On the Stieltjes Integral.

By J. M. WHITTAKER.

(*Received 27th January* 1929. *Read 1st February* 1929.)

1. *Introduction.*

The conception of the integral of one function with respect to another was introduced by Stieltjes in his classical memoir on continued fractions.[1] He defined the integral as

$$\int_a^b f(x)\, dg(x) = \lim_{\lambda \to 0} \sum_{r=1}^{n} f(\xi_r)\{g(x_r) - g(x_{r-1})\}$$

and gave the formula for integration by parts.

$$\int_a^b f(x)\, dg(x) = f(b)\, g(b) - f(a)\, g(a) - \int_a^b g(x)\, df(x).$$

The Stieltjes integral is evidently an extension of the Riemann integral and it might be expected that the scope of the Lebesgue integral could be enlarged without difficulty in the same way. The problem is however by no means easy and has been discussed by many writers.[2] Their conclusions may be indicated roughly by saying that the definition and all the properties of the Lebesgue integral can be extended to the Lebesgue-Stieltjes integral

$$\int_a^b f(x)\, dg(x)$$

provided that $g(x)$ is a function of bounded variation.[3] Viewed as a

---

[1] *Ann. de la Fac. des Sc. de Toulouse*, **8** (1894). A full account of the Riemann-Stieltjes integral has been given by Pollard, *Quarterly Journal*, **49** (1923), 73. In the definition $(x_0 = a \leq \xi_1 \leq x_1 \leq \ldots \leq \xi_n \leq x_n = b)$ is any subdivision of (a, b) and λ is the length of the longest interval (x_{r-1}, x_r).

[2] See especially the account given by Lebesgue, *Leçons sur l'Intégration* (2nd Ed.), Chapter 11 (1928). Also papers by Hildebrandt, *Bull. Amer. Math. Soc.* (2), **24** (1918), 177; and Francis, *Proc. Camb. Phil. Soc.*, **22** (1925), 935.

[3] Another method of defining the integral has been developed by P. J. Daniell, *Annals of Math.*, **19** (1918), 279; **21** (1920), 219; **23** (1923), 169, which can be employed in certain cases when $g(x)$ is not of bounded variation. Starting with $g(x)$, a class of functions $f(x)$ is found for which the integral can be defined.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

development of the Stieltjes integral the new definition is less successful. The Stieltjes integral can exist in cases when (owing to the restriction on $g(x)$) the Lebesgue-Stieltjes integral does not; and as regards integration by parts the former has a distinct advantage. For with the LS-integral formula (1) is only significant when both $f(x)$ and $g(x)$ are of bounded variation.

Another method of defining integration with respect to a function is developed in § 4 of the present memoir. The resulting integral generalises both the Lebesgue integral and the Stieltjes integral. For whenever either of these exists, the "mean-Stieltjes" integral exists and has the same value. Moreover the new definition preserves the most important property of the original Stieltjes integral, that on integration by parts. The definition is suggested by a mode of integration discussed by Denjoy,[1] leading to an integral which he denotes by

$$(B) \int_a^b f(x)\, dx.$$

It is shown in § 4 that Denjoy's condition for the existence of the B-integral can be interpreted as a statement that

$$\lim_{\lambda \to 0} \sum_{r=1}^{n} f(\xi_r)(x_r - x_{r-1}), \qquad \lambda = \max. (x_r - x_{r-1})$$

exists *on the average*. Convergence on the average of general limit operations of this type is defined, and the notion is employed to define the integral of $f(x)$ with respect to $g(x)$ as

$$\lim_{\lambda \to 0} \Sigma\, f(\xi_r)\{g(x_r) - g(x_{r-1})\}$$

whenever the limit exists on the average. It is shown that the integral defined in this way has nearly all the characteristic properties of the Stieltjes integral, including the theorem on change of integrating function.

$$\int_a^b f(x)\, d\psi(x) = \int_a^b f(x)\, g(x)\, d\phi(x), \quad \text{where } \psi(x) = \int_a^x g(t)\, d\phi(t)$$

This theorem was given for the ordinary Stieltjes integral by Hyslop.[2] An analysis of his proof suggests a general theorem on functions of sets, of which use is made later on.

---

[1] *Comptes Rendus*, **169** (1919), 219.

[2] *Proc. Edin. Math. Soc.*, **44** (1926), 79.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

The Stieltjes integral is an example of an incompletely additive function of intervals. It is therefore not possible to give a satisfactory definition of the integral over an arbitrary set of points, as can be done with the Lebesgue-Stieltjes integral owing to the restriction that the function $g(x)$ is of bounded variation. We can however define the integral over some particular sets of points, in particular closed sets. In the special case $f(x) = 1$

$$\int_Q dg(x)$$

so defined can be regarded as the variation of $g(x)$ over the closed set Q. It appears from Theorem 3 that this variation is the same as what Denjoy[1] has called "la variation autour de l'ensemble." The Stieltjes integral over a closed set is discussed in § 3. The most interesting result is the generalised mean value theorem.

$$f'(\xi_1)\, mQ \leqslant sQ \leqslant f'(\xi_2)\, mQ, \qquad \xi_1, \xi_2 \text{ in } Q$$

provided that $f'(x)$ exists at each point of Q and that sQ (the variation of $f(x)$ over Q) exists.

2. *A theorem on functions of sets.*

The principle underlying the theorem is well illustrated in a simple proposition on the Riemann Integral.[2]

If $f(x)$ is Riemann-integrable in (a, b) and if

$$\int_a^x f(t)\, at = 0, \qquad (a \leqslant x \leqslant b)$$

then

$$\int_a^b |f(t)|\, dt = 0.$$

In other words, if for each x in (a, b)

$$\sum_a^x f(\xi_r)(x_r - x_{r-1}) \to 0, \qquad \text{as } \lambda \to 0$$

---

[1] *Ann. de l'École Normale*, **33** (1916), 157. Denjoy does not discuss the conception in detail, another function, "la variation simple," being more convenient in his work. E. C. Francis, *Proc. Camb. Phil. Soc.*, **22** (1925), 924, has given an extension of the mean value theorem in which "la variation simple" is employed, but the theorem given below seems a more direct generalisation.

[2] *Cf.* Hobson, *Functions of a Real Variable* (1921), **1** 451.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

when the points of subdivision divide (a, x) into intervals of greatest length λ, then

$$\sum_a^b |f(\xi_r)|\,(x_r - x_{r-1}) \to 0, \qquad \text{as } \lambda \to 0.$$

This can be generalised as follows.

Let the function $u(E, x)$ be defined for the class K of sets E, and the points x of E. Let E be one of these sets and $E_1, E_2, \ldots E_n$ mutually exclusive sets, all members of K, which together make up E, and let λ be the greatest of $m^* E_1, \ldots m^* E_n$. It is assumed that each E of K contains subsets E_r of arbitrarily small outer measure. Then we are concerned with limit operations of the type

$$(1) \qquad \sum_{r=1}^n u(E_r, \xi_r) \to l, \qquad \text{as } \lambda \to 0$$

familiar in the theory of integration. (1) means that given any positive number ϵ it is possible to find a number η, independent of $\xi_1, \ldots \xi_n$, such that

$$(2) \qquad \left| \sum u(E_r, \xi_r) - l \right| < \epsilon, \qquad \text{provided that } \lambda < \eta$$

$\xi_1, \xi_2, \ldots \xi_n$ being any points of $E_1, E_2, \ldots E_n$ respectively.

The theorem in question is then as follows:—

THEOREM 1. *If for each E in K*

$$S = \sum_{r=1}^n u(E_r, \xi_r) \to 0. \qquad \textit{as } \lambda \to 0$$

then for each E in K

$$Q = \sum_{r=1}^n |u(E_r, \xi_r)| \to 0, \qquad \textit{as } \lambda \to 0.$$

Suppose that this is false. Then there is a set E in K and a positive number, 6ϵ say, and corresponding to each η there is a sum Q^* of the form Q such that

$$(3) \qquad Q^* > 6\epsilon \qquad \text{but } \lambda < \eta$$

Choose η so that (2) is satisfied. Then if P is the sum of those terms of Q^* for which $u(E_r, \xi_r) \geqslant 0$ and N is the sum of the remaining terms,

$$P + N > 6\epsilon$$

so that

$$P > 3\epsilon \qquad \text{or} \qquad N > 3\epsilon.$$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

Take the first alternative so that

$$(4) \qquad P = \sum_r^{(1)} u(E_r, \xi_r) > 3\epsilon$$

$\sum_r^{(1)}, \sum_r^{(2)}$ denoting summations over values of r corresponding to terms in P, N respectively. Now for each E_r we can form a sum S_r like S such that

$$|S_r| < \frac{\epsilon}{s},$$

where s is the number of terms in $\Sigma^{(1)}$. Thus

$$\left|\sum_r^{(1)} S_r\right| < \epsilon$$

and by (4)

$$\sum_r^{(1)} u(E_r, \xi_r) - \sum_r^{(1)} S_r > 2\epsilon.$$

Then by (2)

$$\begin{aligned} \left|\sum_r^{(1)} S_r + \sum_r^{(2)} u(E_r, \xi_r)\right| &= \left|\sum_r^{(1)} u(E_r, \xi_r) - \sum_r^{(1)} S_r - S\right| \\ &\geqslant \left|\sum_r^{(1)} u(E_r, \xi_r) - \sum_r^{(1)} S_r\right| - |S| \\ &> 2\epsilon - \epsilon = \epsilon. \end{aligned}$$

This contradicts (2) since the expression on the left involves a sum over sets of which each has outer measure $< \eta$.

In the same way the hypothesis $N > 3\epsilon$ leads to a contradiction, and the proof of the theorem is complete.

As an example, assume that

$$\int_a^b g(t)\, d\phi(t)$$

exists according to the original definition of Stieltjes. Take K to be the set of intervals $I(x, y)$ in (a, b) and define

$$u(I, \xi) = g(\xi)\{\phi(y) - \phi(x)\} - \int_x^y g(t)\, d\phi(t), \qquad (x \leqslant \xi \leqslant y).$$

Then the theorem shows that

$$(5) \quad \sum_{r=1}^{n} \left| g(\xi_r)\{\phi(x_r) - \phi(x_{r-1})\} - \{\psi(x_r) - \psi(x_{r-1})\}\right| \to 0, \text{ as } \lambda \to 0$$

where $(x_0 = a, x_1, \ldots x_n = b)$ is a subdivision of (a, b) into intervals of greatest length λ and

$$\psi(x) = \int_a^x g(t)\, d\phi(t).$$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

This result leads immediately to the proof of Hyslop's theorem.[1]

If

$$\int_a^b g(t)\, d\phi(t)$$

is well defined and $f(x)$ is bounded in (a, b), then

$$\int_a^b f(x)\, d\psi(x) = \int_a^b f(x)\, g(x)\, d\phi(x)$$

whenever either integral exists.

3. *The Stieltjes integral and the variation of a continuous function over a closed set.*

Let Q be a closed set contained in a finite interval (a, b). Let each point ξ of Q be interior to an interval I_ξ over which

$$\int_{I_\xi} f(x)\, dg(x)$$

exists according to the definition of Stieltjes. By the Heine-Borel theorem,[2] Q can be enclosed in a finite number of the intervals I_ξ.

Inside this finite set of intervals I_ξ, take

$$I = \{(x_1, x'_1), (x_2, x'_2), \ldots (x_n, x'_n)\}$$

a finite set of open intervals enclosing Q narrowly, *i.e.* such that each interval of I contains a point of Q; and let

$$\sigma I = \sum_{r=1}^{n} \int_{xr}^{x'_r} f(x)\, dg(x)$$

Then we define

$$\int_Q f(x)\, dg(x) = \lim_{mI \to mQ} \sigma I$$

whenever the limit exists. A particular case is

$$\int_Q df(x) = \lim_{mI \to mQ} \sum_{r=1}^{n} \{f(x'_r) - f(x_r)\}$$

---

[1] *Proc. Edin. Math. Soc.*, **44** (1926), 79. For the case of $\phi(x)=x$ see J. M. Whittaker, *Proc. Lond. Math. Soc.*, (2) **25** (1926), 213. H. J. Ettlinger, *Journal Lond. Math. Soc.*, **2** (1927), 245, and Miss R. C. Young, *ibid.*, **3** (1928), 117. The latter shows that the theorem is true in space of n dimensions.

[2] Hobson, *op. cit.*, 102.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

and the integral on the left, which will be denoted by

$$s\,(f,\,Q)$$

or sQ if there is no ambiguity, can be interpreted as the variation of $f(x)$ over Q. It will appear later that it is the same as what Denjoy has called "la variation autour de Q."

The integration by parts theorem takes the form

$$\int_Q f\,dg = s\,(fg,\,Q) - \int_Q g\,df$$

and other simple properties of the integral are

$$\int_{Q_1} f\,dg + \int_{Q_2} f\,dg = \int_{Q_1+Q_2} f\,dg, \qquad Q_1 Q_2 = 0,$$

whenever either side exists; and

$$\int_Q f_1\,dg + \int_Q f_2\,dg = \int_Q (f_1+f_2)\,dg$$

$$\int_Q f\,dg_1 + \int_Q f\,dg_2 = \int_Q f\,d\,(g_1+g_2)$$

whenever the integrals on the left exist.

The most interesting results relate to the case when the integral is

$$s\,(f,\,Q) = \int_Q df\,(x)$$

and $f(x)$ is continuous at the points of Q. The latter assumption will be made in the theorems which follow.

THEOREM 2. *Let $(a_1, b_1), (a_2, b_2), \ldots$ be the complementary intervals of Q. Then the necessary and sufficient condition that sQ exists is that there is an I-set I_0 for which*

$$\sum_{n=1}^{\infty} \omega_n$$

converges, where

ω_n = fluctuation of $f(x)$ in common part of (a_n, b_n), I_0.

Moreover

$$sQ = f(b) - f(a) - \sum_{n=1}^{\infty} \{f(b_n) - f(a_n)\}.$$

THEOREM 3. *Let sQ exist. Then if Q_1 is any portion of Q, sQ_1 exists uniformly with respect to Q_1.*

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

A *portion* is the part of Q contained in a closed sub-interval of (a, b). Theorem 3 means that there is a number sQ_1 such that

$$|\sigma I_1 - sQ_1| < \epsilon, \text{ when } mI_1 < mQ_1 + \zeta$$

where ζ depends on ϵ, but not on Q_1. I_1 is any I-set for Q_1, *i.e.* any finite set of open intervals enclosing Q_1 narrowly.

Theorem 2 is of the same type as E. H. Moore's necessary and sufficient condition for the existence of the Harnack-Lebesgue integral,[1] and is proved in much the same way. The same argument establishes theorem 3. The details of the proof may be omitted. It should be added that the existence of sQ does not necessarily imply that sQ_1 exists for every closed subset Q_1 of Q.

Thus, let P_n be a non-dense perfect set in $\left(\frac{1}{n+1}, \frac{1}{n}\right)$ and let $f(x)$ be continuous and constant in the complementary intervals of the sets P_n; and let $f(x)$ increase steadily from the value 0 at $x = \frac{1}{n+1}$ to the value $\frac{1}{n}$ at $x = \frac{1}{2}\left(\frac{1}{n+1} + \frac{1}{n}\right)$ and then decrease steadily to 0 at $x = \frac{1}{n}$. Let

$$Q = Q_0 + \sum_{n=1}^{\infty} P_n, \quad Q_0 = (0, 1, \tfrac{1}{2}, \tfrac{1}{3}, \ldots,).$$

Then $f(x)$ is constant in the complementary intervals of Q so that $sQ = 0$, but sQ_0 does not exist.

THEOREM 4. *Let Q_γ be the perfect component of Q (null if Q is denumerable). Then, if sQ exists,*

$$sQ = sQ_\gamma$$

Find ζ so that

$$(1) \quad |\sigma I - sQ| \leqslant \epsilon, \quad mI < mQ + \zeta$$

and let I_1 be an I-set for Q', the derived set of Q. Then

$$(2) \quad |\sigma I_1 - sQ| \leqslant \epsilon, \; mI_1 < mQ' + \zeta = mQ + \zeta.$$

If not, there is a set I_1 such that

$$|\sigma I_1 - sQ| > \epsilon, \text{ but } mI_1 < mQ + \zeta.$$

[1] Hobson, *op. cit.*, 621. Theorem 2 has been emended in accordance with a suggestion of Dr Hyslop.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

At most a finite number of points of Q are outside I_1. Since $f(x)$ is continuous at each of these points, they can be enclosed in a finite set of intervals I_2 such that

$$|\sigma I_2| < |\sigma I_1 - sQ| - \epsilon$$

and
$$mI_2 < mQ + \zeta - mI_1$$

$I = I_1 + I_2$ is an I-set for Q and

$$|\sigma I - sQ| \geqslant |\sigma I_1 - sQ| - |\sigma I_2| > \epsilon$$
$$mI = mI_1 + mI_2 < mQ + \zeta.$$

These inequalities contradict (1) so that (2) must be true. By induction, if I_p is an I-set for $Q^{(p)}$

(3) $$|\sigma I_p - sQ| \leqslant \epsilon, \qquad mI_p < mQ^{(p)} + \zeta$$

i.e. $$sQ^{(p)} = sQ \qquad \text{for all integers } p.$$

Again, let I_ω be an I-set for $Q^{(\omega)}$. Then

$$\{Q^{(p)} - I_\omega\} \qquad p = 1, 2, 3, \ldots .$$

is a decreasing sequence of closed sets. If each of these sets contains at least one point Cantor's theorem of deduction shows that there is a point ξ common to all the sets. ξ must therefore be a member of $Q^{(\omega)} - I_\omega$. This is false. Hence there must be an integer n such that

$$Q^{(n)} < I_\omega .$$

Thus I_ω is an I-set for $Q^{(n)}$ and by (3)

$$|\sigma I_\omega - sQ| \leqslant \epsilon, \; mI_\omega < mQ^{(\omega)} + \zeta$$

i.e. $$sQ^{(\omega)} = sQ.$$

In this way the result stated is arrived at by transfinite induction.

THEOREM 5. *Let*

(i) *$f'(x)$ exist at every point of a closed set Q in (a, b).*

(ii) *sQ exist.*

Then there are points ξ_1, ξ_2 of Q such that

$$f'(\xi_1)\, mQ \leqslant sQ \leqslant f'(\xi_2)\, mQ.$$

LEMMA. (i) *implies that, given ϵ, it is possible to divide (a, b) into a finite number of sub-intervals such that each sub-interval I_r (supposed*

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

closed at both ends) containing a point of Q contains a point ξ_r of Q with the property

$$\left|\frac{f(x)-f(\xi_r)}{x-\xi_r}-f'(\xi_r)\right|<\epsilon, \qquad x \text{ in } I_r.$$

This is Goursat's lemma with the whole interval (a, b) replaced by the closed set Q. The proof, by repeated bisection, need not be set down in detail.

It will now be shown that (i), (ii) and

(iii) $$f'(x) \geqslant 0 \qquad \text{in } Q$$

imply

$$sQ \geqslant 0.$$

Let ζ be assigned arbitrarily, and find I_0, an I-set for Q, such that

$$mI_0 < mQ + \zeta.$$

Divide each interval of I_0 into sub-intervals with the property of the lemma. Then if such a sub-interval is (x_r, x'_r) and the characteristic point is ξ_r,

$$\begin{aligned} f(\xi_r)-f(x_r) &\geqslant f'(\xi_r)(\xi_r-x_r)-\epsilon(\xi_r-x_r) \\ &\geqslant -\epsilon(\xi_r-x_r). \end{aligned}$$

Similarly

$$f(x'_r)-f(\xi_r) \geqslant -\epsilon(x'_r-\xi_r)$$

so that

$$f(x'_r)-f(x_r) \geqslant -\epsilon(x'_r-x_r)$$

$$\sum_{r=1}^{n}\{f(x'_r)-f(x_r)\} \geqslant -\epsilon\sum_{r=1}^{n}(x'_r-x_r) \geqslant -\epsilon(b-a)$$

$(x_1, x'_1), \ldots . (x_n, x'_n)$ being in order all the sub-intervals, formed by dividing intervals of I_0, which contain points of Q.

On amalgamating any of these which abut we obtain an I-set for Q, contained in I_0 and so such that

$$mI < mQ + \zeta$$

and it has been shown that for this I-set

$$\sigma I \geqslant -\epsilon(b-a).$$

Since ζ is arbitrary

$$sQ \geqslant -\epsilon(b-a)$$

and since this is true for every positive ϵ,

$$sQ \geqslant 0.$$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

The general case (in which (iii) is not assumed) can be reduced to this by considering the function

$$g(x) = f(x) - hx$$

where h is the lower bound[1] of $f'(x)$. The conclusion is

$$sQ \geqslant h \, . \, mQ. \tag{1}$$

Similarly if H is the upper bound of $f'(x)$ in Q

$$sQ \leqslant H \, . \, mQ.$$

To complete the proof it is necessary to shew that if

$$sQ = h \, . \, mQ, \qquad mQ > 0$$

then there is a point ξ_1 such that

$$f'(\xi_1) = h.$$

There is no loss of generality in taking $h = 0$. Thus, given

$$f'(x) \geqslant 0 \text{ in } Q, \quad sQ = 0, \quad mQ > 0$$

we have to show that there is a point ξ_1 of Q such that

$$f'(\xi_1) = 0.$$

If not, let E_n be the subset of Q for which

$$f'(x) \geqslant \frac{1}{n}.$$

Then $E_n \to Q$, and $\therefore$ $mE_n \to mQ > 0$, and $mE_n > 0$ for some n. This set E_n must contain a closed subset Q_1 of positive measure. Thus

$$f'(x) \geqslant \frac{1}{n} \qquad \text{in } Q_1, \qquad mQ_1 > 0.$$

As before, enclose Q in I_0 and divide each interval of I_0 into further sub-intervals with the Goursat property for ϵ, Q_1. This gives an I-set for Q_1. Subtract this set (regarded as an open set) from I_0 and divide up the intervals of the finite set which remains into further sub-intervals with the Goursat property for ϵ, Q. Reject all sub-intervals containing no points of Q and amalgamate any abutting sub-intervals which remain. This gives an I-set for Q and

$$\sigma I = \Sigma^{(1)} \{ f(x'_r) - f(x_r) \} + \Sigma^{(2)} \{ f(x'_r) - f(x_r) \}$$

where $\Sigma^{(1)}$ is taken over the sub-intervals containing points of Q_1 and $\Sigma^{(2)}$ over the remaining sub-intervals.

---

[1] If $f'(x)$ is not bounded below, (1) is certainly true if $mQ > 0$. If $mQ = 0$ and $sQ > 0$ (< 0) it can be inferred that $f'(x)$ is not bounded above (below).

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

For an interval of the first class

$$f(\xi_r) - f(x_r) \geqslant f'(\xi_r)(\xi_r - x_r) - \epsilon(\xi_r - x_r) \geqslant \frac{1}{n}(\xi_r - x_r) - \epsilon(\xi_r - x_r)$$

so that

$$f(x'_r) - f(x_r) \geqslant \frac{1}{n}(x'_r - x_r) - \epsilon(x'_r - x_r)$$

while for an interval of the second class

$$f(x'_r) - f(x_r) \geqslant -\epsilon(x'_r - x_r)$$

so that

$$\begin{aligned} \sigma I &\geqslant \frac{1}{n}\Sigma^{(1)}(x'_r - x_r) - \epsilon\Sigma(x'_r - x_r) \\ &\geqslant \frac{1}{n} mQ_1 - \epsilon(b - a) \\ &> \frac{1}{2n} . mQ_1, \qquad \text{by choice of } \epsilon. \end{aligned}$$

It follows that

$$sQ \geqslant \frac{1}{2n} . mQ_1 > 0.$$

The contradiction implies that $f'(x) = 0$ almost everywhere in Q. This completes the proof of Theorem 5. It is easy to see that the conclusion cannot be replaced by

$$sQ = f'(\xi) . mQ, \qquad \xi \text{ in } Q.$$

For example, let $f(x) = x^2$ and let Q consist of the intervals $(-1, -\frac{1}{2})$ and $(\frac{1}{2}, 1)$.

It is easy to see that if $f(x)$ is continuous at the points of Q the definition of sQ is equivalent to the following.

THEOREM 6. *Let $\Delta(x_0 = a, x_1, .. x_n = b)$ be a subdivision of (a, b) into intervals of greatest length λ and let*

$$S_\Delta = \sum_Q \{f(x_r) - f(x_{r-1})\}$$

the summation being over those intervals which contain points of Q. Then

$$sQ = \lim_{\lambda \to 0} S_\Delta.$$

Again let Q_r be the portion of Q in (x_{r-1}, x_r). Then the *total variation* of $f(x)$ over a closed set Q for which sQ exists can be defined to be the upper bound, for all subdivisions Δ, of

$$\sum_{r=1}^{n} |sQ_r|.$$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

It follows from Theorem 1 that the total variation is the same as

$$\overline{\lim_{\lambda \to 0}} \sum_{Q} |f(x_r) - f(x_{r-1})|.$$

In this case the sets E are the sub-intervals $I(x, y)$ of (a, b) and the simple form of Theorem 1 quoted in § 1 is adequate. Define

$$\left.\begin{aligned} uI &= f(y) - f(x) - sQ_I, && \text{if } (x, y) \text{ contains a point of } Q. \\ &= 0 && , \text{ otherwise.} \end{aligned}\right\}$$

Q_I being the portion of Q in (x, y). If $(\xi_0 = x, \xi_1, \ldots \xi_p = y)$ is a subdivision of (x, y) and $I_s = (\xi_{s-1}, \xi_s)$

$$\sum_{s=1}^{p} uI_s = \sum_{Q_I} \{f(\xi_s) - f(\xi_{s-1}) - sQ_{Is}\} = \sum_{Q_I} \{f(\xi_s) - f(\xi_{s-1})\} - sQ_I$$

$$\to 0, \text{ as } \max.(\xi_s - \xi_{s-1}) \to 0$$

by what has just been said. Thus by Theorem 1

$$\sum_{Q} |f(x_r) - f(x_{r-1}) - sQ_r| \to 0, \text{ as } \lambda \to 0$$

whence

$$\overline{\lim_{\lambda \to 0}} \sum_{Q} |f(x_r) - f(x_{r-1})| = \overline{\lim_{\lambda \to 0}} \sum_{r=1}^{n} |sQ_r|$$

$$= \text{upper bound of } \sum_{r=1}^{n} |sQ_r|$$

4. *Convergence on the average and the Stieltjes integral.*

The limit operations which have occurred in the preceding work have been of the type

$$\sum_{r=1}^{n} u(E_r, \xi_r) \to l, \qquad \text{as } \lambda \to 0, \qquad \lambda = \max.[m^* E_r].$$

The present chapter deals with the existence of these limits *on the average*. The functions $u(E_r, \xi_r, t)$ now involve an additional variable t, being defined for a measurable set T of t, and convergence on the average is defined as follows.

Let k be any positive number and let R_k be the set of values of t, for which

$$\left| \sum_{r=1}^{n} u(E_r, \xi_r, t) - l \right| > k$$

for a fixed subdivision (E_r, ξ_r); then if for each k

$$m^* R_k \to 0, \qquad \textit{as } \lambda \to 0$$

we say that

$$S(t) \equiv \sum_{r=1}^{n} u(E_r, \xi_r, t) \to l \qquad \textit{on the average, as } \lambda \to 0.$$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

THEOREM 7. *If*

$$\sum_{r=1}^{n} u(E_r, \xi_r, t) \to l \qquad \textit{on the average, as } \lambda \to 0$$

then given δ, *a set* T_0 *and a sequence* $\{\Delta_n\}$ *of subdivisions* (E_r, ξ_r) *can be found such that*

$$S_n(t) \to l, \qquad \textit{as } n \to \infty, \qquad t \text{ in } T - T_0$$

and

$$m^* T_0 < \delta$$

$S_n(t)$ *being the* $S(t)$ *corresponding to* Δ_n.

If k is any positive number, a subdivision Δ_1 can be found so that for this Δ_1

$$m^* R_{\frac{k}{2}} < \tfrac{1}{2}\delta.$$

For example, Δ_1 may be found by bisecting (a, b) repeatedly, always taking the point ξ to be the mid point of the corresponding sub-interval. Again, a subdivision Δ_2 can be found so that for it

$$m^* R_{\frac{k}{4}} < \tfrac{1}{4}\delta$$

and so on. Let

$$T_0 = R_{\frac{k}{2}} + R_{\frac{k}{4}} + \ldots.$$

Then

$$m^* T_0 \leqslant m^* R_{\frac{k}{2}} + m^* R_{\frac{k}{4}} + \ldots.$$

$$< \tfrac{1}{2}\delta + \tfrac{1}{4}\delta + \ldots = \delta.$$

Also if t is a point of $T - T_0$

$$S_n(t) \to l, \text{ as } n \to \infty.$$

For, given ϵ, p can be found so that $\frac{k}{2^p} < \epsilon$ and

$$|S_n(t) - l| < \frac{k}{2^p} < \epsilon, \qquad (n \geqslant p).$$

This theorem shows that the (S_m)-integral, defined below, can be represented approximately as a Stieltjes sum; in particular that the Lebesgue integral can be represented approximately as a Riemann sum.[1]

Theorem 1 remains true when the limits are taken on the average. Thus

---

[1] *Cf.* Hobson, *op. cit.*, 585.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

THEOREM 8. *If for each E in K*

$$\sum_{r=1}^{n} u(E_r, \xi_r, t) \to 0 \qquad \textit{on the average as } \lambda \to 0$$

then for each E in K

$$\sum_{r=1}^{n} |u(E_r, \xi_r, t)| \to 0 \qquad \textit{on the average, as } \lambda \to 0.$$

In other words, let R_k be the set of t for which

$$|S(t)| = |\sum_{r=1}^{n} u(E_r, \xi_r, t)| \geqslant k > 0$$

and for each E in K and for each positive k, let

$$m^* R_k \to 0 \qquad \text{as } \lambda \to 0.$$

Let T_k be the set of t for which

$$Q(t) = \sum_{r=1}^{n} |u(E_r, \xi_r, t)| \geqslant k.$$

Then the theorem asserts that for each E, k

$$m^* T_k \to 0 \qquad \text{as } \lambda \to 0.$$

If the assertion is false there is a set E and positive numbers k, δ such that given any positive number η there is a dissection of E for which

(4) $$m^* T_k = m^*[Q \geqslant k] > \delta \qquad \text{but } \lambda < \eta.$$

Let η be such that

(5) $$m^* R_{\frac{k}{6}} = m^*\left[|S| \geqslant \frac{k}{6}\right] < \frac{\delta}{3}, \qquad \lambda < \eta$$

and let Q^* be a Q for which (5) is true with this η. Let Q^* have s terms. For each E_r find S_r like S so that

$$|S_r| < \frac{k}{6s} \qquad \text{except in } R_r,\ m^* R_r < \frac{\delta}{3s}.$$

Then

(6) $$\sum_{r=1}^{s} |S_r| < \frac{k}{6} \qquad \text{except in } R = \Sigma R_r,\ m^* R < \frac{\delta}{3}.$$

Suppose now that t is a point of $T_k{}^* - R - R_{\frac{k}{6}}$ ($T_k{}^*$ is the T_k associated with Q^*). Then

$$Q^*(t) \geqslant k.$$

Either the sum P of the terms in Q^* for which $u(E_r, \xi_r, t)$ is positive, or else the sum N of the terms in Q^* for which $u(E_r, \xi_r, t)$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

is negative must be $\geqslant \frac{1}{2}k$. Take the first alternative; then

$$P = \Sigma^{(1)} u(E_r, \xi_r, t) \geqslant \frac{k}{2}$$

$$\therefore \quad \Sigma^{(1)} u(E_r, \xi_r, t) - \Sigma^{(1)} S_r \geqslant \frac{k}{2} - \frac{k}{6} = \frac{k}{3}, \text{ by (6)}$$

$$\begin{aligned} \therefore \quad |\Sigma^{(1)} S_r + \Sigma^{(2)} u(E_r, \xi_r, t)| &= |S(t) - \{\Sigma^{(1)} u(E_r, \xi_r, t) - \Sigma^{(1)} S_r\}| \\ &\geqslant |\Sigma^{(1)} u(E_r, \xi_r, t) - \Sigma^{(1)} S_r| - |S(t)| \\ &\geqslant \frac{k}{3} - \frac{k}{6} = \frac{k}{6}. \end{aligned}$$

This holds for all t in $T^* - R - R_{\frac{k}{6}}$ and

$$\begin{aligned} m^*(T^* - R - R_{\frac{k}{6}}) &\geqslant m^* T^* - m^* R - m^* R_{\frac{k}{6}} \\ &> \delta - \frac{\delta}{3} - \frac{\delta}{3} \\ &= \frac{\delta}{3}. \end{aligned}$$

Thus

$$m^*\left[\,|\Sigma^{(1)} S_r + \Sigma^{(2)} u(E_r, \xi_r, t)| \geqslant \frac{k}{6}\right] > \frac{\delta}{3}$$

which contradicts (5), since the "λ" of the sum on the left is less than η.

We now define the *mean Stieltjes* integral as the limit on the average, as $\lambda \to 0$, of

$$\sum_{r=1}^{n} f(\xi_r)\{g(x_r) - g(x_{r-1})\}.$$

The precise definition is as follows. Let

$$2a - b = y_0 \leqslant \eta_1 \leqslant y_1 \leqslant \eta_2 \leqslant \cdot \qquad \leqslant y_n = b$$

be a net Δ_1 filling $(2a - b, b)$. Displace Δ_1 through a distance t to the right so that it consists of the points

$$y_0 + t \leqslant \eta_1 + t \leqslant \qquad \leqslant y_n + t.$$

Let ζ_1 be the first η which now lies to the right of a and let ζ_m be the last η which lies to the left of b, and let

$$\zeta_1 \leqslant z_1 \leqslant \zeta_2 \leqslant \cdot\cdot \qquad \leqslant z_{m-1} \leqslant \zeta_m$$

be the intervening part of the net.

$$\Delta(t) = (a = z_0 \leqslant \zeta_1 \leqslant z_1 \leqslant \ldots\ldots \leqslant \zeta_m \leqslant z_m = b), \quad (0 < t < b - a)$$

is then a net filling (a, b). Form the sum

$$S_{\Delta(t)} = \sum_{r=1}^{m} f(\zeta_r)\{g(z_r) - g(z_{r-1})\}.$$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

Then

$$(S)_m \int_a f(x)\, dg(x)$$

will be said to exist and to have the value I if

$$m^* R_k \to 0, \qquad \text{as } \lambda \to 0$$

for each positive number k; R_k being the set of t for which

$$|S_{\Delta(t)} - I| \geqslant k > 0$$

and λ (independent of k) the length of the longest interval (x_{r-1}, x_r) of Δ.

In case $g(x) = x$ the sum $S_{\Delta(t)}$ differs from the sum

$$\sum_{r=1}^{n} f(\xi_r + t)(x_r - x_{r-1})$$

considered by Denjoy only in the terms corresponding to intervals adjacent to a, b and the definitions are substantially equivalent.

It is evident that the (S_m)-integral exists and has the same value as the integral defined by Stieltjes, whenever the latter exists. The (S_m)-integral has moreover nearly all the properties of the original Stieltjes integral.[1] The more simple of these are

(I)
$$\int_a^b 1 \,.\, dg(x) = g(b) - g(a)$$

(II) If $a < c < b$ and either $f(x)$ or $g(x)$ is continuous at c,

$$\int_a^b f\, dg = \int_a^c f\, dg + \int_c^b f\, dg$$

whenever the integrals on the right exist.

(III) If $\int_a^b f\, dg$ exists, so does $\int_a^b cf\, dg$ where c is a constant and

$$\int_a^b cf\, dg = c\int_a^b f\, dg.$$

(IV) If $\int_a^b f_1\, dg$, $\int_a^b f_2\, dg$ exist, so does

$$\int_a^b (f_1 + f_2)\, dg$$

and it is equal to the sum of the given integrals.

[1] *Cf.* Pollard, *loc. cit.*

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

(V) If $\int_a^b f\,dg_1$, $\int_a^b f\,dg_2$ exist, so does

$$\int_a^b f\,d(g_1+g_2)$$

and it is equal to the sum of the given integrals.

(VI) $$\int_a^b g(x)\,df(x) = g(b)f(b) - g(a)f(a) - \int_a^b f(x)\,dg(x)$$

provided that the integral on the right exists and that either $f(x)$ or $g(x)$ is continuous at a, b.

(VII) If $|f(x)| \leqslant K$, $g(x)$ is of bounded variation, and the (S_m)-integral exists, then

$$\left|\int_a^b f(x)\,dg(x)\right| \leqslant K V_a^b g.$$

The proofs of these propositions need not detain us, as they present no points of particular interest or difficulty. The analogue of Hyslop's theorem is also true.

THEOREM 9. *If*

$$\psi(x) = (S_m)\int_a^x g(t)\,d\phi(t)$$

is well defined in (a, b) *and* $f(x)$ *is bounded, then*

$$(S_m)\int_a^b f(x)\,d\psi(x) = (S_m)\int_a^b f(x)\,g(x)\,d\phi(x)$$

whenever either integral exists.

This is deduced from Theorem 8 in the same way that Hyslop's theorem was deduced from Theorem 1.

THEOREM 10.

$$(S_m)\int_a^b f(x)\,dx = (L)\int_a^b f(x)\,dx$$

whenever the integral on the right exists.

This is practically equivalent to the theorem, stated without proof by Denjoy, that integration in accordance with his definition (B) includes Lebesgue integration.

LEMMA. *Let* $\phi(x, h)$ *be measurable in* E *for each* h *and let* $\phi(x, h) \to f(x)$, *as* $h \to +0$, *p.p. in* E. *Then, given* η, *there is a subset* H *of* E *such that*

$$mH < \eta$$

and

$$\phi(x, h) \to f(x), \text{ as } h \to +0, \text{ uniformly in } E - H.$$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

This is a slight extension of Egoroff's theorem[1] and is proved in the same way. "p. p." (presque partout) means "almost everywhere" *i.e.* except in a set of measure zero.

To prove Theorem 10, take first the case

$$|f(x)| \leqslant K.$$

Let

$$F(x) = (L)\int_a^x f(t)\,dt$$

$$\phi(x, h) = \frac{F(x+h) - F(x)}{h}.$$

Then[2]

$$\phi(x, h) \to f(x), \qquad \text{as } h \to +0, \text{ p. p. in } (a, b).$$

Thus, by the lemma, λ_1 can be found so that

$$|\phi(x, h) - f(x)| < \eta, \ (0 < h \leqslant \lambda_1), \qquad x \text{ in } E - H.$$

Again

$$|\phi(x, h) - f(x)| \leqslant 2K, \qquad \text{all } x, h.$$

Thus, if $f(x)$ is defined in the range $b < x < 2b - a$ by

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

then if x is any point of (a, b)

$$\int_a^b |\phi(x+t, h) - f(x+t)|\,dt = \int_a^b |\phi(t, h) - f(t)|\,dt$$

$$= \int_{E-H} |\phi(t, h) - f(t)|\,dt + \int_H |\phi(t, h) - f(t)|\,dt$$

$$< \eta(b - a) + 2K\eta = \delta, \ \text{(say)}$$

hence $\int_a^b |F(x+h+t) - F(x+t) - hf(x+t)|\,dt < \delta h, \qquad (0 < h \leqslant \lambda_1)$

Similarly

$$\int_a^b |F(x+t) - F(x-h'+t) - h'f(x+t)|\,dt < \delta h', \qquad (0 < h' \leqslant \lambda_2)$$

for some λ_2. Thus if λ_0 is the smaller of λ_1, λ_2,

$$\int_a^b |F(x_r + t) - F(x_{r-1} + t) - f(\xi_r + t)(x_r - x_{r-1})|\,dt < \delta(x_r - x_{r-1}),$$
$$x_r - x_{r-1} < \lambda_0$$

where we have written x_r, x_{r-1}, ξ_r in place of $x + h$, $x - h'$, x.

---

[1] Hobson, *op. cit.*, **2**, 144.

[2] Hobson, *op. cit.*, 541.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

Thus if $a \leqslant \xi_1 \leqslant x_1 \leqslant . \leqslant \xi_n \leqslant b$ is any subdivision of (a, b), the longest interval (x_{r-1}, x_r) being of length λ

$$\sum_{r=1}^{n} \int_a^b | F(x_r + t) - F(x_{r-1} + t) - f(\xi_r + t)(x_r - x_{r-1}) | dt < \delta(b-a), \ (\lambda < \lambda_0)$$

hence $\int_a^b \left| \sum_{r=1}^{n} \{F(x_r + t) - F(x_{r-1} + t) - f(\xi_r + t)(x_r - x_{r-1})\} \right| dt < \delta(b-a), \ (\lambda < \lambda_0)$

or $\int_a^b \left| (L)\int_a^b f(x)\,dx - \sum_{r=1}^{n} f(\xi_r + t)(x_r - x_{r-1}) \right| dt \quad < \delta(b-a), \ (\lambda < \lambda_0).$

Thus if R_k is the set of t for which

$$\left| (L)\int_a^b f(x)\,dx - \sum_{r=1}^{n} f(\xi_r + t)(x_r - x_{r-1}) \right| \geqslant k$$

then

$$mR_k \leqslant \frac{\delta(b-a)}{k}, \qquad (\lambda < \lambda_0)$$

and since δ is arbitrary, for fixed $k > 0$

$$mR_k \to 0, \text{ as } \lambda \to 0.$$

It follows that $f(x)$ is integrable in accordance with Denjoy's definition (B).

Suppose now that $f(x)$ is any function which is integrable (L) *i.e.* not necessarily bounded.

Given $\epsilon, k, f(x)$ can be expressed as the sum of a bounded function $f_1(x)$ and a function $f_2(x)$ such that

$$(1) \quad (L)\int_a^b |f_2(x)|\,dx < \frac{\epsilon k}{8(b-a)}.$$

Now if $h(x)$ is any positive integrable-L function

$$m[h(x) \geqslant k] \leqslant \frac{1}{k}(L)\int_a^b h(x)\,dx.$$

Thus, by (1)

$$m\left[\left|\sum_{r=1}^{m} f_2(\zeta_r)(z_r - z_{r-1}) - \int_a^b f_2(x)\,dx\right| \geqslant \frac{k}{2}\right]$$

$$\leqslant \frac{2}{k}\int_a^b \left| \sum_{r=1}^{n} f_2(\zeta_r)(z_r - z_{r-1}) - \int_a^b f_2(x)\,dx \right| dt$$

$$\leqslant \frac{2}{k}\sum_{r=1}^{n}(x_r - x_{r-1})\int_a^b |f_2(t)|\,dt + \frac{2}{k}\int_a^b \frac{\epsilon k}{8(b-a)}\,dt$$

$$(2) \quad < \tfrac{1}{4}\epsilon + \tfrac{1}{4}\epsilon = \tfrac{1}{2}\epsilon.$$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

Again, since $\Sigma f_1(\zeta_r)(z_r - z_{r-1})$ differs from $\Sigma f_1(\xi_r + t)(x_r - x_{r-1})$ only in a finite number of terms,

$$\left| \sum_{r=1}^{m} f_1(\zeta_r)(z_r - z_{r-1}) - \sum_{r=1}^{m} f_1(\xi_r + t)(x_r - x_{r-1}) \right|$$

$=$ sum of a finite number of terms of the form $f(a)(x_r - x_{r-1})$ $= O(\lambda)$

and by what has been proved above

$$m\left[\left|\sum_{r=1}^{n} f_1(\xi_r + t)(x_r - x_{r-1}) - \int_a^b f_1(x)\,dx\right| \geqslant \tfrac{1}{2}k\right] \rightarrow 0, \text{ as } \lambda \rightarrow 0.$$

Thus there exists a positive number η such that

$$(3) \quad m\left[\left|\sum_{r=1}^{m} f_1(\zeta_r)(z_r - z_{r-1}) - \int_a^b f_1(x)\,dx\right| \geqslant \tfrac{1}{2}k\right] < \tfrac{1}{2}\epsilon, \qquad (\lambda < \eta).$$

Thus by (2), (3)

$$m\left[\left|\sum_{r=1}^{m} f(\zeta_r)(z_r - z_{r-1}) - \int_a^b f(x)\,dx\right| \geqslant k\right] < \epsilon, \qquad (\lambda < \eta)$$

and this is the result stated.

THEOREM 11. *If $f(x)$, of period 2π, is integrable (L) and*

$$g(x) = \frac{1}{2\pi}\int_0^{2\pi} f(x+t)\cot\frac{t}{2}\,dt$$

is the function conjugate to $f(x)$, then $g(x)$ is integrable (S_m) with respect to x.

The corresponding theorem for the B-integral has been given by Kolmogoroff,[1] and his proof holds, with slight modifications, in the present case. The next theorem deals with the integration of sequences.[2]

THEOREM 12. *If*

(i) $V_a^b f_n(x) \leqslant K$, *all n*

(ii) $V_a^b\{f_m(x) - f_n(x)\} + |f_m(b) - f_n(b)| \rightarrow 0$, as $m, n \rightarrow \infty$

then there is a function $f(x)$ of bounded variation such that

$$f_n(x) \rightarrow f(x), \qquad \textit{as } n \rightarrow \infty, \qquad (a \leqslant x \leqslant b)$$

and

$$V_a^b\{f_n(x) - f(x)\} \rightarrow 0, \qquad \textit{as } n \rightarrow \infty.$$

---

[1] *Fundamenta Math.*, **11** (1928), 27.

[2] Theorems of the same type involving the Stieltjes integral proper have been given by Hahn, *Monatshefte für Math. u. Physik*, **32** (1922), 84.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

Moreover (i), (ii) *follow from these conclusions and*

$$\int_a^b f_n(x)\,dg(x) \to \int_a^b f(x)\,dg(x), \qquad \text{as } n \to \infty$$

provided that $f_n(x)$ *is continuous at* a, b, *that* $g(x)$ *is bounded in* (a, b) *and that the integral on the left exists* (S_m) *for each* n.

Since

$$|f_m(b) - f_n(b)| \to 0, \qquad \text{as } m, n \to \infty$$

$$f_n(b) \to \text{a limit}, f(b) \text{ (say)}, \qquad \text{as } n \to \infty.$$

Also if x is any point of (a, b).

$$\begin{aligned} |f_m(x) - f_n(x)| &\leqslant |f_m(x) - f_m(b) - f_n(x) + f_n(b)| + |f_m(b) - f_n(b)| \\ &\leqslant V_a^b\{f_m(x) - f_n(x)\} + |f_m(b) - f_n(b)| \\ &\to 0 \qquad \text{as } m, n \to \infty. \end{aligned}$$

$$\therefore \quad f_n(x) \to \text{a limit}, f(x) \text{ (say)} \qquad \text{as } n \to \infty$$

and for any subdivision of (a, b)

$$\begin{aligned} \sum_{r=1}^{p} |f(x_r) - f(x_{r-1})| &= \lim_{n \to \infty} \sum_{r=1}^{p} |f_n(x_r) - f_n(x_{r-1})| \\ &\leqslant \overline{\lim_{n \to \infty}}\, V_a^b f_n(x) \\ &\leqslant K \end{aligned}$$

so that

$$V_a^b f(x) \leqslant K.$$

Again, find N so that

$$V_a^b\{f_m(x) - f_n(x)\} < \epsilon, \qquad (m, n \geqslant N).$$

Then for any fixed subdivision of (a, b)

$$\sum_{r=1}^{p} |f_m(x_r) - f_n(x_r) - f_m(x_{r-1}) + f_n(x_{r-1})| < \epsilon, \quad (m, n \geqslant N).$$

Keep n fixed and let $m \to \infty$

$$\sum_{r=1}^{p} |f(x_r) - f_n(x_r) - f(x_{r-1}) + f_n(x_{r-1})| \leqslant \epsilon, \qquad (n \geqslant N).$$

This is true for all subdivisions of (a, b).

Therefore $V_a^b\{f(x) - f_n(x)\} \leqslant \epsilon, \qquad (n \geqslant N)$

i.e. $\qquad V_a^b\{f(x) - f_n(x)\} \to 0$, as $n \to \infty$.

It is easy to prove that, conversely, these conclusions imply (i), (ii).

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

Now by the integration by parts property (VI)

$$\int_a^b f_n(x)\,dg(x) = \Big[f_n(x)\,g(x)\Big]_a^b - \int_a^b g(x)\,df_n(x)$$

and

$$\left|\int_a^b g(x)\,df_m(x) - \int_a^b g(x)\,df_n(x)\right| = \left|\int_a^b g(x)\,d\{f_m(x) - f_n(x)\}\right|$$

$$\leqslant K\,V_a{}^b(f_m - f_n), \qquad \text{by (VII)}$$

$$\to 0, \text{ as } m, n \to \infty$$

K being the upper bound of $|g(x)|$. Thus

$$\int_a^b g(x)\,df_n(x) \to \text{a limit, } l \text{ (say), as } n \to \infty.$$

Now, given k, N can be found so that

$$V_a{}^b(f - f_N) < \frac{k}{3K}$$

and

$$(1) \quad \left|\int_a^b g(x)\,df_N(x) - l\right| < \frac{k}{3}.$$

Again, given ϵ, η can be found so that

$$(2) \quad m\left[\left|\Sigma\, g(\zeta_r)\{f_N(z_r) - f_N(z_{r-1})\} - \int_a^b g(x)\,df_N(x)\right| \geqslant \frac{k}{3}\right]$$

$$< \epsilon, \; (\lambda < \eta).$$

Now

$$(3) \quad |\Sigma\, g(\zeta_r)\{f_N(z_r) - f_N(z_{r-1})\} - \Sigma\, g(\zeta_r)\{f(z_r) - f(z_{r-1})\}|$$

$$\leqslant K\,V_a{}^b(f - f_N)$$

$$< \frac{k}{3}.$$

Thus by (1), (2), (3)

$$m[|\Sigma\, g(\zeta_r)\{f(z_r) - f(z_{r-1})\} - l| \geqslant k] < \epsilon, \qquad (\lambda < \eta).$$

This proves that

$$\int_a^b g(x)\,df(x) \text{ exists and } = l.$$

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

Hence finally

$$\int_a^b f_n(x)\,dg(x) = \Big[f_n(x)\,g(x)\Big]_a^b - \int_a^b g(x)\,df_n(x)$$
$$\to \Big[f(x)\,g(x)\Big]_a^b - \int_a^b g(x)\,df(x)$$
$$= \int_a^b f(x)\,dg(x).$$

5. *Term by term integration of Fourier series.*

Theorem 12 leads to a theorem of Parseval type.

THEOREM 13. *If*

(i) $f(x)$ *has period* 2π *and is continuous and of bounded variation in* $(-\pi, \pi)$.

(ii) $g(x)$ *is bounded*[1] *in* $(-\pi, \pi)$.

(iii) $a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos nx\, dg(x)$, $b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} \sin nx\, dg(x)$, *exist* (S_m),

then

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dg(x) = \tfrac{1}{2} a_0 a'_0 + \sum_{n=1}^{\infty} (a_n a'_n + b_n b'_n)$$

the series being summable $(C, 1)$[2], a'_n, b'_n *are the Fourier coefficients of* $f(x)$.

LEMMA 1. *If* $f(x)$ *is continuous and of bounded variation in* (a, b).

$$V_{a'}^{b'}\{f(x+y) - f(x)\} \to 0, \text{ as } y \to 0, \ (a < a' < b' < b)$$

$f(x)$ can be expressed[3] as the difference of two continuous increasing functions $P(x)$, $N(x)$. If y is (say) positive and less than $b - b'$,

$$f(x+y) - f(x) = P(x+y) - P(x) - \{N(x+y) - N(x)\}$$

so that

$$V_{a'}^{b'}\{f(x+y) - f(x)\} = P(b'+y) - P(b') + N(b'+y) - N(b')$$
$$- \{P(a'+y) - P(a') + N(a'+y) - N(a')\}$$
$$\to 0, \text{ as } y \to 0$$

since $P(x)$, $N(x)$ are continuous.

---

[1] If $g(x)$ is assumed to be continuous the integrals exist as ordinary Stieltjes integrals, and $f(x)$ need not be continuous. *Cf.* J. M. Whittaker, *Proc. Edin. Math. Soc.* (2), **1** (1928), 169.

[2] *Cf.* Hobson, *op. cit.*, **2**, 579.

[3] Hobson, *op. cit.*, **1**, 317.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press

LEMMA 2. Let $F_n(x)$ *denote the* n^{th} *partial Cesàro sum of the Fourier series of a function* $f(x)$ *satisfying conditions* (i). *Then*

$$V^{\pi}_{-\pi}\{F_n(x)-f(x)\}\to 0, \qquad \textit{as } n\to\infty .$$

For[1]

$$F_n(x)-f(x)=\frac{1}{2n\pi}\int_{-\pi}^{\pi}\{f(x+y)-f(x)\}\left(\frac{\sin\frac{1}{2}ny}{\sin\frac{1}{2}y}\right)^2 dy.$$

Thus for any subdivision of $(-\pi,\pi)$

$$\sum_{r=1}^{p}|\{F_n(x_r)-f(x_r)\}-\{F_n(x_{r-1})-f(x_{r-1})\}|$$

$$\leqslant \frac{1}{2n\pi}\int_{-\pi}^{\pi}\sum_{r=1}^{p}|\{f(x_r+y)-f(x_r)\}-\{f(x_{r-1}+y)-f(x_{r-1})\}|\left(\frac{\sin\frac{1}{2}ny}{\sin\frac{1}{2}y}\right)^2 dy$$

$$\leqslant \frac{1}{2n\pi}\int_{-\pi}^{\pi}V_a{}^b\{f(x+y)-f(x)\}\left(\frac{\sin\frac{1}{2}ny}{\sin\frac{1}{2}y}\right)^2 dy$$

where $a<-\pi\leqslant x\leqslant\pi<b$.

This is true for any subdivision of $(-\pi,\pi)$. Thus

$$V^{\pi}_{-\pi}\{F_n(x)-f(x)\}\leqslant\frac{1}{2n\pi}\int_{-\pi}^{\pi}V_a{}^b\{f(x+y)-f(x)\}\left(\frac{\sin\frac{1}{2}ny}{\sin\frac{1}{2}y}\right)^2 dy.$$

Now, by the preceding lemma, given ϵ, δ can be found such that

$$V^{\pi}_{-\pi}\{f(x+y)-f(x)\}<\epsilon, \qquad |y|<\delta.$$

Thus

$$V^{\pi}_{-\pi}(F_n-f)$$

$$\leqslant\frac{1}{2n\pi}\int_{-\delta}^{\delta}V_a{}^b\{f(x+y)-f(x)\}\left(\frac{\sin\frac{1}{2}ny}{\sin\frac{1}{2}y}\right)^2 dy+\frac{1}{2n\pi}\left(\int_{-\pi}^{-\delta}+\int_{\delta}^{\pi}\right)$$

$$\leqslant\epsilon\cdot\frac{1}{2n\pi}\int_{-\pi}^{\pi}\left(\frac{\sin\frac{1}{2}ny}{\sin\frac{1}{2}y}\right)^2 dy+2V^{\pi}_{-\pi}\cdot\frac{1}{2n\pi}\int_{\delta}^{\pi}\left(\frac{\sin\frac{1}{2}ny}{\sin\frac{1}{2}y}\right)^2 dy$$

$V^{\pi}_{-\pi}=$ total variation of $f(x)$ in $(-\pi,\pi)$

$$<2\epsilon, \qquad (n\geqslant n_0)$$

since

$$\frac{1}{2n\pi}\int_{-\pi}^{\pi}\left(\frac{\sin\frac{1}{2}ny}{\sin\frac{1}{2}y}\right)^2 dy=1,\ \frac{1}{2n\pi}\int_{\delta}^{\pi}\left(\frac{\sin\frac{1}{2}ny}{\sin\frac{1}{2}y}\right)^2 dy\to 0,\ \text{as } n\to\infty .$$

Theorem 13 now follows immediately from Theorem 12.

---

[1] Hobson, *op. cit*, **2**, 557.

https://doi.org/10.1017/S0013091500013614 Published online by Cambridge University Press