ON FUNCTIONS OF CLOSEST APPROXIMATION*

BY

DUNHAM JACKSON

1. **Introduction.** The determination of the polynomial of specified degree which gives the best approximation to a given continuous function $f(x)$ in a given interval (a, b) depends on the meaning attached to the phrase "best approximation." The polynomial for which the maximum of the absolute value of the error is as small as possible is known as the Tchebychef polynomial corresponding to $f(x)$, and has been extensively studied.† The polynomial which reduces the integral of the square of the error to a minimum is obtained by taking the sum of the first terms in the development of $f(x)$ in Legendre's series,‡ and its properties are of course also well known.

The following pages are devoted to a study of the polynomial for which the integral of the mth power of the error is a minimum, where m is any even positive integer, or, more generally, the integral of the mth power of the absolute value of the error, where m is any real number greater than 1. It is found that some of the familiar properties of the approximating function in the case $m = 2$ are carried over to the other values of m. It is shown further, and this is the principal conclusion of the paper, that the polynomial of approximation corresponding to the exponent m approaches the Tchebychef polynomial as a limit when m becomes infinite. The discussion is put in such a form as to apply also to approximation by finite trigonometric sums,§ or more generally to approximate representation by linear combinations of an arbitrary set of linearly independent continuous functions, having such further proper-

* Presented to the Society, April 10, 1920.

† Cf., e.g., Kirchberger, *Ueber Tchebychefsche Annäherungsmethoden*, Dissertation, Göttingen, 1902; Borel, *Leçons sur les fonctions de variables réelles et les développements en séries de polynomes*, pp. 82–92.

‡ Cf., e.g., Gram, *Ueber die Entwickelung reeller Functionen in Reihen mittelst der Methode der kleinsten Quadrate*, Journal für die reine und angewandte Mathematik, vol. 94 (1883), pp. 41–73.

§ For the extension of Tchebychef's theory to the case of trigonometric approximation, see, e.g., Fréchet, *Sur l'approximation des fonctions par des suites trigonométriques limitées*, Comptes Rendus, vol. 144 (1907), pp. 124–125; J. W. Young, *General theory of approximation by functions involving a given number of arbitrary parameters*, these Transactions, vol. 8 (1907), pp. 331–344; Fréchet, *Sur l'approximation des fonctions continues périodiques par les sommes trigonométriques limitées*, Annales de l'École Normale Supérieure, ser. 3, vol. 25 (1908), pp. 43–56.

ties, in the case of the final theorem, as to insure the uniqueness of the best approximating function in the sense of Tchebychef.* It will be apparent that even this general treatment can be extended in various directions, of which nothing more will be said here. The force of the conclusions will be most readily appreciated, on the other hand, if they are made specific by identifying the functions $p_1(x), p_2(x), \cdots, p_n(x)$ of the text with the quantities $1, x, \cdots, x^{n-1}$, and $\phi(x)$ with an arbitrary polynomial of degree $n-1$.

2. **First lemma on bounds of coefficients.** Let

$$p_1(x), p_2(x), \cdots, p_n(x)$$

be n functions of x, continuous throughout the interval

$$a \leqq x \leqq b,$$

and linearly independent in this interval. Let

$$\phi(x) = c_1 p_1(x) + c_2 p_2(x) + \cdots + c_n p_n(x)$$

be an arbitrary linear combination of these functions with constant coefficients, and let H be the maximum of $|\phi(x)|$ in (a, b). Then the following lemma holds:†

LEMMA I. *There exists a constant Q, completely determined by the system of functions $p_1(x), \cdots, p_n(x)$, such that*

$$|c_k| \leqq QH \qquad (k = 1, 2, \cdots, n),$$

for all functions‡ $\phi(x)$.

For each value of k, let the coefficients in the expression

$$\Phi_k(x) = c_{1k} p_1(x) + c_{2k} p_2(x) + \cdots + c_{nk} p_n(x)$$

be determined so that

$$\int_a^b p_i(x)\Phi_k(x)\,dx = 0, \qquad i \neq k; \qquad \int_a^b p_k(x)\Phi_k(x)\,dx = 1.$$

* Cf., e.g., Sibirani, *Sulla rappresentazione approssimata delle funzioni*, Annali di matematica pura ed applicata, ser. 3, vol. 16 (1909), pp. 203–221.

† This lemma is given, with a somewhat different proof, by Sibirani, loc. cit., p. 208. For the polynomial case, a variety of demonstrations have been given: see Kirchberger, loc. cit., pp. 7–9; Borel, loc. cit., pp. 83–84; Tonelli, *I polinomi d' approssimazione di Tchebychev*, Annali di matematica pura ed applicata, ser. 3, vol. 15 (1908), pp. 47–119; pp. 61–62; cf. also Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, vol. I, pp. 374–375, and, for bibliographical references, vol. II, p. 896.

‡ It is evident that the statement is not true if $p_1(x), \cdots, p_n(x)$ are linearly dependent, for, if there is a combination $\phi(x)$, with coefficients not all zero, which vanishes identically, this can be multiplied by a constant so as to give a combination which has arbitrarily large coefficients, and is still identically zero; and this can be added to a combination for which $H \neq 0$ so as to contradict the lemma.

This amounts to subjecting the n coefficients $c_{1k}, \cdots, c_{nk}$ to a set of n simultaneous linear equations. The determinant of the equations is not zero,* for if it were, a set of coefficients, not all zero, could be determined for a function

$$\Phi_0(x) = c_{10}\,p_1(x) + c_{20}\,p_2(x) + \cdots + c_{n0}\,p_n(x)$$

so as to make

$$\int_a^b p_i(x)\Phi_0(x)\,dx = 0 \qquad (i = 1, 2, \cdots, n).$$

It would follow from the last set of equations, however, that

$$\int_a^b [\Phi_0(x)]^2\,dx = \int_a^b [c_{10}\,p_1(x) + \cdots + c_{n0}\,p_n(x)]\Phi_0(x)\,dx = 0,$$

and this would imply that $\Phi_0(x) = 0$ identically, which is impossible, since $p_1(x), \cdots, p_n(x)$ are linearly independent. It is certain, therefore, that the desired functions $\Phi_k(x)$ exist.

Let Q' be the greatest value attained by the absolute value of any $\Phi_k(x)$ in (a, b). Then

$$\left|\int_a^b \phi(x)\Phi_k(x)\,dx\right| \leqq Q'\,H(b-a).$$

On the other hand, from the definition of $\Phi_k(x)$,

$$\int_a^b \phi(x)\Phi_k(x)\,dx = c_k.$$

Consequently, if $Q = Q'(b-a)$,

$$|c_k| \leqq QH.$$

3. **Second lemma on bounds of coefficients.** Let m be a fixed number greater than 1 (not necessarily an integer). Let

$$\Delta_m = \int_a^b |\phi(x)|^m\,dx,$$

and let $\Delta = b - a$. The following lemma is analogous to that already proved:

LEMMA II. *There exists a constant Q_1, completely determined by the system of functions $p_1(x), \cdots, p_n(x)$, and, in particular, independent of m, such that*

$$|c_k| \leqq Q_1(\Delta + \Delta_m) \qquad (k = 1, 2, \cdots, n),$$

for all functions $\phi(x)$.

In the first place, since $m > 1$, $|\phi(x)| \leqq |\phi(x)|^m$ unless $|\phi(x)| < 1$, so

* The non-vanishing of this Gramian determinant is a well-known condition for linear independence; cf., e.g., Kowalewski, *Einführung in die Determinantentheorie*, pp. 320–325.

that, in any case,

$$|\phi(x)| \leqq 1 + |\phi(x)|^m.$$

Hence

$$\int_a^b |\phi(x)|dx \leqq \Delta + \Delta_m,$$

and for any value of x between a and b,

$$\left|\int_a^x \phi(x)\,dx\right| \leqq \int_a^x |\phi(x)|dx \leqq \Delta + \Delta_m. \tag{1}$$

On the other hand, the n functions

$$\int_a^x p_1(x)\,dx,\ \int_a^x p_2(x)\,dx,\ \cdots,\ \int_a^x p_n(x)\,dx \tag{2}$$

are linearly independent, since a linear relation between them would give by differentiation a linear relation connecting $p_1(x), \cdots, p_n(x)$. Therefore, if Q_1 is the constant of Lemma I for the functions (2), it can be inferred from (1), that is, from

$$\left|c_1\int_a^x p_1(x)\,dx + \cdots + c_n\int_a^x p_n(x)\,dx\right| \leqq \Delta + \Delta_m,$$

that

$$|c_k| \leqq Q_1(\Delta + \Delta_m).$$

4. **Third lemma on bounds of coefficients.** In addition to the notation of the preceding sections, let $f(x)$ be a function continuous for $a \leqq x \leqq b$, arbitrary at the outset, but to be kept unchanged throughout the remainder of the discussion; let M be the maximum of $|f(x)|$ in (a, b); and let

$$\delta_m = \int_a^b |f(x) - \phi(x)|^m\,dx.$$

A further development of the ideas of the first two lemmas leads to the following statement:

LEMMA III. *For all functions $\phi(x)$,*

$$|c_k| \leqq Q_1(M\Delta + \Delta + \delta_m) \qquad (k = 1, 2, \cdots, n),$$

where Q_1 is the constant of the preceding lemma.

By an appropriate modification of a remark made in the preceding section, it is recognized that

$$|f(x) - \phi(x)| \leqq 1 + |f(x) - \phi(x)|^m.$$

Hence

$$|\phi(x)| \leqq M + 1 + |f(x) - \phi(x)|^m,$$

and

$$\int_a^b |\phi(x)|dx \leqq (M + 1)\Delta + \delta_m.$$

The concluding steps of § 3, applied to the present case, show that

$$|c_k| \leqq Q_1[(M+1)\Delta + \delta_m].$$

5. **Existence of an approximating function for exponent m.** If the function $f(x)$, the system $p_1, \cdots, p_n$, and the exponent m are given, and the coefficients c_k are regarded as undetermined, the value of δ_m, which is a function of these coefficients, has a lower limit γ_m which is positive or zero. If there is a function $\phi(x)$ for which δ_m actually attains its lower limit, this $\phi(x)$ will be called, for brevity, an *approximating function for the exponent m*. It is readily deduced from Lemma III that such a function will always exist.* For sets of coefficients $c_k^{(j)}$ can be chosen successively, $j = 1, 2, \cdots$, so that, if $\delta_m^{(j)}$ is the corresponding value of δ_m in each case,

$$\lim_{j=\infty} \delta_m^{(j)} = \gamma_m .$$

If $c_1^{(j)}, c_2^{(j)}, \cdots, c_n^{(j)}$ are regarded as the coördinates of a point P_j in space of n dimensions, all the points P_j from a certain value of j on, as soon as $\delta_m^{(j)}$ becomes and remains less than $\gamma_m + 1$, say, will lie in a bounded region,

$$|c_k^{(j)}| \leqq Q_1(M\Delta + \Delta + \gamma_m + 1).$$

The points P_j will have a limit point P in this region, and as the dependence of δ_m on the c's is continuous, the function $\phi(x)$ formed with the coefficients corresponding to the point P will make δ_m equal to γ_m. This approximating function $\phi(x)$ will be denoted by $\phi_m(x)$.

6. **Uniqueness of the approximating function for exponent m.** For each value of m, the approximating function $\phi_m(x)$ is unique. Suppose there were two such functions, $\phi_{\mathrm{I}}(x)$ and $\phi_{\mathrm{II}}(x)$, the subscript m being understood. Let

$$\phi_{\mathrm{III}}(x) = \tfrac{1}{2}[\phi_{\mathrm{I}}(x) + \phi_{\mathrm{II}}(x)],$$

and let $\delta_{\mathrm{I}}, \delta_{\mathrm{II}}, \delta_{\mathrm{III}}$, be the corresponding values of δ_m, so that $\delta_{\mathrm{I}} = \delta_{\mathrm{II}} = \gamma_m$. Furthermore, let

$$r_{\mathrm{I}}(x) = f(x) - \phi_{\mathrm{I}}(x), \qquad r_{\mathrm{II}}(x) = f(x) - \phi_{\mathrm{II}}(x),$$

$$r_{\mathrm{III}}(x) = f(x) - \phi_{\mathrm{III}}(x).$$

Then

$$r_{\mathrm{III}}(x) = \tfrac{1}{2}[r_{\mathrm{I}}(x) + r_{\mathrm{II}}(x)].$$

Since $m > 1$,

$$|r_{\mathrm{III}}(x)|^m \leqq \tfrac{1}{2}|r_{\mathrm{I}}(x)|^m + \tfrac{1}{2}|r_{\mathrm{II}}(x)|^m; \tag{3}$$

if $r_{\mathrm{I}}(x) = X_1$, for any particular value of x, $r_{\mathrm{II}}(x) = X_2$, and $r_{\mathrm{III}}(x) = X_3$, the assertion is that

$$\left|\frac{X_1 + X_2}{2}\right|^m \leqq \frac{|X_1|^m + |X_2|^m}{2},$$

* Cf. Young, loc. cit., p. 335.

which is a consequence of the fact that the graph of the function $Y = |X|^m$ is concave upward.* Moreover, the sign of inequality holds in (3), for any value of x for which $r_{\mathrm{I}} \neq r_{\mathrm{II}}$, that is, whenever $\phi_{\mathrm{I}} \neq \phi_{\mathrm{II}}$. Therefore, if ϕ_{I} and ϕ_{II} are not identically equal,

$$\int_a^b |r_{\mathrm{III}}(x)|^m \, dx < \frac{1}{2}\int_a^b |r_{\mathrm{I}}(x)|^m \, dx + \frac{1}{2}\int_a^b |r_{\mathrm{II}}(x)|^m \, dx,$$

since the integrands are continuous, and the relation (3) is an inequality over a part at least of the interval of integration. That is,

$$\delta_{\mathrm{III}} < \tfrac{1}{2}(\delta_{\mathrm{I}} + \delta_{\mathrm{II}}),$$

or, since $\delta_{\mathrm{I}} = \delta_{\mathrm{II}} = \gamma_m$,

$$\delta_{\mathrm{III}} < \gamma_m.$$

This is inconsistent with the definition of γ_m as the least possible value of δ_m.

Similar reasoning shows that no function $\phi(x)$, other than $\phi_m(x)$, can give even a relative minimum for δ_m as a function of $c_1, \cdots, c_n$. Let $\phi_{\mathrm{II}}(x)$ be any such function $\phi(x)$, let $\phi_{\mathrm{I}}(x) = \phi_m(x)$, and let

$$\phi_{\mathrm{III}}(x) = A\phi_{\mathrm{I}}(x) + B\phi_{\mathrm{II}}(x),$$

where A and B are any two positive constants whose sum is 1. Let $r_{\mathrm{I}}(x)$, $r_{\mathrm{II}}(x)$, $r_{\mathrm{III}}(x)$, and $\delta_{\mathrm{I}}, \delta_{\mathrm{II}}, \delta_{\mathrm{III}}$, be the corresponding values of $f(x) - \phi(x)$ and of δ_m. Then

$$|r_{\mathrm{III}}(x)|^m \leqq A|r_{\mathrm{I}}(x)|^m + B|r_{\mathrm{II}}(x)|^m,$$

the inequality holding whenever $\phi_{\mathrm{I}} \neq \phi_{\mathrm{II}}$. Consequently

$$\delta_{\mathrm{III}} < A\delta_{\mathrm{I}} + B\delta_{\mathrm{II}}, \tag{4}$$

or, since $\delta_{\mathrm{I}} < \delta_{\mathrm{II}}$ and $A + B = 1$,

$$\delta_{\mathrm{III}} < \delta_{\mathrm{II}}.$$

This means, inasmuch as A can be taken arbitrarily small and B arbitrarily near to 1, that it is possible to find functions $\phi(x)$ with coefficients as close to those of $\phi_{\mathrm{II}}(x)$ as may be desired, so that $\delta_m < \delta_{\mathrm{II}}$.

The main conclusions obtained hitherto (not including the last one) can be summarized as follows:

THEOREM I. *For each value of $m > 1$, there exists one and just one approximating function $\phi_m(x)$.*

7. Necessary and sufficient condition† for the approximating function $\phi_m(x)$. Let $\phi_m(x)$ be the approximating function for exponent m as before, and let

$$r_m(x) = f(x) - \phi_m(x).$$

* Analytically, of course, an immediate proof is obtained from the mean value theorem and the fact that dY/dX is an increasing function of X.

† This section is inserted for its own sake, and is not needed for what follows.

When $r_m(x) \neq 0$, let $r_m^{[m-1]}(x)$ be used as an abbreviation for the expression $|r_m(x)|^m/[r_m(x)]$, and let $r_m^{[m-1]}(x) = 0$ when $r_m(x) = 0$ (it is assumed throughout that $m > 1$). Then $r_m^{[m-1]}(x)$ is a quantity having its absolute value equal to $|r_m(x)|^{m-1}$, and having the same algebraic sign as $r_m(x)$ itself; if m is an even integer, $r_m^{[m-1]}(x)$ is simply $[r_m(x)]^{m-1}$. It will be shown that $r_m^{[m-1]}(x)$ must be orthogonal to each of the functions $p_k(x)$ in the interval (a, b):

$$\int_a^b r_m^{[m-1]}(x)\, p_k(x)\, dx = 0 \qquad (k = 1, 2, \cdots, n).$$

To bring out what is essential in the proof, let it be given first for the special case $m = 2$. Let $p(x)$, without subscript, stand for any one of the functions $p_k(x)$, and let h be an arbitrary constant, positive, negative, or zero. Let

$$\phi(x) = \phi_2(x) + hp(x),$$

$$r(x) = f(x) - \phi(x) = r_2(x) - hp(x);$$

then

$$|r(x)|^2 = [r(x)]^2 = [r_2(x)]^2 - 2hr_2(x)p(x) + h^2[p(x)]^2.$$

Hence

$$\delta_2 = \int_a^b |r(x)|^2\, dx = \gamma_2 - 2h\int_a^b r_2(x)p(x)\, dx + h^2\int_a^b [p(x)]^2\, dx,$$

since $r_2(x)$ is understood to be the error of the approximating function for $m = 2$, so that

$$\int_a^b [r_2(x)]^2\, dx = \gamma_2.$$

In the relation

$$\delta_2 = \gamma_2 - h\left[2\int_a^b r_2(x)p(x)\, dx - h\int_a^b [p(x)]^2\, dx\right],$$

suppose that the first of the two terms of the expression in brackets is not zero; it is to be shown that this leads to a contradiction. If h is sufficiently small numerically, the second term will be smaller numerically than the first, and the value of the whole bracket will be different from zero and will have the sign of the first term. If h is given a small value of the same sign as the first term in the bracket, the whole expression to be subtracted from γ_2 will be positive, and the value of δ_2 corresponding to the function $\phi(x)$ will be smaller than γ_2. Since this is contrary to the definition of γ_2, the truth of the assertion is established in the special case.

It is evident that an altogether similar proof can be given if m is any even integer. The demonstration can be modified so as to make it applicable to other cases as well. In general, let

$$\phi(x) = \phi_m(x) + hp(x),$$

$$r(x) = f(x) - \phi(x) = r_m(x) - hp(x),$$

with the understanding that

$$\delta_m = \int_a^b |r(x)|^m dx, \qquad \gamma_m = \int_a^b |r_m(x)|^m dx.$$

Then

$$\frac{d}{dh}\delta_m = \int_a^b \frac{\partial}{\partial h}|r(x)|^m dx.$$

If $r(x) > 0$,

$$\frac{\partial}{\partial h}|r(x)|^m = \frac{\partial}{\partial h}[r(x)]^m = m[r(x)]^{m-1}\frac{\partial}{\partial h}[r(x)]$$

$$= -mp(x)[r(x)]^{m-1} = -mp(x)|r(x)|^{m-1}.$$

If $r(x) < 0$,

$$\frac{\partial}{\partial h}|r(x)|^m = \frac{\partial}{\partial h}[-r(x)]^m = m[-r(x)]^{m-1}\frac{\partial}{\partial h}[-r(x)]$$

$$= mp(x)[-r(x)]^{m-1} = mp(x)|r(x)|^{m-1}.$$

If $r(x) = 0$,

$$\frac{\partial}{\partial h}|r(x)|^m = 0,$$

whether h is given positive or negative increments. In any case,

$$\frac{\partial}{\partial h}|r(x)|^m = -mp(x)|r(x)|^m/[r(x)],$$

a continuous function of x and h, the value of the fraction being taken to be zero when $r(x) = 0$, and

$$\left[\frac{\partial}{\partial h}|r(x)|^m\right]_{h=0} = -mp(x)r_m^{[m-1]}(x),$$

so that

$$\left[\frac{d}{dh}\delta_m\right]_{h=0} = -m\int_a^b p(x)r_m^{[m-1]}(x)\,dx.$$

The last integral must be zero, otherwise it would be possible to give h a small value, positive or negative, so as to make

$$\delta_m(h) < \delta_m(0),$$

that is,

$$\delta_m(h) < \gamma_m,$$

which is inadmissible. So the assertion made at the beginning of the section is true in general.

It is merely another statement of the same conclusion to say that $r_m^{[m-1]}(x)$ must be orthogonal to every function $\phi(x)$.

The necessary condition that has been obtained for $\phi_m(x)$ is also sufficient. This follows from the reasoning in the latter part of § 6, which led up to the remark that no function $\phi(x)$, other than $\phi_m(x)$, can give even a relative minimum for δ_m. Suppose that $\phi_{\mathrm{I}}(x)$ is a linear combination of the functions $p_k(x)$, not identical with $\phi_m(x)$. Let

$$r_{\mathrm{I}}(x) = f(x) - \phi_{\mathrm{I}}(x),$$

and let $r_{\mathrm{I}}^{[m-1]}(x)$ be defined in a manner corresponding to the definition of $r_m^{[m-1]}(x)$ above. It is to be shown that there exists a function $\psi(x)$ which is a linear combination of the p's, such that

$$\int_a^b \psi(x) r_{\mathrm{I}}^{[m-1]}(x)\,dx \neq 0. \tag{5}$$

For any linear combination $\psi(x)$, let

$$\phi(x) = \phi_{\mathrm{I}}(x) + h\psi(x),$$
$$r(x) = f(x) - \phi(x) = r_{\mathrm{I}}(x) - h\psi(x),$$
$$\delta_m = \int_a^b |r(x)|^m\,dx, \qquad \delta_{\mathrm{I}} = \int_a^b |r_{\mathrm{I}}(x)|^m\,dx.$$

By a calculation corresponding to that of the third paragraph preceding, it is seen that

$$\left[\frac{d}{dh}\delta_m\right]_{h=0} = -m\int_a^b \psi(x) r_{\mathrm{I}}^{[m-1]}(x)\,dx. \tag{6}$$

Now let

$$\psi(x) = \phi_m(x) - \phi_{\mathrm{I}}(x);$$

then

$$\phi(x) = h\phi_m(x) + (1-h)\phi_{\mathrm{I}}(x).$$

For positive values of h, the inequality (4) of § 6 is applicable with A, B, δ_{III}, δ_{I}, and δ_{II} replaced by h, $1-h$, δ_m, γ_m, and δ_{I} respectively:

$$\delta_m < h\gamma_m + (1-h)\delta_{\mathrm{I}},$$

that is,

$$\delta_m < \delta_{\mathrm{I}} + h(\gamma_m - \delta_{\mathrm{I}}), \qquad \frac{\delta_m - \delta_{\mathrm{I}}}{h} < \gamma_m - \delta_{\mathrm{I}},$$

the difference $\gamma_m - \delta_{\mathrm{I}}$ being negative. Consequently

$$\left[\frac{d}{dh}\delta_m\right]_{h=0} \leqq \gamma_m - \delta_{\mathrm{I}} < 0,$$

and, because of (6), the inequality (5) is verified.

To summarize, using the symbols $r(x)$ and $r^{[m-1]}(x)$ in a manner corresponding to the previous notation:*

THEOREM II. *In order that $\phi(x)$ be the approximating function for exponent*

* That is, $r(x) = f(x) - \phi(x)$, $r^{[m-1]}(x) = |r(x)|^m/[r(x)]$ when $r(x) \neq 0$, $r^{[m-1]}(x) = 0$ when $r(x) = 0$.

m, it is necessary and sufficient that

$$\int_a^b \psi(x)\, r^{[m-1]}(x)\, dx = 0$$

for all functions $\psi(x)$ which are linear combinations of $p_1(x), \cdots, p_n(x)$.

If the functions $p_k(x)$ are the quantities x^{k-1}, $k = 1, 2, \cdots, n$, it can be inferred further that $r(x)$, if not identically zero, *must change sign at least n times in the interval (a, b)*, for any value of m. Otherwise it would be possible to assign ν points $x_1, x_2, \cdots, x_\nu$, $\nu \leqq n-1$, so that $r(x)$, and hence $r^{[m-1]}(x)$, would be of constant sign (wherever different from zero) in each of the intervals $a \leqq x \leqq x_1$, $x_1 \leqq x \leqq x_2$, $\cdots$, $x_\nu \leqq x \leqq b$, and would take on opposite signs in successive intervals. Then the polynomial*

$$\psi(x) = (x - x_1)(x - x_2) \cdots (x - x_\nu),$$

of degree $\leqq n - 1$, would certainly not be orthogonal to $r^{[m-1]}(x)$, since the product $\psi(x)\, r^{[m-1]}(x)$, continuous and not vanishing identically, would be of constant sign wherever different from zero. Similar reasoning is possible in a class of other cases, including that of approximation by finite trigonometric sums, but of course not in the case of arbitrary functions $p_k(x)$.

8. **Limit of maximum error of $\phi_m(x)$ as m becomes infinite.** In this section and the following one, it will be assumed for convenience that $|f(x)| < 1$ for $a \leqq x \leqq b$. It will turn out that this is no real restriction of generality for the main conclusions, since multiplication of $f(x)$ by any constant corresponds to multiplication of the approximating functions $\phi_m(x)$, and of the other approximating functions to be considered, by the same constant.

For any function $\phi(x)$ (that is, any linear combination of the p's) let l be the maximum value of $|f(x) - \phi(x)|$ in (a, b); let l_m be the maximum of $|f(x) - \phi_m(x)|$, and let l_0 be the lower limit of l for all possible functions $\phi(x)$. It can be inferred from Lemma I that there is at least one $\phi(x)$ for which the limit l_0 is attained.† For l is a continuous function of the coefficients of ϕ, and all the coefficients of any combination ϕ for which l is near l_0 belong to a restricted region in the space of $c_1, \cdots, c_n$, so that there will be some set of values of these parameters for which l reaches its limit. Let the function ϕ corresponding to such a set of coefficients be denoted by‡ $\phi_0(x)$. It may be spoken of as the Tchebychef function, or a Tchebychef function, for $f(x)$; the question of its uniqueness need not be raised until the following section. The purpose of the present section is to prove:

* If $r(x)$ did not change sign at all, it would be understood that $\psi(x) = 1$.

† Cf. Young, loc. cit., p. 335; Fréchet, Annales de l'École Normale Supérieure, loc. cit., p. 45; Sibirani, loc. cit., p. 210.

‡ In view of what follows, it would be more suggestive to represent this function by $\phi_\infty(x)$, and the corresponding maximum error by l_∞, but it is not necessary to anticipate to that extent.

THEOREM III. *As m becomes infinite, l_m approaches the limit l_0.*

From the hypothesis that $|f(x)| < 1$, it follows that γ_m, the lower limit of δ_m for all possible functions $\phi(x)$, is less than $b - a$, for all values of m. For the particular function $\phi(x) = 0$ makes

$$\delta_m = \int_a^b |f(x)|^m \, dx < b - a,$$

and γ_m must be less than or equal to this δ_m. Hence, if c_k is any coefficient of any function $\phi_m(x)$, the term δ_m in the inequality of Lemma III may be replaced by $\Delta = b - a$, while $M < 1$, so that

$$|c_k| \leqq 3Q_1 \Delta. \tag{7}$$

That is, the absolute values of the coefficients have an upper bound which is independent of m.

Let ϵ be any positive quantity, and suppose that $|f - \phi_m| \geqq l_0 + \epsilon$ for some value $x = x_0$ in (a, b):

$$|f(x_0) - \phi_m(x_0)| \geqq l_0 + \epsilon. \tag{8}$$

Since $f(x)$ is continuous for $a \leqq x \leqq b$, it is uniformly continuous there. Let δ' be a positive quantity such that

$$|f(x') - f(x'')| \leqq \tfrac{1}{3}\epsilon$$

for $|x' - x''| \leqq \delta'$; in particular,

$$|f(x) - f(x_0)| \leqq \tfrac{1}{3}\epsilon \tag{9}$$

for $|x - x_0| \leqq \delta'$. Each of the functions $p_k(x)$ is likewise uniformly continuous in (a, b); let δ'' be so small that

$$|p_k(x') - p_k(x'')| \leqq \frac{\epsilon}{9nQ_1 \Delta}$$

whenever $|x' - x''| \leqq \delta''$, for all values of k. In view of (7) and the fact that there are n terms in $\phi_m(x)$, it follows that

$$|\phi_m(x) - \phi_m(x_0)| \leqq \tfrac{1}{3}\epsilon \tag{10}$$

if $|x - x_0| \leqq \delta''$. Let δ be the smaller of the quantities δ', δ''; then, as a consequence of (8), (9), and (10),

$$|f(x) - \phi_m(x)| \geqq l_0 + \tfrac{1}{3}\epsilon, \tag{11}$$

for $|x - x_0| \leqq \delta$, where δ is independent of m, though different values of m may call for different values of x_0. If it be supposed further that $\delta < \frac{1}{2}(b - a)$, then at least one of the intervals $(x_0 - \delta, x_0)$, $(x_0, x_0 + \delta)$ will be wholly contained in (a, b), wherever x_0 may be, and there will certainly be an interval of length δ at least throughout which (11) is satisfied.

Then

$$\int_a^b |f(x) - \phi_m(x)|^m dx \geqq (l_0 + \tfrac{1}{3}\epsilon)^m \delta. \tag{12}$$

On the other hand,

$$\int_a^b |f(x) - \phi_0(x)|^m dx \leqq l_0^m (b - a). \tag{13}$$

But m can be taken so large as to make the right-hand member of (12) larger than the right-hand member of (13). If (12) were still to hold, $\phi_0(x)$ would give a smaller value of δ_m than $\phi_m(x)$, which would be inconsistent with the definition of $\phi_m(x)$ as the function giving the smallest possible value of δ_m. So (12), and with it the hypothesis on which (12) is based, namely the inequality (8), must cease to be true. That is, for all values of m from a certain point on,

$$|f(x) - \phi_m(x)| < l_0 + \epsilon$$

throughout (a, b), and this is equivalent to the assertion of Theorem III.

9. **Limit of $\phi_m(x)$ as m becomes infinite.** In consequence of (7), the coefficients c_k of $\phi_m(x)$, regarded as coördinates of a point in space of n dimensions, must give rise to at least one limit point as m becomes infinite. From Theorem III, with the fact that l, the maximum of $|f(x) - \phi(x)|$, is a continuous function of the coefficients of ϕ, it follows that the value of l for any function ϕ corresponding to such a limiting set of coefficients must be l_0. It is known, however, that in the case of approximation by polynomials* or by finite trigonometric sums,† and in an extensive class of cases generally,‡ there can be only one function $\phi(x)$ for which the limit l_0 is attained. For these circumstances, the statement of Theorem III can be given the more striking form:

THEOREM IV. *If the system of functions $p_k(x)$ is such that the Tchebychef function $\phi_0(x)$ is uniquely determined, then*

$$\lim_{m=\infty} \phi_m(x) = \phi_0(x),$$

in the sense that the coefficients of ϕ_m approach those of ϕ_0, and the value of $\phi_m(x)$ therefore approaches that of $\phi_0(x)$ uniformly for $a \leqq x \leqq b$.

UNIVERSITY OF MINNESOTA,
MINNEAPOLIS, MINN.

* Cf. Kirchberger, Borel, locc. citt.

† Cf. Young, Fréchet, Tonelli, locc. citt.

‡ Cf. Young, Fréchet, Sibirani, locc. citt.