 2}, \quad \ldots, \quad C_2 = \frac{n(n-1)\cdots(n-p+1)}{1 \cdot 2 \cdots p}, \quad \ldots, \quad \ldots;$$

and consequently the general formula may be written

$$d^n(uv) = vd^n u + \frac{n}{1} dvd^{n-1}u + \frac{n(n-1)}{1 \cdot 2} d^2 vd^{n-2}u + \cdots + ud^n v. \tag{1.13}$$

This formula applies for any number of independent variables. In particular, if u and v are functions of a single variable x, we have, after division by dx^n, the expression for the nth derivative of the product of two functions of a single variable.

It is easy to prove in a similar manner formulae analogous to (1.13) for a product of any number of functions.

Another special case in which the general formula reduces to a simpler form is that in which u, v, w are integral linear functions of the independent variables x, y, z.

$$\begin{aligned} u &= ax + by + cz + f, \\ v &= a'x + b'y + c'z + f', \\ w &= a''x + b''y + c''z + f'', \end{aligned}$$

where the coefficients $a, a', a'', b, b', \ldots$ are constants. For then we have

$$\begin{aligned} du &= adx + bdy + cdz, \\ dv &= a'dx + b'dy + c'dz, \\ dw &= a''dx + b''dy + c''dz'', \end{aligned}$$

and all the differentials of higher order $d^n u, d^n v, d^n w$, where $n > 1$, vanish. Hence the formula for $d^n \omega$ is the same as if u, v, w were the independent variables; that is,

$$d^{\omega} = \left(\frac{\partial F}{\partial u} du + \frac{\partial F}{\partial v} dv + \frac{\partial F}{\partial w} dw \right)^{(n)} .$$

We proceed to apply this remark.

18. Homogeneous functions. A function $\phi(x, y, z)$ is said to be homogeneous of degree m, if the equation

$$\phi(u, v, w) = t^m (x, y, z)$$

is identically satisfied when we set

$$u = tx, \quad v = ty, \quad w = tz.$$

Let us equate the differentials of order n of the two sides of this equation with respect to t, noting that u, v, w are linear in t, and that

$$du = xdt, \quad dv = ydt, \quad dw = zdt.$$

The remark just made shows that

$$\left(x \frac{\partial \phi}{\partial u} + y \frac{\partial \phi}{\partial v} + z \frac{\partial \phi}{\partial w} \right)^{(n)} = m(m-1) \cdots (m-n-1) t^{m-n} \phi(x, y, z).$$

If we now set $t = 1$, u, v, w reduce to x, y, z, and any term of the development of the first member,

$$A_{pqr} \frac{\partial^n \phi}{\partial u^p \partial v^q \partial w^r} x^p y^q z^r ,$$

becomes

$$A_{pqr} \frac{\partial^n \phi}{\partial x^p \partial y^q \partial z^r} x^p y^q z^r ,$$

whence we may write, symbolically,

$$\left(x \frac{\partial \phi}{\partial x} + y \frac{\partial \phi}{\partial y} + z \frac{\partial \phi}{\partial z} \right)^{(n)} = m(m-1) \cdots (m-n-1) t^{m-n} \phi(x, y, z),$$

which reduces, for $n = 1$, to the well-known formula

$$m\phi(x, y, z) = x \frac{\partial \phi}{\partial x} + y \frac{\partial \phi}{\partial y} + z \frac{\partial \phi}{\partial z}.$$

Various notations. We have then, altogether, three systems of notation for the partial derivatives of a function of several variables, — that of Leibniz, that of Lagrange, and that of Cauchy. Each of these is somewhat inconveniently long, especially in a complicated calculation. For this reason various shorter notations have been devised. Among these one first used by Monge for the first and second derivatives of a function of two variables is now in common use. If z be the function of the two variables x and y, we set

$$p = \frac{\partial z}{\partial x}, \quad q = \frac{\partial z}{\partial y}, \quad r = \frac{\partial^2 z}{\partial x^2}, \quad s = \frac{\partial^2 z}{\partial x \partial y}, \quad t = \frac{\partial^2 z}{\partial y^2},$$

and the total differentials dz and $d^2 z$ are given by the formulae

$$dz = pdx + qdy,$$
$$d^2 z = rdx^2 + 2sdxdy + tdy^2.$$

Another notation which is now coming into general use is the following. Let z be a function of any number of independent variables $x+1, x_2, x_3, \ldots, x_n$; then the notation

$$p_{\alpha_1 \alpha_2 \ldots \alpha_n} = \frac{\partial^{\alpha_1 + \alpha_2 + \cdots + \alpha_n}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdot \partial x_n^{\alpha_n}}$$

is used, where some of the indices $\alpha_1, \alpha_2, \ldots, \alpha_n$ may be zeros.

19. Applications. Let $y = f(x)$ be the equation of a plane curve C with respect to a set of rectangular axes. The equation of the tangent at a point $M(x, y)$ is

$$Y - y = y'(X - x).$$

The slope of the normal, which is perpendicular to the tangent at the point of tangency, is $-1/y'$; and the equation of the normal is, therefore,

$$(Y - y)y' + (X - x) = 0$$

Let P be the foot of the ordinate of the point M, and let T and N be the points of intersection of the x axis with the tangent and the normal, respectively. The distance PN is called the subnormal; PT, the subtangent; MN the normal; and MT, the tangent.

From the equation of the normal the abscissa of the point N is $x + yy'$, whence the subnormal is $\pm yy'$. If we agree to call the length PN the subnormal, and to attach the sign + or the sign − according as the direction PN is positive or negative, the subnormal will always be yy' for any position of the curve C. Likewise the subtangent is $-y/y'$. The lengths MN and MT are given by the triangles MPN and MPT:

$$MN = \sqrt{\overline{MP}^2 + \overline{PN}^2} = y\sqrt{1 + y'^2},$$

$$MN = \sqrt{\overline{MP}^2 + \overline{PT}^2} = \frac{y}{y'}\sqrt{1 + y'^2}.$$

Various problems may be given regarding these lines. Let us find, for instance, all the curves for which the subnormal is constant and equal to a given number a. This amounts to finding all the functions $y = f(x)$ which satisfy the equation $yy' = a$. The left-hand side is the derivative of $y^2/2$, while the right-hand side is the derivative of ax. These functions can therefore differ only by a constant; whence

$$y^2 = 2ax + C.$$

which is the equation of a parabola along the x axis. Again, if we seek the curves for which the subtangent is constant, we are led to write down the equation $y'/y = 1/a$; whence

$$\log y = \frac{x}{a} + \log C, \quad \text{or} \quad y = Ce^{\frac{x}{a}},$$

which is the equation of a transcendental curve to which the x axis is an asymptote. To find the curves for which the normal is constant, we have the equation

$$y\sqrt{1 + y'^2} = a,$$

or

$$\frac{yy'}{\sqrt{a^2 - y^2}} = 1.$$

The first member is the derivative of $-\sqrt{a^2 - y^2}$; hence

$$-\sqrt{a^2 - y^2} = x + C,$$

or

$$(x + C)^2 + y^2 = a^2,$$

which is the equation of a circle of radius a, whose center lies on the x axis.

The curves for which the tangent is constant are transcendental curves, which we shall study later.

Let $y = f(x)$ and $Y = F(x)$ be the equations of two curves C and C' and let M, M' be the two points which correspond to the same value of x. In order that the two subnormals should have equal lengths it is necessary and sufficient that

$$YY' = \pm yy';$$

that is, that $Y^2 = \pm y^2 + C$, where the double sign admits of the normals' being directed in like or in opposite senses. This relation is satisfied by the curves

$$y^2 = \frac{b^2}{a^2}(a^2 - x^2), \quad Y^2 = \frac{b^2x^2}{a^2},$$

and also by the curves

$$y^2 = \frac{b^2}{a^2}(x^2 - a^2), \quad Y^2 = \frac{b^2x^2}{a^2},$$

which gives an easy construction for the normal to the ellipse and to the hyperbola.

EXERCISES

1. Let $\rho = f(\theta)$ be the equation of a plane curve in polar coordinates. Through the pole O draw a line perpendicular to the radius vector OM, and let T and N be the points where this line cuts the tangent and the normal. Find expressions for the distances OT, ON, MN, and MT in terms of $f(\theta)$ and $f'(\theta)$,

Find the curves for which each of these distances, in turn, is constant.

2. Let $y = f(x)$, $z = \phi(x)$ be the equations of a skew curve Γ, i.e. of a general space curve. Let N be the point where the normal plane at a point M, that is, the plane perpendicular to the tangent at M, meets the z axis; and let P be the foot of the perpendicular from M to the z axis. Find the curves for which each of the distances PN and MN, in turn, is constant.

[*Note*, These curves lie on paraboloids of revolution or on spheres.]

3. Determine an integral polynomial $f(x)$ of the seventh degree in x, given that $f(x)+1$ is divisible by $(x-1)^4$ and $f(x)-1$ by $(x+1)^4$. Generalize the problem.

4. Show that if the two integral polynomials P and Q satisfy the relation

$$\sqrt{1-P^2} = Q\sqrt{1-x^2},$$

then

$$\frac{dP}{\sqrt{1-P^2}} = \frac{ndz}{1-x^2},$$

where n is a positive integer.

[*Note*, From the relation

$$1 - P^2 = Q^2(1-x^2)$$

it follows that

$$-2PP' = Q[2Q'(1-x^2) - 2Qx].$$

The equation (a) shows that Q is prime to P; and (b) shows that P' is divisible by Q.]

5*. Let $R(x)$ be a polynomial of the fourth degree whose roots are all different, and let $x = U/V$ be a rational function of t, such that

$$\sqrt{R(x)} = \frac{P(t)}{Q(t)}\sqrt{R_1(t)},$$

where $R_1(t)$ is a polynomial of the fourth degree and P/Q is a rational function. Show that the function U/V satisfies a relation of the form

$$\frac{dx}{\sqrt{R(x)}} = \frac{kdt}{\sqrt{R_1(t)}}$$

where k is a constant. [Jacobi.]

[*Note.* Each root of the equation $R(U/V) = 0$, since it cannot cause $R'(x)$ to vanish, must cause $UV' - VU'$ and hence also dx/dt to vanish.]

6*. Show that the nth derivative of a function $y = \phi(u)$, where u is a function of the independent variable x, may be written in the form

$$\text{(a)} \qquad \frac{d^n y}{dx^n} = A_1\phi'(u) + \frac{A_2}{1 \cdot 2}\phi''(u) + \cdots + \frac{A_n}{1 \cdot 2 \cdots n}\phi^{(n)}(u),$$

where

$$\text{(b)} \qquad \begin{aligned} A_k = {} & \frac{d^n u^k}{dx^n} - \frac{k}{1}u\frac{d^n u^{k-1}}{dx^n} + \frac{k(k-1)}{1 \cdot 2}u^2\frac{d^n u^{k-2}}{dx^n} + \cdots \\ & + (-1)^{k-1}ku^{k-1}\frac{d^n u}{dx^n} \quad (k = 1, 2, \ldots, n). \end{aligned}$$

[First notice that the nth derivative may be written in the form (a), where the coefficients $A_1, A_2, \ldots, A_n$ are independent of the form of the function $\phi(u)$. To find their values, set $\phi(u)$ equal to $u, u^2, \ldots, u^n$ successively, and solve the resulting equations for $A_1, A_2, \ldots, A_n$. The result is the form (b).]

7*. Show that the nth derivative of $\phi(x^2)$ is

$$\begin{aligned} \frac{d^n\phi(x^2)}{dx^n} = {} & (2x)^n\phi^{(n)}(x^2) + n(n-1)(2x)^{n-2}\phi^{(n-1)}(x^2) + \cdots \\ & + \frac{n(n-1)\cdots(n-2p+1)}{1 \cdot 2 \cdots p}(2x)^{n-2p}\phi^{(n-p)}(x^2) + \cdots, \end{aligned}$$

where p varies from zero to the last positive integer not greater than $n/2$, and where $\phi^{(i)})x^2)$ denotes the ith derivative with respect to x. Apply this result to the functions e^{-x^2}, $\arcsin x$, $\arctan x$.

8*. If $x = \cos u$, show that

$$\frac{d^{m-1}(1-x^2)^{m-\frac{1}{2}}}{dx^{m-1}} = (-1)^{m-1}\frac{1 \cdot 3 \cdot 5 \cdots (2m-1)}{m} sinmu.$$

[Olinde Rodrigues.]

9. Show that Legendre's polynomial,

$$X_n = \frac{1}{2 \cdot 4 \cdot 6 \cdots 2n}\frac{d^n}{dx^n}(x^2-1)^n,$$

satisfies the differential equation

$$(1-x^2)\frac{d^2X_n}{dx^2} - 2x\frac{dX_n}{dx} + n(n+1)X_n = 0.$$

Hence deduce the coefficients of the polynomial.

10. Show that the four functions

$$y_1 = \sin(n \operatorname{arc}\sin x), \quad y_3 = \sin(n \operatorname{arc}\cos x),$$

$$y_2 = \cos(n\mathrm{arc}\sin x), \quad y_4 = \cos(n\mathrm{arc}\cos x),$$

satisfy the differential equation

$$(1-x^2)y'' - xy' + n^2y = 0.$$

Hence deduce the developments of these functions when they reduce to polynomials.

11*. Prove the formula

$$\frac{d^n}{dx^n}(x^{n-1}e^{\frac{1}{x}}) = (-1)^n\frac{e^{\frac{1}{x}}}{x^{n+1}}.$$

[Halphen.]

12. Every function of the form $z = x\phi(y/x) + \psi(y/x)$ satisfies the equation

$$rx^2 + 2sxy + ty^2 = 0,$$

whatever be the functions ϕ and ψ.

13. The function $z = x\phi(x+y) + y\psi(x+y)$ satisfies the equation

$$r - 2s + t = 0,$$

whatever be the functions ϕ and ψ.

14. The function $z = f[x+\phi(y)]$ satisfies the equation $ps = qr$, whatever be the functions f and ϕ.

15. The function $z = x^n\phi(y/x) + y^{-n}\psi(y/x)$ satisfies the equation

$$rx^2 + 2sxy + ty^2 + px + qy = n^2z,$$

whatever be the functions ϕ and ψ.

16. Show that the function

$$y = |x-a_1|\phi_1(x) + |x-a_2|\phi_2(x) + \cdots + |x-a_n|\phi_n(x),$$

where $\phi_1(x), \phi_2(x), \ldots, \phi_n(x)$ together with their derivatives, $\phi_1'(x), \phi_2'(x), \ldots, \phi_n'(x)$, are continuous functions of x, has a derivative which is discontinuous for $x = a_1, a_2, \ldots, a_n$.

17. Find a relation between the first and second derivatives of the function $z = f(x_1, u)$, where $u = \phi(x_2, x_3)$; x_1, x_2, x_3 being three independent variables, and f and ϕ two arbitrary functions.

18. Let $f'(x)$ be the derivative of an arbitrary function $f(x)$. Show that

$$\frac{1}{u}d^2udx^2 = \frac{1}{v}\frac{d^2v}{dx^2}.$$

where $u = [f'(x)]^{-\frac{1}{2}}$ and $v = f(x)[f'(x)]^{-\frac{1}{2}}$.

19*. The nth derivative of a function of a function $u = \phi(y)$, where $y = \Psi(x)$, may be written in the form

$$D^n x\phi = \sum \frac{n!}{i!j!\cdots k!} D_y^p \phi \left(\frac{\Psi'}{1}\right)^i \left(\frac{\Psi''}{1\cdot 2}\right)^j \left(\frac{\Psi'''}{1\cdot 2\cdot 3}\right)^h \cdots \left(\frac{\Psi^{(l)}}{1\cdot 2\cdots l}\right)^k,$$

where the sign of summation extends over all the positive integral solutions of the equation $i + 2j + 3h + \cdots + lk = n$, and where $p = i + j + \cdots + k$.

[Faa de Bruno, *Quarterly Journal of Mathematics*, Vol. I, p. 869.]