THE UNIVERSITY OF WISCONSIN
*madison, wisconsin*

# THREE NOTES ON DERIVATIVES AND INTEGRALS

I. On derivatives and double differences
II. Stieltjes integrals without bounded variation
III. Remark on stochastic integrals

L. C. Young

MRC Technical Summary Report #677
July 1966

UNITED STATES ARMY

MATHEMATICS RESEARCH CENTER

MATHEMATICS RESEARCH CENTER, UNITED STATES ARMY
THE UNIVERSITY OF WISCONSIN

Contract No.: DA-31-124-ARO-D-462

# THREE NOTES ON DERIVATIVES AND INTEGRALS

I. On derivatives and double differences
II. Stieltjes integrals without bounded variation
III. Remark on stochastic integrals

L. C. Young

MRC Technical Summary Report #677
July 1966

Madison, Wisconsin

ABSTRACT

1. On derivatives and double differences. Let $F, \varphi, \psi$ denote real functions defined in a neighborhood of the origin on the positive real half-line, and suppose, for small positive values of the variables, that (i) $\varphi$ is Borel measurable, $\psi$ is continuous and monotone increasing, and $\varphi(0) = \psi(0) = 0$; (ii) $u < v$ implies $v^{-1}\varphi(v) \leq 2u^{-1}\varphi(u)$ and $v^{-1}\psi(v) \leq 2u^{-1}\psi(u)$; (iii) $\int u^{-1}\varphi(u)\,d\psi(u)$ converges at $0$; and moreover that (iv) $|F(h+k)-F(h)-F(k)+F(0)| \leq \varphi(h)\psi(k)$. Then $F'(0)$ exists, and an estimate is given for the error in replacing it by a finite difference ratio.

2. Stieltjes integrals without bounded variation. The generalized Stieltjes integral $\int f\,dg$ (where $f, g$ are supposed of period one for simplicity and integration is over a period) is defined as the derivative $F'(0)$ of the convolution $F(x) = \int f(t)g(x+t)\,dt$. It is shown to exist (and an estimate for the error in replacing it by a finite difference ratio is given) provided that

$$\int_0^1 [f(x+h)-f(x)][g(x)-g(x-k)]\,dx \leq \varphi(h)\psi(k)$$

where $\varphi, \psi$ are subject to (i)(ii)(iii) in the abstract of 1.

3. Remark on stochastic integrals. A definition of stochastic integral similar to that of 2 is given.

# THREE NOTES ON DERIVATIVES AND INTEGRALS

L. C. Young

## I. ON DERIVATIVES AND DOUBLE DIFFERENCES

1. Introduction. We discuss in this note an inequality which connects a derivative with double differences, or more precisely, with two auxiliary functions $\phi, \psi$, termed estimate-functions, which play, for double differences, a part similar to that of moduli of continuity for a single difference. This was originally machinery devised for defining a Stieltjes integral without bounded variation [2, 4], a topic to which we return in a companion note [5]. As first presented at an invited address [3], the inequality involved derivation along the binary sequence $2^{-\nu}$ ($\nu = 1, 2, \ldots$). M. Riesz and J. G. van der Corput suggested at the time that it be reformulated to include a criterion for the existence of an ordinary derivative, but a later refinement actually involved derivation along a still more sparse sequence [4]. Here we show that the reformulation can, nevertheless, be effected even in the more refined case. This reformulation has been partly influenced by a recent note of Beurling [1]. We give also a theorem on term by term derivation.

Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin under Contract No.: DA-11-022-ORD-2059, and in part by NASA Contract No.: 144-6184, and NSF Contract No.: 144-4773.

2. Estimate-functions. We consider in this section a pair of functions $\varphi(u)$, $\psi(u)$ $(0 \leq u \leq 1)$ which vanish for $u = 0$ and are otherwise positive, and which are subject, for $0 < u < v \leq 1$, to the inequality

$$(2.1) \qquad \frac{\varphi(v)}{v} \leq 2\frac{\varphi(u)}{u}, \quad \frac{\psi(v)}{v} \leq 2\frac{\psi(u)}{u}.$$

In addition, we suppose $\psi$ continuous and monotone increasing, and $\varphi$ Borel measurable. Further, we denote by $\{h\}$ a decreasing sequence $h_\nu$ $(\nu = 1, 2, \ldots)$ with limit $0$, whose terms are less than an initial number $h = h_0 \leq 1$, and we write

$$S_{\{h\}} = \sum_{\nu=1}^{\infty} h_\nu^{-1} \varphi(h_\nu) \psi(h_{\nu-1}),$$

$$S(h) = \int_0^h \frac{\varphi(u)}{u} d\psi(u),$$

$$S \quad = S(1).$$

We say further that $\{h\}$ is subject to the condition $(c_1)$, if, for each $\nu$, the ratio $h_{\nu-1}/h_\nu$ is an integer expressible as a power of $2$, and to the condition $(c_2)$, if

$$2\psi(h_\nu) \leq \psi(h_{\nu-1}) \leq 8\psi(h_\nu).$$

The pair of functions will be termed estimate-functions if $S < \infty$. The following lemma provides alternative forms for this condition:

(2.2) Lemma. (i) The following conditions (a), (b), (c) are all equivalent: (a) $S < \infty$; (b) $S_{\{h\}} < \infty$ for some $\{h\}$; (c) $S_{\{h\}} < \infty$ for some $\{h\}$ subject to $(c_1)$ and $(c_2)$. (ii) We have, for any $\{h\}$, (2.3) $$S(h_0) \le 2S_{\{h\}} ;$$

and moreover, for any $\{h\}$ subject to $(c_1)$ and $(c_2)$,

(2.4) $$h_1^{-1}\varphi(h_1)\psi(h_1) \le 16\, S(h_1) ,$$

(2.5) $$S_{\{h\}} \le 128\, S(h_1) .$$

Proof. We write for short

$$s_\nu = \int_{h_\nu}^{h_{\nu-1}} \frac{\varphi(u)}{u} d\psi(u) \; ; \qquad s_\nu^* = h_\nu^{-1}\varphi(h_\nu)\psi(h_{\nu-1}) .$$

By (2.1) for $\varphi$,

$$s_\nu \le 2h_\nu^{-1}\varphi(h_\nu)[\psi(h_{\nu-1}) - \psi(h_\nu)] ,$$

$$2s_{\nu+1} \ge h_\nu^{-1}\varphi(h_\nu)[\psi(h_\nu) - \psi(h_{\nu+1})] .$$

From the first of these relations we derive $s_\nu \le 2s_\nu^*$, and so (2.3). From the second, by successively using the first and second inequality of $(c_2)$ with $\nu+1$ in place of $\nu$, and then the second inequality of $(c_2)$ itself, we find that

$$2s_{\nu+1} \ge h_\nu^{-1}\varphi(h_\nu)\psi(h_{\nu+1}) \ge h_\nu^{-1}\varphi(h_\nu)\psi(h_\nu)/8 \ge s_\nu^*/64 .$$

Evidently (2.4) and (2.5) follow from these relations.

To establish (i) it will now suffice to show that there exists an $\{h\}$ subject to $(c_1)$ and $(c_2)$. For this purpose, we attach to any $h$, where $0 < h \le 1$ and $h^*$ in the same interval such that $\psi(h) = 8\psi(h^*)$; for definiteness we take $h^*$ to be the smallest such number: $h^*$ clearly exists and is $< h$, since $\psi$ is continuous and monotone increasing. We now denote by $\hat{h}$ the smallest number of the form $h2^{-k}$ $(k = 1, 2, \ldots)$ such that $\hat{h} \geq h^*$. Clearly $\psi(\hat{h}) \geq \psi(h^*) = \psi(h)/8$. Further $\hat{h} < 2h^*$ and so, by (2.1) for $\psi$, we find that $\psi(\hat{h}) \leq (\hat{h}/h^*) \cdot 2\psi(h^*) \leq 4\psi(h^*) = \psi(h)/2$. Hence, by setting $h_0 = h$, $h_1 = \hat{h}$, and generally $h_\nu = \hat{h}_{\nu-1}$, we define by induction a sequence $\{h\}$ subject to $(c_1)$ and $(c_2)$, and this completes the proof.

3. Double difference identities and inequalities. We denote by $F(x)$ $(0 \leq x \leq 1)$ a continuous real-valued function, and if necessary we shall suppose it extended suitably to the interval $(0 \leq x \leq 2)$, although we shall really only be concerned with its values for small $x$. We write $\Delta_2 F$ for the expression

$$F(h+k) - F(h) - F(k) + F(0)$$

where $0 \leq h \leq 1$, $0 \leq k \leq 1$, so that $\Delta_2 F$ is a function $\Delta_2 F(h,k)$ symmetric in $h,k$. We say that $F$, or more precisely its double difference $\Delta_2 F$, admits the estimate-functions $\varphi, \psi$, if the latter are subject to the conditions of the preceding section and we have

$$|\Delta_2 F| \leq \varphi(h)\,\psi(k)\,. \tag{3.1}$$

We write further for short, if $N$ is an integer $> 1$ and $0 < Nh \leq 1$,

$$a(N,h) = \frac{F(Nh) - F(0)}{Nh} - \frac{F(h) - F(0)}{h}\,,$$

and we note, substantially as in [4], that

$$a(N,h) = [F(Nh) - NF(h) + (N-1)F(0)]/(Nh)$$

$$= \sum_{\mu=1}^{N-1} [F((\mu+1)h) - F(\mu h) - F(h) + F(0)]/(Nh)$$

$$= \frac{1}{N} \sum_{\mu=1}^{N-1} h^{-1} \Delta_2 F(\mu h, h) ,$$

and hence that

$$|a(N,h)| \leq h^{-1} \varphi(h) \psi(Nh) . \tag{3.2}$$

Finally, given the sequence $\{h\}$, we term derivative along this sequence, and we denote by $F'_{\{h\}}$, the limit for $h = h_\nu$, $\nu \to \infty$, of $[F(h)-F(0)]/h$. If we set $h_{\nu-1} = N_\nu h_\nu$ and suppose $\{h\}$ subject to $(c_1)$ and $(c_2)$ of the preceding section, we have then

$$-F'_{\{h\}} = -\frac{F(h_0)-F(0)}{h_0} + \sum_{\nu=1}^{\infty} a(N_\nu, h_\nu) ,$$

and hence, by (3.2) and (2.2), $F'_{\{h\}}$ then exists and satisfies

$$\left|\frac{F(h_0)-F(0)}{h_0} - F'_{\{h\}}\right| \leq 128\, S(h_1) . \tag{3.3}$$

4. Existence of the derivative $F'(0)$. We first prove the existence of a binary derivative, i.e. of a derivative along any binary sequence $h \cdot 2^{-\nu}$. We denote such a derivative, which may possibly depend on the first term $h$, by $F'_h$.

To this effect let $\{h\}$ be a sequence subject to $(c_1)$ and $(c_2)$, and let $h = h_0$ be its first term. We denote by $h^*$ any number of the form $h2^{-n}$, where $n$ is a positive integer. There will then exist an integer $\nu$ such that $h_\nu \leq h^* < h_{\nu-1}$ and an integer $N$, which is a power of 2, such that $Nh_\nu = h^*$. We shall suppose $N \neq 1$ since the excluded case is trivial, and we then have

$$\frac{F(h^*) - F(0)}{h^*} - \frac{F(h_\nu) - F(0)}{h_\nu} = a(N, h_\nu) .$$

This is in absolute value at most $h_\nu^{-1} \varphi(h_\nu)\psi(Nh_\nu)$ and therefore at most $h_\nu^{-1}\varphi(h_\nu)\psi(h_{\nu-1})$, and this last quantity tends to 0 as $\nu \to \infty$. It follows at once that $F'_h$ exists and has the same value as $F'_{\{h\}}$. It follows further from (3.3) that

$$\left| \frac{F(h) - F(0)}{h} - F'_h \right| \leq 128\, S(h) . \tag{4.1}$$

It only remains to prove that $F'_h$ is independent of $h$. To this effect, we first consider the difference $F'_h - F'_k$ where $k$ is a rational multiple $p/q$ of $h$. We write $h = q\delta$, $k = p\delta$, $h2^{-\nu} = h_\nu$, $k2^{-\nu} = k_\nu$, $\delta 2^{-\nu} = \delta_\nu$. Then

$$\frac{F(h_\nu) - F(0)}{h_\nu} - \frac{F(\delta_\nu) - F(0)}{\delta_\nu} = a(q, \delta_\nu)$$

where, by (3.2), (2.1) for $\psi$, and (2.4), the right-hand side is in absolute value

$$\leq q\varphi(\delta_\nu)\frac{\psi(q\delta_\nu)}{q\delta_\nu} \leq 2q\varphi(\delta_\nu)\frac{\psi(\delta_\nu)}{\delta_\nu} \leq 32qS(\delta_\nu) .$$

By making $\nu \to \infty$, we derive $F'_h - F'_\delta = 0$. Similarly $F'_k - F'_\delta = 0$, and therefore

$$F'_h = F'_k .$$

Hence, if we substitute in (4.1) for $h$ a rational $k$, and we make $k$ tend to an irrational $h$, we see, by continuity of $F$, that there is a number $\lambda = F'_k$ independent of $h$ such that

$$\left|\frac{F(h)-F(0)}{h} - \lambda\right| \leq 128\,S(h) .$$

Clearly this requires $\lambda = F'(0)$, so that this derivative exists. Incidentally we find that

$$\left|\frac{F(h)-F(0)}{h} - F'(0)\right| \leq 128\,S(h) . \tag{4.2}$$

We thus have the following result:

(4.3) Theorem. Let $F(x)$ admit the estimate-functions $\varphi, \psi$. Then the derivative $F'(0)$ exists and satisfies (4.2).

At the same time we can derive the following passage to the limit theorem:

(4.4) Theorem. Let $\sigma(u)$ $(0 \leq u \leq 1)$ denote a monotone increasing function, continuous at $0$, and let $\rho(u)$ be a Borel measurable function, which takes positive values $\leq +\infty$ for $0 < u \leq 1$, and for which $u^{-1}\rho(u)$ is Lebesgue-Stieltjes integrable with respect to $\sigma(u)$. Further let $F_n(x)$ $(0 \leq x \leq 1)$ be functions which tend to a limit function $F(x)$ as $n \to \infty$, for each sufficiently small $x$, and suppose that $F_n$ admits estimate-functions $\varphi_n, \psi_n$, where $\varphi_n \leq \rho$, and where, for every interval $(a,b)$ in $(0,1)$, $\psi_n(b) - \psi_n(a) \leq \sigma(b) - \sigma(a)$. Then the derivative $F'(0)$ exists and we have

$$F'(0) = \lim F_n'(0) .$$

Proof. If $\lambda$ denotes any limit of a subsequence of the numbers $F_n'(0)$, we find, for fixed small $h$, by taking limits in (4.2) applied to the corresponding $F_n$, that

$$\left|\frac{F(h) - F(0)}{h} - \lambda\right| \leq 128 \int_0^h \frac{\rho(u)}{u} \, d\sigma(u) .$$

Here the right-hand side is arbitrarily small with $h$, and therefore $\lambda = F'(0)$. This completes the proof.

In a number of applications, $\psi_n$ is a function $\psi$ independent of $n$, and we can set $\sigma = \psi$.

LITERATURE


1. Beurling A., Analyse harmonique de pseudo-mesures. Intégration par rapport aux pseudo-mesures. C. R. Acad. Sc. Paris 258 (1964) 1984-7.

2. Young L. C., Inequalities connected with binary derivation, integrated Lipschitz conditions and Fourier series. Math. Zeitschr. 43 (1937) 449-467.

3. ________, Intégrales de Stieltjes et convergence des séries de Fourier. Soc. Math. de France, Conférences de la Réunion internat. de math. Paris (1937).

4. ________, A further inequality for Stieltjes integrals. Journ. London Math. Soc. 18 (1943) 78-82.

5. ________, Stieltjes integrals without bounded variation, p. 10 of this report.

6. ________, Remark on stochastic integrals, p. 15 of this report.

## II. STIELTJES INTEGRALS WITHOUT BOUNDED VARIATION

5. Introduction. This brief note consists of a few remarks, which constitute, as the numbering of the sections indicates, a sequel to the note [22] on derivatives and double differences, and which indicate the connection between the machinery there treated and Stieltjes integrals without bounded variation. Such integrals were first introduced more than thrity years ago [18] and the machinery of double differences originated from their study [2]. They are related to such topics as fractional integration, convergence of Fourier series, stochastic integrals. In addition to the papers [1-6] cited in the previous note [22], we list consecutively [7-23] a number of further relevant papers at the end of this note. The theorems of the preceding note enable us to give a slightly simpler form to a generalized Stieltjes integral introduced in [4], and to show that it includes a definition recently proposed by Beurling [1].

6. The generalized Stieltjes integral. We denote by $f, g$ two real-valued functions of period one on the real line, such that the convolution

$$F(x) = \int_0^1 f(x+t)g(t)\,dt$$

exists. The periodicity assumption, which amounts to subtracting suitable linear functions, represents no real loss of generality. By the generalized Stieltjes integral $\mathcal{J}(f,g)$ over the unit period, we shall mean the quantity $-F'(0)$, where $F'(0)$ is the derivative at $0$ of the function $F$, considered for small positive $x$. It is shown in [2] that this definition includes, as special cases, the classical Riemann-Stieltjes integral $\int f dg$, and also the Lebesgue integral of $fg'$ when $g'$ is the derivative of an absolutely continuous $g$ and the functions $f, g'$ belong to conjugate Lebesgue classes.

In [2] and [3], the above definition was then modified by interpreting derivation in a special manner, as derivation along a binary sequence $2^{-\nu}$, or along a still more sparse sequence. This slight artificiality can now be removed, since it turns out that, in the cases there considered, the ordinary derivative exists.

Our definition allows us to apply the machinery of the preceding note. The double difference $\Delta_2 F$ can now be written

$$-\int_0^1 [f(t+h)-f(t)][g(t)-g(t-k)]\,dt$$

and we therefore suppose that there exist a pair $\varphi, \psi$ of estimate-functions, such that

$$\left|\int_0^1 [f(t+h)-f(t)][g(t)-g(t-k)]\,dt\right| \leq \varphi(h)\,\psi(k) \;. \tag{6.1}$$

Under these circumstances, it then follows that $\mathcal{J}(f,g)$ exists. We recall, however, that $\varphi, \psi$ are to satisfy the conditions of section 2 of the preceding note, and in particular the condition

$$\int_0^1 \frac{\varphi(u)}{u}\, d\psi(u) < \infty \;. \tag{6.2}$$

The conditions for the existence of $\mathcal{J}(f,g)$ can be specialized by attaching $\varphi, \psi$ in some particular manner to $f, g$. Thus if $g$ is continuous, we can choose

$\varphi, \psi$ to be the moduli of "continuity" of $f, g$, provided that (6.2) then holds. The corresponding definition of $\mathcal{J}(f, g)$ then agrees with a definition of Beurling [1], except that Beurling supposes $f$ continuous as well as $g$.

Beurling's definition seems at first quite different, and we shall not repeat it here. What is relevant here is simply that his integral is the limit as $\epsilon \to 0$ of a classical Stieltjes integral $\int f_\epsilon \, dg$, and therefore of our $\mathcal{J}(f_\epsilon, g)$, where $f_\epsilon$ is a specially constructed approximation to $f$, consisting of a function which is monotone in each of a finite number of parts of $(0,1)$, and which possesses a modulus of continuity $\varphi_\epsilon$ not exceeding the modulus of continuity $\varphi$ of $f$. In this case, theorem (4.4) of the preceding note clearly shows that $\mathcal{J}(f_\epsilon, g)$ tends to $\mathcal{J}(f, g)$, and this ensures that Beurling's integral coincides with $\mathcal{J}(f, g)$.

More general sufficient conditions for the existence of the generalized Stieltjes integral $\mathcal{J}(f, g)$ are obtained as in [2, 20, and 4], by choosing, subject to an interchange of $f$ and $g$ if preferred, the functions $\varphi(h)$, $\psi(k)$ to be, for instance, the suprema for $0 \leq u \leq h$ and $0 \leq v \leq k$ of the expressions

$$\left( \int_0^1 |f(t+u) - f(t)|^p \, dt \right)^{1/p}, \qquad \left( \int_0^1 |g(t) - g(t-v)|^q \right)^{1/q},$$

where $p, q$ are conjugate indices in Hölder's inequality (i.e. where $p^{-1} + q^{-1} = 1$). The functions $\varphi, \psi$ are then again supposed to verify (6.2).

Interesting questions arise from the connection of our integrals with pseudo-measures [1] and also with stochastic integrals.

## FURTHER LITERATURE

[7] Burkill J. C., Differential properties of the Young-Stieltjes integral, Journ. London Math. Soc. 23 (1948) 22-28.

[8] Hardy G. H. and Littlewood J. E., Some properties of fractional integrals, II, Math. Zeitschr. 34 (1931) 403-439.

[9] Gehring F. W., A study in $\alpha$-variation, Trans. Amer. Math. Soc. 76 (1954) 420-443.

[10] Lévy Paul, Propriétés intrinsèques des fonctions et Intégrales de Stieltjes Bull. Soc. Math. France 69 (1941) 5-9 (Communications et Conferences).

[11] _________, Intégrales de Stieltjes généralisées, Ann. Soc. Pol. Math. 25 (1952-3) 17-26.

[12] Love E. R., Fractional integration and almost periodic functions, Proc. London Math. Soc. (2) 44 (1938) 363-397.

[13] _________, A generalization of absolute continuity, Journ. London Math. Soc. 26 (1951) 1-13.

[14] Love E. R. and Young L. C., Sur une classe de fonctionnelles linéaires, Fund. Math. 28 (1937) 243-257.

[15] _________, On fractional integration by parts, Proc. London Math. Soc. (2) 44 (1938) 1-35.

[16] Orlicz W., Généralization d'un théorème de Banach et Mazur, Ann. Soc. Pol. Math. 19 (1946-7) 62-65.

[17] Wiener N., The quadratic variation of a function and its Fourier coefficients, Journ. Mass. Inst. Tech. 3 (1924) 73-94.

[18] Young L. C., An inequality of the Hölder type connected with Stieltjes integration, Acta. Math. 67 (1936) 251-282.

[19] __________, General inequalities for Stieltjes integrals and the convergence of Fourier series, Math. Ann. 115 (1938) 581-612.

[20] __________, Inequalities connected with bounded p-th power variation in the Wiener sense and with integrated Lipschitz conditions, Proc. London Math. Soc. (2) 43 (1937) 449-467.

[21] __________, A convergence criterion for Fourier series, Quart. Journ. (Oxford) 9 (1938) 115-120.

[22] __________, On derivatives and double differences, p. 1 of this report.

[23] __________, Remark on stochastic integrals, p. 15 of this report.

## III. REMARK ON STOCHASTIC INTEGRALS

7. Introduction. In this brief stochastic postscript to the preceding notes [iii,iv] we continue to use the language of analysis rather than probability, since we merely wish to point out how stochastic integrals can be introduced (in the spirit of those notes, but without the material there developed), as derivatives of corresponding convolutions. However, these derivatives are understood, not as in classical analysis, but in a sense which is natural in a stochastic context, as limits in the mean. Moreover, the stochastic integral, considered here, is the special type of integral, which goes back in effect to Wiener, and which is attached to a "stochastic process with orthogonal increments" [b, Chap. IX]. A more general concept, although highly desirable, has yet to be formulated in a conclusive manner, and our remarks do not concern it, except in so far as they may prove helpful in its formulation.

8. Background. We shall now be dealing with complex-valued functions, and $\bar{z}$ denotes the conjugate of the complex value $z$. Moreover, the same symbol $\Delta$ will be used for an interval $(a, b)$ on the real line, and for a difference operator $\Delta f = f(b) - f(a)$ on a function $f$. Finally, given any such interval $\Delta$, we shall denote by $\delta$ in this note the characteristic function $\delta(\lambda)$ of $\Delta$, defined to be 1 for $a < \lambda < b$, and 0 otherwise.

On the Cartesian product $T \times \Omega$ of a real line $T$ (time axis) and a set $\Omega$ (probability space), we consider a fixed complex-valued function $x = x(t, \omega)$ (termed stochastic process), and we attach to $\Omega$ a fixed unit measure $d\omega$ (termed probability) defined on a suitable Boolean $\sigma$-ring of subsets. Differences of the form $\Delta x = x(b,\omega) - x(a,\omega)$ will be supposed square integrable in $d\omega$.

The function $x$ will be said to have orthogonal increments, if, for each pair of non-overlapping intervals $\Delta_1, \Delta_2$ of $T$, we have

$$\int (\Delta_1 x)(\Delta_2 \overline{x})\, d\omega = 0 \ . \tag{8.1}$$

It is a well-known consequence [ii, Chap. II §10] of this relation, that there exists an increasing function $\mu = \mu(t)$ on $T$, such that, for an arbitrary pair of intervals $\Delta_1, \Delta_2$ on $T$, the left-hand side of (8.1) has the value

$$\Delta\mu = \mu(b) - \mu(a) \ ,$$

where $\Delta = (a,b)$ is the intersection of $\Delta_1$ and $\Delta_2$. This strengthened form of (8.1) may therefore conveniently be taken as the definition of the property of having orthogonal increments in the sequel. It then follows that, if the end points of $\Delta_1, \Delta_2$ avoid the at most countable set of discontinuities of $\mu$ on $T$, we have

$$\int (\Delta_1 x)(\Delta_2 \overline{x})\, d\omega = \int \delta_1(\lambda)\delta_2(\lambda)\, d\mu(\lambda) \ . \tag{8.2}$$

This exclusion of discontinuities could be avoided at the expense of considering neutral intervals, instead of open ones, and interpreting $\delta_1, \delta_2$ as place-functions, in the sense of Carathéodory [ i ] . In this type of context, there is always a slight artificiality in the notions of sets of points and real functions, which is a little unpleasant. However the formula (8.2) in the form stated above, is sufficient for our purposes.

9. <u>The basic identity.</u> We shall suppose that $x$ has orthogonal increments and we fix a corresponding monotone increasing $\mu$. Further, by subtracting $x(0,\omega)$ from $x$ if necessary, we shall suppose that $x(0,\omega) = 0$ and therefore $x$ is now square integrable in $d\omega$. In addition, we shall suppose that $x$ is locally integrable in the Lebesgue measure $dt$. (We are not trying to set up a minimum system of assumptions under which the formula to be established is valid. However we return to this question below.)

We denote further by $y$ a complex-valued function $y(t)$, $t \in T$, of which, for the moment, we suppose that it is bounded, Borel measurable, and of compact support, and we write $\psi(t)$ for its indefinite integral. We now form the convolution

$$\mathfrak{Z}(t,\omega) = \int_T y(u)x(t+u,\omega)\,du$$

and we fix two positive real numbers $h, k$. We shall verify the identity

$$(9.1)\quad \begin{cases} \displaystyle\int_\Omega \left|\frac{\mathfrak{Z}(h,\omega)-\mathfrak{Z}(0,\omega)}{h} - \frac{\mathfrak{Z}(k,\omega)-\mathfrak{Z}(0,\omega)}{k}\right|^2 d\omega \\ \displaystyle\qquad = \int_T \left|\frac{\psi(\lambda)-\psi(\lambda-h)}{h} - \frac{\psi(\lambda)-\psi(\lambda-k)}{k}\right|^2 d\mu(\lambda) . \end{cases}$$

It will clearly suffice to prove, for arbitrary real $h > 0$, $k > 0$ the identity of the expressions

$$\mathfrak{a}(h,k) = \int_\Omega \{\mathfrak{Z}(h,\omega)-\mathfrak{Z}(0,\omega)\}\{\mathfrak{Z}(k,\omega)-\mathfrak{Z}(0,\omega)\}\,d\omega$$

and

$$\mathfrak{B}(h,k) = \int_T \{\psi(\lambda)-\psi(\lambda-h)\}\{\overline{\psi}(\lambda)-\overline{\psi}(\lambda-k)\}\,d\mu(\lambda) .$$

We write $\Delta_h, \Delta_k^*$ for the intervals $(t, t+h)$, $(t^*, t^*+k)$, and $\delta_h, \delta_k^*$ for the characteristic functions, of a parameter $\lambda$, of these intervals. The functions $\delta_h, \delta_k^*$ depend also on $t$, $t^*$ respectively, and we note that the expressions

$$(9.2) \qquad \int_T y(t)\,\delta_h(\lambda)\,dt, \qquad \int_T \overline{y}(t^*)\,\delta_k^*(\lambda)\,dt^*$$

are the integrals of $y(t)$, $\overline{y}(t^*)$ on the intervals $\lambda-h < t < \lambda$ and $\lambda-k < t^* < \lambda$, and therefore have the values

$$\psi(\lambda)-\psi(\lambda-h), \qquad \overline{\psi}(\lambda)-\overline{\psi}(\lambda-k) .$$

This being so, we write $\mathfrak{G}(h,k)$ in the form

$$(9.3) \qquad \iiint y(t)\,\overline{y}(t^*)\,U\,U^*\,dt\,dt^*\,d\omega ,$$

where $U = \Delta_h x$, $U^* = \Delta_k \overline{x}$. By Fubini's theorem, we may integrate first in $\omega$. In so doing we shall ignore a countable set of $t$, $t^*$ (depending on $h$ or $k$) and we can then apply (8.2). We find that

$$\int UU^* d\omega = \int \delta_h(\lambda)\,\delta_k^*(\lambda)\,d\mu(\lambda) .$$

Consequently, when we multiply by $y(t)\,\overline{y}(t^*)$ and integrate in $dt\,dt^*$, we obtain, by another application of Fubini's theorem, the integral in $d\mu(\lambda)$ of the product of the expressions (9.2) and therefore the integral $\mathfrak{B}(h,k)$. This completes the proof.

It also follows, of course, that, for instance,

$$(9.4) \qquad \int_{\Omega} \left|\frac{\mathfrak{F}(h,\omega) - \mathfrak{F}(0,\omega)}{h}\right|^2 d\omega = \int_T \left|\frac{\psi(\lambda) - \psi(\lambda - h)}{h}\right|^2 d\mu(\lambda) .$$

Moreover, the function $y(\lambda)$, $\lambda \in T$, is clearly the limit in mean square, with measure $d\mu$, of the sequence of difference ratios $k^{-1}\{\psi(\lambda) - \psi(\lambda - k)\}$, as $k$ describes a sequence tending to $0$. It follows at once from (9.1) that the corresponding limit in mean square in $\Omega$ of the difference ratios

$$\frac{\mathfrak{F}(k,\omega) - \mathfrak{F}(0,\omega)}{k}$$

exists and has a value $\mathfrak{J}(\omega)$ independent of the choice of the sequence of $k$. Moreover

$$(9.5) \qquad \begin{cases} \displaystyle\int_{\Omega} \left|\frac{\mathfrak{F}(h,\omega) - \mathfrak{F}(0,\omega)}{h} - \mathfrak{J}(\omega)\right|^2 d\omega \\ \qquad\qquad = \displaystyle\int_T \left|\frac{\psi(\lambda) - \psi(\lambda - h)}{h} - y(\lambda)\right|^2 d\mu(\lambda) . \end{cases}$$

The quantity $\mathfrak{J}(\omega)$ is the stochastic integral

$$\int_T y(t)\, dx(t, \omega) ,$$

and we see that it is thus the derivative (in $t$ at $t = 0$) in mean square (in $\Omega$) of the convolution $\mathfrak{F}$.

This definition of the stochastic integral is here subject to unnecessarily restrictive hypotheses as compared with the standard definition, which goes back to Wiener. Local integrability in $t$ of $x(t, \omega)$ is there not required, and the assumption that y(t) be bounded and of compact support is replaced by that of square integrability in $d\mu(t)$. This, however, would not allow us to define the convolution directly, but at best only as a limit in the mean, which would detract from the simplicity of our definition. It is sometimes useful to look for an alternative formulation at the expense of secondary restrictions, with the ultimate object of achieving greater generality ("Reculer pour mieux avancer"). At any rate, it would be interesting to see whether a combination of our definition with the methods of the preceding notes could lead to a more generally applicable concept of stochastic integral.

## REFERENCES IN THIS NOTE


( i) Carathéodory, C., Mass und Integral und ihre Algebraisierung, (Birkhäuser, Basel 1956) available also in English translation: Algebraic Theory of Measure and Integration (Chelsea, New York, 1963).

( ii) Dobb, J. L., Stochastic processes, Wiley, New York, 1953.

( iii) Young, L. C., Derivatives and double differences, p. 1 of this report.

( iv) ___________, Stueltjes integrals without bounded variation, p. 10 of this report.