THE FIRST VARIATION OF A FUNCTIONAL

June 1929,

W.G.Smiley Jr.

Professor P.J.Daniell(1) has proved that the first variation of a functional is a linear functional of the variation of the argument, and by a theorem(2) of Frédéric Riesz, is expressible as a Stieltjes integral

$$1) \qquad \delta F[\overset{b}{\underset{a}{f}}(x)] = D(f; \varphi) = \int_a^b \varphi(\xi)\, d\alpha(\xi)$$

where $F[\overset{b}{\underset{a}{f}}]$ is a functional of the continuous function, $f(x)$ in the interval $a \leq x \leq b$; and $D(f;\varphi) = \lim_{\epsilon \to 0} \frac{F[f+\epsilon\varphi]-F[f]}{\epsilon}$
The proof is not entirely complete, but by slightly strengthening one of the assumptions the result is obtained.

Professor G.C.Evans(3) had previously demonstrated, that by assumptions related to those of Professor Daniell, the first variation takes the form of Professor Volterra

$$2) \qquad D(f;\varphi) = \int_a^b F'[f|\xi]\,\varphi(\xi)\,d\xi \qquad .$$

In this paper it is proposed to show that by adding certain postulates of Professor Evans' to the assumptions sufficient to give a complete proof of Professor Daniell's result, the first variation is reduced from the form 1) to the form 2). It will also be shown that by assumptions very similar to those of Professor

---

(1) Bulletin of the Am.Math.Soc., 2nd Series. Vol. XXV, No. 9 pp. 414-416, June 1919 (Derivative of a Functional)
(2) (Demonstration a un Theoreme Concernant Les Operation Fonctionelle Leneaires) Annales de L'Ecole Normale Superieue V 31, p. 10.
(3) Bull. Am.Math.Soc., 2nd Series Vol XXL No. 8, pp. 387-397. May 1915. A Note on the Derivative and the Variation of a Function. Depending on all the Value of Another Function 1.

Daniell's, the first variation takes the form of M. Frechet(1)

$$3) \qquad D(f;\varphi) = \int_a^b \varphi(\xi)\beta(\xi)\,d\xi$$

Professor Daniell makes the following assumptions:

$$(I) \qquad |F[f_1] - F[f_2]| \leqq M \max |f_1 - f_2|$$

(II) The first variation $D(f';\varphi)$ exists for all continuous $\varphi'(x)$, and all continuous $f'(x)$ in the neighborhood of $f(x)$.

By means of (II) it is proved that $D(f,\varphi)$ is distributive in i.e.:

$$4) \qquad D(f;\varphi_1+\varphi_2) = D(f;\varphi_1) + D(f;\varphi_2)$$

This proof depends on the fact that the limit

$$D(f+\epsilon m\varphi_2;\varphi_1) = \lim_{\epsilon\to 0} \frac{F[f+\epsilon m\varphi_2+\epsilon\varphi_1]-F[f+\epsilon m\varphi_2]}{\epsilon}$$

exists. It must be shown then that given $\eta > 0$

$$5) \qquad \left| D(f+\epsilon m\varphi_2;\varphi_1) - \frac{F[f+\epsilon m\varphi_2+\epsilon\varphi_1]-F[f+\epsilon m\varphi_2]}{\epsilon} \right| < \eta$$

for $0<\epsilon<\delta$. But this result does not follow from II unless the limit $D(f';\varphi)$ exists uniformly for all $f'(x)$ in the neighborhood of $f(x)$. This gives us the condition necessary to

---

(1) M. Frechet, Transactions Am.Math.Soc., Vol 15, 1914 p. 139.

complete the proof.

(II') The limit $D(f';\varphi)$ of (II) exists uniformly for all $f'(t)$ in the neighborhood of $f(t)$.

Proof of the Distributive Property of the First Variation.

THEOREM:- $$D(f;\, c_1\varphi_1 + c_2\varphi_2) = c_1 D(f;\varphi_1) + c_2 D(f;\varphi_2)$$

To obtain this result from the conditions I, and (II') by an alternative method, there is need of several lemmas.

Lemma a) $F[f(t)]$ is continuous in $f(t)$.

This fact is an immediate consequence of (I) since given any $\eta > 0$ we have

$$|F[f_1] - F[f_2]| \leqq M \max|f_1 - f_2| \leqq \eta$$

by taking

$$\max|f_1 - f_2| \leqq \frac{\eta}{M} \qquad \therefore$$

$$|F[f_1] - F[f_2]| \leqq \eta \qquad |f_1 - f_2| \leqq \delta$$

Q.E.D.

Lemma b) $D(f,\varphi)$ is continuous in $f(t)$.

This can be proved directly. Given any number $\eta > 0$ we can find δ such that

$$\left|\frac{F[f_0 + \epsilon_0\varphi] - F[f_0]}{\epsilon_0} - D(f_0;\varphi)\right| \leqq \frac{\eta}{4}\;; \qquad \left|\frac{F[f' + \epsilon_0\varphi] - F[f']}{\epsilon_0} - D(f';\varphi)\right| \leqq \frac{\eta}{4}$$

for ϵ_0 fixed $< \delta$.

consequently

$$7)\qquad \left| D(f_0;\varphi) - D(f';\varphi) \right| \leq \frac{\eta}{2} + \left| \frac{F[f_0+\epsilon_0\varphi] - F[f'+\epsilon_0\varphi]}{\epsilon_0} + \frac{F[f'] - F[f_0]}{\epsilon_0} \right|$$

from the uniformity of condition II' we can take $|f'-f_0|$ as small as we please without altering the validity of 6); take $|f'-f_0|$ so small that

$$\left| \frac{F[f_0+\epsilon_0\varphi] - F[f'+\epsilon_0\varphi]}{\epsilon_0} \right| \leq \frac{\eta}{2}$$

and

$$\left| \frac{F[f'] - F[f_0]}{\epsilon_0} \right| \leq \frac{\eta}{2}$$

(Lemma a); from 7) then

$$\left| D(f_0;\varphi) - D(f';\varphi) \right| \leq \eta$$

and $D(f;\varphi)$ is therefore continuous in $f(x)$.

(Lemma c) $$D(f;\varphi_2) = \lim_{\epsilon\to 0} \frac{F[f+\epsilon\varphi_1+\epsilon\varphi_2] - F[f+\epsilon\varphi_1]}{\epsilon}$$

This result follows at once from the fact that given $\eta > 0$;

$$8)\qquad \left| D(f;\varphi_2) - \frac{F[f+\epsilon\varphi_1+\epsilon_2\varphi_2] - F[f+\epsilon\varphi_1]}{\epsilon} \right| \leq \eta$$

for all ϵ some δ small enough. We obtain 8) as follows:

$$9)\qquad \left| D(f+\epsilon_1\varphi_1;\varphi_2) - D(f;\varphi_2) \right| \leq \frac{\eta}{2}$$

for all $\epsilon_1 \leq \delta_1$, by lemma b).

$$\left| D(f+\epsilon_1\varphi_1;\varphi_2) - \frac{F[f+\epsilon_1\varphi_1+\epsilon\varphi_2] - F[f+\epsilon_1\varphi_1]}{\epsilon} \right| \leq \frac{\eta}{2}$$

by II' for all $\epsilon, \epsilon_1 <$ some number δ; and consequently we may take, in this expression $\epsilon_1 = \epsilon < \delta$, obtaining:

10) $$\left| D(f+\epsilon_1\varphi_1;\varphi_2) - \frac{F[f+\epsilon\varphi_1+\epsilon\varphi_2]-F[f+\epsilon_1\varphi_1]}{\epsilon} \right| \leq \frac{\eta}{2}$$

From 9) and 10) we obtain:

8) $$\left| D(f;\varphi_2) - \frac{F[f+\epsilon\varphi_1+\epsilon\varphi_2]-F[f+\epsilon\varphi_1]}{\epsilon} \right| \leq \eta$$

for all ϵ such that $0 < \epsilon \leq \delta$ small enough. Q.E.D.

We have, given η, δ, such that, for $\epsilon < \delta$

$$\left| D(f;\varphi_1+\varphi_2) - \frac{F[f+\epsilon\varphi_1+\epsilon\varphi_2]-F[f]}{\epsilon} \right| \leq \frac{\eta}{6}$$

$$\left| D(f;\varphi_1) - \frac{F[f+\epsilon\varphi_1]-F[f]}{\epsilon} \right| \leq \frac{\eta}{6}$$

$$\left| D(f;\varphi_2) - \frac{F[f+\epsilon\varphi_2]-F[f]}{\epsilon} \right| \leq \frac{\eta}{6}$$

consequently

10) $$\left| D(f;\varphi_1+\varphi_2) - D(f;\varphi_1) - D(f;\varphi_2) \right| \leq \frac{\eta}{2} + \left| \frac{F[f+\epsilon\varphi_1+\epsilon\varphi_2]-F[f+\epsilon\varphi_1]}{\epsilon} - \frac{F[f+\epsilon\varphi_2]-F[f]}{\epsilon} \right|$$

but by lemma c) for ϵ small enough

$$\left| \frac{F[f+\epsilon\varphi_1+\epsilon\varphi_2]-F[f+\epsilon\varphi_1]}{\epsilon} - D(f;\varphi_2) \right| \leq \frac{\eta}{4}$$

and we have

$$\left| \frac{F[f+\epsilon\varphi_2]-F[f]}{\epsilon} - D(f;\varphi_2) \right| \leq \frac{\eta}{4}$$

for ϵ small enough, and we can take ϵ so small that

$$\left| \frac{F[f+\epsilon\varphi_1+\epsilon\varphi_2]-F[f+\epsilon\varphi_1]}{\epsilon} - \frac{F[f+\epsilon\varphi_2]-F[f]}{\epsilon} \right| \leq \frac{\eta}{2}$$

holds. Finally from 10) we get

$$\left| D(f;\varphi_1+\varphi_2) - D(f;\varphi_1) + D(f;\varphi_2) \right| \leq \eta$$

and the theorem is proved.

From the definition of the first variation we have

$$D(f;c\varphi) = c\,D(f;\varphi)$$

since

$$\begin{aligned} D(f;c\varphi) &= \lim_{\epsilon\to 0} \frac{F[f+\epsilon c\varphi]-F[f]}{\epsilon} \\ &= c \lim_{\mu\to 0} \frac{F[f+\mu\varphi]-F[f]}{\mu} \\ &= c\,D(f;\varphi) \end{aligned}$$

From this fact follows the full distributive property

11) $$D(f;c_1\varphi_1+c_2\varphi_2) = c_1 D(f;\varphi_1) + c_2 D(f;\varphi_2)$$

From (I) with Professor Daniell we get

$$|F[f+\epsilon\varphi]-F[f]| \leq \epsilon M \max|\varphi|$$

and

12) $$|D(f;\varphi)| \leq M \max|\varphi|$$

From 11) and 12) we see that $D(f;\varphi)$ is a linear functional and by the theorem of M. Frederic Riesz(1), is of the form 1) where $\alpha_f(x)$ is a function of limited variation in (a,b), depending only on $f(x)$. (and F)

From the results obtained by Professor Evans(2), by adding to (I), and (II'), the following hypotheses:

(α') $F[f(x)]$ is defined for continuous functions $f(x)$ having at most a finite number of simple discontinuities.

$$(\gamma') \qquad F'[f/\xi] = \lim_{\substack{\omega \to 0 \\ \alpha \to \xi \\ \beta \to \xi}} \frac{F[f + \omega\, {}_\alpha\psi_\beta] - F[f]}{\omega \int_\alpha^\beta {}_\alpha\psi_\beta \, dx}$$

where $\psi(x)$ is any continuous function of one sign in (a,b); and the function

$$\begin{aligned} {}_\alpha\psi_\beta(x) &= \psi(x), & (\alpha \leqq x \leqq \beta) \\ &= 0, & (a \leqq x < \alpha;\ \beta < x \leqq b) \end{aligned}$$

it follows at once that the first variation is reduced from the form 1) to 2).

But an analogous result can be obtained from assumptions much less stringent. Let us replace the postulate (I) by

$$\text{(III)} \qquad |F[f_1] - F[f_2]| \leqq M \int_a^b |f_1 - f_2| \, dx$$

The postulate (I) follows immediately from (III) so that from (II'), and (III) we obtain as before

$$D(f; \varphi) = \int_a^b \varphi(\xi) \, d\alpha_f(\xi)$$

---

(1) Loc.cit. P 11
(2) Loc.cit. P 397, #12.

The definition of $F[f(x)]$ will now be extended to a broader functional field. Consider the functional $F[f(x)]$ defined for every continuous function, $f(x)$, in (a,b), and within the region R bounded by the functions $\Phi_1 < \Phi_2$ i.e.

$$R: \quad \Phi_1(x) \leqq f(x) \leqq \Phi_2(x), \qquad (a \leqq x \leqq b)$$
$$\Phi_1, \Phi_2 \text{ continuous}$$

also satisfyng

$$\text{(III)} \qquad \left|F[f_1] - F[f_2]\right| \leqq M\int_a^b |f_1 - f_2|\,dx$$

Then $F[f(x)]$ is defined for all the functions of Baire.

For let $\{f_n\}$ be any sequence of continuous functions in R, convergent in the mean of order 1. This sequence converges to a limiting function $\Psi(x)$. Now consider the sequence of functional values

$$13) \qquad F[f_1] + (F[f_2] - F[f_1]) + \cdots$$

corresponding to the sequence of functions

$$f_1 + (f_2 - f_1) + \cdots$$

The series 13) converges to a finite unique limit for on considering the remainders we see

$$|R_{m,1}| = |F[f_{n+1}] - F[f_n]| \leqq M \int_a^b |f_{n+1} - f_n| \, dx$$

$$|R_{m,2}| = |F[f_{n+2}] - F[f_n]| \leqq M \int_a^b |f_{n+2} - f_n| \, dx$$

$$\cdots \qquad \cdots \qquad \cdots$$

$$|R_{m,p}| = |F[f_{n+p}] - F[f_n]| \leqq M \int_a^b |f_{n+p} - f_n| \, dx$$

$$\cdots \qquad \cdots \qquad \cdots$$

may all be made less than η, arbitrarily small by taking n large enough; this result follows from the hypothesis that the sequence $\{f_n\}$ converges in the mean. Consequently the series 13) converges to a finite limit which is denoted by $\Psi(x)$, and we have

$$\lim_{n \to \infty} F[f_n] = F[\Psi(x)]$$

This limit, moreover, is unique; suppose $\{f_n\}$, and $\{g_n\}$ are two such sequences converging in the mean to $\Psi(x)$.

Then

$$\lim_{n \to \infty} F[f_n] = F[\Psi] \quad ; \quad \lim_{n \to \infty} F[g_n] = F_1[\Psi]$$

but we have

$$|F[\Psi] - F_1[\Psi]| \leqq (1) + (2) + (3)$$

where

$$(1) = |F[\psi] - F[f_n]|$$

$$(2) = |F_1[\psi] - F[g_n]|$$

$$(3) = |F[f_n] - F[g_n]|$$

but (3) $\leq M \int_a^b |f_n - g_n|\, dx$, and we can take $n > n_0$ such that (1) $\leq \epsilon/3$, $n_* > n_0$ such that (2) $\leq \epsilon/3$ $n_0 > n_0$ great enough such that (3) $\leq \epsilon/3$ and consequently we have

$$|F[\psi] - F_1[\psi]| \leq \epsilon$$

or

$$\lim_{n\to\infty} F[f_n] = \lim_{n\to\infty} F[g_n] = F[\psi]$$

Q.E.D.

Also it is apparent that (III) is satisfied for such functions $\psi(x)$; for let $\{f_n\}$, $\{g_n\}$ be two sequences of continuous functions converging in the mean to $\psi(x)$, and $\varphi(x)$ respectively, in the regiòn R. Then we have

$$|F[f_n] - F[g_n]| \leq M \int_a^b |f_n - g_n|\, dx$$

for all n, and in the limit we get, since the f_n, g_n are bounded,

$$\text{(III')} \qquad |F[\psi] - F[\varphi]| \leq M \int_a^b |\psi - \varphi|\, dx$$

From (III') we see that $F[\psi]$ is continuous in the limit functions .

The above reasoning suffices to show the functional may be extended to the first class of functions (Baire), for which it satisfies the hypothesis (III), and the condition of continuity. In an exactly analogous way the functional may be defined for functions of all finite classes of Baire, in the region R.

To extend the definition of $F[f(x)]$ further, consider the sequence:

$$f_1, f_2, \cdots f_n, \cdots$$

which converges in the mean to the function $\psi(x)$, in R, and in which each f_n is of class n. By the preceeding results $F[f_n]$ is defined for all n, is continuous in f_n, and satisfies (III) for all f_n. Then exactly as before in considering the sequence:

$$F[f_1] + (F[f_2] - F[f_1]) + \cdots$$

we ascertain that $\lim_{n \to \infty} F[f_n] = F[\psi]$ is unique and finite. Thus $F[f]$ is defined for functions of the class ω, the first transfinite number, and as before it follows that $F[f(x)]$ satisfies (III) on these functions also.

In summing up it is clear that the definition of may be extended to all the functions of Baire by the method of transfinite induction.

For suppose we assume $F[f(x)]$ defined for all functions of class $< \alpha$, contained in the region R, satisfying (III) and consequently continuous in such functions. Then given any function, $\psi(x)$ in R, of class α and consequently the limit of a sequence of functions

$$f_1, f_2, \cdots f_n, \cdots$$

of classes $< \alpha$, we have, as before,

$$\lim_{n \to \infty} F[f_n] = F[\psi]$$

is defined and unique. Also

$$|F[\psi] - F[\varphi]| \leqq M \int_a^b |\psi - \varphi|\, dx$$

by taking the limit in

$$|F[f_n] - F[g_n]| \leqq M \int_a^b |f_n - g_n|\, dx$$

where $\psi(x)$, and $\varphi(x)$ are two functions of class α, limits respectively of $\{f_n\}$, $\{g_n\}$ of class $< \alpha$, and finally $F[\psi]$ is continuous in $\psi(x)$ of class α.

Since we have already shown that $F[\psi(x)]$ is defined, for functions of class I, satisfies (III) for such functions, and is consequently continuous, by the principle of transfinite recurrence these properties exist for functions of all classes.

It has been shown that by replacing postulate (I) by (III), the result 3) of Professor Daniell is still obtained. It will now be demonstrated that by means of this assumption, the variation takes the form

$$3) \qquad D(f;\varphi) = \int_a^b \varphi(\xi)\,\beta(\xi)\,d\xi$$

THEOREM:- $\alpha_f(\lambda)$ is an absolutely continuous function of λ in the interval (a,b).

For suppose $\alpha_f(\lambda)$ is not absolutely continuous, then given $\mu > 0$, there is a sequence of sets of non-overlapping intervals

$$\left\{ \left(a_1^{(m)}, b_1^{(m)}\right), \left(a_2^{(m)}, b_2^{(m)}\right), \ldots \left(a_{j_m}^{(m)}, b_{j_m}^{(m)}\right)\right\}$$

such that

$$14) \qquad \left| \sum_1^{j_m} \left\{ \alpha\left(b_i^{(m)}\right) - \alpha\left(a_i^{(m)}\right)\right| > \mu$$

where

$$\lim_{m\to\infty} \sum_1^{j_m} \left(b_i^{(m)} - a_i^{(m)}\right) = 0$$

Consider the function

$$\begin{aligned} {}_{a'}C_{b'} &= C, \text{ a positive constant, } (a' \le \lambda \le b') \\ &= 0, \qquad (a \le \lambda < a',\ b' < \lambda \le \end{aligned}$$

M. Riesz has shown that a linear functional may be extended to functions of this type(1). Consequently we have

$$\lim_{m\to\infty} D\left(f;\ \sum_1^{i_m} {}_mC_{a_i^{(m)}b_i^{(m)}}\right) = T\left[\sum_1^{i_m} {}_{a_i^{(m)}}C_{b_i^{(m)}}\right] = \int_a^b \sum_1^{i_m} {}_{a_i^{(m)}}C_{b_i^{(m)}}\, d\,g(x)$$

(where $T[{}_{a'}C_{b'}]$ is a linear functional of ${}_{a'}C_{b'}$ in $\mathcal{C}_{a',b'}$ defined in 16))

$$= C \sum_1^{i_m} \left\{ \alpha(b_i^{(m)}) - \alpha(a_i^{(m)}) \right\}$$

and

15) $$\left| T\left[\sum_1^{i_m} {}_{a_i^{(m)}}C_{b_i^{(m)}}\right] \right| \geqq C\nu$$

for $m > m_0$ large enough, by 14)

16.) Consider the functions

$$\begin{aligned} {}_nC_{a'b'} &= C &&, \quad (a' \leqq x \leqq b') \\ &= 0 &&, \quad (a \leqq x \leqq a';\ b' \leqq x \leqq b) \\ &= \text{linear} &&, \quad (a' - \tfrac{1}{n} \leqq x \leqq a';\ b' \leqq x \leqq b' + \tfrac{1}{n}) \end{aligned}$$

letting

$$\Delta F = F\left[f + \epsilon \sum_1^{i_m} {}_mC_{a_i^{(m)}b_i^{(m)}}\right] - F[f]$$

$$I_m = \int_a^b \left| \sum_1^{i_m} {}_mC_{a_i^{(m)}b_i^{(m)}} \right| dx$$

we have

$$|\Delta F| \leq \epsilon M I_m \quad \text{or} \quad \left|\frac{\Delta F}{\epsilon}\right| \leq M I_m$$

---

(1) Loc.cit., page 11 et seq.

and letting $\epsilon \to 0$

$$\left| D\left(f;\ \sum_{1}^{k_m} n\, C_{a_i^{(m)} b_i^{(m)}}\right) \right| \leqq M\, I_n$$

and as $n \to \infty$ we obtain

$$\left| T\left[\sum_{1}^{k_m} \alpha_i^{(m)} C_{a_i^{(m)} b_i^{(m)}} \right] \right| \leqq M \int_a^b \left| \sum_{1}^{k_m} C_{a_i b_i^{(m)}} \right| dx$$

$$\leqq M\, C \sum_{1}^{k_m} \left(b_i^{(m)} - a_i^{(m)} \right)$$

which is argitrarily small, and therefore by taking m large enough, can be mdde

$$\leqq \rho$$

This gives a contradiction with 15) and $\alpha(x)$ must be absolutely continuous.

Finally we have(1)

$$D(f;\varphi) = \int_a^b \varphi(\xi)\, \beta(\xi)\, d\xi$$

for if $\alpha(x)$ is absolutely continuous, it is equal to $\int_a^x \alpha'(x)\, dx + C$ and the Stieltjes integral $\int_a^b \varphi(\xi)\, d\alpha(\xi)$ reduces to the Lesbesque integral, $\int_a^b \varphi(\xi)\, \alpha'(\xi)\, d\xi$

Hence if $\beta(x)$ denotes $\alpha'(x)$ where it exists and is, say, 0 otherwise, then we have the above result.

G. C. EVANS —

(1) Functionals and their Applications, #39, page 59. Am.Math.Soc. Colloquium Lectures, Vol V, The Cambridge Colloquium 1916, Part I.