INFINITE SYSTEMS OF FUNCTIONS.

BY PROFESSOR W. E. MILNE.

THE study of infinite systems of functions is approached in this paper from an elementary point of view, and by easy steps there are derived results of considerable generality. It is found that every enumerable system of real functions whose squares are integrable in the sense of Lebesgue either is orthogonal, or possesses an adjoint, or is essentially linearly dependent. Corresponding to every normalized system of functions $\varphi_1, \varphi_2, \varphi_3, \cdots$, is a set of constants $\lambda_1, \lambda_2, \lambda_3, \cdots$, where $1 \geqq \lambda_i \geqq 0$ for every i, with the properties:

A necessary and sufficient condition that the system

(*a*) be orthogonal is that $\lambda_i = 1$ for every i,

(*b*) possess an adjoint is that $\lambda_i > 0$ for every i,

(*c*) be essentially linearly dependent is that $\lambda_i = 0$ for some i.

§ 1.

For simplicity the discussion is limited to real functions of a single real variable.* The symbol Ω denotes the class of all such functions whose squares are integrable in the sense of Lebesgue in the interval (a, b). Functions of Ω, as well as sums and products of such functions are integrable in (a, b). The word "function" in this paper will always mean a function of class Ω, and all properties stated of a system of class Ω will be understood to hold throughout the interval (a, b). It is assumed that the reader is familiar with the terms norm of a function, normalized system, orthogonal system, biorthogonal systems, complete systems, essential linear dependence (of a finite number of functions), convergence in the mean, etc.†

§ 2.

Let there be given an enumerable system $[\varphi]$ of normalized functions $\varphi_1, \varphi_2, \varphi_3, \cdots$, of class Ω. First we investigate the

* The reader will see that the methods may be extended to more general systems.

† Definitions of all these terms are given by Brand "On infinite systems of linear integral equations," *Annals of Math.*, vol. 14 (1913), p. 101.

dependence of a finite set of these functions $\varphi_1, \varphi_2, \cdots, \varphi_n$, denoted by $[\varphi]^{(n)}$, and look not merely for a test to distinguish between essential linear dependence and independence, but for a measure of the independence, or of the nearness to dependence. To formulate this precisely we assume that $\varphi_1, \varphi_2, \cdots, \varphi_n$ are independent and ask how nearly can any given function φ_i of the set be expressed linearly in terms of the remaining functions of $[\varphi]^n$, where the nearness is measured by the integral of the square of the remainder. Then write

$$\xi_i^{(n)} = \varphi_i - (c_1\varphi_1 + \cdots + c_n\varphi_n), \tag{1}$$

in which $c_i = 0$ and the remaining c's are chosen to minimize the norm

$$\int_a^b [\xi_i^{(n)}]^2 dx = \lambda_i^{(n)}. \tag{2}$$

When the c's are determined in the usual manner* and substituted into (1) the result may be put in the form†

$$\xi_i^{(n)} = \frac{1}{G_i^{(n)}} \begin{vmatrix} 1 & \int\varphi_1\varphi_2 & \cdots & \int\varphi_1\varphi_n \\ \int\varphi_1\varphi_2 & 1 & \cdots & \int\varphi_2\varphi_n \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_1 & \varphi_2 & \cdots & \varphi_n \\ \cdots & \cdots & \cdots & \cdots \\ \int\varphi_1\varphi_n & \int\varphi_2\varphi_n & \cdots & 1 \end{vmatrix}, \tag{3}$$

in which $G_i^{(n)}$ denotes the gramian of $[\varphi]^{(n)}$ with φ_i omitted, and in which the non-integrated terms $\varphi_1, \varphi_2, \cdots, \varphi_n$ occupy the ith row of the determinant.

This set of functions $\xi_i^{(n)}$ has a number of interesting properties. For from equation (3) it is readily seen that

$$\int_a^b \varphi_i\xi_j^{(n)} dx = 0 \tag{4}$$

if i is not equal to j, while

$$\int_a^b \varphi_i\xi_i^{(n)} dx = \frac{G^{(n)}}{G_i^{(n)}}, \tag{5}$$

* See, e.g., Gram, "Uber die Entwickelung reeller Funktionen in Reihen mittelst der Methode der kleinsten Quadrate," *Crelle*, vol. 94 (1883), pp. 41–73. Byerly, "Approximate representation," *Annals of Math.*, vol. 12 (1911), pp. 128–148. Bôcher, "Introduction to the theory of Fourier's series," *Annals of Math.*, vol. 7 (1905), p. 81. Brand, loc. cit.

† As usual, $\int\varphi_i\varphi_j$ stands for $\int_a^b \varphi_i\varphi_j dx$.

in which $G^{(n)}$ denotes the gramian of the system $[\varphi]^n$. From (2), (3), and (5) we find

$$\lambda_i^{(n)} = \int_a^b [\xi_i^{(n)}]^2 dx = \frac{G^{(n)}}{G_i^{(n)}}. \tag{6}$$

From (5) and (6) follows

$$\int_a^b \varphi_i \xi_i^{(n)} dx = \lambda_i^{(n)}. \tag{7}$$

From (3), (4), and (7) we also get

$$\int_a^b \xi_i^{(m)} \xi_i^{(n)} dx = \lambda_i^{(n)} \qquad (m < n). \tag{8}$$

From (6) and (8) we have finally

$$\int_a^b [\xi_i^{(m)} - \xi_i^{(n)}]^2 dx = \lambda_i^{(m)} - \lambda_i^{(n)} \qquad (m < n). \tag{9}$$

§ 3.

So far it has been assumed that the functions of $[\varphi]^{(n)}$ are essentially linearly independent. If that is not the case, the right hand sides of equations (3), (5), and (6) may become indeterminate through the vanishing of both numerator and denominator. But the $\xi_i^{(n)}$ and $\lambda_i^{(n)}$ are still determinate. For if the function φ_m $(m < n)$ is expressible linearly in terms of the functions of $[\varphi]^{(n)}$ with φ_m and φ_i omitted, the value of $\lambda_i^{(n)}$ is obviously unchanged if φ_m be dropped from the set. Therefore in forming equations (3) and (5) we shall omit every function φ_m which is expressible linearly in terms of the functions (excepting φ_i) which precede φ_m in the system $[\varphi]$. Then equations (3) to (9) hold without exception. It is evident that a necessary and sufficient condition for essential linear dependence of $[\varphi]^{(n)}$ is $\lambda_i^{(n)} = 0$ for some i.

§ 4.

Consider now the case of an infinite system of functions. From (9) it is apparent that

$$\lambda_i^{(m)} \geqq \lambda_i^{(n)} \quad \text{if } m < n. \tag{10}$$

This relation, together with the fact that the λ's are never negative, insures the convergence of $\lambda_i^{(n)}$ for each i as n

becomes infinite,

(11) $$\lim_{n \doteq \infty} \lambda_i^{(n)} = \lambda_i \qquad (i = 1, 2, \cdots),$$

where

$$1 \geqq \lambda_i \geqq 0 \qquad (i = 1, 2, \cdots).$$

Recall now the following theorem of Fischer. A necessary and sufficient condition that a sequence of functions $\theta_1, \theta_2, \theta_3, \cdots$ of class Ω converge in the mean to a function of this class is that to every positive ϵ however small there correspond a number N such that $\int_a^b [\theta_m - \theta_n]^2 dx < \epsilon$ whenever m and $n > N$.* Since $\lambda_i^{(n)}$ converges we see from (9) that the condition for mean convergence is satisfied for the sequence of functions $\xi_i^{(n)}$ when i is held fast and n becomes infinite. Therefore $\xi_i^{(n)}$ converges in the mean to a function ξ_i of class Ω. This system of functions $[\xi_i]$ has the property that

(12) $$\int_a^b \varphi_i \xi_j dx = \begin{cases} \lambda_i & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

For, by Schwarz's inequality,

$$\left[\int_a^b \varphi_i(\xi_j^{(n)} - \xi_j)dx\right]^2 \leqq \int_a^b (\xi_j^{(n)} - \xi_j)^2 dx,$$

since φ_i is normalized. Since the right-hand side approaches zero as n becomes infinite we have

$$\int_a^b \varphi_i \xi_j dx = \lim_{n=\infty} \int_a^b \varphi_i \xi_j^{(n)} dx,$$

from which, by (4), (7), and (11), we obtain (12).

§ 5.

Suppose that the system $[\varphi]$ is orthogonal. Then every gramian is unity, hence by (6) every $\lambda_i^{(n)}$ is unity, and therefore every λ_i is unity. Conversely, if every λ_i is unity, it follows from (10) that every $\lambda_i^{(n)}$ is unity, and in particular $\lambda_1^{(2)} = 1$. But this can be true only if

$$\begin{vmatrix} 1 & \int\varphi_1\varphi_2 \\ \int\varphi_1\varphi_2 & 1 \end{vmatrix} = 1,$$

* Fischer, "Sur la convergence en moyenne," *Comptes Rendus*, May 1907, p. 1022.

from which we find that

$$\int_a^b \varphi_1\varphi_2 dx = 0.$$

Now any two functions of the system may be placed first, and therefore the system must be orthogonal. Therefore we have

THEOREM I. *A necessary and sufficient condition that a normalized system* $[\varphi]$ *be orthogonal is that* $\lambda_i = 1$ *for every* i.

Suppose next that $[\varphi]$ has an adjoint system $[\psi]$. For sake of symmetry we shall assume that $[\psi]$ is normalized, which requires a slight change in the usual definition of an adjoint system, for we simply require that the integral $\int_a^b \varphi_i\psi_i dx$ be greater than zero instead of requiring it to be unity. Now multiply equation (3) by ψ_i and integrate. The result is

$$\int_a^b \psi_i\xi_i^{(n)} dx = \int_a^b \varphi_i\psi_i dx.$$

Hence, by Schwarz's inequality,

$$\left[\int_a^b \varphi_i\psi_i dx\right]^2 \leqq \left[\int_a^b \psi_i^2 dx\right]\left[\int_a^b \xi_i^{(n)2} dx\right],$$

from which it follows that

$$\left[\int_a^b \varphi_i\psi_i dx\right]^2 \leqq \lambda_i^{(n)}$$

for every value of i. Therefore $\lambda_i > 0$ for every i. Conversely if $\lambda_i > 0$ for every i the system $[\psi]$ defined by the equations

$$\psi_i = \xi_i/\sqrt{\lambda_i} \qquad (i = 1, 2, \cdots), \tag{13}$$

is a normalized system adjoint to $[\varphi]$, as we see from (2) and (12). This proves

THEOREM II. *A necessary and sufficient condition that a normalized system* $[\varphi]$ *have an adjoint is that* $\lambda_i > 0$ *for every* i.

It is of interest to construct the functions ζ_i and the numbers μ_i which bear to the system $[\psi]$ defined in (13) the same relations respectively that ξ_i and λ_i bear to $[\varphi]$. It is found upon carrying through the computations that

$$\zeta_i = \varphi_i \sqrt{\lambda_i}$$

and

$$\mu_i = \lambda_i,$$

as might have been expected from considerations of symmetry.

Another interesting property of the adjoint defined above is that of all normalized systems $[\psi]$ adjoint to $[\varphi]$ none gives greater value to the integrals $\int_a^b \varphi_i\psi_i dx$ than does the system defined by (13). For any adjoint whatever gives

$$\left[\int_a^b \varphi_i\psi_i dx\right]^2 \leqq \lambda_i$$

but the adjoint defined by (13) requires the equality sign only.

Definition: A function θ is said to be *expressible linearly* in terms of a system $[\varphi]$ if there exists a sequence of linear combinations of φ's converging in the mean to θ. It is evident that the functions ψ_i defined by (13) are all expressible linearly in terms of $[\varphi]$.

Consider a system $\varphi_0, \varphi_1, \varphi_2, \cdots$, where φ_0 is any function of Ω and the system $\varphi_1, \varphi_2, \varphi_3, \cdots$ is complete. Form the function ξ_0. By (12) ξ_0 will be orthogonal to $\varphi_1, \varphi_2, \varphi_3, \cdots$, but as this system is complete ξ_0 must be essentially zero. From the definition of $\xi_0^{(n)}$ in (1) it follows that φ_0 must be expressible linearly in terms of $\varphi_1, \varphi_2, \varphi_3, \cdots$. This gives a theorem almost identical with one proved by Brand.*

THEOREM III. *If a system* $[\varphi]$ *is complete, every function of class* Ω *is expressible linearly in terms of* $[\varphi]$.

Definition: If any function φ_i of a system $[\varphi]$ is expressible linearly in terms of the remaining functions of the system, then $[\varphi]$ is said to be *essentially linearly dependent.*

If any λ_i is zero it is apparent that the system is essentially linearly dependent. On the other hand if the system is dependent, and φ_i is therefore expressible in terms of the remaining functions, the number λ_i must be zero. For if it were equal to $\epsilon > 0$, we could select a linear combination Σ_i, for which

$$\int_a^b [\varphi_i - \Sigma_i]^2 dx < \epsilon.$$

* Loc. cit., Theorem 12.

Then we may take n so great that $\xi_i^{(n)}$ contains every function in Σ_i, and then, since $\lambda_i^{(n)}$ is the least norm, we must have $\lambda_i^{(n)} < \epsilon$, and consequently $\lambda_i < \epsilon$. Hence

THEOREM IV. *A necessary and sufficient condition that a normalized system* $[\varphi]$ *be essentially linearly dependent is that* $\lambda_i = 0$ *for some* i.

Theorems II and IV give

THEOREM V. *A necessary and sufficient condition that a system* $[\varphi]$ *have an adjoint is that it be essentially linearly independent.*

THE UNIVERSITY OF OREGON,
November, 1919.

ON CERTAIN RELATED FUNCTIONAL EQUATIONS.

BY DR. W. HAROLD WILSON.

(Read before the American Mathematical Society December 27, 1917.)

§ 1. *Introduction.*

THIS paper treats of the relationships which exist between certain functional equations. In § 2, the equations

$$(1) \qquad S(x-y) = S(x)C(y) - C(x)S(y),$$

and

$$(2) \qquad C(x-y) = C(x)C(y) - k^2S(x)S(y)$$

are considered individually and as a system. It is shown that (1) and (2) have their solutions in common if $C(x)$ is an even function and $S(x) \not\equiv 0$. As a consequence, it is shown that if $k \neq 0$, then

$$S(x) = [F(x) - F(-x)]/2k, \text{ and } C(x) = [F(x) + F(-x)]/2,$$

where $F(x+y) = F(x)F(y)$. If $k = 0$ and $S(x) \not\equiv 0$, $C(x) \equiv 1$ and

$$S(x+y) = S(x) + S(y).$$

The work at this point is very closely allied to that of