Cambridge University Press
978-1-107-49362-9 - The Fundamental Theorems of the Differential Calculus
W. H. Young
Excerpt
More information

I. PRELIMINARY NOTIONS.

1. Independent variables. Interval. Right and left. In the present tract we shall deal with one or more real independent variables x, each of which assumes all values in some interval, which may be *closed* ($a \leqslant x \leqslant b$) or *completely open* ($a < x < b$) or *half-open*. The values of x may be supposed represented in any of the usual ways on a straight line, which will be laid horizontal and so that, as x increases, the representative point moves to the right. This straight line will be imagined closed at each end by a point, the point $+\infty$ on the right and $-\infty$ on the left, and when we speak of the interval (a, b), the left-hand end-point may, unless the contrary is stated, be $-\infty$, and the right-hand end-point $+\infty$.

If there are two or more independent variables, these will be represented on rectangular axes, and the ensemble of them (x, y), or $(x_1, x_2, \ldots, x_n)$, will be represented by a point in the plane or higher space. The correlative of an interval will then be a rectangle, or n-dimensional parallelepiped, and may be closed or completely or partially open. In what immediately follows x will, for brevity, be used equally for a single variable and for the ensemble of several variables, and the word *interval* will be used with the understanding that in higher space the proper interpretation is to be put upon it.

2. Function of one or more variables. Finite. Bounded. If to each x there is an unique value f, this is said to be a *function* of x, and we write $f(x)$ for it. It is found convenient to include the two distinct infinite numbers $+\infty$ and $-\infty$ as among the values which the most general kind of function may assume[(a)].

If $f(x)$ has at each point x a finite value, it is said to be a *finite function*. If there is some finite number greater (less) than any value of the function, $f(x)$ is said to be *bounded above* (*below*), and if bounded both above and below, $f(x)$ is a *bounded function*.

Ex. 1. Let $f(0)=0$, $f(x)=\frac{1}{x}$, $(x \neq 0)$. Then f is a finite function, unbounded both above and below.

II. LIMITS.

3. Upper and lower bound. Upper and lower limit. Unique limit. The least (greatest) number which is not less (greater) than any value of a function is called its *upper* (*lower*) *bound*. It is therefore the same thing to say a function is unbounded above (below) or that the upper (lower) bound is $+\infty$ $(-\infty)$.

If we take a sequence of intervals* each inside the preceding,

$$d_1, d_2, \ldots, d_n, \ldots,$$

and having one and only one common internal point a, the upper (lower) bounds of $f(x)$ for values of x other than a in these successive intervals will not increase (decrease), so that they form a monotone descending (ascending) series

$$U_1 \geqslant U_2 \geqslant \ldots \geqslant U_n \geqslant \ldots,$$
$$L_1 \leqslant L_2 \leqslant \ldots \leqslant L_n \leqslant \ldots.$$

The lower bound ϕ of the former and the upper bound ψ of the latter series are therefore unaltered if we omit any finite number of its constituents, or if we interpolate the upper (lower) bound taken with respect to any intermediate interval. Hence it is easily perceived that ϕ and ψ do not depend on the particular sequence of intervals chosen, provided only the same point a is in each case the sole and only common internal point. Thus ϕ and ψ depend only on the point a and the function f, and are called respectively *the upper and lower limits of f at the point a*. As we vary a, ϕ and ψ become functions $\phi(x)$ and $\psi(x)$, and are called *the associated upper and lower limiting functions* [(β)] *of $f(x)$*.

It may sometimes be convenient in considering the behaviour of $f(x)$ in the neighbourhood of the point a to omit other values of x besides the value a, which is always omitted. The values of x which are retained must form a set S with a as limiting point, and should in each case be expressly defined. We then get, by a precisely similar method, *upper and lower limits at the point a with respect to the set S*. Each such limit is said to be *a limit* of $f(x)$ at the point a, and there will always be *a plurality of limits*, except when the upper and lower limits at a coincide, in which case $f(x)$ is said to have *an unique limit* at the point a, and its value is the common value of the upper and lower limits. When there is a single independent variable x, it is important to consider the case in which the set S consists of all

* See Appendix III.

points of an interval on one side only of the point a. We thus get upper and lower and possibly intermediate *limits on the right and on the left* respectively of the point a, and, when these coincide, we get *an unique limit on the right or on the left.*

4. Plurality of Limits. The concept of a plurality of limits can only be clearly grasped by reference to the elementary facts of the Theory of Sets of Points. As the function $y=f(x)$ is supposed to be any whatever, the mode of distribution of the set of points G_d on the y-axis, which represents the values of the function at all points except a of some interval d containing a on the x-axis, is any whatever. Certain of these points may be repeated an infinite number of times, since the function may assume the same value over and over again. If this is the case, we add such points to the first derived set* of G_d. The set H_d so formed is still a closed set, since the addition of points of a set to its first derived set introduces no new limiting points. In particular the set H_d includes the point Q_d which represents the upper bound of $f(x)$ in the interval d considered.

If we now let the interval d shrink up to the point a, as in the preceding article, the successive closed sets H_d so obtained lie each inside the preceding sets, and therefore, by Cantor's Theorem of Deduction*, determine a closed set H, consisting of all their common points, which is easily seen to be the same, however the interval d shrinks up to the point a, and includes the unique limiting point Q of the points Q_d. *This closed set H of values of y is said to constitute the set of limits of the function $f(x)$ at the point a.* Since the set H certainly includes the point Q, the set of limits includes the upper (and similarly the lower) limit as defined in the preceding article, and this whether these limits are taken with respect to the continuum or any other set.

On the other hand, since every set of points contains at least one sequence* having as unique limiting point any required limiting point of the original set, it follows that, if l be any limit of $f(x)$ at $x=a$, there is a sequence $x_1, x_2, \dots x_n, \dots$ having a as unique limiting point, such that, passing along this sequence, $f(x)$ has l as unique limit. This important property proves at the same time, in conjunction with the preceding paragraph, that the limits(γ) as defined in the present article are the same as in the preceding article.

We use the notation

$$\underset{x=a}{\text{Llt}}\, f(x)$$

* See Appendix III.

to denote the set of all the limits of $f(x)$ at the point a, while, if it is known that there is an unique limit, we write $\underset{x=a}{\mathrm{Lt}}\, f(x)$ for that limit. Thus the equation

$$y = \underset{x=a}{\mathrm{Lt}}\, f(x)$$

must be understood to mean both that $f(x)$ has an unique limit at $x=a$, and also that that limit is y.

5. Double and repeated limits. Any repeated limit is a double limit. If there are two independent variables x and y, the limits at (a, b), obtained in the manner explained in §§ 3, 4, with the two-dimensional interpretation given in § 1, are called *double limits* of $f(x, y)$ at (a, b).

Similarly if there are n variables, the corresponding limits are called n-ple limits at the point.

If $f(x, y)$ is a function of two independent variables x and y, it becomes a function of x alone when we keep y constant and has a corresponding set of simple limits

$$\underset{x=a}{\mathrm{Llt}}\, f(x, y).$$

If there is only one such limit for each value of y, this limit defines a function of y, and has as such a set of limits for $y=b$; these are called *the repeated limits* of $f(x, y)$ first with respect to x and then with respect to y, and written

$$\underset{y=b}{\mathrm{Llt}}\ \underset{x=a}{\mathrm{Lt}}\, f(x, y).$$

Similarly, if there is an unique limit when y is kept constant,

$$\underset{x=a}{\mathrm{Llt}}\ \underset{y=b}{\mathrm{Lt}}\, f(x, y)$$

denotes *the repeated limits* of $f(x, y)$, first with respect to y and then with respect to x.

Sometimes it is desirable to consider only such double limits as result from values of $f(x, y)$ in a neighbourhood of the point (a, b) other than points on the axial cross through (a, b), that is on $x=a$ and on $y=b$. Such a neighbourhood is called *a non-axial neighbourhood*. In particular we have the following theorem.

THEOREM. *Any repeated limit is a double limit, taken with respect to a non-axial neighbourhood of the point considered.*

For by the definition of a repeated limit, say u, whatever sequence of values different from a, say $x_1, x_2 \ldots, \ldots x_n, \ldots$ be taken, having a as unique limit, $f(x_n, y)$ has, for fixed y different from b, an unique limit, say $v(y)$, when n is indefinitely increased, and the quantities $v(y)$ have at $y=b$ the repeated limit u in question as one of their limits; that is

$$v(y) = \underset{n=\infty}{\text{Lt}}\, f(x_n, y), \quad (y \neq b),$$

$$u = \text{one of the } \underset{y=b}{\text{Llt}}\, v(y).$$

Represent the values of these functions on a straight line, as in § 4,

$f(x_n, y)$ by the point $P_{n,y}$,

$v(y)$ by the point Q_y,

u by the point Q.

Then, unless the points $P_{n,y}$, for fixed y, all coincide with the exception of a finite number of them, Q_y is the unique limiting point of the points $P_{n,y}$, while in the excluded case, it is the repeated point itself. In either case, denoting as in § 4 these points $P_{n,y}$, which as n and y vary, lie in a neighbourhood d of the point a, by the set G_d, each point Q_y is a point of the closed set H_d, consisting of the first derived set of G_d together with its repeated points, if any.

Again Q is either a repeated point of the set H_d or one of its limiting points, and therefore in any case it is for all neighbourhoods d a point of the set H_d, since that set is closed. Thus Q represents by § 4 a double limit of $f(x, y)$ at the point (a, b). Moreover, since the points (x_n, y) all lie off the axial cross, this double limit is taken with respect to a non-axial neighbourhood of the point (a, b).

It is an immediate consequence that *if $f(x, y)$ has an unique double limit, and the simple limit*

$$\underset{h=0}{\text{Lt}}\, f(a+h, b+k)$$

is unique for all values of k in a certain neighbourhood of $k=0$, then the repeated limit

$$\underset{k=0}{\text{Lt}}\ \underset{h=0}{\text{Lt}}\, f(a+h, b+k)$$

exists and is equal to the double limit.

It may evidently happen that, even when $\underset{x=a}{\text{Llt}}\, f(x, y)$ do not all coincide, the upper and lower limits, and therefore all intermediate limits, have an unique limit, which is the same for each, as y approaches b. If we agree to call this an unique repeated limit, it is evident that the above reasoning holds, with the small modification that the values $x_1, x_2, \ldots$ are not independent of y, but form for each value of y a sequence with the same property as before. Thus such an unique repeated limit is also a double limit.

It need hardly be remarked that the existence of an unique limit of $f(x, y)$ for each fixed value of y, does not, of course, involve that of an unique limit when y varies with x, which would be concomitant to the existence of an unique double limit.

The most general definition of a repeated limit, when neither of the simple limits involved is unique, and the corresponding theorem, which still holds, are not required in the present connexion [23].

III. CONTINUITY AND SEMI-CONTINUITY.

6. Upper and lower semi-continuity. If the upper limit at a is $\leqslant$ the value of f at that point, that is if (§ 3)

$$\phi(a) \leqslant f(a),$$

$f(x)$ is said to be *upper semi-continuous* at the point a. If on the other hand

$$f(a) \leqslant \psi(a)$$

$f(x)$ is said to be *lower semi-continuous* at a. If at every point of an interval, open or closed, $f(x)$ is upper (lower) semi-continuous it is said to be *an upper* (*lower*) *semi-continuous function*. In particular it is easily proved that $\phi(x)$ is an upper and $\psi(x)$ a lower semi-continuous function [δ].

Ex. 2. Let $f(x)=0$, when x is zero or irrational, and

$$f(x)=\frac{1}{q}, \text{ when } x=\frac{p}{q} \text{ is rational, } (0<x\leqslant 1),$$

p and q being integers prime to one another.

Then $f(x)$ is an upper semi-continuous function, and its $\psi(x)=0$, $\phi(x)=0$. By changing the sign of f we get a lower semi-continuous function.

It follows immediately from the definitions that *the points, if any, at which an upper* (*lower*) *semi-continuous function assumes values* $\geqslant k$ ($\leqslant k$) *form a closed set.*

Another important property is the following:—*An upper* (*lower*) *semi-continuous function assumes in every closed interval its upper* (*lower*) *bound in that interval.* For let $k_1 < k_2 < \ldots < k_n < \ldots$ be an ascending sequence of numbers having the upper bound U of an upper semi-continuous function $f(x)$ for upper bound. Then if G_n denote the closed set of points at which $f(x) \geqslant k_n$, there is, by Cantor's Theorem of Deduction, at least one point a common to all the successive sets G_n. Hence $k_n \leqslant f(a) \leqslant U$ for all integers n. But U is the upper bound of the k's, therefore $f(a) = U$.

Hence it easily follows[e] that *an upper (lower) semi-continuous function which is finite in* (a, b) *is bounded in some interval inside* (a, b).

THEOREM. *A monotone decreasing sequence of functions* $f_1 \geqslant f_2 \geqslant \dots$ *which are upper semi-continuous at a point P has for limit a function f which is also upper semi-continuous at P.*

For at a point where $f = +\infty$, it is, of course, upper semi-continuous, we have therefore only to prove the theorem at a point where

$$f(P) < A,$$

A being a finite quantity.

Since $f(P)$ is the limit of $f_n(P)$, we can determine m so that

$$f_m(P) < A,$$

and, since f_m is upper semi-continuous, we can find an interval d, containing P as internal point, such that *throughout it*

$$f_m(x) < A.$$

Since the sequence of functions is monotone decreasing, it follows that, for all values of $n \geqslant m$,

$$f_n(x) < A,$$

so that, throughout the interval d,

$$f(x) \leqslant A.$$

Since A was any quantity greater than $f(P)$, this shews that f is upper semi-continuous at P, which proves the theorem.

In the same way we have the corresponding theorem :—*A monotone increasing sequence of functions* $f_1 \leqslant f_2 \leqslant \dots$ *which are lower semi-continuous at P, has for limit a function f which is also lower semi-continuous at P.*

7. Continuity. If a function $f(x)$ is both upper and lower semi-continuous at the point a, that is, if it has an unique limit at a whose value is the same as $f(a)$, the function is said to be *continuous* at the point a. If it is continuous at every point of an interval, open or closed, it is said to be *a continuous function* in that interval.

We already know therefore that a continuous function has the following properties :—

The points, if any, at which a continuous function assumes values $\leqslant k$ *and those at which it is* $\geqslant k$, *and therefore those at which it is* $= k$, *all form closed sets.*

A continuous function assumes its upper and lower bounds in every closed interval.

It has also the following important property, which, as will be seen in the sequel, is not confined to continuous functions, and is shared by all differential coefficients, whether or no they are continuous :—

A continuous function assumes all values between its upper and lower bounds.

In fact since a continuum cannot be divided into two closed sets, the two closed sets of points at which respectively $f \geqslant k$ and $f \leqslant k$ (k being any value between the upper and lower bounds of f), must have a common point, at which accordingly $f = k$.

It follows from the second of the above properties that *a finite continuous function is bounded.* The word *continuous* is therefore often used as synonymous with *finite and continuous*, and this usage will in the present tract be adopted, except where the contrary is stated. It is in this sense that the word *continuous* must be understood in the following alternative definition which may be called the ϵ-definition of continuity.

A function is said to be continuous at a point if, given any positive quantity ϵ, however small, we can find a closed interval with that point as internal point, such that the difference between the upper and lower bounds of the function—the so-called oscillation of the function—in that interval is less than ϵ.

The theorem of § 6 gives us a standard test for continuity, viz. *a function which is at the same time the limit of a monotone descending and of a monotone ascending sequence of continuous functions is a continuous function.* It may be remarked that this condition is not only sufficient but also necessary[7].

The set of limits at a point of discontinuity of a function which is continuous throughout an open interval ending at that point is of the simplest character, namely a closed interval. To prove this we merely have to remark that if k is any value less than $\phi(a)$ and greater than $\psi(a)$, $f(x)$ assumes values both $> k$ and $< k$, and therefore assumes the value k at a point of the open interval. Thus k is certainly one of the limits, since it is a repeated value (§ 4). Therefore the set of limits consists of every value from $\phi(a)$ to $\psi(a)$ inclusive.

8. Pointwise discontinuous function. A function which, without being necessarily continuous, has in *every* interval a point of continuity is called a *pointwise discontinuous function.* In other words, the points of continuity of such a function are *everywhere dense*, without necessarily filling up any interval. It will now be proved that a finite *semi-continuous function is pointwise discontinuous*[ζ].

For simplicity of wording, we shall suppose $f(x)$ to be finite and lower semi-continuous, so that inside any chosen interval we can choose an interval in which $f(x)$ is bounded (§ 6) and

$$f(x) \leqslant \psi(x) \leqslant \phi(x) \quad\ldots\ldots\ldots\ldots(1).$$

We remark first, that, whatever be the nature of a bounded function $f(x)$, the points at which

$$\phi(x) - f(x) \geqslant k \quad\ldots\ldots\ldots\ldots(2)$$

can in no case fill up an interval. For, if a were an internal point of such an interval, we should have

$$\text{upper} \underset{x=a}{\mathrm{Lt}}\, \phi(x) \geqslant \text{upper} \underset{x=a}{\mathrm{Lt}}\, f(x) + k \geqslant \phi(a) + k,$$

which is not true, since, by § 6, $\phi(x)$ is upper semi-continuous and is bounded, since $f(x)$ is bounded.

In our case $\phi(x) - f(x)$ is the excess of an upper over a lower semi-continuous function and is therefore upper semi-continuous. Hence the points at which (2) holds, form a closed set, which is therefore, by what precedes, *dense nowhere*. There must therefore be an interval throughout which

$$\phi(x) - f(x) < k.$$

Repeating this process in this interval with $\frac{1}{2}k$ for k, and so on, we arrive at a point, internal to all the successive intervals, at which

$$\phi(x) - f(x) < \frac{k}{2^n}$$

for every value of the integer n. Therefore, since the left-hand side is by (1) not negative, it must, at this point, be zero. Hence, using (1) again, at this point,

$$f(x) = \psi(x) = \phi(x),$$

which proves the function to be continuous there, and therefore to possess a point of continuity in every interval.

We have seen that the limit of a monotone sequence of continuous functions is a semi-continuous function, it follows therefore from the above that it is a pointwise discontinuous function. Baire has proved[(5), (8)] that *the limit of a sequence of continuous functions is always pointwise discontinuous*, and not only so but that it *is pointwise discontinuous with respect to every perfect set*; that is to say, approaching a certain point a of any perfect set by means of points of that set in any manner, we shall get for $f(x)$ the unique limit $f(a)$.

The following example shews that, even when the sequence of continuous functions is monotone, the limiting function need not be continuous.

Ex. 3. Let $f_n(x)$ have at all the points $\frac{p}{q}$, where $q<n+2$, the value $\frac{1}{q}$, and between these points be linear, so that the locus $y=f_n(x)$ is a broken line. These functions are continuous, and they form a monotone descending sequence. Their limiting function is the function $f(x)$ of Ex. 2, § 6.

IV. DIFFERENTIATION.

9. The Incrementary Ratio. Derivates. Differential coefficient. Second and higher differential coefficients; these are repeated limits. If $f(x)$ is a (finite) continuous function of a single real variable x throughout a closed interval (a, b), *the incrementary ratio*

$$m(x, y) \equiv \frac{f(x)-f(y)}{x-y}$$

is a function of the ensemble (x, y) defined at all points of a certain closed square except on the diagonal $x=y$, and is continuous at every point at which it is defined. It has, like a proper continuous function, the property of *assuming every value between its upper and lower bounds*, as may be proved without difficulty(15).

The limits with respect to y of $m(x, y)$ at any point on the diagonal are called the *derivates* of $f(x)$ at the point x; they are called *right or left-hand derivates*, according as $y>x$ or $y<x$, and, in particular, the upper and lower limits on the two sides are called *the upper and lower derivates* on the two sides(7).

If the derivates all coincide, their common value is called *the differential coefficient of* $f(x)$. In other words $f(x)$ has at the point x a differential coefficient provided

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

has as h approaches zero in any manner an unique limit, and this limit is the value of the differential coefficient; if this is true when h is positive (negative) the value is that of the *right-hand* (*left-hand*) *differential coefficient**.

Here we are expressly assuming that $f(x)$ is finite and continuous. At a point at which $f(x)$ is not finite, or is discontinuous, we shall say that a differential coefficient does not exist. It may also not exist at a point at which $f(x)$ is continuous. It should, however, be remarked

* Notice that by the differential coefficient at the left (right) hand end-point of the interval considered we mean the right (left) hand differential coefficient.