Since $\|U_{n_k}(\xi)\| > 0$, the sphere $\|z - U_{n_k}(\xi)\| \leq \|U_{n_k}(\xi)\|/2$, is non-vacuous. That such a sphere is a p-set was demonstrated in §3. The sphere K we were required to construct has therefore been shown to exist, and Theorem 3 is proved.

LAFAYETTE COLLEGE

ON THE APPROXIMATION OF FUNCTIONS BY SUMS OF ORTHONORMAL FUNCTIONS

EDWIN N. OBERG

1. **Introduction.** The main object of this paper is to derive, in a simple manner, upper bounds for the norms of the derivative of

$$(1) \qquad\qquad \sum_{i=0}^{n} a_i \phi_i(x)$$

in C and L^2 spaces, where the a_i are arbitrary constants, and $\{\phi_i(x)\}$ is any set of functions on a given finite or infinite interval (a, b). We apply our method, properly modified, first to the case where the $\phi_i(x)$ are characteristic solutions of conjugate sets of integral equations, then to other classes of functions whose first derivatives $\{\phi_i'(x)\}$ are orthogonal with respect to a weight function $\sigma(x)$. Finally, we apply our results to the question of convergence of sums[1] of type (1) that minimize

$$\int_a^b \rho(x) \left| f(x) - \sum_{i=0}^n a_i \phi_i(x) \right|^m dx, \qquad\qquad m > 0.$$

The leading results of our investigation may be summarized briefly as follows:

$$(A) \qquad \left| \frac{d}{dx} \sum_{i=0}^n a_i \phi_i(x) \right| \leq \lambda_n k(x) \left(\int_a^b \left[\sum_{i=0}^n a_i \phi_i(x) \right]^2 dx \right)^{1/2}$$

where λ_n is a positive number that increases with n and $k(x)$ is a func-

[1] For the specialized cases when the approximating functions are trigonometric sums or polynomials, see D. Jackson, *The theory of approximation*, Amer. Math. Soc. Colloquium Publications vol. 11, 1930, pp. 86–89, 96–101.

tion, independent of n, that is less than or equal to a bound M almost everywhere.

(B)
$$\left(\int_a^b \sigma(x) \left[\frac{d}{dx} \sum_{i=0}^n a_i\phi_i(x) \right]^2 dx \right)^{1/2}$$
$$\leqq N_n \left(\int_a^b \left[\sum_{i=0}^n a_i\psi_i(x) \right]^2 dx \right)^{1/2}$$

where $\{\psi_i(x)\}$ may be *any* set of orthonormal functions on (a, b), and N_n, for a given n and arbitrary coefficients a_i, is the least number that relates the left- and right-hand sides of (B).

Somewhat similar investigations have been made by Hille, Szegö, and Tamarkin,[2] McEwen,[3] and Shohat[4] but in each case with more restricted functions and by methods that are different from ours.

2. **Rough bounds for the derivative of sums of characteristic solutions.** Let $K(x, t)$ be a given continuous function on an interval $a \leqq x \leqq b$, $a \leqq \tau \leqq b$, and let

$$(2) \qquad \phi(x) = \lambda \int_a^b K(x, t)\psi(t)dt, \qquad \psi(x) = \lambda \int_a^b K(t, x)\phi(t)dt$$

be conjugate integral equations which have $K(x, t)$ as kernel. It is well known[5] that, except for the special case when $K(x, t) = \sum_{i=0}^n \phi_i(x)\psi_i(t)/\lambda_i$, there exists an infinite set of characteristic values for λ, and corresponding sets of characteristic functions,

$$\phi_0(x), \phi_1(x), \cdots, \phi_n(x), \cdots, \psi_0(x), \psi_1(x), \cdots, \psi_n(x), \cdots$$

that satisfy these equations. The set of functions $\{\psi_i(x)\}$ can be adjusted so that all of the characteristic numbers are positive and it can be assumed without loss of generality that $\lambda_0 \leqq \lambda_1 \leqq \cdots \leqq \lambda_n \leqq \cdots$. Moreover, the sets $\{\phi_i(x)\}$ and $\{\psi_i(x)\}$ can be regarded as separately orthonormal on (a, b). In addition to the hypotheses of continuity on $K(x, t)$ we assume differentiability with respect to x under the integral sign of $\int_a^b K(x, t)dt$ and that

[2] Hille, Szegö, and Tamarkin, *On some generalizations of a theorem of A. Markoff*, Duke Math. J. vol. 3 (1937) pp. 729–739.

[3] W. H. McEwen, *A note on an extension of Bernstein's theorem.* Amer. J. Math. vol. 60 (1938) pp. 309–319.

[4] J. Shohat. *On a general formula in the theory of Tschebycheff polynomials and its application*, Trans. Amer. Math. Soc. vol. 29 (1927), p. 569.

[5] See E. Schmidt. *Zur theorie der linearen und nichtlinearen Integralgleichungen*, Math. Ann. vol. 63 (1907) pp. 459–463.

$$\int_a^b \left[\frac{\partial}{\partial x} K(x,\, t) \right]^2 dt$$

as a function of x exists almost everywhere on $(a,\, b)$.

Let $a_0, a_1, \cdots, a_n$ be any set of coefficients and let

$$s_n(x) = \sum_{i=0}^{n} a_i \phi_i(x), \quad \bar{s}_n(x) = \sum_{i=0}^{n} a_i \psi_i(x), \quad \bar{s}_{n\lambda}(x) = \sum_{i=0}^{n} a_i \lambda_i \psi_i(x).$$

From (2) it follows that

$$\frac{d}{dx} s_n(x) = \int_a^b \frac{\partial}{\partial x} K(x,\, t) \bar{s}_{n\lambda}(t) dt$$

whence, as a consequence of Schwarz's inequality, we have

$$\left| \frac{d}{dx} s_n(x) \right| \leq \left(\int_a^b \left[\frac{\partial}{\partial x} K(x,\, t) \right]^2 dt \right)^{1/2} \left(\int_a^b \bar{s}_{n\lambda}^2(t) dt \right)^{1/2}.$$

By the property of orthonormality of the set $\{\psi_i(x)\}$, and the fact that the λ's increase with n,

$$\int_a^b \bar{s}_{n\lambda}^2(t) dt = \sum_{i=0}^{n} a_i^2 \lambda_i^2 \leq \lambda_n^2 \sum_{i=0}^{n} a_i^2 = \lambda_n^2 \int_a^b \bar{s}_n^2(t) dt.$$

The above relation, in view of the fact that the a's are arbitrary, is the best possible bound for $\int_a^b \bar{s}_{n\lambda}^2(t) dt$ in terms of $\int_a^b \bar{s}_n^2(t) dt$ since the equality holds when a_n is the only nonzero coefficient. However, the last integral may be replaced by the integral of the square of a sum of any $n+1$ orthonormal functions on the interval $(a,\, b)$ if the coefficients of this sum are $a_0, a_1, \cdots, a_n$. In particular, $\int_a^b \bar{s}_n^2(t) dt$ may be replaced by $\int_a^b s_n^2(t) dt$.

The previous discussion may be summarized as follows.

THEOREM A. *If the kernel $K(x,\, t)$ satisfies the hypotheses of the preceding discussion, and if $s_n(x)$ is any arbitrary sum of $\phi_0(x)$, $\phi_1(x)$, $\cdots$, $\phi_n(x)$, then*

$$\left| \frac{d}{dx} s_n(x) \right| \leq M \lambda_n \left(\int_a^b s_n^2(t) dt \right)^{1/2}$$

where M is a constant, over the parts of the interval $(a,\, b)$ for which $\int_a^b [\partial K(x,\, t)/\partial x]^2 dt$ is bounded. Moreover, if $(a,\, b)$ is finite and u_n is the maximum of $|s_n(x)|$ on $(a,\, b)$, then

(3) $$\left| \frac{d}{dx} s_n(x) \right| \leq M' \lambda_n u_n, \qquad M' = M(b-a)^{1/2}.$$

For special sets of characteristic functions and kernels, closer bounds then (3) are obtainable by other means[6] than ours. Thus, for example, if we consider the particular kernel[7]

$$K(x, t) = x(1 - t), \qquad\qquad 0 \leq x \leq t,$$
$$= t(1 - x), \qquad\qquad t \leq x \leq 1,$$

on the interval $0 \leq x \leq 1$, $0 \leq t \leq 1$, the integral equation $u(x) = \lambda \int_0^1 K(x, t)u(t)dt$ has as characteristic solutions and corresponding characteristic numbers the sets $\{\sin n\pi x\}$ and $\{n^2\pi^2\}$, respectively. The application of (3) gives a bound of order $n^2 u_n$ for the derivative of a sum of sine functions, which is rough when compared with the well known Bernstein's theorem for trigonometric sums.[8] However, the bound given by (3) is of interest since it applies to a very wide class of functions for which it seems impossible to obtain better results unless more restrictive hypotheses are assumed on the kernel or on the characteristic functions themselves.

3. **Bernstein's theorem in L^2 space for a sum of functions whose derivatives are orthogonal with respect to a weight function.** Hille, Szegö, and Tamarkin[9] have established the corresponding Markoff's theorem for L^m spaces when $m \geq 1$, but only when the sums involved are polynomials or trigonometric sums. Our method of proof is entirely different from theirs and we are not able to obtain the results of this section from their paper.

Let $\{\phi_i(x)\}$ be any set of functions, not necessarily orthonormal, but let the set of first derivatives of these functions be orthogonal with respect to a positive weight function $\sigma(x)$ on the interval (a, b). Assume that the normalization factor $N_i = (\int_a^b \sigma(x)[\phi_i'(x)]^2 dx)^{1/2}$ increases with the subscript i. Let $K_n(x, t) = \sum_{i=0}^n \phi_i(x)\psi_i(t)$ where $\{\psi_i(x)\}$ is any set of orthonormal functions on (a, b). Let $s_n(x)$ denote any arbitrary sum of the functions $\phi_0(x), \phi_1(x), \cdots, \phi_n(x)$, and let $s_n'(x)$ be its first derivative. Let $\bar{s}_n(x) = \sum_{i=0}^n a_i\psi_i(x)$ where $a_0, a_1, \cdots, a_n$ are the same coefficients as in $s_n(x)$.

Since

$$s_n'(x) = \int_a^b \frac{\partial}{\partial x} K_n(x, t)\bar{s}_n(t)dt$$

---

[6] See McEwen, op. cit. pp. 295–296.

[7] The kernel $K(x, t)$ in this instance is the Green's function of the system $d^2u/dx^2 + \lambda u = 0$, $u(0) = u(1) = 0$, and thus highly restricted when compared with any kernel of a system of form (2).

[8] See Jackson, op. cit. p. 80.

[9] Hille, Szegö, and Tamarkin, op. cit. Footnote 2.

it follows from the hypotheses placed on the sets $\{\phi_i'(x)\}$, $\{N_i\}$, and $\{\psi_i(x)\}$ that if this equation is multiplied by $\sigma(x)s_n'(x)$ and integrated over (a, b), then

$$\int_a^b \sigma(x)[s_n'(x)]^2 dx = \int_a^b \left[\int_a^b \sigma(x)s_n'(x)K_n'(x, t)dx\right] \bar{s}_n(t)dt$$

$$= \int_a^b \left[\sum_{i=0}^n a_i N_i^2 \psi_i(t)\right]\left[\sum_{i=0}^n a_i \psi_i(t)\right]dt$$

$$= \sum_{i=0}^n a_i^2 N_i^2 \leq N_n^2 \sum_{i=0}^n a_i^2 = N_n^2 \int_a^b \bar{s}_n^2(t)dt.$$

The factor N_n^2 is obviously the least number that relates the right- and left-hand members of the above inequality since the equality holds when all but a_n of the coefficients $a_1, a_2, \cdots, a_n$ are equal to zero. As a result we have this theorem.

THEOREM B. *If $\{\phi_i(x)\}$ is a given set of functions, and if there exists a positive weight function $\sigma(x)$ such that the set of functions*

$$\sigma^{1/2}(x)\phi_0'(x), \ \sigma^{1/2}(x)\phi_1'(x), \cdots, \sigma^{1/2}(x)\phi_n'(x)$$

are orthogonal on (a, b) and if $s_n(x) = \sum_{i=0}^n a_i\phi_i(x)$ is any arbitrary sum of the set $\phi_0(x), \phi_1(x), \cdots, \phi_n(x)$ and $\bar{s}_n(x) = \sum_{i=0}^n a_i\psi_i(x)$ is the sum of any arbitrary set of orthonormal functions on the interval (a, b), the coefficients $a_0, a_1, \cdots, a_n$ being the same in both $s_n(x)$ and $\bar{s}_n(x)$, then

(4) $$\left(\int_a^b \sigma(x)[s_n'(x)]^2 dx\right)^{1/2} \leq N_n \left(\int_a^b [\bar{s}_n(x)]^2 dx\right)^{1/2}$$

gives the closest relation between the integrals

$$\left(\int_a^b \sigma(x)[s_n'(x)]^2 dx\right)^{1/2} \quad and \quad \left(\int_a^b [\bar{s}_n(x)]^2 dx\right)^{1/2} when$$

$$N_n = \left(\int_a^b \sigma(x)[\phi_n'(x)]^2 dx\right)^{1/2}.$$

As a first application of the above theorem consider the case when both of the sets $\{\phi_i(x)\}$ and $\{\psi_i(x)\}$ are the normalized Legendre[10] polynomials $\{(2i+1/2)^{1/2}P_i(x)\}$. The set of first derivatives of these functions are orthogonal[11] with respect to the weight function $1-x^2$

---

[10] See, for example, D. Jackson, *Fourier series and orthogonal polynomials*, The Carus Monograph no. 6, pp. 45–68.

[11] See Jackson op. cit. Footnote 10 p. 149.

on the interval $(-1, 1)$ so that the hypothesis on $\{\phi_i'(x)\}$ holds in accord with the above theorem. Since

$$N_n^2 = \int_{-1}^{+1} (1 - x^2) [P_n'(x)]^2 dx = \frac{2n+1}{2} \cdot \frac{2(n+1)!}{(n-1)!} \cdot \frac{1}{2n+1} = n(n+1)$$

we have[12]

$$\left(\int_{-1}^{+1} (1 - x^2) [s_n'(x)]^2 dx\right)^{1/2} \leqq (n(n+1))^{1/2} \left(\int_{-1}^{+1} s_n^2(x) dx\right)^{1/2}$$

$$= O(n) \left(\int_{-1}^{+1} s_n^2(x) dx\right)^{1/2}.$$

Similar results can be obtained with other Jacobi, the Laguerre, and the Hermite polynomials.[13] In the last instance we take $\{\phi_i(x)\} = \{H_i(x)/(2\pi)^{1/4}(i!)^{1/2}\}$ and $\{\psi_i(x)\} = \{e^{-x^2/4} H_i(x)/(2\pi)^{1/4}(i!)^{1/2}\}$. Since $H_i'(x) = iH_{i-1}(x)$, then $\sigma(x) = e^{-x^2/2}$ makes the set of functions $\{\sigma^{1/2}(x)\phi_i'(x)\}$ orthogonal on $(-\infty, +\infty)$. The normalization factor $N_n = (\int_a^b \sigma(x) [\phi_n'(x)]^2 dx)^{1/2} = n^{1/2}$ so that

$$\left(\int_{-\infty}^{+\infty} e^{-x^2/2} [s_n'(x)]^2 dx\right)^{1/2} \leqq n^{1/2} \left(\int_{-\infty}^{+\infty} e^{-x^2/2} [s_n(x)]^2 dx\right)^{1/2}.$$

Another example is of interest in that it gives an application of the above theorem to functions which do not involve polynomials. Let $J_0(\mu_i x)$ denote the Bessel functions of zero order where $\{\mu_i\}$ are the roots[14] of the Bessel functions of the first order, $J_1(\mu) = 0$. Let $\phi_i(x) = J_0(\mu_i x)/k_i$ and $\psi_i(x) = x^{1/2} J_0(\mu_i x)/k_i$, where $k_i^2 = \int_0^1 x J_0^2(\mu_i x) dx = J_0^2(\mu_i)/2$. Since $J_0'(\mu_i x) = -\mu_i J_1(\mu_i x)$ it follows that the set of derivatives, $\{\phi_i'(x)\}$, are orthogonal on the interval $(0, 1)$ with respect to the weight function $\sigma(x) = x$. Thus

$$\int_0^1 x\phi_i'(x)\phi_j'(x) dx = (1/k_i k_j) \int_0^1 x J_0'(\mu_i x) J_0'(\mu_j x) dx$$

$$= (\mu_i \mu_j/k_i k_j) \int_0^1 x J_1(\mu_i x) J_1(\mu_j x) dx = 0$$

---

[12] It is of interest to compare this result with Bernstein's theorem for polynomials in C space. See D. Jackson, *Bernstein's theorem and trigonometric approximation*, Trans. Amer. Math. Soc. vol. 40 (1936) pp. 225–226.

[13] See Jackson op. cit. Footnote 10 p. 225 Problem 5; p. 227 Problem 7; p. 180 Formula 6.

[14] See, for example, Watson, *Theory of Bessel functions*, pp. 477–521, 576–596. We are making use of the well known formulas, $\int_0^1 x J_k^2(\mu x) dx = (1/2)[J_k'^2(\mu) + J_{k+1}^2(\mu)] - (k/\mu)[J_k(\mu) J_{k+1}(\mu)]$ and $\int_0^1 x J_k(\mu_i x) J_k(\mu_j x) dx = [J_k(\mu_j) J_k'(\mu_i) - J_k(\mu_i) J_k'(\mu_j)]/(\mu_j^2 - \mu_i^2)$ when $k = 0$ and $k = 1$.

when $i \neq j$. The normalization factor of $x^{1/2}J_1(\mu_n x)$ is $\left| J_2(\mu_n) \right|/2^{1/2}$ so that $N_n^2 = \int_0^1 x \left[\phi_n'(x) \right]^2 dx = \mu_n^2 J_2^2(\mu_n)/J_0^2(\mu_n)$. But by the well known[15] identity, $(2/x)J_1(x) = J_0(x) + J_2(x)$, it is seen that when μ_n is a root of $J_1(\mu) = 0$ then $J_0(\mu_n) = -J_2(\mu_n)$ so that the right-hand side of the above equation is μ_n^2 which is known[16] to be of the order n^2. Hence if

$$s_n(x) = \sum_{i=0}^{n} a_i \phi_i(x) = \sum_{i=0}^{n} a_i \frac{J_0(\mu_i x)}{k_i},$$

it follows that

$$\left(\int_0^1 x \left[s_n'(x) \right]^2 dx \right)^{1/2} \leq \mu_n \left(\int_0^1 x \left[s_n(x) \right]^2 dx \right)^{1/2}$$

$$= O(n) \left(\int_0^1 x \left[s_n(x) \right]^2 dx \right)^{1/2}.$$

4. **Degree of convergence of sums of characteristic solutions.** Let $\{\phi_i(x)\}$ and $\{\psi_i(x)\}$ be closed infinite sets of orthonormal solutions of a system of integral equations of form (2) in which the kernel $K(x, t)$ is continuous[17] on an interval $a \leq x \leq b$, $a \leq t \leq b$. Let the corresponding characteristic numbers be positive and arranged as a nondecreasing sequence with respect to the subscript i. Let $f(x)$ be a continuous function on $a \leq x \leq b$ and let

$$\sum_{i=1}^{\infty} \left(\lambda_i^p \int_a^b f(x) \phi_i(x) dx \right)^2$$

be convergent for some value of $p \geq 1$. Then sufficient hypotheses have been assumed so that the equation of the first kind

$$(5) \qquad f(x) = \int_a^b K(x, t) u(t) dt$$

has a solution $u(x)$ in L^2 space.[18]

Let $b_0, b_1, \cdots, b_n$ be arbitrary coefficients and let

$$\tau_{n\lambda}(x) = \frac{b_0 \phi_0(x)}{\lambda_0} + \frac{b_1 \phi_1(x)}{\lambda_1} + \cdots + \frac{b_n \phi_n(x)}{\lambda_n},$$

$$\bar{\tau}_n(x) = b_0 \psi_0(x) + b_1 \psi_1(x) + \cdots + b_n \psi_n(x).$$

[15] See Watson op. cit. p. 17. Take $n=1$.

[16] See Watson op. cit. p. 506. Take $\nu=0$, $\alpha=0$.

[17] This hypotheses on $K(x, t)$ can be lightened to include kernels that are summable and of summable square. See Footnote 18.

[18] See E. Picard, *Sur un théorème général relatif aux équations intégrales de première espèce et sur quelques problèmes de physiques mathématique*, Rend. Circ. Math. Palermo vol. 29 (1930) pp. 79–83.

By (2) it follows that $\tau_{n\lambda}(x) = \int_a^b K(x, t)\bar{\tau}_n(t)dt$ so that by subtracting $\tau_{n\lambda}(x)$ from $f(x)$ we have with the aid of Schwarz's inequality,

$$(6) \quad |f(x) - \tau_{n\lambda}(x)| \leq \left(\int_a^b K^2(x, t)dt\right)^{1/2} \left(\int_a^b [u(t) - \bar{\tau}_n(t)]^2 dt\right)^{1/2}.$$

The above expression shows that any set of the b's that will make $\bar{\tau}_n(t)$ converge in the mean to $u(t)$ will make $\tau_{n\lambda}(x)$ converge uniformly to $f(x)$ as n becomes infinite. In particular, if $b_i = \int_a^b u(t)\psi_i(t)dt$, then

$$\int_a^b [u(t) - \bar{\tau}_n(t)]^2 dt = \sum_{i=n+1}^{\infty} \left(\int_a^b u(t)\psi_i(t)dt\right)^2$$

$$= \sum_{i=n+1}^{\infty} \left(\lambda_i \int_a^b f(t)\phi_i(t)dt\right)^2.$$

But this in turn is at most equal to

$$(7) \qquad \frac{1}{\lambda_{n+1}^{2p-2}} \sum_{i=n+1}^{\infty} \left(\lambda_i^p \int_a^b f(t)\phi_i(t)dt\right)^2$$

for any $p \geq 1$ when n is sufficiently large so that $\lambda_{n+1} > 1$.

From (6) and (7) we have the following consequence.

THEOREM C. *If $\{\phi_i(x)\}$ and $\{\psi_i(x)\}$ are closed infinite sets of orthonormal solutions of the system of equations*

$$\phi(x) = \lambda \int_a^b K(x, t)\psi(t)dt, \qquad \psi(x) = \lambda \int_a^b K(t, x)\phi(t)dt$$

corresponding to a set of positive characteristic numbers $\lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_n \leq \cdots$, and if $K(x, t)$ is continuous on $a \leq x \leq b$, $a \leq t \leq b$, then, if $f(x)$ is continuous on $a \leq x \leq b$ and

$$\sum_{i=0}^{\infty} \left(\lambda_i^p \int_a^b f(x)\phi_i(x)dx\right)^2$$

is convergent for some $p \geq 1$, there exist sums $q_n(x)$ of the set $\{\phi_i(x)\}$ such that

$$|f(x) - q_n(x)| \leq \frac{I_1 \nu_n}{\lambda_n^{p-1}}$$

where I_1 is an absolute constant and $\lim_{n=\infty} \nu_n = 0$.

In connection with Theorem C a further comment is of interest if in (5) we can differentiate $f(x)$ and $\int_a^b K(x, t)dt$ s times, the latter under the integral sign, and if the set of functions $\{\phi^{(s)}(x)\}$ are orthogonal

with respect to a positive weight function $\sigma(x)$ on (a, b). In that case it follows if $f^{(s)}(x)$ is multiplied by $\sigma(x)\phi_i^{(s)}(x)$ and integrated over (a, b),

$$\int_a^b \sigma(x)f^{(s)}(x)\phi_i^{(s)}(x)dx = \int_a^b \left[\int_a^b \sigma(x)K^{(s)}(x, t)\phi_i^{(s)}(x)dx \right] u(t)dt$$

$$= \frac{\overline{N}_i^2}{\lambda_i} \int_a^b \psi_i(t)u(t)dt$$

where $\overline{N}_i$ is the normalizing factor of $\sigma^{1/2}(x)\phi_i^{(s)}(x)$. The series $\sum_{i=0}^\infty ((\overline{N}_i/\lambda_i)\int_a^b \psi_i(t)u(t)dt)^2$ converges since

$$\sum_{i=0}^\infty \left(\int_a^b \sigma^{1/2}(x)f^{(s)}(x)\sigma^{1/2}(x) \frac{\phi_i^{(s)}(x)}{\overline{N}_i} dx \right)^2$$

is convergent by virtue of Bessel's inequality. But if $\lim_{i=\infty} \overline{N}_i/\lambda_i$ is not zero, then $\sum_{i=0}^\infty (\int_a^b \psi_i(t)u(t)dt)^2$ converges and as a consequence $\sum_{i=0}^\infty (\lambda_i^p \int_a^b f(x)\phi_i(x)dx)^2$ is convergent for some $p \geq 1$.

5. **Convergence of the minimizing sums of the integral**[19] $g_n = \int_a^b \rho(x)|f(x) - Q_n(x)|^m dx$, $m > 0$ **when** (a, b) **is a finite interval.** Let $f(x)$ be a continuous function on (a, b), and let the weight function $\rho(x)$ be summable and have a positive lower bound v on (a, b). Let $\{\phi_i(x)\}$ be a set of characteristic solutions of conjugate sets of integral equations of form (2). Let these solutions correspond to a set of characteristic numbers $\{\lambda_i\}$ that are in the order $0 < \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_n$. Let $\int_a^b K(x, t)dt$ be differentiable with respect to x under the integral sign and let $\int_a^b [\partial K(x, t)/\partial x]^2 dt$ as a function of x exist almost everywhere. Let $Q_n(x) = \sum_{i=0}^n a_i\phi_i(x)$ be any sum of the first $n+1$ functions of the set $\{\phi_i(x)\}$, and let $\Phi_n(x)$ denote any one[20] of these that makes g_n the least. Let ϵ_n be the maximum for $|f(x) - Q_n(x)|$ on (a, b). By following the steps Jackson has taken to investigate convergence of minimizing trigonometric sums and polynomials[21] we are led, by virtue of (3) of Theorem A to the following result.

THEOREM D. *If $Q_n(x)$ is an arbitrary sum of the first $n+1$ functions of the set $\{\phi_i(x)\}$, and $\Phi_n(x)$ is a corresponding minimizing sum of*

---

[19] For a proof of the existence of minimizing sums see D. Jackson, *On functions of closest approximation*, Trans. Amer. Math. Soc. vol. 22 (1921) pp. 117–128. See also vol. 25 (1923) pp. 333–337.

[20] If $m \geq 1$, the minimizing function is unique. See Jackson op. cit. Footnote 19.

[21] See Jackson op. cit. Footnote 1.

g_n *there exists a positive factor* L *independent of the characteristic numbers* $\{\lambda_i\}$ *and* ϵ_n *such that*

$$| f(x) - \Phi_n(x) | \leq L\lambda_n^{1/m} \epsilon_n.$$

Moreover, since L *depends upon* $(\int_a^b [\partial K(x, t)/\partial x]^2 dt)^{1/2}$, *it is bounded on* $a \leq x \leq b$ *except perhaps on a set of points of measure zero.*

Thus, since ϵ_n is the maximum of $|f(x) - Q_n(x)|$ on (a, b), the question of convergence of $\Phi_n(x)$ to $f(x)$ on the intervals of $a \leq x \leq b$, where L exists, is made to depend upon whether sums $Q_n(x)$ exist for an infinite sequence of values of n such that

$$\lim_{n = \infty} \lambda_n^{1/m} \epsilon_n = 0.$$

This is insured by Theorem C if the series

$$\sum_{i=0}^{\infty} \left(\lambda_i^p \int_a^b f(x)\phi_i(x)dx \right)^2$$

is convergent for some $p \geq (1/m) + 1$.

Theorem D can be strengthened[22] considerably through the application of Theorem B if in addition to the property of orthonormality on the set $\{\phi_i(x)\}$, the set of derivatives $\{\phi_i'(x)\}$ are orthogonal on (a, b) with respect to a continuous non-vanishing weight function $\sigma(x)$. The essential point is that λ_n is usually large in comparison with the normalizing factor N_n, as for example in the integral equation that was cited earlier (p. 71). Here $\lambda_n = O(n^2)$ while N_n of the set of derivatives of $\{\sin nx\}$, that is $\{n \cos nx\}$, is of order n.

In the subsequent discussion we shall also assume that $\int_a^b |d\sigma^{1/2}(x)/dx| dx$ exists. Using the same definitions and notations as in the preceding discussion, we let $\pi_n(x) = \Phi_n(x) - Q_n(x)$ and $r_n(x) = f(x) - Q_n(x)$. Let η be the maximum of $|\pi_n(x)|$ on (a, b) and let x_0 be the abscissa at which this maximum is attained. Since

$$\sigma^{1/2}(x)\pi_n(x) - \sigma^{1/2}(x_0)\pi_n(x_0)$$

$$= \int_{x_0}^x \frac{d}{dx} [\sigma^{1/2}(x)\pi_n(x)] dx = \int_{x_0}^x \sigma^{1/2}(x)\pi_n'(x)dx + \int_{x_0}^x \pi_n(x)\frac{d}{dx}\sigma^{1/2}(x)dx$$

for any x on the interval (a, b), we have with the aid of Schwarz's inequality,

---

[22] In this section of the paper we can extend our discussion to the case when the interval (a, b) is infinite. The treatment requires further hypotheses on the functions $\sigma(x)$, $\{\phi_i(x)\}$ and $\{\phi_i'(x)\}$. In general the details are similar and are therefore omitted.

$$\left| \sigma^{1/2}(x)\pi_n(x) - \sigma^{1/2}(x_0)\pi_n(x_0) \right|$$

$$\leqq \left| x - x_0 \right|^{1/2} \left(\int_{x_0}^{x} \sigma(x) \left[\pi_n'(x) \right]^2 \left| dx \right| \right)^{1/2}$$

$$+ \int_{x_0}^{x} \left| \pi_n(x) \right| \left| \frac{d}{dx} \sigma^{1/2}(x) \right| \left| dx \right|.$$

But since x_0 and x are on (a, b) by Theorem B it follows that

$$\left(\int_{x_0}^{x} \sigma(x) \left[\pi_n'(x) \right]^2 \left| dx \right| \right)^{1/2} \leqq \left(\int_{a}^{b} \sigma(x) \left[\pi_n'(x) \right]^2 dx \right)^{1/2}$$

$$\leqq N_n \left(\int_{a}^{b} \pi_n^2(x) dx \right)^{1/2}$$

so that

$$\left| x - x_0 \right|^{1/2} \left(\int_{x_0}^{x} \sigma(x) \left[\pi_n'(x) \right]^2 \left| dx \right| \right)^{1/2}$$

(8) $$+ \int_{x_0}^{x} \left| \pi_n(x) \right| \left| \frac{d}{dx} \sigma^{1/2}(x) \right| \left| dx \right|$$

$$\leqq \left| x - x_0 \right|^{1/2} N_n (b - a)^{1/2} \eta + \left| x - x_0 \right| \eta H$$

where H is a constant independent of x and N_n. If N_n is large and

$$\left| x - x_0 \right| < \frac{\sigma(x_0)}{16 N_n^2 (b - a)}$$

then

$$\left| \sigma^{1/2}(x)\pi_n(x) - \sigma^{1/2}(x_0)\pi_n(x_0) \right|$$

$$\leqq \frac{\eta}{4} \sigma^{1/2}(x_0) + \frac{\eta H \sigma(x_0)}{16 N_n^2 (b - a)} < \frac{\eta}{2} \sigma^{1/2}(x_0),$$

whence $\left| \sigma^{1/2}(x)\pi_n(x) \right| > (\eta/2)\sigma^{1/2}(x_0)$. If w is the maximum of $\sigma^{1/2}(x)$ on (a, b), then $\left| \pi_n(x) \right| > (\eta/2)\sigma^{1/2}(x_0)/w$.

Let γ_n be the minimum of g_n. For a value of x on (a, b) at a distance not greater than $\sigma(x_0)/16 N_n^2 (b-a)$ from x_0,

$$\gamma_n > v \left| \pi_n(x) - r_n(x) \right|^m \frac{\sigma(x_0)}{16 N_n^2 (b - a)}.$$

Thus, for the case when $(\eta/4)(\sigma^{1/2}(x_0)/w) > \epsilon_n$ we have, since $\gamma_n < I_2 \epsilon_n^m$, $I_2 = \int_a^b \rho(x) dx$,

$$(9) \qquad \eta < \frac{4w}{\sigma^{1/2}(x_0)} \left(\frac{16[b-a]}{v} \gamma_n N_n^2 \right)^{1/m} < s_1 N_n^{2/m} \epsilon_n$$

where s_1 is an absolute constant.

On the other hand, if $(\eta/4)(\sigma^{1/2}(x_0)/w) < \epsilon_n$, we have

$$\eta < \frac{4w}{\sigma^{1/2}(x_0)} \epsilon_n$$

so that for an N_n sufficiently large, (9) is a universal bound for η.

Hence, from the inequality

$$| f(x) - \Phi_n(x) | \leq | f(x) - Q_n(x) | + | \Phi_n(x) - Q_n(x) |$$
$$= | r_n(x) | + | \pi_n(x) |$$

we have the following result.

THEOREM E. *If in addition to the hypotheses placed on* $\rho(x), f(x),$ *and the set* $\{\phi_i(x)\}$ *in Theorem* D *it is assumed that the set of functions* $\{\phi_i'(x)\}$ *are orthogonal with respect to a continuous positive weight function* $\sigma(x)$, *and if* $\int_a^b | d\sigma^{1/2}(x)/dx | dx$ *exists, then there exists a constant* s_2 *independent of* N_n, ϵ_n, *and also of* x *such that*

$$| f(x) - \Phi_n(x) | \leq s_2 N^{2/m} \epsilon_n.$$

Theorem E is subject to further refinement if certain further hypotheses are assumed on the set $\{\phi_i(x)\}$. If the set of functions $\{\phi_i'(x)\}$ are continuous on the interval (a, b), (the condition of orthogonality only requires that the set $\{\phi_i'(x)\}$ with respect to $\sigma(x)$ be integrable L^2), and if at the maximum of $| \pi_n(x) |$, $d\pi_n(x)/dx = 0$, then by a law of the mean[23] for integrals it is seen that

$$\frac{\int_{x_0}^x \sigma(x) [\pi_n'(x)]^2 | dx |}{| x - x_0 |} \leq \frac{\int_a^b \sigma(x) [\pi_n'(x)]^2 dx}{b - a}$$

when x is sufficiently close to x_0. Thus the first integral on the left-hand side of (8) can be replaced by

$$\frac{| x - x_0 |^{1/2}}{(b-a)^{1/2}} \left(\int_a^b \sigma(x) [\pi_n'(x)]^2 dx \right)^{1/2}$$

---

[23] By the law of the mean, there exist points ξ and c on the intervals $| x, x_0 |$ and (a, b), respectively, such that $\sigma(\xi) [\pi_n'(\xi)]^2 = (\int_{x_0}^x \sigma(x) [\pi_n'(x)]^2 | dx |)/| x - x_0 |$ and $\sigma(c) [\pi_n'(c)]^2 = (\int_a^b \sigma(x) [\pi_n'(x)]^2 dx)/(b-a)$. Since $\pi'(x)$ is continuous and $\pi_n'(x_0) = 0$ it follows, since $\sigma(x)$ is positive and bounded on (a, b), that $\sigma(\xi) [\pi_n'(\xi)]^2$ can be made smaller than $\sigma(c) [\pi_n'(c)]^2$ by taking x sufficiently close to x_0.

and this leads to a bound of order $N^{1/m}\epsilon_n$ for $|f(x) - \Phi_n(x)|$ in contrast with $N^{2/m}\epsilon_n$ of Theorem E. As an example, a relative maximum of $|\pi_n(x)|$ is assured if the following boundary conditions[24] are assumed on $\Phi_n(x)$: $\Phi_n(a) = Q_n(a)$, and $\Phi_n(b) = Q_n(b)$. In that case $\pi_n(a) = \pi_n(b) = 0$ and as a consequence of the hypotheses on the functions $\{\phi_i'(x)\}$, $\pi_n'(x) = 0$ at the maximum of $|\pi_n(x)|$.

The hypothesis on $\sigma(x)$ may be lightened to include weight functions that vanish at some points along the interval. The only restriction required by the proof is that $\sigma(x)$ shall not vanish at the point where $|\pi_n(x)|$ takes on its maximum, and this in turn is assured by appropriate boundary conditions on the minimizing function $\Phi_n(x)$, such as were mentioned in the last paragraph. This last observation is of importance in working out details for the convergence of sums in terms of the Bessel, Legendre, Laguerre, and the Hermite functions, among a great number of others that involve weight functions. The details of these discussions follow the above general outline and are therefore omitted.

THE UNIVERSITY OF IOWA

---

[24] The placing of a boundary condition on $\Phi_n(x)$ does not affect the existence of minimizing sums. See for example, D. Jackson, *Problems of approximation with integral boundary conditions*, Amer. J. Math. vol. 55 (1933) pp. 153–166.