A Set of Continuous Orthogonal Functions.

Von

Philip Franklin in Cambridge (Mass.) (U. S. A.).

1. Introduction.

There exist continuous functions for which, at some points of the interval of orthogonality the classical Fourier series fails to converge. The analogous expansions in orthogonal functions arising from the simpler boundary value problems seem to share this property with the Fourier expansion[1]). This led A. Haar to ask if the property was common to all sets of orthogonal functions. He showed that it was not by exhibiting a set of orthogonal functions giving, as the expansion of any continuous function, a series converging uniformly to the function throughout the fundamental interval. The individual functions of his set, however, are *discontinuous*, so that his example does not exclude the possibility of the property being common to all sets of *continuous* orthogonal functions. In this paper we construct a set of continuous orthogonal functions similar to Haar's set in that the expansion of any continuous function represents the function everywhere.

2. Definition of the functions.

Consider the set of functions defined for $0 \leqq x \leqq 1$ by

$$
(1) \qquad
\begin{array}{llll}
v_0 = 1, & v_1 = x, & & \\
v_2 = 0, & x \leqq \frac{1}{2} \quad \text{and} & v_2 = x - \frac{1}{2}, & x \geqq \frac{1}{2}, \\
\cdots & \cdots & \cdots & \cdots \\
v_n = 0, & x \leqq a_n \quad \text{and} & v_n = x - a_n, & x \geqq a_n
\end{array}
$$

where $a_n = (2n - 1 - 2^k)/2^k$, k integral and such that the highest power of 2 contained in $2n - 1$ is 2^k.

[1]) A. Haar, Math. Annalen 69 (1910), pp. 331—371.

Thus a_n is the n^{th} term of the series

$$0, \tfrac{1}{2}, \tfrac{1}{4}, \tfrac{3}{4}, \tfrac{1}{8}, \tfrac{3}{8}, \tfrac{5}{8}, \tfrac{7}{8}, \tfrac{1}{16}, \ldots$$

and all the a_n are distinct.

Each of these functions v_i is continuous for $0 \leqq x \leqq 1$, which we take as our fundamental interval. Furthermore, these functions are linearly independent, since the first two obviously are so, and if any v_n $(n \geqq 2)$ were linearly dependent on a finite number of functions of the set not including v_n, it could not have a discontinuity in its derivative at a_n. Thus we may apply the process of orthogonalization[2] to this set and so obtain the normal and orthogonal set f_n, where

$$f_0 = 1,$$

$$f_{n+1} = \left[v_{n+1} - \sum_{h=1}^{n} f_h \int_0^1 f_h(t)\, v_{n+1}(t)\, dt\right] \Big/ \pm \sqrt{\int_0^1 v_{n+1}(t)^2\, dt}. \tag{2}$$

We find, for example

$$\begin{aligned}
&f_0 = 1, \qquad f_1 = \sqrt{3}(1-2x),\\
&f_2 = \sqrt{3}(1-4x), \qquad x \leqq \tfrac{1}{2} \text{ and } \sqrt{3}(4x-3), \qquad x \geqq \tfrac{1}{2},\\
(3)\quad &f_3 = \sqrt{3}(10-76x)/19, \qquad x \leqq \tfrac{1}{4},\\
&\phantom{f_3 = {}}\sqrt{3}(52x-22)/19, \qquad \tfrac{1}{4} \leqq x \leqq \tfrac{1}{2},\\
&\phantom{f_3 = {}}\sqrt{3}(10-12x)/19, \qquad x \geqq \tfrac{1}{2}.
\end{aligned}$$

The sign of these functions is not determined by (2), we take that sign before the radical which makes the first sign change of each f_n one from plus to minus. This defines in a unique way a set of continuous functions normal and orthogonal on the unit interval.

3. Convergence Properties.

The series expansion of a given function $F(x)$ in terms of the functions f_n is

$$\sum_{i=0}^{\infty} A_i f_i(x), \quad \text{where} \quad A_i = \int_0^1 F(t)\, f_i(t)\, dt. \tag{4}$$

As we shall derive the convergence properties from approximation considerations, we recall that, for any linear combination of the first $(n+1)\, f_i$, $T_n = \sum_{i=0}^{n} C_i f_i$, we have:

[2] cf. e. g. Courant-Hilbert, Methoden der math. Physik, Berlin 1924, p. 34.

$$(5)\qquad E(T_n)=\int_0^1 (F(x)-T_n(x))^2\,dx$$
$$=\int_0^1 F(x)^2\,dx-\sum_{i=0}^{n} A_i^2+\sum_{i=0}^{n}(C_i-A_i)^2.$$

This shows that, for n fixed, the mean square error $E(T_n)$ is least when $C_i=A_i$, and T_n is a partial sum of the series (4), $S_n=\sum_{i=1}^{n} A_i f_i$, and also that, as n increases, $E(S_n)$ decreases.

We shall now show that, if $F(x)$ is continuous, for any positive ε there exists an $N(\varepsilon)$ such that, if $n>N$, $E(S_n)<\varepsilon$. We first define a broken line function of type n, B_n, to be a continuous function, linear in each of the intervals bonded by consecutive points of the set (b_j),

$$(6)\qquad 0,\frac{1}{2^k},\frac{2}{2^k},\ \ldots\ a_n,\ a_n+\frac{1}{2^k},\ a_n+\frac{3}{2^k},\ \ldots\ 1.$$

These (except for 1) are simply the a_i, $i\leqq n$, arranged in order of magnitude. Every function B_n is clearly a linear combination of the first $(n+1)\,v_i$, for we may reduce it to zero by subtracting the constant times v_0 which makes it zero at $x=0$, then the constant times v_1 which coincides with the remainder in the first interval, then the constant times v_i $\left(i=2^{k-1}+1\right.$ so that $\left.a_i=\frac{1}{2^k}\right)$ which coincides with the new remainder in the first two intervals, and so on. As it is evident from (2) that each v_i is linearly dependent on the first $(i+1)f_j$, we see that each B_n is a linear combination of the first $(n+1)f_j$, i. e. it is a T_n.

From the continuity of $F(x)$ in the closed interval $0\leqq x\leqq 1$, we infer the existence of a $\delta(\varepsilon)$ such that $|F(x_1)-F(x_2)|<\varepsilon$, whenever $|x_1-x_2|<\delta$. If K is an integer for which

$$(7)\qquad \frac{1}{2^K}<\delta(\varepsilon),\qquad N=2^K$$

may be taken as $N(\varepsilon)$. For, as all the points $m/2^K$ $(m=1,2,3\ldots 2^K-1)$ are points a_i with $i\leqq N$, the broken line function which agrees with $F(x)$ at 0, 1 and these points, and is linear in the intervals determined by them is a B_n and hence a T_n for any $n>N$. But, for this function, from (7) we have

$$(8)\qquad |F(x)-T_n(x)|<\varepsilon,$$

so that $E(T_n)$ and hence $E(S_n)<\varepsilon$.

We shall next prove that when n exceeds the N of (7), $S_n(x)$ uniformly approximates $F(x)$. A characteristic property of $S_n(x)$ is the fact that it is the T_n which minimizes $E(T_n)$. To apply this property, we need a lower limit for $E_{12}(T_n)$, the contribution to $E(T_n)$ from an interval $x_1 \leqq x \leqq x_2$, in which T_n is a linear function of x. We assume that

$$(9) \qquad x_2 - x_1 = D < \delta(\varepsilon).$$

Consequently, if $L(x)$ is the linear function for which

$$(10) \qquad L(x_1) = F(x_1) \quad \text{and} \quad L(x_2) = F(x_2),$$

we will have throughout the interval $x_1 x_2$

$$(11) \qquad |F(x) - L(x)| < \varepsilon \qquad (x_1 \leqq x \leqq x_2).$$

It follows from this that, in this interval

$$(12) \qquad |F(x) - T_n(x)| > |L(x) - T_n(x)| - \varepsilon.$$

Recalling that $L(x)$ und $T_n(x)$ are both linear in the interval $x_1 x_2$ and putting

$$(13) \qquad H_1 = |F(x_1) - T_n(x_1)|, \qquad H_2 = |F(x_2) - T_n(x_2)|,$$

we find for the right member of (12):

$$(14) \quad |L(x) - T_n(x)| - \varepsilon = \frac{1}{D}|(H_2 \pm H_1)x - H_2 x_1 \mp H_1 x_2| - \varepsilon,$$

the upper and lower signs corresponding to the cases in which $T_n(x_1)$ and $T_n(x_2)$ are on opposite sides of, or on the same side of $L(x)$. If we take the upper signs, and integrate the square of this expression over the parts of the interval $x_1 x_2$ where the expression is positive, we find

$$(15) \qquad E_{12}(T_n) > \frac{D[(H_1-\varepsilon)^3 + (H_2-\varepsilon)^3]}{3(H_1+H_2)},$$

when H_1 and H_2 are both $\geqq \varepsilon$. If either of these is $< \varepsilon$ the corresponding parenthesis does not appear in (15), but as it is negative, it may be inserted without destroying the inequality. If at least one of the H's is $> 4\varepsilon$, we shall have

$$(16) \qquad \frac{1}{H_1 + H_2} > \frac{1}{2[(H_1-\varepsilon)+(H_2-\varepsilon)]},$$

and in view of this, (15) becomes:

$$(17) \quad E_{12}(T_n) > \frac{D}{12}[(H_1 - \varepsilon)^2 + (H_2 - \varepsilon)^2], \qquad \text{Max}(H_1, H_2) > 4\varepsilon.$$

When the lower signs in (14) are used, we have in place of (15):

$$(18) \qquad E_{12}(T_n) > \frac{D[(H_1-\varepsilon)^3 - (H_2-\varepsilon)^3 - \varepsilon^2(H_1 - H_2)]}{3(H_1 - H_2)}.$$

When H_1 and H_2 are both $\geqq \varepsilon$, the last term in the numerator does not appear. As after division by the denominator it is negative, its presence merely strengthens the inequality. We have inserted it so that (18) will apply to the case in which one of the H's, say H_1 is $< \varepsilon$. Here the first term does not appear, and may not be supplied by itself since it is a positive multiple of the denominator. However, in view of our second condition, when $H_1 < \varepsilon$, $H_2 > 4\varepsilon$, and

$$\frac{(H_1-\varepsilon)^3}{H_1-H_2} - \varepsilon^2 < 0, \tag{19}$$

so that these two terms may be inserted together. As it is easily seen that (18) implies (17), this last inequality gives a lower limit for $E_{12}(T_n)$ which holds in all cases.

Let us now consider S_n, which is linear in the intervals bounded by the b_j of (6). We put

$$H_j = |F(b_j) - S_n(b_j)|. \tag{20}$$

Let M be the greatest of the quantities H_j, and b_g one of the points at which $H_g = M$. Suppose that $M > 4\varepsilon$, and let b_p be the first point to the left of b_g at which $H_p \leqq 4\varepsilon$ (or 0, if no such point exists) and b_q (or 1) the first such point to the right. We will compare S_n with a particular T_n, obtained from it by the following process. For values of x outside the interval $b_p \leqq x \leqq b_q$, $T_n(x)$ coincides with $S_n(x)$. Inside this interval, we put

$$\begin{aligned} (21)\quad & T_n(b_p) = S_n(b_p) \;\; (\text{or } F(0)), \quad T_n(b_q) = S_n(b_q) \;\; (\text{or } F(1)), \\ & T_n(b_i) = F(b_i) \qquad (p < i < q). \end{aligned}$$

In view of the linear character of T_n in the intervals between the b_j, it is defined by these values.

To compare the mean square errors $E(S_n)$ and $E(T_n)$, we need merely compare the contributions from the interval $b_p b_q$. Applying (17), we find

$$E_{pq}(S_n) > \frac{3}{2}(b_q - b_p)\varepsilon^2 + \frac{D}{12}(M-\varepsilon)^2 - 4D\varepsilon^2. \tag{22}$$

Here $D = \frac{1}{2^k}$, so that the elementary b_j intervals are at least D, and at most $2D$ in width. For T_n, we obviously have

$$E_{pq}(T_n) < (b_q - b_p)\varepsilon^2 + 100 D\varepsilon^2, \tag{23}$$

since the numerical error is at most ε, cf. (8), except in the end intervals, where it is at most 5ε. From (22) and (23) we find:

(24) $E_{pq}(S_n) - E_{pq}(T_n) > \frac{D}{12}(M - \varepsilon)^2 - 104\,D\,\varepsilon^2 + \frac{1}{2}(b_q - b_p)\,\varepsilon^2,$

and since

(25) $$E_{pq}(S_n) - E_{pq}(T_n) = E(S_n) - E(T_n) < 0,$$

we must have[3])

(26) $$M < 37\,\varepsilon.$$

Thus all the $H_j < 37\,\varepsilon$, and in consequence

(27) $$|F(x) - S_n(x)| < 38\,\varepsilon \qquad (n > N(\varepsilon)),$$

which proves our contention about uniform approximation. We may express this result as

Theorem I. *If any function $F(x)$, continuous in the interval $0 \leqq x \leqq 1$, be expanded in a series of constant multiples of the functions f_j* (3), *the resulting series* (4) *will converge to $F(x)$ at all points of the unit interval, and, in fact, uniformly.*

Let us next treat the case in which $F(x)$ is a measurable function with summable square in the unit interval. Such a function can be approximated to an arbitrary degree of exactness, in the sense of least squares, by a continuous function[4]), $C(x)$, so that

(28) $$\int_0^1 [F(x) - C(x)]^2\,dx < \eta^2.$$

As we have shown above, (8), that there exists a T_n for which

(29) $$\int_0^1 [C(x) - T_n(x)]^2\,dx < \eta^2,$$

for this function $E(T_n) < 4\,\eta^2$. This implies the same relation for $E(S_n)$, and shows that for the functions now treated the expansion will converge in the mean.

If, at any point of the unit interval, $F(x)$ is continuous, we may show that the expansion converges there in the ordinary sense. Let, then, x_c be a point of continuity of $F(x)$, and $x_1\,x_2$ an interval such that

(30) $$|F(x) - F(x_c)| < \varepsilon, \quad x_1 \leqq x \leqq x_2; \quad x_2 - x = x - x_1.$$

Let us take n so large that $\frac{1}{2^k}$ is small compared with $x_2 - x_1$, and

---

[3]) By a longer calculation, which treats the cases more in detail, we may show that, in fact, $M < 17\,\varepsilon$.

[4]) cf. e. g. E. W. Hobson; Theory of functions of a Real Variable, 2nd ed., Cambridge 1921, vol. 1, p. 584.

for S_n consider the H_j for the b_j of the interval $x_1 x_2$ here defined by

$$H_j = | F(x_c) - S_n(b_j) | \qquad (x_1 \leqq b_j \leqq x_2). \tag{31}$$

If more than $1/P$ of these H_j were $\geqq 4\varepsilon$, from (17) we would have

$$E(S_n) \geqq E_{12}(S_n) > (x_2 - x_1)\varepsilon^2/P. \tag{32}$$

As $E(S_n)$ approaches zero as n increases, we may take n so large that $1/P < \frac{1}{4}$. When this is done, $\frac{3}{4}$ of the $H_j < 4\varepsilon$, and in particular b_j for which this holds exist on both sides of x_c in the interval $x_1 x_2$. Calling b_p the first such point to the left of x_c, and b_q the first one to the right, we may use the argument given above to establish (27), to show that each of the H_j for b_j adjacent to x_c (or for x_c itself if that is a b_j) $< 37\varepsilon$, and hence

$$| F(x_c) - S_n(x_c) | < 38\varepsilon \qquad (n > N_1(\varepsilon)). \tag{33}$$

This proves the convergence of S_n to $F(x)$ at any point of continuity, and an obvious modification of the argument shows that the convergence is uniform for any closed interval in which $F(x)$ is continuous. We have thus established:

Theorem II. *If any measurable function $F(x)$, with summable square in the interval $0 \leqq x \leqq 1$, be expanded in a series of constant multiples of the functions f_j (3), the resulting series (4) will converge to $F(x)$ at all points of continuity. Further, in any closed interval of continuity the convergence will be uniform. Over the whole interval, the series converges in the mean to $F(x)$.*

The consideration of simple examples shows that if the function $F(x)$ is continuous in each of the intervals $x_1 x_2$ and $x_2 x_3$, but discontinuous at x_2, the series will in general oscillate between two finite values at x_2.

4. Other Functions.

We have based the set of function used, f_i (3), on the functions v_i (1) with breaks at a_n, the proper fractions with denominators integral powers of 2. Similarly, we could obtain an orthogonal set of broken line functions from any other set of points p_n. For such a set to have the convergence properties of theorems I and II, it is sufficient that the points p_n be everywhere dense on the unit interval and enumerated in such a way that, of the intervals marked off at any stage by these points, the ratio of the greatest interval to the least remains uniformly bounded for the entire set. With this restriction on the p_n, the theorems can be proved essentially as above.

5. Application to the Haar set.

The set of discontinuous functions used by Haar may be obtained by applying the process of orthogonalization (2) to the derivatives of our functions v_i (1). As the linear combinations of them are step functions, the deductions made from (5) show that the set is complete. As, in a step function, any one step may be altered without disturbing the rest of the function, it is obvious from the minimum property that in any interval between adjacent points of discontinuity b_j, $S_n(x)$ must lie between the greatest and least values of $F(x)$ in the interval. From this remark most of the convergence theorems given by Haar may be deduced immediately without further calculation.

Massachusetts Institute of Technology.

(Eingegangen am 20. 12. 1927.)