IV

THE INTEGRAL AND ITS GENERALIZATIONS

THE definite integral can be considered from several points of view. In college text-books it is defined for functions of simple type (continuous and with a finite number of maxima and minima) as the limit of a certain sum; but in its applications to mechanics it is expressed frequently in a form which does not correspond rigorously to the definition. A moment of inertia, for example, may be defined in the form $\int r^2 dM$, where M is not equal to r, nor is r a function of M. The idea of integral is broader than the elementary type.

It is of some interest to consider the broader notion, to analyze it and reduce it as far as possible to a logical form amenable to the methods of mathematics. For this purpose it is helpful to consider some of the less mathematical and less logical concepts which are felt to be akin to integrals. The first to occur to an applied mathematician is that of a moment. We have a fairly clear "a priori" idea of what is meant by the moment of a mechanical structure about some axis. It bears the same relation to rotation as does a force to parallel motion. It can easily be measured experimentally. The same is true of moment of inertia or rotational inertia. A similar idea, though not so simple, is that of the total potential due to an electrostatic or gravitational system.

In all these cases we have the familiar space and time as a background. In a more general and vaguer field of thought there is the average. Statistical averages are obtained for many kinds of entities, numerically measurable by necessity, but entities dependent on objects or qualities which are not always numerical in their essence. These averages are, in practice, merely ratios of sums of a finite number of terms, but there is a feeling in many problems that there is a ghostly "correct" average to which we are but approximating by our rough methods. It is true that in some cases this ghost is misleading and illogical, as in the case of mortality statistics (unless we are considering the biological problem of the average mortality of all living tissue). Nevertheless one feels frequently that if it could only be grasped there is a limiting true average even in problems and sciences in which space and time or even pure number enter in a secondary manner, in which these appear as conditions rather than as fundamental characteristics. In mathematics itself the modern tendency is to consider not only collections of numbers and relations between numbers, but collections of functions, of curves, of logical classes, and so on. And there appears no absolute objection to such ideas as that of the average maximum of a collection of functions or the average area of a collection of closed curves, except that these concepts are vague and had not until recently been defined in a satisfying form. We shall see that at the present time this objection has fallen to the ground.

How can we characterize this broader concept of integral? From the examples given and others which the reader may invent there emerge several characters which are universally found in any concepts which are akin to integrals or moments or averages.

There must be a background of entities which may be denoted by certain marks. In a simple average or in the moment of a finite set of particles these marks may be suffixes 1, 2, 3, .. which distinguish one individual from another. In the moment of a continuous structure the positive integers may not be sufficient. The number of "individuals" may be too great. We therefore denote the general mark by a noncommittal symbol "p," which, for imaginative purposes, can be thought of as a "point" in some general geometry. A number or a point in space, a function or a curve, a color or a quality, may be such a mark.

The collection of these marks considered in the particular case is called P_0. There is again a numerical property of these marks, determinate when a particular mark p is chosen, that is to say, a function of p denoted by f(p). For a moment of inertia p is a particle or material point, f(p) is the square of the distance of that point from the axis. For an average error, p stands for a particular case considered, f(p) for the error involved in that case. For an ordinary integral the mark p is a number x lying in an interval a to b and f(p) = f(x) is some function of x. For a functional space p is a function belonging to a class P_0, f(p) is a functional of the function p.

In the case of closed curves p may be a closed curve of some simple type, f(p) the Green's function for that curve and two fixed points A, B. [A Green's function is usually regarded as a function of one variable point for a fixed curve and fixed point in the enclosed area, but it is convenient at times to hold the two points fixed and vary the curve.[1]]

The next step is, somehow, to sum f(p) over all marks p belonging to P_0. If the number of marks is limited we have an ordinary algebraic sum; if the collection P_0 has the power of the denumerable the sum is an infinite series; if P_0 is an interval in a number-space of a finite number of dimensions the sum is an integral. If P_0 is a continuous material body the sum may be a total moment or total potential at a fixed point.

[1] G. C. Evans, "Cambridge Colloquium," New York, 1918, p. 15.

But the summation cannot take place as a simple sum of the values of $f(p)$ for all p belonging to P_0 if P_0 has a power greater than the denumerable, because that would, in general, lead to an infinite result. It is necessary to proceed by a process of approximation. P_0 is subdivided into a finite number of subgroups $P_1, P_2, \ldots P_n$, to each of which some weight is assigned. In each subgroup P_r, a typical mark p_r is chosen and the value $f(p_r)$ of f at p_r is determined and multiplied by the weight m_r of the group. These products are added in the form

$$S_1 = m_1 f(p_1) + \ldots + m_n f(p_n).$$

A further subdivision is made with the result S_2 and so on. The sequence $S_1, S_2, \ldots$ may approach a limit which may be the same under many different processes of subdivision of P_0 and many different methods of choice of the representative of each group. The limiting sum is then called the integral of $f(p)$ over P_0 with the weighting system $m(P)$, P being a variable subgroup of P_0.

The integral depends, we see, not only on the existence of a function $f(p)$, but also on a method of weighting $m(P)$ a function of collections P of the marks p, picked out from P_0.

We now introduce a notation for our integral in order to be able to speak of it. Since it is a number which is determinate when the function $f(p)$ (p = element of P_0) and the weighting $m(P)$ (P = subgroup of P_0) are given, we call it

$$\int f(p)dm(P),$$

or $\int fdm$, or, when we are not so interested in the weighting, $S(f)$.

If there is to be any analogy with mechanics or statistics, the weighting $m(P)$ must be such that the sum of the weights of any number of distinct collections $P_1, P_2, \ldots P_n$ is equal to the weight of the group consisting of all elements belonging to any one of them, that is, if $P = P_1 + P_2 + \ldots + P_n$,

$$m(P) = m(P_1) + m(P_2) + \ldots + m(P_n).$$

Now in such a subject as this, limiting processes and infinite series will naturally occur and difficulties will constantly appear unless we allow the weighting to be additive even for an infinite set of subgroups. That is to say, when

$$P = P_1 + P_2 + \ldots,$$
$$m(P) = m(P_1) + m(P_2) + \ldots,$$

provided there is no p common to two groups P_n.

Integrals which do not satisfy this relation are useful in some fields, but each requires its special treatment. Here we confine our attention to cases in which the additive property holds for an infinite as well as for a finite sum.

Now it happens that the sum of a set of classes or collections is independent of the order of the summation; for, by its definition, the sum of a set of classes is the class of elements belonging to some one class of the set. It follows that the sum $m(P_1) + m(P_2) + \ldots$, which is a numerical series, must be independent of the order of its terms. In the usual case the weights m(P) are essentially positive (or zero), and then the above series has the same sum in any order if it is convergent at all. But in electrostatics the charge in a volume can be regarded as a weight of the collection of points contained in the volume, and it will satisfy all the usual requirements for a weight except that of constant sign. To be as general as possible we should allow the possibility that m(P) is sometimes negative. Then the series $m(P_1) + m(P_2) + \ldots$ must be absolutely convergent.

It is of considerable value to distinguish in some way integrals which are based on essentially positive weight from the more general type, in the first place because the former are more common in practice and satisfy more completely the properties of ordinary integrals; secondly, because, as we shall show, the other apparently more general sum can be regarded as the difference of two integrals with positive weighting. For an integral with positive weighting we use the symbol I(f) instead of S(f).

Returning to our attempted definition of the integral in terms of the weights, we have

$$S(f) = \lim S_n,$$

$$S_n = \sum_{i=1}^{i_n} f(p_{i,n})m(P_{i,n}).$$

It follows that if, for every p in P_o,

$$f(p) = f_1(p) + f_2(p) + \ldots + f_n(p),$$
$$S(f) = S(f_1) + S(f_2) + \ldots + S(f_n).$$

Again as with the weight m(P) so here with the function f, our integral will be more amenable to analysis if it satisfies this additive law not only for a finite but also for a denumerably infinite set of functions; that is, if when

$$f(p) = f_1(p) + f_2(p) + \ldots,$$
$$S(f) = S(f_1) + S(f_2) + \ldots.$$

Now, since S(f) is defined by means of a limit, this statement involves the validity of a change in order of a double limit. Again this double limit can be expressed as a double series in the following manner:

In the process by which S_n is obtained we subdivide P_o successively into parts so that $P_{i,n}$ is further subdivided into a set of distinct collections $P_{j,n+1}$ where j takes on a set of values depending on i. Or, again, $P_{i,n}$ itself is a part of one of the collections $P_{j,n-1}$ of the previous stage. Let

$$T_o(f) = f(p_{1,o})m(P_o),$$

$$T_n(f) = \sum_{i=1}^{i_n} [f(p_{i,n}) - f(p_{j,n-1})]m(P_{i,n}), n = 1, 2, \ldots.$$

Then

$$S(f_r) = \sum_{n=o}^{\infty} T_n(f_r),$$

$$S(f_1) + S(f_2) + \ldots = \Sigma_r \Sigma_n T_n(f_r),$$
$$S(f) = \Sigma_n T_n(f) = \Sigma_n \Sigma_r T_n(f_r).$$

Such a change of order of summation in a double series we know to be legitimate if the double series is absolutely convergent. This suggests the restriction of the series $f_1 + f_2 + \ldots$ to absolutely convergent series.

We are now in a position to state the following general property of our integral:

If $f_1(p) + f_2(p) + \ldots$ is absolutely convergent and equal to $f(p)$ for all marks p in P_0, then

$$S(f) = S(f_1) + S(f_2) + \ldots .$$

Since the sum $f_1 + f_2 + \ldots$ is the same under any derangement of the terms, the sum $S(f_1) + S(f_2) + \ldots$ will be equal to $S(f)$ for any derangement and the series is absolutely convergent.

To distinguish this type of general integral from other types, for example that of E. H. Moore,[1] we call it an absolute integral.

Our problem is now explained in general terms. It remains to give it a more detailed and more logical examination.

Given a body of definitions it would be possible to examine the whole as a logical system for the consistency and independence of its postulates; but before this can be done it is necessary to determine some process of definition.

At first there may only be a small class of functions for which the initial defining processes, such as the method of subgroups and their representatives outlined above, give a single result. The analysis must be developed by stages proceeding from a smaller to ever wider classes of functions for which the integral is defined. And this extension will be

[1] E. H. Moore, "Bulletin of the American Mathematical Society," Vol. XVII (1912), p. 334.

most naturally obtained through sums of absolutely convergent series. We shall also use the principle that when a function lies between two other functions whose integrals are defined and equal, then this common value is also the integral of the given function. This principle, however, can only be used in the case of the positive integral I(f).

Before we proceed any further we introduce two concepts which are extremely important in connection with integrals of absolute type.

To any collection P contained in P_0 there corresponds a function f(p) equal to 1 when p belongs to P, and to 0 when p does not belong to P. Given the collection P, f(p) is completely and uniquely determined. Vice versa if a function of this type (equal everywhere either to 1 or 0) is given, it determines the collection P of marks p for which it is equal to 1. For such a function the integral will equal the weight of the collection P. For example, in a statistical problem the weight of any subgroup is also the average of the numbers equal to 1 in the subgroup and to 0 for all other marks or subgroups. Hence to state the weighting of the integral is the same thing as to state the value of the integral for a certain simple class of functions.

To express the problem in this way, in terms of an initial class of functions, has several advantages. On the one hand we may prefer to use for the purposes of our exposition a different class of initial functions; for example, functions defined by certain series or polynomials; on the other hand the first step from these special functions to the next class may suggest the processes of extension to even wider classes.

The second concept is related to monotone sequences. If a series is absolutely convergent it is the difference of two convergent series of positive (strictly, non-negative) terms, in which the partial sums form monotone sequences. Monotone sequences and absolute convergence are therefore intimately related and this implies that monotone sequences are essential items in the theory of integrals of absolute type. Now in order to obtain such sequences it is valuable, if not necessary, to consider the greater of two numbers.

Let A be the class of numbers not greater than the number a, B the class of numbers not greater than b, and so on. In logic there are two combinations of classes, the logical sum $A \vee B$ which consists of all numbers not greater than one of the two numbers, that is, not greater than the greater of a and b, and the logical product $A \wedge B$, which consists of all numbers not greater than either of the two numbers, that is, not greater than the lesser of a and b. This suggests the following symbols and the corresponding properties:

$$a\gamma b = \text{the greater of a and b},$$
$$= \tfrac{1}{2}(a + b) + \tfrac{1}{2} \mid a - b \mid.$$
$$a\lambda b = \text{the lesser of a and b},$$
$$= \tfrac{1}{2}(a + b) - \tfrac{1}{2} \mid a - b \mid.$$
$$-(a\lambda b) = (-a)\gamma(-b), \quad -(a\gamma b) = (-a)\lambda(-b),$$
$$(a\gamma b)\gamma c = a\gamma(b\gamma c), \quad (a\lambda b)\lambda c = a\lambda(b\lambda c),$$

if $a \leqq c$, $(a\gamma b)\lambda c = a\gamma(b\lambda c)$, and we have the right to use the notation $a\gamma b\gamma c$, $a\lambda b\lambda c$, and (when $a \leqq c$) $a\gamma b\lambda c$. Moreover

$$a\gamma b + a\lambda b = a + b,$$
$$a = a\lambda b + (a - b)\gamma 0.$$

Many other similar identities could be proved by inspection or by correlation with the logical classes A, B, ...

Again, if $a \leqq b$, $a\gamma c \leqq b\gamma c$, $a\lambda c \leqq b\lambda c$, and if $\lim a_n = a$, $\lim a_n\gamma c = a\gamma c$, $\lim a_n\lambda c = a\lambda c$.

We can extend the same notation to functions so that $f\gamma g$ is that function which for each p is equal to the greater of f(p) and g(p). The above equalities and inequalities will be immediately applicable to functions as well as real numbers. One point may be noticed, that if f_1 is the function corresponding to the collection P_1 and f_2 to P_2, then $f_1\gamma f_2$ corresponds to $P_1 + P_2$ and $f_1\lambda f_2$ to P_1P_2.

Now if an initial class of functions is given which possess integrals in such a way that the additive postulate is satisfied, then we can proceed to define the integral of any linear combination of these functions and this will yield consistent results. The chief difficulty consists in the next stage, which is not immediately obvious. One important requirement is that a distinctly dynamic and constructive process should be present in order that the integral may be defined for a far wider class of functions than the initial class. E. H. Moore has proceeded by considering limits of relatively uniform convergent sequences. This leads to an interesting development to which the reader is referred.[1] It is, however, a process distinctly different—as a generative process—from the processes invented by Borel, Lebesgue, and W. H. Young. We prefer to follow the latter development. In this there is maintained on the one hand a constant correlation with the weight of subgroups, and on the other hand with a type of absolute convergence. These two are related more closely than might at first be imagined. For just as a series Σu_n is absolutely convergent when and only when the series $\Sigma \mid u_n \mid$ is convergent, so in the theory of Lebesgue integrals, f(x) is summable if, and only if, $\mid f(x) \mid$ is summable.

In the case where p is a real number between 0 and 1 let us assume, as an example to show the relations between the weights of intervals and absolute summability, that the initial class of functions is such that any linear combination of functions of the class and the modulus of any function of the class are members of the class. Suppose also that $f(x) = 1$ $(0 \leqq x \leqq 1)$ and $f(x) = x$ $(0 \leqq x \leqq 1)$ are among

[1] E. H. Moore, "Bulletin of the American Mathematical Society," Vol. XVII (1912), p. 334.

the members of the initial class. Then it will be possible to assign a weight to any interval. For the function

$$0 \gamma (mx - b) \lambda 1$$

is equal to 0 from $x = 0$ to $x = b/m$, thence linear to $x = (b + 1)/m$ where it is equal to 1 and remains equal to 1 up to $x = 1$. Since $f\gamma g$ is a linear combination of f, g, and $| f - g |$, $f\gamma g$ is a member of the initial class if f and g are members. Therefore the function above defined possesses an integral. Now make m and b increase indefinitely but in such a way that $(b + 1)/m$ remains equal to c $(0 \leqq c \leqq 1)$. The limiting function will be equal to 0 for $x < c$ and to 1 when $x \geqq c$. If the limiting process is valid the integral of this function can be regarded as the weight (measure) of the closed interval $c \leqq x \leqq 1$. Subtracting another interval $d < x \leqq 1$ (this time keeping b/m fixed and equal to d) we obtain the weight of any closed interval (c, d).

Reciprocally, if the original aspect is that of the weighting of subgroups, the set of marks p for which $| f(p) |$ has a given value is simply the sum of those for which f(p) and for which $-f(p)$ have that value. Hence the determination of the integral of $| f |$ will be of the same type as for f and $-f$.

With this general somewhat diffuse and nonlogical analysis of the foundations performed we are in a position to make a careful attack of the problem.

We assume, first, that there are certain marks p belonging to a class P_0; secondly, that there is a class T_0 of functions f(p) of these marks defined for all p in P_0; thirdly, that this class is such that when f, f_1 belong to the class so also do $| f |$, cf (where c is any real constant), and $f + f_1$.

For each of these functions of the fundamental class T_0 there is defined in some way an integral S(f) which satisfies the following postulate:

(A) If $f_1(p)$, $f_2(p)$, . . . form a finite or denumerable set of functions belonging to T_0, if their sum is f(p) and the sum $|f_1| + |f_2| + \ldots$ is g(p) for every p in P_0, where f(p), g(p) are also members of T_0, then

$$S(f) = S(f_1) + S(f_2) + \ldots .$$

This postulate can be analyzed into at least three postulates.

(C) $S(cf) = cS(f)$, if c is a real constant and if f belongs to T_0.

(S) $S(f_1 + f_2) = S(f_1) + S(f_2)$, if f_1, f_2 belong to T_0.

(L) $\lim S(f_n) = 0$, if $f_1(p) \geqq f_2(p) \geqq \ldots \geqq 0 = \lim f_n(p)$ for every p in P_0 where $f_1, f_2, \ldots$ belong to T_0.

The postulate (A) implies these three together for (S) is a special case of (A), (L) is a special case if we substitute $f_n' = g - (|f_1| + |f_2| + \ldots + |f_n|)$ where

$$g = |f_1| + |f_2| + \ldots \text{ in (A)}.$$

Finally (C) is a consequence of (S) when the constant is rational, and of this together with (A) when c is the sum of an absolutely convergent series of rational numbers, which is true for any real value.

But (A) implies more than is implied by (C S L) together. Let $f(p) \geqq 0$ for all p in P_0 and consider some finite set of functions $\phi_1, \phi_2, \ldots \phi_n$ such that $\phi_i(p) \geqq 0$ for all p in P_0 ($i = 1, 2, \ldots n$) and such that

$$\phi_1 + \phi_2 + \ldots + \phi_n = f.$$

Then, when the ϕ's are varied, postulate (A) implies that $|S(\phi_1)| + |S(\phi_2)| + \ldots + |S(\phi_n)| \leqq I(f)$ where I(f) depends only on f and not on the particular subdivision of f.

For, if not, given any integer s we could choose the ϕ's in such a way that

$$\phi_{1,s} + \ldots + \phi_{n_s,s} = f, \ \phi_{i,s} \geqq 0, \ i = 1, 2, \ldots n_s,$$

$$|S(\phi_{1,s})| + \ldots + |S(\phi_{n_s,s})| > 2^s,$$

or if $\psi_{i,s} = \phi_{i,s}/2^s$,

$$\psi_{1,s} + \ldots + \psi_{n_s,s} = f/2^s,$$

$$| S(\psi_{1,s}) | + \ldots + | S(\psi_{n_s,s}) | > 1.$$

Now if $f = \Sigma_s f/2^s$

$$= \Sigma_s \Sigma_{i=1}^{n_s} \psi_{i,s}$$

which can be expressed as a single series of positive terms which is absolutely convergent, then postulate (A) asserts that

$$\Sigma_s \Sigma_i | S(\psi_{i,s}) |$$

is convergent, in contradiction to our result that each partial sum of terms summed for a particular s is greater than 1. Hence there does exist such an upper bound I(f).

Let us express this inequality in a simpler form. If f is a member of T_0 (all our functions have so far been assumed to be of this class), $|f|$ is a member of T_0, and if $f_1 = f \gamma 0$, $f_2 = - f \lambda 0$,

$$f = f_1 - f_2, \; | f | = f_1 + f_2.$$

Now f_1 and f_2 are non-negative and therefore by the analysis immediately preceding,

$$\begin{aligned} | S(f) | &= | S(f_1) - S(f_2) | \\ &\leqq | S(f_1) | + | S(f_2) | \\ &\leqq I(| f |). \end{aligned}$$

A further consequence of (A) is therefore

(M) $$| S(f) | \leqq I(| f |)$$

if f is a member of class T_0, where $I(\phi)$ is a non-negative number defined for all non-negative functions ϕ belonging to T_0 such that if $\phi \leqq \psi$, $I(\phi) \leqq I(\psi)$.

The latter part of this statement follows immediately from the definition of $I(\phi)$ as an upper bound (by upper bound we mean "least upper bound").

It will later be proved that the postulate system C S L M together imply (A) and therefore (A) and (C S L M) are logically equivalent.

Let us first consider positive integrals. A positive integral is such that the weight of any subgroup P is non-negative. This implies the more general assertion that if $f(p) \geqq 0$ for all p in P_0, $I(f) \geqq 0$. This we can designate postulate P. Taken in conjunction with (S) or (A) it implies that if $f \leqq g$, $I(f) \leqq I(g)$; and this again that

$$|I(f)| \leqq I(|f|),$$

which is a special form of (M).

Then AP imply C S L P while C S L P imply C S L M. We can now prove that C S L imply A, so that the two latter systems are equivalent, and since (A) implies (M), (M) is not, after all, independent of C S L, but is really implied by them.

For let $f_1, f_2, \ldots$ be a denumerably infinite set of functions satisfying the conditions of postulate (A). Let

$$h_n = g - (|f_1| + |f_2| + \ldots + |f_n|),$$

$$k_n = (g - f) - [(|f_1| - f_1) + (|f_2| - f_2) + \ldots + (|f_n| - f_n)].$$

Then h_n, k_n satisfy the conditions of postulate (L) and if this is assumed, $\lim S(h_n) = 0$, $\lim S(k_n) = 0$. Combining these with C S we see that (A) is satisfied. Furthermore if c is a real constant, it is not difficult to prove that S L imply C. On the other hand S and L are independent, for if

$S(f) = \lim f(x)$ as x approaches 0 through positive values,

S is satisfied but not L in general; while if

$$S(f) = f^2(0),$$

L is satisfied but not S.

Again P is independent of L and S. Therefore we can write our postulates as an irreducible minimum in the form (A) for the general integral S(f), or in the equivalent form (S L), and then add for the positive integral I(f) the postulate (P) to either system. But it is helpful to remember that C, M are consequences of either system.

We shall show how the general S-integral can be expressed as the difference of two positive I-integrals, then consider how the positive integral is extended and state a number of examples illustrating some applications of the general theory. After that we shall state some of the more important theorems which are needed in applications.

If $f(p) \geqq 0$, for all p in P_0 and if f is of class T_0, define $I_1(f)$ as the upper bound of $S(\phi)$ for all functions ϕ of class T_0 such that $0 \leqq \phi \leqq f$. This upper bound exists, since by M,

$$S(\phi) \leqq I(|\phi|) = I(\phi) \leqq I(f).$$

[At present there is no relation between I_1 and I, and neither has been proved to be a positive or I-integral.]

The function $\phi = 0$ is a member of T_0 and its S-integral is 0, and therefore

$$I_1(f) \geqq 0,$$

so that I_1 satisfies postulate P.

Let f_1, f_2 be two non-negative functions of class T_0. If $0 \leqq \phi_1 \leqq f_1$, $0 \leqq \phi_2 \leqq f_2$, then $0 \leqq \phi_1 + \phi_2 \leqq f_1 + f_2$. However ϕ_1, ϕ_2 may be varied, and they can be varied independently,

$$S(\phi_1) + S(\phi_2) = S(\phi_1 + \phi_2) \leqq I_1(f_1 + f_2),$$

$$I_1(f_1) + I_1(f_2) \leqq I_1(f_1 + f_2).$$

On the other hand, if $0 \leqq \phi \leqq f_1 + f_2$, $\phi - f_1 \leqq f_2$.

Now $\phi = \phi \lambda f_1 + (\phi - f_1)\gamma 0 = \phi_1 + \phi_2,$

where $0 \leqq \phi_1 \leqq f_1$, $0 \leqq \phi_2 \leqq f_2$. Hence

$$S(\phi) = S(\phi_1) + S(\phi_2) \leqq I_1(f_1) + I_1(f_2).$$

But this is true however we vary ϕ and therefore

$$I_1(f_1 + f_2) \leqq I_1(f_1) + I_1(f_2).$$

When these two inequalities are combined we see that $I_1(f)$ satisfies postulate S, at least if $f_1, f_2 \geqq 0$.

For any function of class T_0 we define

$$I_1(f) = I_1(f\gamma 0) - I_1(-f\lambda 0).$$

If $f = g - h$ where g, h are non-negative, $g \geqq f\gamma 0$, $g = f\gamma 0 + k$, $h \geqq -f\lambda 0$, $h = k - f\lambda 0$, where $k \geqq 0$.
By the case already considered,

$$I_1(g) - I_1(h) = I_1(f\gamma 0) + I_1(k) - I_1(k) - I_1(-f\lambda 0)$$
$$= I_1(f).$$

By this means it is possible to show that postulate S is satisfied by all functions of class T_0.

Instead of proving L directly we shall prove A in the case where the functions summed are non-negative. Let

$$f_1 + f_2 + \ldots = f$$

where all these functions are of class T_o and non-negative. Let ϕ be of class T_0 and such that $0 \leqq \phi \leqq f$. Take

$$\phi_1 = \phi\lambda f_1,\ \psi_1 = (\phi - f_1)\gamma 0,$$
$$\phi_2 = \psi_1\lambda f_2,\ \psi_2 = (\psi_1 - f_2)\gamma 0,$$

and so on.

Then $0 \leqq \phi_n \leqq f_n$, and $\phi = \phi_1 + \psi_1$,
$= \phi_1 + \phi_2 + \psi_2$,
$= \ldots\ldots$

Also $\psi_1 - f_2 \leqq \psi_1$, $0 \leqq \psi_1$, and therefore $\psi_2 \leqq \psi_1$. Similarly $\psi_3 \leqq \psi_2$, $\psi_4 \leqq \psi_3$, ...
Now $\phi - f_1 \leqq f - f_1$, so that $\psi_1 \leqq f - f_1$; $\psi_1 - f_2 \leqq f - f_1 - f_2$, so that $\psi_2 \leqq f - f_1 - f_2$, and so on. Then since $\psi_n \geqq 0$, $\lim \psi_n = 0$. Therefore

$$S(\phi) = S(\phi_1) + S(\phi_2) + \ldots + S(\phi_n) + S(\psi_n)$$
$$\leqq I_1(f_1) + I_1(f_2) + \ldots + I_1(f_n) + S(\psi_n).$$

But if we choose any functions $g_1, g_2, \ldots$ such that $0 \leqq g_n \leqq f_n$, $g_1 + g_2 + \ldots + g_n \leqq f$, and

$$S(g_1) + S(g_2) + \ldots + S(g_n) \leqq I_1(f),$$
$$I_1(f_1) + I_1(f_2) + \ldots + I_1(f_n) \leqq I_1(f).$$

Therefore the series of non-negative terms

$I_1(f_1) + I_1(f_2) + \ldots$ is convergent and its sum is not greater than $I_1(f)$. To return, since $\lim S(\psi_n) = 0$,

$$S(\phi) \leqq I_1(f_1) + I_1(f_2) + \ldots,$$

and since ϕ can be varied at random this proves a second inequality which with the other proves that

$$(I_1(f) = I_1(f_1) + I_1(f_2) + \ldots .$$

To prove L itself we substitute f_1 for f, f_2 for $f - f_1, \ldots$ in the above. Then we have proved that $I_1(f)$ is an integral satisfying SL and it is also of positive type.

Now we define $I_2(f) = I_1(f) - S(f)$, and I_2 will also be an integral of positive type, for among the functions ϕ we can choose f itself in the definition of $I_1(f)$.

Then we also take $I(f) = I_1(f) + I_2(f)$ and this is a positive integral. I_1 is called the positive, I_2 the negative and I the modular integral associated with S(f). The relations between them are

$$S(f) = I_1(f) - I_2(f), \; I(f) = I_1(f) + I_2(f).$$

We have now proved that any S-integral is the difference of two positive integrals.

$$\begin{aligned} |S(f)| &= |S(f\gamma 0) - S(-f\lambda 0)| \\ &\leqq |S(f\gamma 0)| + |S(-f\lambda 0)| \\ &\leqq I_1(f\gamma 0) + I_2(f\gamma 0) + I_1(-f\lambda 0) + I_2(-f\lambda 0) \end{aligned}$$

since all these integrands are positive. But $I = I_1 + I_2$, $|f| = f\gamma 0 - f\lambda 0$ and

$$|S(f)| \leqq I(|f|).$$

If $f \geqq 0, \phi_1, \phi_2, \ldots \phi_n \geqq 0$ and if

$$\begin{aligned} &\phi_1 + \phi_2 + \ldots + \phi_n = f, \\ &|S(\phi_1)| + |S(\phi_2)| + \ldots + |S(\phi_n)| \\ &\qquad \leqq I(\phi_1) + I(\phi_2) + \ldots + I(\phi_n), \\ &\qquad = I(f). \end{aligned}$$

Again, given any positive e and given $f \geqq 0$, we can choose ϕ so that $0 \leqq \phi \leqq f$ and $S(\phi) > I_1(f) - e/2$.

If we choose $e < 2I_1(f)$, $S(\phi) > 0$ so that $|S(\phi)| = S(\phi)$.

$$\begin{aligned} |S(f - \phi)| &= |I_1(f) - S(\phi) - I_2(f)| \\ &\geqq I_2(f) - [I_1(f) - S(\phi)] \\ &> I_2(f) - e/2. \end{aligned}$$

Therefore

$$|S(\phi)| + |S(f - \phi)| > I_1(f) + I_2(f) - e$$
$$= I(f) - e.$$

From these two inequalities we draw the conclusion that this $I(f)$ defined in terms of I_1 and I_2 is the same upper bound of $|S(\phi_1)| + |S(\phi_2)| + \dots + |S(\phi_n)|$, where $\phi_1, \phi_2, \dots$ are non-negative functions whose sum is f, as that upper bound which we called $I(f)$ previously in our analysis of postulate (A).

In extending the definition of the integral to functions outside the initial class the natural step is to functions which are the sums of absolutely convergent series of functions of class T_0. Since such a series is the difference of two series of positive terms, and to the latter correspond increasing sequences it is simpler, as W. H. Young has shown, to consider the latter first.

If $f_1 \leqq f_2 \leqq \dots$ is a non-decreasing sequence of functions of class T_0, the initial class, then $\lim f_n$ exists in any case if we allow $+\infty$ as a possible value, and we say that $\lim f_n = f$ is of class T_1. It follows that

$$I(f_1) \leqq I(f_2) \leqq \dots,$$

if I is some positive integral, and $\lim I(f_n)$ exists if, again, we allow $+\infty$ as a possible value.

This limit has been shown by the author[1] to be independent of the actual sequence defining f and to depend only on the limit f itself. We have a right to call the limit $\lim I(f_n)$ the integral of f, provided it is finite. If this is the case we define $I(f) = \lim I(f_n)$ and say that f is a summable member of T_1. It can then be shown that this new integral also satisfies postulates SLP. If f is any function, we define the upper semi-integral of f, $\dot{I}(f)$ as the lower bound of $I(g)$ for all functions g of class T_1 such that $g \geqq f$. In other

[1] P. J. Daniell, "Annals of Mathematics," Vol. XIX (1918), p. 279.

words, we strive to define the integral of f by means of numbers which it cannot exceed and call the least possible of such upper boundaries for its value a semi-integral until we can ascertain that the integral itself is definite and unique. At the same time we define the lower semi-integral of f, $\underset{\cdot}{I}(f)$ as the negative of the upper semi-integral of $-f$. If it happens that the two semi-integrals are equal then f is said to be summable and its integral is defined to be the common value of the two semi-integrals.

The class of summable functions has the same properties as the initial class T_0. Indeed if f, f_1 are summable so are $|f|$, cf where c is a real constant, and $f + f_1$. Also the postulates SLP and with them A and C are satisfied where in A it is necessary, in general, to retain the condition that $|f_1| + |f_2| + \ldots$ is not greater than some summable function, such as a summable member of T_1, though any summable function serves the same purpose and gives no greater generality.

The fundamental theorem of this class of integrals is that if $f_1, f_2, \ldots$ is a sequence of summable functions converging to a limit f, and if a summable function g exists such that $|f_n| \leqq g$ for all values of n, then f is summable and lim $I(f_n) = I(f)$.

Another interesting property which suggests another method of development is that the necessary and sufficient condition that f be summable is that, given any positive e, it is possible to find a function f_e of the initial class T_0 such that

$$\dot{I}(|f - f_e|) < e.$$

In other words, a summable function must be within any "distance" however short of some member of T_0, measuring "distance" in the function space by means of the upper semi-integral of the modular difference of the functions. This geometrical analogy is very fruitful but cannot here be entered into.

For the general definition of S(f) we take

$$S(f) = I_1(f) - I_2(f),$$

provided f is summable with respect to both I_1 and I_2, which will happen whenever f is summable with respect to I the modular integral associated with S.

Sometimes an operation J(f) is given in some way for an already wide class of functions. It is an important problem to decide whether it coincides with an integral I(f) as defined above by extension. Evidently, in the first place, the two must be identical for members of the class T_0. If J(f) satisfies also, for all functions f for which it happens to be defined, the inequality

$$J(f) \leqq J(g), \text{ when } f \leqq g,$$

then it is sufficient to prove the identity of J(f) and the extension of I(f) for all summable members of class T_1. Or, what is the same thing, that when f is the limit of a non-decreasing sequence of functions f_n of class T_0,

$$J(f) = \lim J(f_n).$$

Again, since, of necessity, $\lim J(f_n) \leqq J(f)$, f_n being not greater than f, it is sufficient to show that $J(f) \leqq \lim J(f_n)$.

Illustrative examples.

(1) Absolutely convergent series. Let P_0 be the class of positive integers 1, 2, 3, . . . with weights $w_1, w_2, \ldots$ assigned of such a character that $|w_1| + |w_2| + \ldots$ is convergent. Let f_n be a function of the integral variable n. Then in some cases, for example when f_n is uniformly limited, the series

$$S(f) = f_1w_1 + f_2w_2 + \ldots$$

will be absolutely convergent.

By what is known of absolutely convergent double series postulate A is satisfied. In this case we can define S(f) without the necessity of proceeding by successive stages.

However we could let T_0 be the class of functions f_n equal to a number other than 0 for a finite set of values of n and 0 for all the others. A function f_n which is non-negative for all sufficiently large values of n would then be of class T_1.

$I_1(f)$ will be the sum of $f_n w_n$ for all non-negative w_n and $I(f)$ the sum $\Sigma_n f_n \mid w_n \mid$.

(2) Let P_0 be the class of real numbers, x, where $0 \leqq x \leqq 1$. Let weighting be assigned to each interval equal to its length.

For T_0 we choose the class of step-functions f(x) of the type

$$f(x) = c_i, \; x_{i-1} \leqq x \leqq x_i,$$

where (0, 1) has been divided into some finite set of intervals

$$0 = x_0 < x_1 < \ldots < x_n = 1.$$

If we wish it is sufficient, and leads finally to no loss in generality, to confine T_0 to be the class of functions of this type restricting the division points $x_1, x_2, \ldots x_n$ to be those obtained by dividing the interval (0, 1) into $n = 2^r$ equal parts, r being some integer. We then define

$$I(f) = \int f(x)dx = \Sigma_i c_i(x_i - x_{i-1}).$$

This integral is of positive type, for when $f(x) \geqq 0$, every $c_i \geqq 0$ and $I(f) \geqq 0$. This is the usual integral when the integrand f is sufficiently restricted. The extension given in this paper leads to the Lebesgue integral, as W. H. Young has shown. In place of the step-functions the class T_0 could be taken to be the class of continuous functions, but then the proof of the existence of the integral is not immediate as it is for step-functions.

(3) To discuss the generalization of the ordinary integral, let P_0 and T_0 be as before and consider a definition of the general integral in the form

$$S(f) = \Sigma_i c_i[\alpha(x_i) - \alpha(x_{i-1})]$$

where α (x) is some function of x, $0 \leqq x \leqq 1$.

Postulate S is satisfied but not L necessarily. For example let

$$f_n(x) = 1, c_n \leqq x < d,$$
$$= 0 \text{ otherwise}$$

where c_n is made to approach d from below as n increases indefinitely.

The conditions $f_1(x) \geqq f_2(x) \geqq \ldots \geqq f_n(x) \geqq \ldots \geqq 0 = \lim f_n(x)$ for all x, are satisfied. But

$$S(f_n) = \alpha(d) - \alpha(c_n),$$
$$\lim S(f_n) = \alpha(d) - \alpha(d - 0).$$

Unless $\alpha(x)$ is continuous this difference may not be 0 as the postulate L demands. Either we must restrict $\alpha(x)$ to be continuous or reconstruct our definition. Let us study the problem rather from the point of view of weight. If we allow weight (measure) to be discontinuous it becomes necessary to distinguish between closed and open intervals, it is necessary to consider the weight of even a single point. As a matter of fact the problem can be solved satisfactorily if we let T_0 be the class of functions which are constant ($= c_i$) over each of a finite set of intervals Δ_i into which (0, 1) can be divided, if we define

$$S(f) = \Sigma_i c_i m(\Delta_i),$$

where $m(\Delta)$ is the measure or weight of the interval Δ. For this we define

$$m(\Delta) = \alpha(d+0) - \alpha(c-0), \Delta \text{ is } (c \leqq x \leqq d), \text{ closed;}$$
$$= \alpha(d+0) - \alpha(c+0), \Delta \text{ is } (c < x \leqq d)$$
$$= \alpha(d-0) - \alpha(c+0), \Delta \text{ is } (c < x < d), \text{ open;}$$
$$= \alpha(d-0) - \alpha(c-0), \Delta \text{ is } (c \leqq x < d).$$

By $\alpha(c+0)$ we mean the limit (assumed existent) of $\alpha(c+\epsilon)$ as ϵ approaches 0 through positive values. Similarly for $\alpha(c-0)$.

If in particular $\alpha(x)$ is a non-decreasing function of x, denoted by $\beta(x)$, then the above limits exist and all the postulates can be proved. In this case $M(\Delta) \geqq 0$, and the integral is an I-integral

$$I(f) = \int f(x)dm(\Delta)$$
$$= \int f(x)d\,\beta(x).$$

Again if $\alpha(x)$ is the difference of two non-decreasing functions $\beta_1(x)$, $\beta_2(x)$ the integral using α can be defined as the difference of the integrals using β_1, β_2,

$$\int f(x)d\,\alpha(x) = \int f(x)d\,\beta_1(x) - \int f(x)d\,\beta_2(x).$$

It is known that if $\alpha(x)$ is of limited variation, that is, such that $\Sigma_i \mid \alpha(x_i) - \alpha(x_{i-1}) \mid$ is limited by a number K independently of the method of subdivision, then $\alpha(x)$ can be expressed as the difference of two non-decreasing functions. The converse is easily proved since if $\alpha(x) = \beta_1(x) - \beta_2(x)$,

$$\mid \alpha(x_i) - \alpha(x_{i-1}) \mid \leqq \beta_1(x_i) - \beta_1(x_{i-1}) + \beta_2(x_i) - \beta_2(x_{i-1})$$

and the number K may be chosen to be $\beta_1(1) - \beta_1(0) + \beta_2(1) - \beta_2(0)$. The least function $\beta_1(x) + \beta_2(x) - \beta_1(0) - \beta_2(0)$ is called the total variation function $\omega(x)$ corresponding to $\alpha(x)$. This case leads to the Stieltjes integral with respect to a function of limited variation

$$\int f(x)d\,\alpha(x).$$

The first extension of the definition leads to integrals of continuous functions f(x), which are the Stieltjes integrals proper. Further extensions of the class of integrands on the lines of the Lebesgue integral lead to the general Radon-Young integral.[1] A frequently more useful notation is

$$\int f(x)dm(e)$$

where m(e) is an additive (for a denumerable infinity as well as for a finite number) function of sets e of values of the mark p = x. And then the corresponding modular integral can be expressed, after Radon, as

$$\int f(x) \mid dm(e) \mid.$$

1 J. Radon, "Sitzungsberichte der Akademie der Wissenschaften, Wien" (1913), p. 1295.
W. H. Young, "Proceedings of the London Mathematical Society," Vol. XIII (1914), p. 109.

(4) Let P_0 be the set of all real numbers, T_0 the class of functions f(x) equal to a constant other than 0 only over a finite number of bounded (compact) intervals, and 0 elsewhere. For the sake of simplicity choose $\beta(x)$ as a function which is non-decreasing from $-\infty$ to $+\infty$ and which may or may not approach a finite limit as x approaches $+\infty$ or $-\infty$. As in case (3) we may define the weight of a bounded interval Δ as

$$m(\Delta) = \beta(d+0) - \beta(c-0)$$

for a closed interval (c, d) and in a similar manner for other types of interval. Then for a member of T_0 define

$$I(f) = \Sigma_i c_i m(\Delta_i)$$

where c_i is the constant value of f over the subinterval Δ_i. If it happens that $\beta(-\infty)$, $\beta(+\infty)$ both exist as finite numbers, the remaining analysis differs in no essential from that of case (3). If, however, $\beta(-\infty) = -\infty$, $\beta(+\infty) = +\infty$, for example, it appears that only functions f(x) which approach 0 with more than a certain rapidity of convergence at $\pm\infty$ will be summable. It still holds that f is summable when and only when $|f|$ is summable. Cases in which

$$\int_{-\infty}^{\infty} f(x)\,d\beta(x)$$

is conditionally convergent will not appear as direct cases of our general integral, but will require separate handling as limiting cases beyond our immediate aim. For example, it will be possible to consider

$$\int_{-\infty}^{\infty} f(x)\,dx$$

as a case of an absolute integral if the functions satisfy a relation of type

$\lim |f(x) / x^\lambda| = 0$, as $|x|$ increases indefinitely for some $\lambda > 1$.

But an integral like

$$\int_{-\infty}^{\infty} \sin x \, dx / x$$

will need separate treatment.

(5) Let P_0 be the class of real numbers, x, such that $0 \leqq x \leqq 1$, let T_0 be the class of step-functions constant over each of a finite set of subintervals, but restricted to be 0 in a neighborhood of $x = 0$ and in a neighborhood of $x = 1$. Let $\beta(x)$ be a non-decreasing function of x which may be $-\infty$ at $x = 0$ and $+\infty$ at $x = 1$. For any interval Δ not containing nor abutting on 0 or 1, define

$$m(\Delta) = \beta(d+0) - \beta(c-0)$$

for a closed interval $\Delta = (c, d)$ and similarly for the other types of interval. Define

$$I(f) = \Sigma_i c_i m(\Delta_i),$$

where $f(x) = c_i$ on the interval Δ_1. In particular, let $\beta(x) = -\operatorname{ctn} \pi x$. Then, provided f(x) is measurable in the sense of Borel and approaches 0 with sufficient rapidity at $x = 0, x = 1$, it will be possible to define

$$I(f) = \int_0^1 f(x) d\,\beta(x).$$

In the particular case given it is sufficient if

$$| f(x) / \sin^2 x |$$

is uniformly limited in the interval $(0 < x < 1)$ and if $f(x) = 0$ at $x = 0, x = 1$.

(6) Let T_0 be the class of real numbers $-1 \leqq x \leqq 1$, let T_0 be the class of functions constant over each of a finite set of subintervals but restricted by the condition that no end-points of intervals are at $x = 0$. In other words, $x = 0$ is contained strictly within one of the intervals. For a particular example let $\alpha(-1) = \alpha(1) = 0$, and let α increase from either point towards $x = 0$ in such a way that lim $[\alpha(\epsilon) - \alpha(-\epsilon)]$ exists as ϵ approaches 0.

transforming the cases already given. In particular, we can extend (7) to any finite number of dimensions. It is only necessary to have some foundation in an additive function of intervals $\triangle$. Again examples (1) to (6) could be generalized to several dimensions.

(8) This example due to G. C. Evans[1] suggests a different form of application of the general theory of integrals and of additive functions of sets. Let f(e) be an additive function of plane sets, that is, additive for an infinite sum as well as for a finite sum. Then a function of curves F(s) can be defined so that if s_1 is a particular curve,

$$F(s_1) = -\int_{\Sigma} \psi(P)\, df(e)$$

where P is a point of the fundamental set Σ and where

$$\psi(P) = \int_{s_1} \frac{\cos nr}{r}\, ds_1,$$

nr being the angle between the inward drawn normal to s_1 and r, r being the vector P_1P drawn from a point P_1 on s_1. It is then shown that $F(s_1) = f(e)$ where e is the set of points within s_1, if $F(s_1)$ is a "continuous" function of curves. This means that in some cases a function of curves can be used as a weight of the sets of points within the curves.

(9) If $(x_1, x_2, \ldots x_n \ldots.)$ is a point in a space of a denumerable infinity of dimensions such that $0 \leqq x_i \leqq 1$ $(i = 1, 2, \ldots)$, then we can define the weight $m(\triangle)$ of an interval $\triangle$ such as

$$a_i \leqq x_i \leqq 1 - b_i \ (i = 1, 2, \ldots)$$

as equal to

$$m(\triangle) = \operatorname{Prod}_{i=1}^{\infty} (1 - b_i - a_i),$$

an infinite product which may diverge to 0 or converge to a value not greater than 1.

[1] G. C. Evans, "Rendiconti della Reale Accademia dei Lincei," Vol. XXVIII (1919) pp. 262-5.

Then if f(p) is a continuous function of $p = (x_1, x_2, \dots)$ in the sense of Fréchet we can define

$$\int f(p)dm(e)$$

an integral of an infinite number of dimensions.[1] The author has also considered functions of limited variation in an infinite number of dimensions. This type of integral might possibly be useful in connection with probability of sets of functions defined by means of Fourier constants or by the coefficients of a series expression.

(10) Recently N. Wiener[2] has investigated the preliminary problem of weighting in general integrals and in his example (d) defines an integral in a space of continuous functions. Wiener proves that every bounded continuous functional is summable in accordance with his definition of an integral. Further papers on this subject are to be published soon.

This is but a beginning of a new field.

In a further paper[3] by the author it is proved that not only is the general integral S(f) expressible as a difference of two positive integrals, but that a function λ everywhere equal to 1 or -1 exists such that for all summable f

$$S(f) = I(\lambda f)$$

where I is the modular integral associated with S, provided that there exists at least one summable function $h > 0$ except at marks p for which every summable f vanishes. For the simple Stieltjes integral this means that we can find a function λ equal everywhere to ± 1 such that if f is summable with respect to the weighting m(e) then

$$\int f(x)dm(e) = \int f(x)\lambda(x) \mid dm(e) \mid .$$

[1] P. J. Daniell, "Annals of Mathematics," Vol. XX (1919), p. 281; Vol. XXI (1919), p. 30.

[2] N. Wiener, "Annals of Mathematics," Vol. XXI (1920), December.

[3] P. J. Daniell, "Annals of Mathematics," Vol. XXI (1920), p. 203.

The same theorem can be applied to any number of dimensions. The function λ corresponds to a derivative of the function of sets m(e) with respect to its modular function of sets

$$\int_e 1 \mid dm(e) \mid .$$

This suggests the further problem of generalized derivatives. If m(e) is an additive function of sets e and if M(e) is additive and positive, if also m(e) = 0 whenever M(e) is 0, then we may expect to find a function D(p) summable with respect to M(e) such that

$$m(e) = \int_e D(p)dM(e).$$

At the same time it is to be expected that if f(p) is summable with respect to m(e) then

$$\int f(p)dm(e) = \int f(p)D(p)dM(e).$$

All this, however, is without rigorous justification, at present.

P. J. DANIELL.

THE · DE VINNE · PRESS
IMPRIMATUR