GENERAL THEORY OF APPROXIMATION BY FUNCTIONS INVOLVING A GIVEN NUMBER OF ARBITRARY PARAMETERS*

BY

JOHN WESLEY YOUNG

Introduction.

The following is a special case of the problem to be considered in this paper: Given a function $\phi(x)$ of the real variable x, continuous on a finite interval (a, b); to determine the polynomial $p(x)$ of *given* degree n, which gives the closest approximation to the given function ϕ on the interval (a, b). This problem becomes definite only when the meaning of the phrase "closest approximation" has been precisely stated, and the meaning adopted will depend on the ultimate object in view.

TCHEBYCHEV seems to have been the first to consider this problem.† He regarded that polynomial as giving the best approximation, which rendered the maximum of $|p(x) - \phi(x)|$, as x varied over (a, b), as small as possible. A different point of view would lead one to seek a polynomial of the given degree which rendered as small as possible the expression

$$\int_a^b (p - \phi)^2 dx,$$

or the expression

$$\int_a^b |p - \phi| dx,$$

etc. In all of these cases, and in the more general ones to be referred to presently, the problem consists in the determination of a set of parameters a_i of a function $f(x; a_0, a_1, \cdots, a_n)$ of the real variable x, such that the maximum of $|f(x)|$, as x varies over a given interval, shall be as small as possible.

This problem TCHEBYCHEV himself (loc. cit.) seems to have regarded as the *general type of problem connected with the approximate representation of functions*, although he seems to have done little with this general problem beyond stating it. He confines himself to a detailed discussion of the case where the function f takes the form $p(x) - \phi(x)$, or $r(x) - \phi(x)$, where $p(x)$ is as before a polynomial of given degree, and where $r(x)$ is a rational function of which the degrees of the numerator and denominator are prescribed. The parameters a_i referred to above are then, of course, the coefficients of the various powers of x. The fact that the degrees of the polynomials are prescribed, i. e. the number of arbitrary parameters, is essential to this type of problem.

Certain generalizations of the problems treated by TCHEBYCHEV at once suggest themselves. On the one hand, the functions $p(x)$ or $r(x)$ might be replaced by any function of a given *class* $\mathfrak{C}$ of functions $g(x; a_0, a_1, \cdots, a_n)$. On the other hand, the form of the function f to be considered might be made more general. In the case actually discussed by TCHEBYCHEV, f was simply $g - \phi$, where ϕ was a given continuous function. Both generalizations are included, if we identify f with a function Ug, where Ug denotes the result of operating on g with some *functional operation* U. If then there exists a set of parameters a_i which renders the maximum of $|Ug|$, as x varies over a certain given finite interval, as small as possible, we will call the resulting function g *a function of approximation in the class* $\mathfrak{C}$ *with reference to* U.

The fundamental theoretical problems that now present themselves are as follows: 1) Given an operation U and a class $\mathfrak{C}$, does a function of approximation in $\mathfrak{C}$ with reference to U exist? 2) What are necessary and sufficient conditions that a function in $\mathfrak{C}$ be a function of approximation in $\mathfrak{C}$ with reference to U?

TCHEBYCHEV has given answers to these problems (loc. cit.) for the special cases already attributed to him, where Ug has the form $g - \phi$, and where g is either a polynomial or a rational fractional function of a specified kind. His methods, which lack the degree of rigor required at this day, have recently been revised by KIRCHBERGER,* and still more recently BOREL† has given elegant proofs of those theorems of TCHEBYCHEV which relate to *polynomials* of approximation.

It is the object of the present paper to give a two-fold generalization of TCHEBYCHEV's theorems; the latter will be found as special cases of our theory (§ 5, *a*). The method followed is largely that of BOREL. In § 1 we state the general problem with the necessary precision and prove the existence of a solution under very general conditions. In §§ 2, 3, 4, restricting ourselves to the

---

*KIRCHBERGER, *Über Tchebychefsche Annäherungsmethoden*, Dissertation, Göttingen, 1902; in abstract, Mathematische Annalen, vol. 57 (1903), pp. 509–540.

† BOREL, *Leçons sur les fonctions de variables réelles, etc.*, 1905, pp. 82–88.

case in which the parameters enter linearly into the functions g and to the case where we have $Ug = g - \phi$, we prove again the existence of a solution, derive necessary and sufficient conditions which a solution must satisfy, and prove that the solution is unique. The results of these sections are summarized in Theorems 4 and 5. Theorem 4 will be found to express a characteristic property of a general class of "functions of approximation," which seems to have been recognized hitherto only for the case of polynomials and rational functions of approximation. After considering in § 5 certain special cases of the theory developed in the preceding sections, we extend the theory further in § 7 to include the case in which Ug has the form $Vg - \phi$, where V is any one-valued, distributive operation. This makes possible the application of the theory to the approximate representation of functions restricted only to satisfy certain functional equations of a general form; some of these are referred to in § 8, where we obtain as an illustration a theorem concerning the polynomial of approximation of given degree for a linear differential equation with constant coefficients. In § 6 we show how the problem may be formulated analytically.

I. The problem in general.

§ 1. *The general existence theorem.*

Let $f = f(x) = f(x; a_0, a_1, a_2, \cdots, a_n)$ be a function of the real variable x and $n + 1$ parameters a_i, subject to the following conditions A:

$A1)$ $f(x; a_0, a_1, a_2, \cdots, a_n)$ shall be defined as a one-valued and continuous function of its arguments for every x of a finite interval (a, b) and for all real finite values of the parameters a_i.*

$A2)$ For every positive number M there shall exist an N, such that if the relation

$$|f(x; a_0, a_1, a_2, \cdots, a_n)| \leqq M$$

be satisfied for all values of x on (a, b), then the parameters a_i all satisfy the relations

$$|a_i| \leqq N \qquad (i = 0, 1, 2, \cdots, n).$$

Now suppose the parameters a_i to be undetermined but fixed. As x varies over (a, b), $|f|$ will attain its maximum value m at least once. This maximum $m(a_0, a_1, a_2, \cdots, a_n)$ by $A1)$ is a continuous function of the a_i. Hence as the a_i vary over all real finite values, m certainly has a lower bound μ which is either positive or zero. Our problem is to ascertain whether there exists a

---

* Throughout this paper the word interval will always imply that the endpoints are included. This condition $A1)$ could, moreover, for the purposes of this section be stated more generally by defining $f(x)$ merely on a set of points E on (a, b), with the restriction that E be *perfect*.

finite set of parameters α_i for which this lower bound is actually attained, i. e., such that we have

$$m(\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_n) = \mu.$$

If such a set exists, we will call it a *minimizing set of parameters for f.* The problem of determining such a set of parameters would seem to arise in connection with most, if not all, problems of approximation by means of functions with a *given* number of arbitrary parameters.

The question concerning the existence of a minimizing set of parameters can be answered in the affirmative. For, if we choose any particular set of parameters, say $a'_0, a'_1, a'_2, \cdots, a'_n$, and denote $m(a'_0, a'_1, a'_2, \cdots, a'_n)$ by M, we know that μ is at most equal to M. We can then without loss of generality confine ourselves to functions f such that we have throughout (a, b)

$$|f| \leqq M.$$

But condition $A2)$ requires, then, that there exist a number N such that all the parameters of f satisfy the relations

(1) $|a_i| \leqq N$ $(i = 0, 1, 2, \cdots, n).$

Hence the parameters may be restricted to the finite closed domain defined by the relations (1), and in this domain the function $m(a_0, a_1, a_2, \cdots, a_n)$ attains its lower bound μ for at least one set of values $\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_n$. Hence we have

THEOREM 1. *There exists at least one minimizing set of parameters for any function satisfying conditions A.*

As an immediate corollary of this theorem, we obtain an existence theorem for the very general class of approximation problems referred to in the *Introduction.* Suppose there is given a class of functions, $g(x; b_0, b_1, b_2, \cdots, b_n)$, with the arbitrary parameters b_i, such that a function g is fully determined as soon as the values of the parameters b_i are fixed. Then as we have seen, the problem of determining, on a finite interval, that function g which will give the *best* approximation to a given function (or to a function satisfying a given functional equation) leads to the determination of a minimizing set of parameters for a function Ug, where Ug denotes the function obtained from g by some functional operation U. If then we identify the function Ug with the function f just considered, Theorem 1 gives the desired information concerning the existence of a solution. The parameters b_i of g will in general be the parameters a_i of $f = Ug$. However, it should be noted that in performing the operation U some of the parameters b_i may disappear, and new arbitrary parameters may be introduced. The parameters b_i which actually appear in Ug we will call the *effective* parameters of g with reference to U. Obviously in applying the conditions A to Ug the effective parameters b_i, together with whatever new param-

eters may have been introduced, play the rôle of a_i's. Any function g, the effective parameters of which have the values belonging to a minimizing set for Ug, we call a *function g of approximation with reference to U*. We have, then,

THEOREM 2. *There exists a function g of approximation with reference to the operation U, provided $f = Ug$ satisfies conditions A.*

II. FUNCTIONS OF APPROXIMATION FOR GIVEN FUNCTIONS.

§ 2. *Another form of the existence theorem.*

We now assign to the function f of the preceding section a more explicit but still very general form. We suppose given a set of $n + 1$ functions,

$$s_0, s_1, s_2, \cdots, s_n,$$

of the real variable x. The function s_i we call the elementary function of rank i. By means of some or all of these functions s_j we form a new function,

$$S_k = a_0 s_0 + a_1 s_1 + \cdots + a_k s_k,$$

which we call an S-function of rank k, if s_k is the elementary function of highest rank occurring in S_k. By allowing the parameters a_j to assume all real finite values we obtain the class of S-functions of rank not higher than k, which we will denote by $\mathfrak{S}_k$. It is clear that $\mathfrak{S}_l$ is contained in $\mathfrak{S}_k$, if l is less than k.

We suppose, further, that there is given a function ϕ of x. The function $y = S_n - \phi$ is to play the rôle of the function f of the preceding section. The functions s_j and ϕ are subject to the following conditions B:

$B1$) The functions s_j and ϕ are one-valued and continuous at every point of (a, b).

$B2$) If one parameter a_j is different from zero, the function S_k does not vanish at more than k points of (a, b).

$B3$) For no set of parameters a_j does the function $y = S_n - \phi$ vanish at every point of (a, b).

Condition $B2$) requires a word of explanation. It implies that each point x_1 of (a, b) at which S_k vanishes can be enclosed within a finite interval throughout which (except at x_1) S_k is different from zero. There are then two cases to consider. If S_k changes sign as x passes through x_1, we call x_1 a *simple* zero of S_k; if S_k does not change sign as x passes through x_1, we call x_1 a *double* zero. In applying condition $B2$) *every double zero must be counted twice.*[*]

We restate the discussion in terms of the new symbols as follows. As x varies over (a, b), the function $|y| = |S_n - \phi|$, continuous on (a, b), attains its maximum value m at least once; $m = m(a_0, a_1, a_2, \cdots, a_n)$ is a continuous

[*] If a or b is a zero of S_k, it counts as simple.

function of the parameters a_j. For we may vary the a_j by a sufficiently small amount to insure that throughout (a, b), where we have

(1) $$y = a_0 s_0 + a_1 s_1 + a_2 s_2 + \cdots + a_n s_n - \phi,$$

y, and hence m, will change by less than ϵ. We seek to determine a function y in $\mathfrak{S}_n$,

$$\Sigma_n = \alpha_0 s_0 + \alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_n s_n,$$

such that $m(\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_n)$ shall be equal to the lower bound μ of m; in other words, we seek a minimizing set of parameters for y. If such a minimizing set exists, we call Σ_n a *function of approximation in $\mathfrak{S}_n$ for ϕ*.

We have thought it desirable to state these conditions B independently of conditions A, in order that the remainder of the discussion might gain in unity. It is not difficult to show now, however, that conditions B imply A, and hence establish the existence of a solution to our problem. This we shall do by showing that for all functions (1) satisfying a relation $|y| \leqq M$ (M being finite and greater than μ), each of the parameters satisfies a relation

$$|a_j| \leqq N_j,$$

where the N_j are finite numbers.

We note first that a function

$$S_n(x) = \sum_{j=0}^{n} a_j s_j(x)$$

may be determined and is indeed uniquely determined by its values c_h at $n+1$ distinct points ξ_h ($h = 0, 1, \cdots, n$) on (a, b), i. e., by the equations

$$c_h = \sum_{j=0}^{n} a_j s_j(\xi_h) \qquad (h = 0, 1, \cdots, n).$$

For the determinant

$$|s_j(\xi_h)| \qquad (j, h = 0, 1, \cdots, n),$$

does not vanish; otherwise a function S_n, not identically zero, would exist which vanishes at the $n+1$ points ξ_h, in contradiction to $B2$). Denoting by

$$S_n^{(i)}(x) = \sum_{j=0}^{n} a_j^{(i)} s_j(x)$$

that function in $\mathfrak{S}_n$ which has at the ξ_h the values $c_h^{(i)}$ given by the equations

$$c_h^{(i)} = 0 \qquad (i \neq h),$$
$$c_i^{(i)} = 1,$$

we can express any function S_n having any values c_h at ξ_h by the formula*

$$S_n(x) = \sum_{i=0}^{n} c_i S_n^{(i)}(x) = \sum_{i=0}^{n} \sum_{j=0}^{n} c_i a_j^{(i)} s_j(x).$$

---

*This is a generalization of LAGRANGE's formula of interpolation.

Now any function S_n satisfying the relation

$$| S_n(x) - \phi(x) | \leqq M$$

satisfies the relation

$$| S_n(x) | \leqq M + \lambda,$$

where λ is the maximum of $| \phi(x) |$ on (a, b). But, by the formula just given, every function S_n can be written in the form

$$S_n(x) = \sum_{i=0}^{n} S_n(\xi_i) S_n^{(i)}(x),$$

or

$$\sum_{j=0}^{n} a_j s_j(x) = \sum_{j=0}^{n} s_j(x) \sum_{i=0}^{n} S_n(\xi_i) a_j^{(i)}.$$

Hence we have

$$a_j = \sum_{i=0}^{n} S_n(\xi_i) a_j^{(i)},$$

and so

$$| a_j | \leqq (M + \lambda) \sum_{i=0}^{n} | a_j^{(i)} | \leqq N_j,$$

provided

$$| S_n(x) - \phi(x) | \leqq M.$$

This gives the desired condition.   Hence we·have the following result :

THEOREM 3.   *There exists a function of approximation in* $\mathfrak{S}_n$ *for* ϕ, *provided conditions* B *are satisfied.*

§ 3. *A necessary condition for a function of approximation in* $\mathfrak{S}_n$.

We have just shown (Theorem 3) that there exists at least one function of approximation in $\mathfrak{S}_n$ for ϕ under conditions B. Let this function be $\sum_n$, and as before denote $\sum_n$ by y. Then $|y|$ attains its maximum μ in at least one point of (a, b). Let x' be any point for which $y = \mu$ and x'' any point for which $y = -\mu$. The sets $[x']$ and $[x'']$, consisting of all points x' and x'' respectively, are evidently closed since y is continuous on (a, b). Further, we may determine a number 2δ such that the oscillation of y in any interval of length $\leqq 2\delta$ will be less than a positive ϵ, less than μ. The distance of any point x' from the nearest x'' will then be greater than 2δ. Hence *it is possible to divide the interval* (a, b) *into a finite number of adjoining subintervals,*

$$L_1, L_2, L_3, \cdots, L_p,$$

such that no L_q *contains points of both the sets* $[x']$ *and* $[x'']$, *and such that one of two consecutive intervals* L_q *will contain points of* $[x']$ *and the other will contain points of* $[x'']$. Moreover, the points $\zeta_1, \zeta_2, \zeta_3, \cdots, \zeta_{p-1}$ which

divide (a, b) into the intervals L_q must be chosen distinct from any of the points x' or x'', and this evidently is always possible.

The above discussion will apply to any function in $\mathfrak{S}_n$, and it should be noted that the number p of intervals L_q is fixed for a given S_n. But we can now prove the following condition:

A necessary condition that a given function $\sum_n$ in $\mathfrak{S}_n$ be a function of approximation in $\mathfrak{S}_n$ for ϕ is that the number of intervals L_q exceed $n+1$.

We shall show that under the hypothesis $p \leqq n+1$ we can so construct a function S_n in $\mathfrak{S}_n$ that the maximum of $|S_n - \phi|$ on (a, b) is less than μ. Let

$$T_{p-1} = c_0 s_0 + c_1 s_1 + c_2 s_2 + \cdots + c_{p-1} s_{p-1}$$

be a function in $\mathfrak{S}_{p-1}$ which vanishes at each of the points $\zeta_i \, (i = 1, 2, \cdots, p-1)$. That such a function exists is shown by an argument similar to the one given on p. 336, which makes it clear that only the ratios $c_0 : c_1 : \cdots : c_{p-1}$ are determined by the conditions thus far imposed on T_{p-1}. We may then write

$$T_{p-1} = \eta \cdot T'_{p-1},$$

where T'_{p-1} is fully and uniquely determined, but where η is any arbitrary constant.*

Now, since T_{p-1} vanishes at $p-1$ different points on (a, b) and since by $B2$) it vanishes nowhere else, it follows that T_{p-1}, being continuous, changes sign as x passes from one interval L_q into the next, and only then. We then determine the sign of η so that T_{p-1} is positive within every interval L_q containing a point x' and negative within every interval L_q containing a point x''.

Make every x' the center of an interval I of length δ and every x'' the center of an interval J of length δ, and let K be the set of points on (a, b) not included in any I or J. Then the upper bound μ' of $|\sum_n - \phi|$ on K is certainly less than μ. Now choose $|\eta|$ so small that $|T_{p-1}| = |\eta| |T'_{p-1}|$ is less than the smaller of the numbers $\mu - \epsilon$ and $\mu - \mu'$ at every point of (a, b).

In any interval I we now have

$$\mu - \epsilon < \sum_n - \phi \leqq \mu$$

and

$$0 < T_{p-1} < \mu - \epsilon,$$

since no I contains a point ζ_i. Whence throughout any I the relation

$$0 < \sum_n - T_{p-1} - \phi < \mu$$

is satisfied, and similarly in any interval J

$$0 > \sum_n - T_{p-1} - \phi > -\mu.$$

---

* If $p = 1$, simply take T_{p-1} equal to η.

Finally, at every point of K we have

$$|\Sigma_n - \phi| \leqq \mu' \qquad \text{and} \qquad |T_{p-1}| < \mu - \mu',$$

and therefore

$$|\Sigma_n - T_{p-1} - \phi| < \mu.$$

Hence the function $S_n = \Sigma_n - T_{p-1}$, which is in $\mathfrak{S}_n$ under the hypothesis $p \leqq n + 1$, is such that at every point of (a, b) the relation

$$|S_n - \phi| < \mu$$

holds, which is the desired result.

§ 4. *The sufficiency of the condition and the uniqueness of the solution.*

Conversely, let Σ_n be any function in $\mathfrak{S}_n$ for which the necessary condition of § 3 is satisfied. There exist then at least $n + 2$ points

$$x_1 < x_2 < x_3 < \cdots < x_{n+2}$$

on (a, b) at which we have

$$|\Sigma_n(x_\kappa) - \phi(x_\kappa)| = \mu \qquad (\kappa = 1, 2, \cdots, n + 2),$$

and

$$\Sigma_n(x_\kappa) - \phi(x_\kappa) = -[\Sigma_n(x_{\kappa+1}) - \phi(x_{\kappa+1})] \qquad (\kappa = 1, 2, \cdots, n + 1).$$

Now suppose there were another function

$$\Sigma'_n = \alpha'_0 s_0 + \alpha'_1 s_1 + \cdots + \alpha'_n s_n$$

in $\mathfrak{S}_n$, such that on (a, b) the maximum value of $|\Sigma'_n - \phi|$ were less than, or even merely equal to, μ. Then the function

$$(\Sigma_n - \phi) - (\Sigma'_n - \phi) = \Sigma_n - \Sigma'_n$$

would be alternately positive (or zero) and negative (or zero) at the points x_κ. But this would require the continuous function $\Sigma_n - \Sigma'_n$ to have on (a, b) at least $n + 1$ simple zeros (or their equivalent, if double zeros occur), which by $B2$) is impossible, unless all the parameters of

$$\Sigma_n - \Sigma'_n = (\alpha_0 - \alpha'_0) s_0 + (\alpha_1 - \alpha'_1) s_1 + \cdots + (\alpha_n - \alpha'_n) s_n$$

vanish. But this requires Σ_n and Σ'_n to be identical. Hence we reach, in connection with the result of § 3, the following conclusions:

THEOREM 4. *A necessary and sufficient condition that a function Σ_n in $\mathfrak{S}_n$ be a function of approximation in $\mathfrak{S}_n$ for ϕ is that the number of intervals L_q exceed $n + 1$.*

THEOREM 5. *There exists one and only one function of approximation in* $\mathfrak{S}_n$ *for a function* ϕ.*

§ 5. *Special cases of the general theory.*

a) The theorems of Tchebychev on approximation by polynomials and rational functions.

If in the preceding theory we place $s_i = x^i$, that is, if we identify the class $\mathfrak{S}_n$ with the class of polynomials of degree not higher than n, we are led at once to the results of TCHEBYCHEV already referred to in the *Introduction*. Theorems 4 and 5 then give:

There exists one and only one polynomial of approximation of degree n for any function ϕ *on any given finite interval throughout which* ϕ *is continuous; and a necessary and sufficient condition that a given polynomial of degree n be a polynomial of approximation of that degree is that the number of intervals* L_q *exceed* $n + 1$.

A similar result was proved originally by TCHEBYCHEV † for the case of a rational function $R_n(x)$, in which the denominator $D(x)$ is fixed and the numerator is of prescribed degree n. We need here only to place $s_i = x^i / D(x)$ to obtain his theorem, which is analogous to the one just given. Conditions *B* will be satisfied provided no root of $D(x) = 0$ lies on (a, b).

b) Approximation by finite trigonometric series.

If we place $s_i = \cos ix$, our functions S_n assume the form

$$C_n = a_0 + a_1 \cos x + a_2 \cos 2x + \cdots + a_n \cos nx.$$

Such a function satisfies the condition *B*2) concerning the number of zeros, provided the length of the interval (a, b) does not exceed π. For every function C_n can be expressed as a polynomial in $\cos x$ of degree not higher than n. If ϕ also satisfies *B*3), *Theorems 4 and 5 will apply to the approximate representation of a given continuous function by means of functions of the form* C_n. Other types of trigonometric series having similar properties are as easily obtainable.

c) Other special types of functions having similar properties might be discussed, but it hardly seems worth while to go any further into details. Series of *Legendre polynomials* and series of *hyperbolic functions* are perhaps the simplest that suggest themselves, for which the crucial condition *B*2) can be readily satisfied. The possible application of Theorems 4 and 5 appears to be wide, in view of the general character of the available classes $\mathfrak{S}_n$.

---

* Since the above was written, a special case of this theorem, concerning the representation of functions by finite trigonometric series, has been announced by M. MAURICE FRÉCHET, *Sur l'approximation des fonctions par des suites trigonométriques limitées,* Comptes Rendus de l'Académie des Sciences de Paris, vol. 144, no. 3 (January 21, 1907), p. 124. Cf. also § 5*b*) above.

† Loc. cit.

§ 6. *Analytic formulation of the problem.*

Conditions B are sufficient to give an analytic formulation to the problem of determining the function of approximation in $\mathfrak{S}_n$ for a given continuous function ϕ, in case the functions S_n and ϕ possess derivatives as to x which are one-valued and continuous everywhere except at a finite number of points $\tau_1, \tau_2, \cdots, \tau_\lambda$ on (a, b), with the property that for every τ_i there exists an e_i such that

$$\lim_{x=\tau_i} (x - \tau_i)^{r_i} \cdot \frac{d}{dx} [S_n(x) - \phi(x)] = 0.$$

For Theorem 4 implies the existence of $n + 2$ points

$$x_1, x_2, \cdots, x_{n+2}$$

on (a, b) for which we have

$$|\Sigma_n(x_i) - \phi(x_i)| = \mu \qquad (i=1, 2, \cdots, n+2).$$

Since at these points $S_n - \phi$ is either a maximum or a minimum, we have also

$$\frac{d}{dx} [\Sigma_n - \phi] = 0 \quad \text{when } x = x_1,$$

except possibly when one of the points x_i coincides with a, $\tau_1, \tau_2, \cdots, \tau_\lambda$, or b.

If then we place

$$g(x) = (x - a)(x - \tau_1)^{e_1}(x - \tau_2)^{e_2} \cdots (x - \tau_\lambda)^{e_\lambda}(x - b),$$

a function of approximation in $\mathfrak{S}_n$ would have to satisfy the following system of equations:

$$(S_n - \phi)^2 - \mu^2 = 0, \qquad g(x) \frac{d}{dx}(S_n - \phi) = 0, \qquad (x=x_i ; i=1, 2, \cdots, n+2).$$

These $2n + 4$ equations contain just $2n + 4$ unknowns, namely, the $n + 1$ parameters a_j, the $n + 2$ quantities x_i, and the quantity μ. Among the sets of parameters satisfying this system must occur the set α_j giving the function of approximation.

III. FUNCTIONS OF APPROXIMATION FOR FUNCTIONAL EQUATIONS.

§ 7. *Extension of the theory.*

The results obtained in the previous paragraphs may be extended readily to the problem of finding functions of approximation for a certain general class of functional equations. If we have given a functional equation $U(y) = 0$ and a class of functions $g(x; a_0, a_1, \cdots, a_n)$, we define a *function of approximation for $U(y) = 0$ in the given class* to be a function g which *renders the maximum of $|U(g)|$ on (a, b) as small as possible.*

Theorems 4 and 5 may be extended to include this type of problem under the following conditions:

We suppose given a set of $m + 1$ *elementary* functions,

$$t_0, t_1, t_2, \cdots, t_m,$$

and form, in a way precisely analogous to the one previously used (§ 2), the various classes $\mathfrak{T}_k$ of functions T_k of rank not higher than k,

$$T_k = a_0 t_0 + a_1 t_1 + \cdots + a_k t_k \qquad (k = 0, 1, 2, \cdots, m).$$

We suppose given further a function $\phi(x)$, and a functional operation V which is effective in a single-valued distributive* way on the functions of $\mathfrak{T}_m$, the resulting values $V(T_m)$ being themselves functions of one variable. *Our treatment in §§ 2, 3, 4 applies to the case now under consideration in which, in view of the distributive character of V, the function $U(g)$ has the form*

$$V(T_m) - \phi = a_0 V(t_0) + a_1 V(t_1) + \cdots + a_m V(t_m) - \phi.$$

But this function has the same form as the one already considered in §§ 2, 3, 4, if we identify $V(t_i)$ with s_i, and $V(T_k)$ with S_k. If then we restate conditions B with the substitutions just indicated, and call the new conditions C, we may extend Theorems 4 and 5 to the new type of problem. Before doing so, however, the following remarks should be made regarding the details of the contemplated extension.

If any of the $V(t_i)$ $(i = 0, 1, \cdots, m)$ are zero identically on (a, b), the corresponding parameters a_i are *non-effective* (cf. § 1). Only the *effective* parameters of a T_k play any essential part in the discussion, and hence in stating condition $C2)$ (corresponding to $B2)$) the number k must evidently be replaced by $e_k - 1$, where e_k is the number of effective parameters of T_k with respect to V.

If in any function T_k we place equal to zero all the non-effective parameters with respect to V, we call the resulting function *reduced* with respect to V. A reduced function contains just e_k terms and all the parameters occurring in it are effective. In extending our theory we must be careful to specify that the functions T_k with which we operate are to be reduced functions, whenever this is necessary; as, for instance, in the determination of the functions $T_k^{(i)}$ satisfying the relations $V(T_k^{(i)}(\xi_j)) = 1$ or 0, according as we have $i = j$ or $i \neq j$ in the extended existence theorem (cf. p. 336).

---

* A functional operation A is said to be *distributive*, if we have $A(\alpha + \beta) = A(\alpha) + A(\beta)$ and $A(c\alpha) = cA(\alpha)$, for every constant value c and for every two functions α, β in the realm of definition of A. Cf. PINCHERLE, *Funktionaloperationen und -gleichungen*, Encyklopädie der Mathematischen Wissenschaften, vol. 2 A 11 (1906), p. 767. In our case we have the notion of an operation which is one-valued and distributive *with reference to a class* of functions, nothing being specified *concerning its behavior or existence* for functions not in the class.

The treatment of this more general case, which evidently reduces to the one first considered when V is the identical operation (i. e., $V(T_m) = T_m$), is throughout essentially identical with the treatment already given in detail for the simpler case. The conditions C may now be stated thus:

C 1) The functions $V(t_i)$ $(i = 0, 1, 2, \cdots, m)$ and ϕ are one valued and continuous at every point of (a, b).

C 2) If one effective parameter a_j is different from zero, the function $V(T_k)$ does not vanish at more than $e_k - 1$ points on (a, b).

C 3) For no set of parameters a_j does the function $y = V(T_m) - \phi$ vanish at every point of (a, b).

A function T_m which renders as small as possible the maximum of $|V(T_m) - \phi|$ on (a, b) we call a function of approximation in $\mathfrak{T}_m$ *with reference to V and ϕ.* The extended theorems then are as follows:

THEOREM 6. *A necessary and sufficient condition that a function T_m in $\mathfrak{T}_m$ be a function of approximation in $\mathfrak{T}_m$ with reference to V and ϕ is that the number of intervals M_q exceed the number of effective parameters in T_m.*

Here the intervals M_q are constructed with reference to the function $V(T_m) - \phi$, just as the intervals L_q were constructed with reference to the function $S_m - \phi$.

THEOREM 7. *The effective parameters of a function of approximation in $\mathfrak{T}_m$ with reference to V and ϕ are uniquely determined; i. e., there exists one and only one reduced function of approximation in $\mathfrak{T}_m$ with reference to V and ϕ.*

One or two further details concerning the proof of Theorem 6 require mention. If as in the previous treatment we denote by $\zeta_1, \zeta_2, \cdots, \zeta_{p-1}$ the points which divide (a, b) into the p intervals M_q, it is necessary to construct a function T_l which vanishes at the $p - 1$ points ζ_i and nowhere else on (a, b) (cf. p. 338). This is clearly possible, provided there exists a number $e_l = p$. That such a number does exist is readily seen from the fact that the sequence of numbers e_i $(i = 0, 1, \cdots, m)$ contains each of the integers from 1 to e_m at least once; for we evidently have $e_{i+1} = e_i$ or $e_{i+1} = e_i + 1$ according as a_{i+1} is not or is an effective parameter. The proof, following exactly the lines previously laid down, then leads to the determination of a function T_l satisfying the relation $|V(T_m) - V(T_l) - \phi| < \mu$. But since V is distributive, this implies the relation $|V(T_m - T_l) - \phi| < \mu$, and $T_m - T_l$ under the hypothesis $p \leqq e_m$ is a function in $\mathfrak{T}_m$, which serves to establish the necessity of the condition under consideration. The latter remark applies also to the proof of the sufficiency of the condition. To go through all the details of the proofs of the last theorems appears superfluous.

It is readily seen also that the analytic formulation of the problem given in § 6 may be extended at once to the present case.

8. *Application of the extended theorems.*

In seeking functional equations to which Theorems 6 and 7 apply, we should notice that a particular form of equation is prescribed, viz., $V(y) - \phi = 0$, where y is the unknown function of x and where V must be distributive. Beyond this it is only necessary that conditions C be satisfied. The variety of distributive operations is very large. As examples of interest we may mention those which give rise to linear differential equations with constant or variable coefficients, linear difference equations, linear integral equations, etc.

As a special case, we may consider the following:

Let

$$(1) \qquad c_0 \frac{d^\nu y}{dx^\nu} + c_1 \frac{d^{\nu-1} y}{dx^{\nu-1}} + \cdots + c_\nu y - \phi(x) = 0$$

be a linear differential equation with constant coefficients c_i, $(c_\nu \neq 0)$, and let the class $\mathfrak{T}_m$ be the class of polynomials of degree m. It is clear that conditions C will be satisfied, provided $\phi(x)$ is not identical on (a, b) with any polynomial of degree equal to (or lower than) m. We have then, from Theorems 6 and 7,

THEOREM 8. *There exists one and only one polynomial of approximation of degree m for any linear differential equation* (1) *under the conditions specified, and a necessary and sufficient condition that a given polynomial of degree m be a polynomial of approximation of this degree is that the number of intervals M_q exceed $m + 1$.*

PRINCETON UNIVERSITY,
September, 1906.