NOTES ON GENERALIZED FOURIER SERIES WITH APPLICATION TO GRAVITATIONAL FIELD DETERMINATION

by

Walter Thomas Blackshear

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute

in candidacy for the degree of

MASTER OF SCIENCE

in

MATHEMATICS

APPROVED: ______________________________
Chairman

______________________ ______________________

______________________ ______________________

July 1967

Blacksburg, Virginia

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

SYMBOLS AND NOTATIONS

$a \in A$	a is an element of the set A
$m(A)$	the measure of the set A
$M \subseteq A$	the set M is a subset of the set A
$M \subset A$	the set M is a proper subset of the set A
$A \cap B$	the intersection of the sets A and B
$A \setminus B$	the complement of the set B with respect to the set A
C_{L^2}	set of continuous square integrable functions
L^2	set of square integrable functions
ρ	denotes a metric or a weight function
$\{\varphi_n\}$	the set of functions $\varphi_1, \varphi_2, \ldots$
$\|f\|$	the norm of the function f
(f, g)	the inner product of the functions f and g
$\Gamma(\alpha)$	the Gamma function
$p_n(x)$	an orthogonal polynomial of degree n
$p_n^{(\alpha,\beta)}(x)$	a normed Jacobi polynomial
$J_n^{(\alpha,\beta)}(x)$	a non-normed Jacobi polynomial
$P_n(x)$	a non-normed Legendre polynomial
$C_n(x)$	a non-normed Chebysheff polynomial
$B_n(x)$	a non-normed polynomial defined by equation 7.1d
$K_n(x,t)$	the n^{th} kernel of an orthogonal system
$L_n(x)$	the n^{th} Lebesque function of an orthogonal system

$(B)_{lj}$	denotes the l,j element of the matrix B
$p_n'(x)$	$\frac{dp_n(x)}{dx}$
inf	greatest lower bound
sup	least upper bound
$\left.\frac{\partial f(t_i)}{\partial \bar{A}}\right\rvert_{\bar{A} = \bar{A}_0}$	the row vector of partial derivatives of f at time t_i with respect to the parameters in the vector $\bar{A}$, evaluated at $\bar{A} = \bar{A}_0$

INTRODUCTION

The purpose of study is to investigate the conditions under which an arbitrary member of a certain class of functions may be associated with a series and in what manner the associated series represents the given function. The class of functions is taken to be L^2, the set of all functions whose square is integrable in the Lebesgue sense. The method of series association is that of generalized Fourier series defined as follows. Let $\{\varphi_n(x)\}$ be an orthonormal system in L^2 and let $f \in L^2$. Then the generalized Fourier series of f with respect to $\{\varphi_n(x)\}$ is the series whose general term is $a_n \varphi_n(x)$, where a_n is the inner product of f and $\varphi_n(x)$ over the interval of orthogonality of the system $\{\varphi_n(x)\}$.

The relationship is denoted

$$f \sim \sum_{n=1}^{\infty} a_n \varphi_n(x)$$

without regard to convergence.

The first five sections of this paper consists of background material, a proof of the existence of a complete orthonormal system in L^2 and a discussion of some of the properties of generalized Fourier series corresponding to rather general orthogonal systems.

Sections 6 and 7 restrict the discussion to systems of orthogonal polynomials arising from the Gram-Schmidt orthonormalization of the system $\{x^n\}$, in the interval $[-1, 1]$, with respect to a given weight function. The Jacobi polynomial systems are treated as a subset of orthogonal polynomial systems.

The remainder of the paper consists of considering the solution to the potential equation, with boundary conditions, in terms of Jacobi polynomials. Under simplifying assumptions this problem is solved in terms of three special Jacobi systems. The method of solution is least-squares, differential correction, applied to discrete observations of the boundary conditions. This method is analogous to the techniques frequently used in gravitational field determination from observations of an artificial satellite.

1. INTRODUCTORY MATERIAL

It is assumed that the applicability of series expansion for arbitrary functions to the field of applied mathematics is well known. For the interested reader an excellent brief historical summary of Fourier series is presented in the Carslaw reference. The purpose of this paper is to investigate the conditions under which an arbitrary member of a certain class of functions may be associated with a series and in what manner the associated series represents the given function. Interest will be in defining a class of functions and selecting from this class a proper subset which may be combined in such a way as to approximate all members of the class. Consideration will be limited to real valued functions of a real variable.

When dealing with a certain class of functions having a common domain of definition and certain specified properties it is conceptually helpful to consider this class as a space and the individual functions as elements or points in this space. Assume for the present that a function space is given. A method of combining elements of this space is defined as follows.

Definition 1.1: A linear combination of a set $f_1, f_2, \cdots$ of functions defined on a set A is an expression of the form:

$$a_1 f_1 + a_2 f_2 + \cdot\cdot$$

where $a_1, a_2, \cdots$ are real numbers and addition and multiplication are defined:

$$\left(a_1 f_1 + a_2 f_2 + \cdots\right) (x) = a_1 f_1(x) + a_2 f_2(x) + \cdots .$$

This definition permits construction of series of the form $\sum_{n=1}^{\infty} a_n f_n(x)$. Since the aim is to associate series of this form with functions of the given space, the concept of a function space being closed under finite linear combinations is introduced. The following theorem provides conditions under which a function space is also a linear space.

Theorem 1.1: A function space S is a linear space if and only if (1) whenever $f, g \in S$ so is $f + g$ and (2) whenever $f \in S$, so is af for any real number a.

It is generally desirable for the class of functions under consideration to constitute a linear space since any finite combination yields a function having the same properties as the original elements of the space.

Theorem 1.1 gives an indication of the possible existence of a proper subset of a linear space S such that any member of S may be expressed as a linear combination of the elements of this subset. If such a subset exists its elements would be likely candidates for use in series expansions of the remaining functions in S. However, there are additional properties which will prove desirable for an expansion set. These properties are introduced by the following definitions.

Definition 1.2: The system of functions $f_1, f_2, \cdots, f_n$ are linearly dependent if there exists constants $c_1, c_2, \ldots, c_n$, not all zero, such that

$$c_1 f_1 + \cdots + c_n f_n = 0$$

If, however, this equality implies $c_i = 0$, $i = 1, 2, \cdots, n$, the system is linearly independent.

Definition 1.3: The infinite system of functions $f_1, f_2, \cdots$ is linearly independent if every finite subset is linearly independent.

Definition 1.4: If $f_1, f_2, \cdots$ are elements of a linear function space S, the subspace of S which consists of all linear combinations of these functions is called the linear closure of the functions $f_1, f_2, \cdots$ denoted $M\{f_n\}$. The closure of the set $M\{f_n\}$ is called the closed linear closure and is denoted $\overline{M}\{f_n\}$.

Definition 1.5: The system of functions $\{f_n\}$ is called complete if $\overline{M}\{f_n\}$ is the whole space.

In light of these definitions a more appropriate set of functions for expansion purposes would be any linearly independent complete set of functions in the space S since it is a minimal set which spans S.

In order to state how well a series represents a given function a concept of distance is needed, the properties of which are now defined.

Definition 1.6: A distance functional is a real valued function defined for pairs of functions f and g of a space S, denoted d(f,g), which has the following properties:

1. $d(f,g) \geq 0$
2. $d(f,g) = d(g,f)$
3. $d(f,f) = 0$
4. $d(f,h) \leq d(f,g) + d(g,h)$.

A space with a distance function is called a distance space or a pseudo-metric space. If the distance function also satisfies:

5. $d(f,g) > 0$ for $f \neq g$

then the space is called a metric space.

This definition of distance also incorporates a concept of magnitude which is made precise by the following definition.

Definition 1.7: Let f be an element of a linear distance space S. The distance from the 0 element of S to f is called the norm and denoted $\|f\|$, that is; $d(0,f) = \|f\|$.

This brief introduction of general concepts provides a basis for describing specific examples which will have application to the theory of generalized Fourier series.

2. THE NORMED LINEAR FUNCTION SPACE L^2

The class of functions which will be of primary interest is the set of all measurable real valued functions f defined on a bounded measurable subset A of the real continuum such that $f^2(x)$ is integrable (summable) in the Lebesgue sense over A. The set A will generally be restricted to be a finite interval $[a, b]$. This class of functions shall be denoted by L^2. That L^2 is a linear space is proven on the strength of theorem 1.1 and the following two theorems whose proofs follow from the definition of the Lebesgue integral.

<u>Theorem 2.1</u>: If $f, g \in L^2$ then $f + g \in L^2$.

<u>Theorem 2.2</u>: If $f \in L^2$ and a is an arbitrary real number, then $a\,f(x) \in L^2$.

The distance function for L^2 is defined:

Definition 2.1: $$d(f,g) = \left\{ \int_A \left[f(x) - g(x)\right]^2 dx \right\}^{1/2}$$

It follows from the properties of the Lebesgue integral that this function satisfies the properties of a distance function as enumerated in definition 1.6, although it does not constitute a metric if one distinguishes between functions which differ on a set of zero Lebesque measure. In order to avoid this difficulty two functions f and g will be considered equivalent in A if the subset of A upon which f and g differ is of zero measure. Distinction will not be made between equivalent functions, so that each point in L^2 becomes a class of equivalent functions. In particular the zero function θ is the class of all functions f such that $f(x) = 0$ almost everywhere in A.

The norm for the space L^2 is defined:

Definition 2.2: If $f \in L^2$ the norm of f is:

$$\|f\| \equiv \left\{ \int_A \left[f(x)\right]^2 dx \right\}^{1/2} .$$

Additional concepts which will later prove extremely useful are those of a "scalar" or "inner product," orthogonality, and normality.

Definition 2.3: The inner product of $f, g \in L^2$ is:

$$(f,g) = \int_A f(x)\, g(x)\, dx .$$

Definition 2.4: If $f, g \in L^2$, they are orthogonal on A if and only if $(f,g) = 0$.

Definition 2.5: If $f \in L^2$, f is normal if and only if $\|f\| = 1$.

Definition 2.6: A set of functions $\{\varphi_n\}$ which are both orthogonal and normal will be called an orthonormal set of functions.

Provided now with the linear metric space L^2, definitions of norm and orthogonality, the discussion is directed to the problem of associating with each element in L^2 a series expansion in terms of a subset of L^2.

3. DEFINITION OF GENERALIZED FOURIER SERIES

Assume for the present that a function $f \in L^2$ and an orthonormal system of continuous functions $\{\varphi_n(x)\} \subset L^2$ are given and it is known that $f(x)$ can be expanded in a uniformly convergent series,

$$\sum_{n=1}^{\infty} a_n \varphi_n(x) ,$$

with respect to $\{\varphi_n(x)\}$. For this special case a procedure can be described for determining the constant coefficients a_n.

By hypothesis

$$f(x) = \sum_{n=1}^{\infty} a_n \varphi_n(x) .$$

Multiplication by an arbitrary $\varphi_n(x)$ does not affect the property of uniform convergence. It is also known from the theory of infinite series that termwise integration is permitted. Thus

$$\int_A f(x) \varphi_m(x) dx = \sum_{n=1}^{\infty} a_n \int_A \varphi_m(x) \varphi_n(x) dx .$$

Due to the orthonormality of the $\{\varphi_n(x)\}$ this equation reduces to

$$\int_A f(x) \varphi_m(x) dx = a_m . \qquad (3.1)$$

The coefficient a_m determined by this formula is called the Fourier coefficient with respect to the orthonormal system $\{\varphi_n(x)\}$.

The hypotheses of the preceeding discussion severely limit the class of functions which can be expanded in a series having coefficients given by equation (3.1). In particular the requirement that

$$\sum_{n=1}^{\infty} a_n \varphi_n(x)$$

converge uniformly forces $f(x)$ to be continuous. However, this example does provide some motivation for extending the association to arbitrary elements of L^2.

Definition 3.1: Let $\{\varphi_n(x)\}$ be an orthonormal system in L^2 and let $f \in L^2$. Then the generalized Fourier series of f with respect to $\{\varphi_n(x)\}$ is the series whose general term is $a_n \varphi_n(x)$ where a_n is the Fourier coefficient (f, φ_n). f is said to be expanded in a generalized Fourier series with respect to $\{\varphi_n(x)\}$ and the relationship is denoted,

$$f \sim \sum_{n=1}^{\infty} a_n \varphi_n$$

without regard to convergence.

Under definition 3.1 the existence of the associated series is guaranteed if the system $\{\varphi_n(x)\}$ exists and φ_n is Lebesque integrable, $n = 1, 2, \cdots$. Since by the Hölder inequality if $f \in L^p$, for the integral $\int_A f(x) \varphi_n(x)\, dx$ to have meaning it is necessary and sufficient for $\varphi_n \in L^q$ where $\frac{1}{p} + \frac{1}{q} = 1$. (ref. 1, page 18.) However, there is no a priori reason why the series in definition 3.1 should converge, or if it converges, should represent the given function f. Thus there are three fundamental areas of investigation; (1) the existence of a denumerable, orthonormal complete set of functions in L^2, (2) convergence of the associated generalized Fourier series and (3) the way in which a convergent series represents the associated function.

4. EXISTENCE OF A COMPLETE ORTHONORMAL SYSTEM IN L^2

Proof of the existence of a complete orthonormal system $\{\varphi_n\}$ in L^2 will be conducted along the following lines.

1. Prove the set of continuous functions in L^2, denoted C_L2, is separable; that is, contains a denumerable everywhere dense subset.
2. Prove that C_L2 is dense in L^2.
3. Prove that if a set E contains an everywhere dense separable subset E_1, then E is separable, and thus conclude that L^2 is separable.
4. Prove that given an everywhere dense denumerable subset in L^2 a complete orthonormal system can be constructed in L^2.

The proof of proposition 1 begins with the following lemma.

Lemma 4.1: The space C of all continuous functions f on the finite interval $[a, b]$, with metric

$$\rho_c\ (f,g) \equiv \max\ |f(x) - g(x)|;\ x \epsilon [a, b]$$

is separable.

Proof: By a well known theorem due to Weierstrass (ref. 2, page 571) every function $f(x)$ continuous on a finite interval may be represented as a limit of a uniformly convergent sequence of polynomials

$$Q_n(x) = \sum_{i=0}^{n} a_i\ x^i$$

where a_i is a real number. Note that uniform convergence is convergence in the metric of C. Thus for any $\epsilon > 0$ a $Q_n(x)$ can be found such that

$$\rho_c\ \left(f,\ Q_n\right) < \frac{\epsilon}{2}\ .$$

Take as a denumerable set the set D of all polynomials

$$P_n(x) = \sum_{i=0}^{n} r_i \, x^i$$

where the r_i are rational numbers. Since the rational numbers are dense in the real numbers a $P_n(x)$ exists such that

$$\rho_c \left(Q_n, P_n\right) < \frac{\epsilon}{2} .$$

Therefore,

$$\rho_c \left(f, P_n\right) \leq \rho_c \left(f, Q_n\right) + \rho_c \left(Q_n, P_n\right) < \epsilon .$$

Hence the denumberable set D is dense in C and therefore C is separable.

For the space $C_L 2$ the metric is:

$$\rho_{C_L 2} \left(f, g\right) = \left\{ \int_a^b \left[f(x) - g(x)\right]^2 dx \right\}^{1/2}$$

If f and g are continuous on a finite interval, f - g is bounded there. Let $M = \max \left|f(x) - g(x)\right|$. Then

$$\rho_{C_L 2} (f, g) \leq \left[M^2 (b - a)\right]^{1/2} = M(b - a)^{1/2} .$$

Now if $g \to f$ in the sense of the metric ρ_c, this implies $M \to 0$. Hence $g \to f$ in the sense of the metric $\rho_{C_L 2}$. Thus the set D is dense in $C_L 2$ and therefore $C_L 2$ is separable.

The proof of proposition 2 will be initiated with the introduction of the Lebesque integral in terms of simple functions, that is, measurable functions which take on a finite or countable number of values.

Definition 4.1: Let $f(x)$ be a simple function which assumes the values $y_1, y_2 \ldots, y_n \ldots$. The integral of $f(x)$ over A is defined by the equation

$$\int_A f(x)\,dx \equiv \sum_{n=1}^{\infty} y_n\, m\left\{x : x \epsilon A,\ f(x) = y_n\right\} \qquad (4.1)$$

$f(x)$ is called integrable over A if the sequence of partial sums of the series 4.1 converges absolutely. If $f(x)$ is integrable, then the sum 4.1 is called the integral of $f(x)$ over A.

Definition 4.2: The function $g(x)$ is integrable over the set A if there exists a sequence of simple functions $f_n(x)$ which are integrable over A and converge uniformly to $g(x)$. Thus

$$\int_A g(x)\,dx = \lim_{n \to \infty} \int_A f_n(x)\,dx\ .$$

It was previously shown that uniform convergence implied convergence in the metric of C_L^2, which is also the metric of L^2. Thus it follows immediately from definition 4.2 that the simple functions belonging to L^2 are dense in L^2.

Now let A be a metric space having a measure which satisfies the condition that all open and all closed sets in A are measurable and for any set $M \subseteq A$, $m(A) = \inf_{M \subseteq G} m(G) = \sup_{F \subseteq M} m(F)$, where the lower bound is taken over all open sets G containing M and the upper bound is taken over all closed sets F which are contained in M. These conditions are satisfied by the Lebesgue measure. (ref. 3, chapter 8.) Proposition 2 is restated and proven as follows.

<u>Theorem 4.1</u>: The set C_L^2 is dense in L^2. (ref. 4, page 98.)

Proof: It is first shown that any simple function $f \in L^2$, and hence any function in L^2, can be approximated as closely as desired by simple functions which take on a finite number of values.

Let $f(x)$ take on the values $y_1, \cdots, y_n, \cdots$ on the sets $E_1, \cdots, E_n, \cdots$. Since f is integrable, the series

$$\sum_n y_n^2 \, m(E_n) = \int_A f^2(x) \, dx$$

converges. Thus for any $\epsilon > 0$ there exists on N such that

$$\sum_{n > N} y_n^2 \, m(E_n) < \epsilon .$$

Let

$$f_N(x) = \begin{cases} f(x) & \text{for} \quad x \in E_i \, , \, i \leq N \\ 0 & \text{for} \quad x \in E_i \, , \, i > N \end{cases} .$$

Then

$$\int_A \left[f(x) - f_N(x)\right]^2 dx = \sum_{n > N} y_n^2 \, m(E_n) < \epsilon .$$

Therefore the simple functions in L^2 which take on only a finite number of values are dense in L^2.

It is sufficient now to show that the continuous functions are dense in the simple functions which take on a finite number of values. Furthermore, since every simple function of this type is a linear combination of the characteristic function $\psi_M(x)$ of measurable sets, it suffices to give the proof for these characteristic functions. Let

M be a measurable set in the metric space A. Then for any $\epsilon > 0$ a closed set F_M and an open set G_M can be found such that $F_M \subset M \subset G_M$ and $m(G_M) - m(F_M) < \epsilon$. Define the function $\varphi_\epsilon(x)$ by

$$\varphi_\epsilon(x) = \frac{\rho(x, A\backslash G_M)}{\rho(x, A\backslash G_M) + \rho(x, F_M)}$$

where $A\backslash G_M$ is the complement of G_M with respect to A and $\rho(x, A\backslash G_M)$ and $\rho(x, F_M)$ are the distances from x to the sets $A\backslash G_M$ and F_M respectively.

The function $\varphi_\epsilon(x) = 0$ for $x \epsilon A\backslash G_M$ and equals 1 for $x \epsilon F_M$. It is continuous since the functions $\rho(x, A\backslash G_M)$ and $\rho(x, F_M)$ are continuous and since their sum is never equal to zero. The function $\psi_M(x) - \varphi_\epsilon(x)$ is not greater than unity of $G_M\backslash F_M$ and is zero outside this set. Therefore

$$\int_A \left[\psi_M(x) - \varphi_\epsilon(x)\right]^2 dx < \epsilon ,$$

which completes the proof of the theorem.

Proposition 3 is now restated and proved.

<u>Theorem 4.2</u>: If an everywhere dense subset E_1 of the metric space E (with metric ρ) is a separable space, then E is separable. (ref. 5, page 98.)

Proof: Let A be a denumerable everywhere dense subset of E_1. Take $x \epsilon E$ and assume $\epsilon > 0$. Since E_1 is everywhere dense in E there exists $x' \epsilon E_1 \cap E$ such that $\rho(x, x') < \frac{\epsilon}{2}$; and since A is everywhere dense in E_1 there exists $x'' \epsilon (A \cap E_1) \cap E$ such that

$\rho(x', x'') < \frac{\epsilon}{2}$. Hence $\rho(x, x'') < \epsilon$. Due to the arbitrariness of ϵ it follows that A is everywhere dense in E. Hence E is separable.

From theorems 4.1 and 4.2 and the separability of C_L^2 it is concluded that L^2 is separable. Thus let $\{g_n\}$ be a countable everywhere dense set of functions in L^2. Discarding from this set those functions which are linearly dependent with the preceeding ones yields a set $\{f_n\}$ which is linearly independent and complete since each g_n is a linear combination of the elements in $\{f_n\}$ and $\{g_n\}$ is dense in L^2.

The discussion has now led to the principal theorem of this section.

<u>Theorem 4.3</u>: (Gram-Schmidt Process) Let the system of functions $\{f_n\}$ be linearly independent. Then there exists a system of functions $\{\varphi_n\}$ satisfying the following conditions.

1. The system $\{\varphi_n\}$ is orthonormal;

2. every function φ_n is a linear combination of the functions $f_1, f_2, \cdots, f_n$:

$$\varphi_n = a_{n1} f_1 + \cdots + a_{nn} f_n ,$$

where $a_{nn} \neq 0$,

3. every function f_n is a linear combination of the functions $\varphi_1, \varphi_2, \cdots, \varphi_n$:

$$f_n = b_{n1} \varphi_1 + \cdots + b_{nn} \varphi_n$$

where $b_{nn} \neq 0$. Every function of the system $\{\varphi_n\}$ is uniquely determined (up to the sign) by conditions 1-3. (ref. 7, page 151.)

Proof: Let

$$d_1 = \int_A \left[f_1(x)\right]^2 dx, \quad \varphi_1(x) = \frac{f_1(x)}{\sqrt{d_1}} .$$

Let

$$G_2(x) = f_2(x) - C_{21}\, \varphi_1(x), \quad C_{21} = \int_A f_2(x)\, \varphi_1(x)\, dx ,$$

$$d_2 = \int_A \left[G_2(x)\right]^2 dx, \quad \varphi_2(x) = \frac{G_2(x)}{\sqrt{d_2}} .$$

Then

$$\int_A G_2(x)\, \varphi_1(x)\, dx = 0 = \int_A \varphi_2(x)\, \varphi_1(x)\, dx$$

and

$$\int_A \left[\varphi_2(x)\right]^2 dx = 1 .$$

In general, let $\varphi_3(x), \varphi_4(x), \cdots$ be defined successively by the relations:

$$G_n(x) = f_n(x) - \sum_{k=1}^{n-1} c_{nk}\, \varphi_k(x), \quad c_{nk} = \int_A f_n(x)\, \varphi_k(x)\, dx$$

$$d_n = \int_A \left[G_n(x)\right]^2 dx, \quad \varphi_n(x) = \frac{G_n(x)}{\sqrt{d_n}} .$$

It follows from the definition that each φ_n is orthogonal to $\varphi_1 \cdots, \varphi_{n-1}$ and is normalized. By construction each φ_n is a linear combination of the functions $f_1, f_2, \cdots, f_n$, and conversely. Also $d_n \neq 0$ since if $d_n = 0$ then

$$\int_A \left[f_n(x) - \sum_{k=1}^{n-1} c_{nk}\, \varphi_k(x) \right]^2 dx = \int_A \left[f_n(x) - \sum_{k=1}^{n-1} e_{nk}\, f_k(x) \right]^2 dx = 0 \ .$$

Thus

$$f_n(x) = \sum_{k=1}^{n-1} e_{nk}\, f_k(x)$$

almost everywhere in A, which contradicts the hypothesis that $\{f_n(x)\}$ is a linearly independent set. Hence the theorem if proven.

Since the existence of a countable dense set in L^2 has been shown, theorem 4.3 establishes the existence of a complete countable orthonormal system in L^2.

5. CONVERGENCE OF GENERALIZED FOURIER SERIES

Before such a vague concept as convergence can have mathematical application it is essential that the intuitive idea of nearness be made precise. Two forms of convergence have already been encountered in section 4; (1) uniform convergence as convergence in the metric of the space C, and (2) convergence in the mean as convergence in the metric of L^2 (more exactly, convergence in the mean of order two). Two additional concepts are introduced by the following definitions.

Definition 5.1: The sequence of functions $f_n(x)$ defined on some space with measure S, is said to converge almost everywhere to the function $F(x)$ if

$$\lim_{n \to \infty} f_n(x) = F(x)$$

except on a point set of zero measure.

Definition 5.2: The sequence of measurable functions $f_n(x)$ converges in measure to the function $F(x)$, if for any $\epsilon > 0$

$$\lim_{n \to \infty} m\left\{x : \left|f_n(x) - F(x)\right| \geq \epsilon\right\} = 0 .$$

The following four theorems, whose proofs are given in reference 4, give the relations between these forms of convergence.

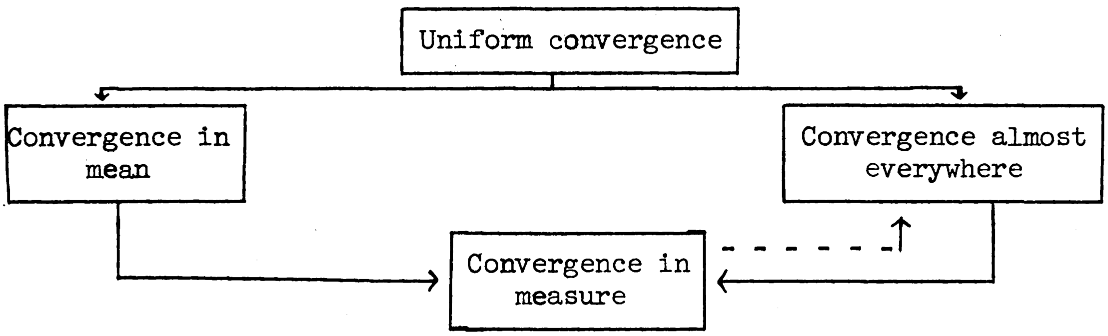
<u>Theorem 5.1</u>: If the sequence $\left\{f_n(x)\right\}$ of functions in L^2 converges uniformly to $f(x)$, then $f(x) \in L^2$ and $\left\{f_n(x)\right\}$ converges to $f(x)$ in the mean.

<u>Theorem 5.2</u>: If the sequence $\left\{f_n(x)\right\}$ converges to $f(x)$ in the mean, then one can select from it a subsequence $\left\{f_{nk}(x)\right\}$ which converges to $f(x)$ almost everywhere.

Theorem 5.3: Let the sequence of measurable functions $f_n(x)$ converge in measure to $f(x)$. Then one can select from the sequence $\{f_n(x)\}$ a subsequence $\{f_{nk}(x)\}$ which converges to $f(x)$ almost everywhere.

Theorem 5.4: If a sequence of measurable functions $f_n(x)$ converges almost everywhere to some function $F(x)$ then it converges to the same limit function $F(x)$ in measure.

The above convergence relations are summarized by the following scheme (ref. 4):



where the dashed arrow means that from a sequence which converges in measure a subsequence can be obtained which converges almost everywhere.

It is most appropriate to consider first the question of convergence in the mean of an associated generalized Fourier series since this concept incorporates the metric of L^2.

This question is answered by the next theorem.

Theorem 5.5: For any function with an integrable square, the Fourier series with respect to any orthogonal system converges (in the mean) in the metric space L^2.

To prove theorem 5.5 the following two theorems are presented.

Theorem 5.6: In order that a sequence $\{f_n(x)\}$ of L^2 functions should converge in the mean to a function $f(x)$ in L^2, it is necessary and sufficient that $\|f_m - f_n\| \to 0$ for $m, n \to \infty$.

Proof: See reference 8, page 12.

Theorem 5.7: Bessel's Inequality. Of all the polynomials of the nth order with respect to an orthogonal system $\{\varphi_n(x)\}$, the best approximation in the metric space L^2 for $f(x) \in L^2$ is given by the nth partial sum of its Fourier series with respect to this system.

Proof: Let $(\varphi_1, \varphi_2, \cdots, \varphi_n)$ be a finite orthonormal family of functions in L^2. Let $\varphi = \beta_1 \varphi_1 + \beta_2 \varphi_2 + \cdots + \beta_n \varphi_n$ and $f \in L^2$. The objective is to select the β_i in such a way as to minimize $[d(f, \varphi)]^2 = \|f - (\beta_1 \varphi_1 + \cdots + \beta_n \varphi_n)\|^2$. Expand the right-hand side as the inner product $(f - \varphi, f - \varphi)$, taking into account the orthonormality of the φ_i to obtain:

$$[d(f, \varphi)]^2 = \|f\|^2 + \sum_i^n \beta_i^2 - 2\left[\beta_1 (f, \varphi_1) + \cdots + \beta_n (f, \varphi_n)\right]$$

$$= \sum_i^n [\beta_i - (f, \varphi_i)]^2 + \left\{\|f\|^2 - \sum_i^n (f, \varphi_i)^2\right\}.$$

Since the β_i occur only in the first summation it follows that $d(f, \varphi)$ is a minimum for $\beta_i = (f, \varphi_i)$. Hence $d(f, \varphi)$ is a minimum when the β_i are the Fourier coefficients of f with respect to $\{\varphi_n\}$. Note that when the β_i are so chosen the following inequality is obtained for each n:

$$\sum_i^n \beta_i^2 \leq \|f\|^2 .$$

Returning now to the proof of theorem 5.5 it will be shown that the partial sums $S_n(s)$ of the Fourier series for $f(x) \in L^2$ satisfy the conditions of theorem 5.6. For any integer n and any $p \geq 1$:

$$\left\| S_{n+p}(x) - S_n(x) \right\|^2 = \left\| \sum_{k=n+1}^{n+p} \beta_k \, \varphi_k(x) \right\|^2$$

$$= \int_a^b \left| \sum_{k=n+1}^{n+p} \beta_k \, \varphi_k(x) \right|^2 dx$$

$$= \sum_{k=n+1}^{n+p} \beta_k^{\,2} ,$$

since the system $\{\varphi_n(x)\}$ is orthonormal. By virtue of Bessel's inequality it is known that if $f(x) \in L^2$, then $\sum_{k=1}^{\infty} \beta_k^{\,2} < +\infty$.

Therefore for any $\epsilon > 0$ it is possible to find an N such that $\sum_{k=n+1}^{n+p} \beta_k^{\,2} < \epsilon^2$ for $n \geq N$ and hence

$$\left\| S_{n+p}(x) - S_n(x) \right\| < \epsilon .$$

Thus by theorem 5.6 the Fourier series for $f(x) \in L^2$ converges in L^2.

It should be noted that only the convergence in L^2 has been proven. It does not follow that this sum should be equal to $f(x)$ in the sense of the L^2 metric. To resolve this question it is necessary to consider the completeness of the system $\{\varphi_n(x)\}$. Definition 1.5 of a complete system is rephrased in the following form.

Definition 5.3: The system $\{\varphi_n(x)\}$ is complete in L^2 if for any $f(x) \in L^2$ and any $\epsilon > 0$ it is possible to choose the numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$ so that

$$\left\| f(x) - \sum_{k=1}^{n} \alpha_k \varphi_k(x) \right\| < \epsilon .$$

For future reference the concept of a closed system is introduced here. (These definitions of closed and complete systems are frequently interchanged in the literature.)

Definition 5.4: The system $\{\varphi_n(x)\}$ is closed in $L^2 [a, b]$ if the only function $f(x) \in L^2 [a, b]$ which is orthogonal to every $\varphi_n(x)$ is the function $f(x) = 0$ almost everywhere in $[a, b]$.

<u>Theorem 5.8</u>: In the space L^2 the completeness and closure of a system are equivalent.

Proof: See reference 1, page 62.

Now let $\{\varphi_n(x)\}$ be a complete system in L^2 and $f(x)$ be any element of L^2. Then given $\epsilon > 0$ it is possible to select numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that;

$$\left\| f(x) - \sum_{k=1}^{n} \alpha_k \varphi_k(x) \right\| < \epsilon .$$

But in proving Bessel's inequality it was shown that the best approximation to $f(x)$ in the L^2 metric is given by the polynomial $\sum_{k=1}^{n} \beta_k \varphi_k(x)$, where the β_k are the Fourier coefficients of $f(x)$.

Therefore

$$\left\| f(x) - \sum_{k=1}^{n} \beta_k \varphi_k \right\| \leq \left\| f(x) - \sum_{k=1}^{n} \alpha_k \varphi_k \right\| < \epsilon .$$

Since

$$\left\| f(x) - \sum_{k=1}^{n} \beta_k \varphi_k \right\|^2 = \|f\|^2 - \sum_{k=1}^{n} \beta_k^{\,2}$$

it follows that

$$0 \leq \|f\|^2 - \sum_{k=1}^{n} \beta_k^{\,2} < \epsilon^2 .$$

Hence in the limit

$$\sum_{k=1}^{n} \beta_k^{\,2} = \|f\|^2 .$$

This equality is known as Parseval's equality.

The foregoing discussion can be considerably strengthened by the next theorem, which also assures the uniqueness of the Fourier series.

<u>Theorem 5.9</u>: Riesz-Fischer: Let c_n $(n = 1, 2, \ldots)$ be any sequence of numbers for which $\sum_{n=1}^{\infty} c_n^{\,2} < +\infty$ and $\{\varphi_n(x)\}$ be any orthonormal system. Then there exists an $f(x) \in L^2$ such that the c_n are its Fourier coefficients with respect to $\{\varphi_n(x)\}$; if the system is complete, then there exists only one such $f(x)$.

Proof: Let the sequence $S_n = \sum_{k=1}^{n} c_k \varphi_k$ be set up. Since $\sum_{n=1}^{\infty} c_n^{\,2}$ converges then for any $\epsilon > 0$, an N can be chosen such

$\sum_{N+1}^{\infty} c_n^2 < \epsilon$. But then

$$\left\| S_{n+p}(x) - S_n(x) \right\|^2 = \sum_{k=n+1}^{n+p} c_k^2 < \epsilon \qquad (n \geq N,\ p > 0)\ .$$

Hence there exists an $f(x) \in L^2$ such that $f(x) = \lim_{n \to \infty} S_n(x)$, that is, $f(x) = \sum_{n=1}^{\infty} c_n \varphi_n(x)$. Multiplying both sides of this equation by φ_m and integrating it is seen, due to the orthonormality of the $\{\varphi_n\}$ that $c_m = \int_a^b f(x)\, \varphi_m(x)\, dx$.

To prove the uniqueness of $f(x)$, assume $g(x) = \sum_{n=1}^{\infty} c_n \varphi_n(x)$, $g(x) \neq f(x)$. Then $\int_a^b [f - g]\, \varphi_n\, dx = c_n - c_n = 0$, for every n. Hence $f - g \neq 0$ is orthogonal to every φ_n, which contradicts the hypothesis that $\{\varphi_n\}$ is closed, that is, complete.

Being assured that for any $f \in L^2$ there exists a unique expansion which converges to it in the sense of the L^2 metric the question arises as to what can be said with regard to other more stringent concepts of convergence. Consider then the question of the convergence properties of a general orthogonal series

$$\sum_{n=0}^{\infty} c_n \varphi_n(x); \tag{5.1}$$

where nothing more than orthonormality in $[a, b]$ is assumed about the system $\{\varphi_n(x)\}$. In this event the problem of convergence almost everywhere appears to be the only appropriate consideration. To clarify this, note that the values of the functions $\varphi_n(x)$ can be chosen arbitrarily on a set $N \subset [a, b]$ of zero measure without effecting the orthonormality of the system. Obviously such an arbitrary selection of values could also make the series $\sum c_n \varphi_n(x)$ divergent on N. Suppose then that only orthonormality of the system is assumed and it is desired to determine convergence features of the series (5.1) from properties of the coefficients $\{c_n\}$. In this study the coefficients $\{c_n\}$ are assumed to be the Fourier coefficients associated with some $f \in L^2$. Thus from Bessel's inequality it follows that $\sum_{n=0}^{\infty} c_n^2 < \infty$; even though this sum may be arbitrarily large as is asserted by the following theorem.

<u>Theorem 5.10</u>: If $\{\varphi_n(x)\}$ is any orthogonal system for which $|\varphi_n(x)| \leq A$, then a continuous function can be found for which the Fourier coefficients C_n with respect to $\{\varphi_n(x)\}$ satisfy

$$\sum_{n=1}^{\infty} |C_n|^{2-\epsilon} = +\infty$$

for any $\epsilon > 0$.

Proof: Reference 1, page 341.

Now assume the condition $\sum_{n=0}^{\infty} |C_n| < \infty$. This condition implies the absolute convergence of (5.1) almost everywhere, since by the Cauchy-Schwartz inequality

$$\sum_{n=0}^{\infty} \int_a^b |C_n \varphi_n(x)| \, dx \leq (b - a) \sum_{n=0}^{\infty} |C_n| \sqrt{\int_a^b \varphi_n^2(x) \, dx}$$

$$= (b - a) \sum_{n=0}^{\infty} |C_n| < \infty . \qquad (5.2)$$

The desired conclusion can now be drawn with the aid of the following theorem due to B. Levi (ref. 8, page 11).

Theorem 5.11: If $\{f_n(x)\}$ is a monotone increasing sequence of L-integrable functions and

$$\left|\int_a^b f_n(x)\,dx\right| \leq C \qquad (n = 0, 1, \cdots)$$

then the limiting function $f(x) = \lim_{n\to\infty} f_n(x)$ is also L-integrable and

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx .$$

If in particular $\{u_n(x)\}$ are L-integrable functions such that

$$\sum_{n=0}^{\infty} \int_a^b |u_n(x)|\,dx < \infty$$

then the series $\sum_{n=0}^{\infty} u_n(x)$ is absolutely convergent almost everywhere.

Proof: The sequence $\{f_n(x) - f_0(x)\}$ is monotone increasing, $f(x) - f_0(x)$ is its limiting function and $f_n(x) - f_0(x) \geq 0$. Now the hypotheses imply

$$\int_a^b \left[f_n(x) - f_0(x)\right] dx \leq \lim_{n\to\infty} \int_a^b \left[f_n(x) - f_0(x)\right] dx$$

$$\leq C + \int_a^b |f_0(x)|\,dx .$$

By Fatou's theorem (ref. 8, page 10) $f(x)$ is also L-integerable and

$$\int_a^b f(x)\,dx \leq C .$$

Hence

$$\int_a^b \left[f(x) - f_0(x)\right] dx \leq \lim_{n \to \infty} \int_a^b \left[f_n(x) - f_0(x)\right] dx .$$

The reverse inequality follows from:

$$f(x) - f_0(x) \geq f_n(x) - f_0(x) .$$

Thus

$$\int_a^b \left[f(x) - f_0(x)\right] dx = \lim_{n \to \infty} \int_a^b \left[f_n(x) - f_0(x)\right] dx ,$$

Which is equivalent to the first assertion of the theorem. The second assertion is proven by letting

$$f_n(x) = \sum_{k=0}^{n} \left|u_k(x)\right| .$$

Applying the second part of theorem 5.11 to the relationship (5.2) completes the proof that

$$\sum_{n=0}^{\infty} \left|c_n\right| < \infty$$

is a sufficient condition for the orthogonal series (5.1) to converge. Thus it follows that convergence tests on the coefficients lie between the conditions $\sum c_n^2 < \infty$ and $\sum \left|c_n\right| < \infty$. Due to the considerable number of preliminary theorems required, some of the following theorems are presented without proof.

Theorem 5.12: Rademacher-Menchoff. Let $\{\varphi_n(x)\}$ be an orthonormal system. The orthogonal series

$$\sum_{n=0}^{\infty} c_n \varphi_n(x) \tag{5.3}$$

is convergent almost everywhere if the condition

$$\sum_{n=1}^{\infty} c_n^{\,2} \log^2 n < \infty \tag{5.4}$$

is fulfilled.

Proof: Reference 8, page 80.

The question of whether condition (5.4) can be relaxed is answered in part by the next theorem.

<u>Theorem 5.13</u>: If $\{c_n\}$ is a positive monotone decreasing sequence of numbers for which

$$\sum_{n=1}^{\infty} c_n^{\,2} \log^2 n = \infty$$

holds then there exists in $[a, b]$ an orthonormal system $\{\Phi_n(x)\}$ dependent on $\{c_n\}$ such that the orthogonal series

$$\sum_{n=0}^{\infty} c_n \Phi_n(x)$$

is divergent everywhere in $[a, b]$.

Proof: Reference 8, page 88.

Therefore if complete generality is permitted in the choice of an orthogonal system and the function to be expanded it follows from theorems 5.12 and 5.13 that condition (5.4 is both necessary and sufficient for convergence almost everywhere. The following theorem can be derived as a consequence of theorem 5.12.

Theorem 5.14: If the series

$$\sum_{n=1}^{\infty} (\log n)^2 \sum_{k=m_n+1}^{m_{n+1}} c_k^{\,2}$$

is convergent, then the sequence $\left\{S_{m_n}(x)\right\}$ of the partial sums of the orthogonal series (5.3) is convergent almost everywhere.

Proof: Reference 8, page 83.

Since the above theorems are based on the properties of the coefficients it is natural to ask what properties are to be expected for the Fourier coefficients for an $f \in L^2$. From Bessel's inequality it follows that the Fourier coefficients tend to zero. What is of primary interest is the rate at which they tend to zero, however little additional can be said unless conditions are imposed upon either the orthogonal system or the functions f to be expanded. For the present assume only that the orthonormal system $\left\{\varphi_n(x)\right\}$ consists of bounded functions. Then the next two theorems imposing conditions on the functions f can be proven.

Theorem 5.15: Mercer's Theorem. If for an orthonormal system $\left\{\varphi_n(x)\right\}$ the functions are all bounded; that is

$$|\varphi_n(x)| \leq M \qquad a \leq x \leq b \qquad (n = 1, 2, \cdots)$$

then the Fourier coefficients of any summable function with respect to this system tend to zero.

Proof: Let $f(x)$ be summable and $\epsilon > 0$ be given. First find a function $F(x)$ for which $\int_a^b |f(x) - F(x)| \, dx < \epsilon$, while $F(x)$ is bounded. This is possible by the very definition of the Lebesque integral.

Since any bounded function belongs to L^2 it follows that the Fourier coefficients of $F(x)$ tend to zero, that is, there exists an N such that for $n > N$, $\left| \int_a^b F(x) \varphi_n(x) \, dx \right| < \epsilon$. Also

$$\left| \int_a^b [f(x) - F(x)] \varphi_n(x) \, dx \right| \leq M\epsilon,$$

hence

$$\left| \int_a^b f(x) \varphi_n(x) \, dx \right| < \epsilon(1 + M)$$

for $n > N$. Thus

$$\int_a^b f(x) \varphi_n(x) \, dx \to 0 \quad \text{as} \quad n \to \infty .$$

<u>Theorem 5.16</u>: Riesz's Theorem. Let $\{\varphi_n(x)\}$ be an orthonormal system of bounded functions.

1. If $f \in L^p(a, b)$, then the Fourier coefficients of f with respect to $\{\varphi_n\}$ satisfy the condition

$$\|C\|_q \leq M^{\frac{2}{p} - 1} \|f\|_p .$$

2. If for a sequence of numbers C_n, $\|C\|_p < + \infty$ then there exists a function $f(x) \in L^p(a, b)$ satisfying

$$C_n = \int_a^b f(x) \varphi_n(x) \, dx$$

for all n such that

$$\|f\|_q \leq M^{\frac{2}{p}-1} \|c\|_p \; .$$

where

$$\|c\|_q = \left\{ \sum_{n=0}^{\infty} |c_n|^q \right\}^{\frac{1}{q}}$$

$$\|f\|_q = \left\{ \int_a^b |f|^q \, dx \right\}^{\frac{1}{q}}$$

and

$$\frac{1}{p} + \frac{1}{q} = 1 \; .$$

Proof: Reference 1, page 219.

It appears that convergence features for orthogonal expansions might be improved if more was assumed about the expanded function f than that $f \in L^2$. However this is not necessarily the case since Banach (ref. 9) has proved the following deep theorem.

Theorem 5.17: There is a continuum of orthonormal systems $\{\varphi_n(x)\}$ such that the Fourier expansion

$$f(x) \sim \sum_{n=0}^{\infty} c_n \, \varphi_n(x)$$

for every differentiable function $f(x)$ is divergent almost everywhere.

It is apparent then that to continue the discussion requires more restrictive conditions on the system $\{\varphi_n(x)\}$. Due to the limited scope of this study, further discussion will be restricted to the particular orthonormal systems obtained by orthogonalization of the linearly independent system $\{x^n\}$ in the interval $[a, b]$.

6. ORTHOGONAL POLYNOMIALS

Let $\rho(x)$ be a nonnegative function integrable over (a, b) and positive on a set of points such that $\int_a^b \rho(x)\, dx > 0$. Let the system

$$\left\{ \left[\rho(x)\right]^{\frac{1}{2}} x^n \right\} \tag{6.1}$$

be taken as the linearly independent set in the Gram-Schmidt Process (theorem 4.3). The results of orthonormalizing the system (6.1) yields functions of the form $\left[\rho(x)\right]^{\frac{1}{2}} p_n(x)$ where $p_n(x)$ is a polynomial of degree n in x, since these are linear combinations of the elements of the system (6.1). It is assumed that the coefficient of the highest degree term in $p_n(x)$ has been chosen positive. The $p_n(x)$ are orthonormal polynomials with respect to the weight function $\rho(x)$, that is,

$$\int_a^b \rho(x)\, p_m(x)\, p_n(x)\, dx = 0 \qquad m \neq n,$$

$$\int_a^b \rho(x) \left[p_n(x)\right]^2 dx = 1 .$$

The polynomials $p_n(x)$ thus generated are unique in the sense of the following theorem.

Theorem 6.1: The only orthonormal system $\{\varphi_n(x)\}$ with respect to the weight function $\rho(x)$ in which $\varphi_n(x)$ is a polynomial of degree n with positive leading coefficient is $\{p_n(x)\}$.

Proof: Obviously every power of x^k can be represented in the form

$$x^k = \beta_0\, p_0(x) + \beta_1\, p_1(x) + \cdots + \beta_k\, p_k(x) .$$

Hence any polynomial $\varphi_n(x)$ of degree n can be represented in the form

$$\varphi_n(x) = \sum_{k=0}^{n} \gamma_k\, p_k(x) \qquad (6.2)$$

Conversely the polynomials $p_k(x)$ can be represented as a linear combination of $\varphi_0, \varphi_1, \cdots, \varphi_k$. Thus it follows that for $k < n$, $p_k(x)$ is orthogonal to φ_n. Multiplying both sides of (6.2) by $\rho(x)\, p_\nu(x)$, $\nu < n$ and integrating yields;

$$\sum_{k=0}^{n} \gamma_k \int_a^b \rho(x)\, p_\nu(x)\, p_k(x)\, dx = \gamma_k = \int_a^b \rho(x)\, \varphi_n(x)\, p_\nu(x) = 0$$

for $\nu = 0, 1, \cdots, n - 1$. Hence $\varphi_n(x) = \gamma_n\, p_n(x)$ and therefore

$$1 = \int_a^b \rho(x)\, \varphi_n^2(x)\, dx = \gamma_n^2 \int_a^b \rho(x)\, p_n^2(x)\, dx = \gamma_n^2 .$$

Since both φ_n and p_n have positive leading coefficients it follows that $\varphi_n = p_n$.

The next theorem follows immediately from the completeness of the polynomials in $[a, b]$ (see Lemma 4.1) and the fact that any polynomial can be represented as a linear combination of the $p_n(x)$.

<u>Theorem 6.2</u>: The orthonormal system of polynomials $p_n(x)$ with respect to the weight function $\rho(x)$ is complete in L^2.

The polynomials $p_n(x)$ will be used in the expansion of functions $f \in L^2$ in a manner entirely analogous with that of definition 3.1, that is, $f(x)$ is said to be expanded in a series with respect to $p_n(x)$ if

$$f(x) \sim \sum_{n=0}^{\infty} a_n \, p_n(x) \tag{6.3}$$

where

$$a_n = \int_a^b \rho(x) \, f(x) \, p_n(x) \, dx \, . \tag{6.4}$$

The selection of the system $p_n(x)$ is in part motivated by the following theorem.

<u>Theorem 6.3</u>: Among all the polynomials $\pi_n(x)$ of degree not greater than n the integral

$$\int_a^b \left[f(x) - \pi_n(x) \right]^2 \rho(x) \, dx$$

attains its minimum value for $\pi_n(x) = S_n(x)$ where $S_n(x)$ is the nth partial sum of the expansion of $f(x)$ in the polynomials $p_n(x)$.

Proof: The proof follows that of theorem 5.7 after noting that $\pi_n(x)$ can be written as a linear combination of the $p_k(x)$, $k = 0, 1, \cdots, n$.

Before returning to the question of convergence it is of interest to consider some of the properties of the $p_n(x)$. In particular a recursion relationship will now be developed. Let a_k and b_k denote the coefficients of the kth and k - 1st powers of x in $p_k(x)$ and consider the expression of $xp_n(x)$ as a finite sum of the $p_k(x)$; $k = 0, 1, \cdots, n + 1$, that is,

$$xp_n(x) = \sum_{k=0}^{n+1} c_{nk} \, p_k(x) \tag{6.5}$$

where

$$c_{nk} = \int_a^b \rho(x)\ xp_n(x)\ p_k(x)\ dx\ . \qquad (6.6)$$

If $k < n - 1$, $xp_k(x)$ is of degree less than n, hence $p_n(x)$ is orthogonal to $xp_k(x)$ with respect to the weight function, thus from (6.6) it follows that $c_{nk} = 0$ for $k < n - 1$. Comparison of the coefficients of x^{n+1} in (6.5) yields $c_{n,n+1} = \frac{a_n}{a_{n+1}}$. Also from (6.6) $c_{nk} = c_{kn}$ hence $c_{n,n-1} = c_{n-1,n} = \frac{a_{n-1}}{a_n}$. Therefore (6.5) reduces to the following recursion relationship:

$$xp_n(x) = \frac{a_n}{a_{n+1}} p_{n+1}(x) + c_{nn}\ p_n(x) + \frac{a_{n-1}}{a_n} p_{n-1}(x) \qquad (6.7)$$

Comparison of the coefficients of x^n in (6.7) yields

$$c_{nn} = \frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}$$

If the assumption is made that $p_{-1}(x) = 0$, $a_{-1} = 0$, then equation (6.7) holds for all nonnegative n.

Consider next the case in which the interval $[a, b]$ of orthogonality is symmetric with respect to the origin, that is, $[a, b] = [-b, b]$ and $\rho(x)$ is an even function. Let $g(x)$ be an arbitrary polynomial of degree less than n, and consider the integral;

$$I = \int_b^{-b} \rho(x)\ p_n(-x)\ g(x)\ dx\ .$$

Change the variable of integration by letting $t = -x$. Then

$$I = \int_b^{-b} \rho(-t)\, p_n(t)\, g(-t)\, dt = \int_b^{-b} \rho(t)\, p_n(t)\, g(-t)\, dt = 0$$

since $p_n(t)$ is orthogonal to every polynomial of degree less than n. Due to the arbitrariness of $g(x)$ it follows that $p_n(-x)$ has the same orthogonality property as $p_n(x)$. Also it is readily seen that $p_n(-x)$ is normalized and $(-1)^n\, p_n(-x)$ has a positive leading coefficient. Hence by theorem 6.1, $(-1)^n\, p_n(-x) = p_n(x)$. Thus $p_n(x)$ can consist of only even powers of x or only odd powers of x, depending on whether n is even or odd. Hence in this case the coefficient b_n of $p_n(x)$ is zero and equation (6.7) reduces to:

$$xp_n(x) = \frac{a_n}{a_{n+1}}\, p_{n+1}(x) + \frac{a_{n-1}}{a_n}\, p_{n-1}(x) \ .$$

In addition the following theorem holds.

Theorem 6.4: If the set of polynomials $\{p_n(x)\}$ is orthogonal with respect to $\rho(x) > 0$ over $[a, b]$, the zeros of $p_n(x)$ are all distinct and all lie in the open interval (a, b).

Proof: Reference 10, page 149.

If additional highly restrictive assumptions are made about the weight function $\rho(x)$, the next theorem can be proven.

Theorem 6.5: If the weight function $\rho(x)$ satisfies the differential equation

$$\frac{\rho'(x)}{\rho(x)} = \frac{D + Ex}{A + Bx + Cx^2}$$

in which A, B, C D, and E are constants and $\left(A + Bx + Cx^2\right) \rho(x)$ vanishes at the end points of $[a, b]$ then the $p_n(x)$ satisfy a differential equation of the form:

$$\alpha(x)\ p_n''(x) + \beta(x)\ p_n'(x) + \gamma(n)\ p_n(x) = 0,$$

in which $\alpha(x)$ is a polynomial of second degree, specifically $A + Bx + Cx^2$, $\beta(x)$ is a polynomial of the first degree, specifically $[(B + D) + (2C + E)x]$ and $\gamma(n) = -[Cn\ (n + 1) + En]$.

Proof: Reference 7, page 164.

Returning to the question of convergence, let $S_n(x)$ denote the nth partial sum of the series (6.3); that is,

$$S_n(x) = \sum_{k=0}^{n} a_k\ p_k(x) \tag{6.8}$$

Letting t be the variable of integration in (6.4) and substituting the expressions for the a_k into (6.8) yields:

$$S_n(x) = \int_a^b \rho(t)\ f(t) \sum_{k=0}^{n} p_k(t)\ p_k(x)\ dt$$

$$= \int_a^b \rho(t)\ f(t)\ K_n(x,t)\ dt\ .$$

The sum $K_n(x,t)$ is called the nth kernel of the orthogonal system $\{p_n(x)\}$. For later reference, the nth Lebesque function of the system is defined:

$$L_n(x) = \int_a^b \rho(t)\ \left|K_n(x,t)\right|\ dt\ .$$

A relatively simple expression for the n^{th} kernel of the system can be developed as follows. Let equation (6.7) be multiplied by $p_n(t)$:

$$xp_n(x)\ p_n(t) = \frac{a_n}{a_{n+1}}\ p_{n+1}(x)\ p_n(t) + c_{nn}\ p_n(x)\ p_n(t) + \frac{a_{n-1}}{a_n}\ p_{n-1}(x)\ p_n(t).$$

If this equation is subtracted from the corresponding form with t and x interchanged, the term $c_{nn}\ p_n(t)\ p_n(x)$ cancels and the result may be expressed in the form:

$$(t - x)\ p_n(t)\ p_n(x) = \frac{a_n}{a_{n+1}}\left[p_{n+1}(t)\ p_n(x) - p_n(t)\ p_{n+1}(x)\right]$$
$$- \frac{a_{n-1}}{a_n}\left[p_n(t)\ p_{n-1}(x) - p_{n-1}(t)\ p_n(x)\right].$$

If this equation be written with n replaced successively by $n-1, n-2, \ldots, 0$, addition of the $n+1$ equations thus obtained yields:

$$(t - x)\sum_{k=0}^{n} p_k(t)\ p_k(x) = \frac{a_n}{a_{n+1}}\left[p_{n+1}(t)\ p_n(x) - p_n(t)\ p_{n+1}(x)\right] \tag{6.9}$$

This identify is known as the Christoffel-Darboux formula. Thus the n^{th} kernel of the system can be expressed;

$$K_n(x,t) = \frac{a_n}{a_{n+1}}\ \frac{p_{n+1}(t)\ p_n(x) - p_n(t)\ p_{n+1}(x)}{t - x} \tag{6.10}$$

From equation (6.10) it is possible to deduce a convergence theorem having application to Fourier series. First, a function $f(x)$ is said to satisfy a Dini-Lipschitz condition of order α at the point ξ if for sufficiently small values of h the inequality

$$|f(\xi + h) - f(\xi)| < \frac{C}{|\log|h||}\ \alpha \qquad (C = \text{constant})$$

holds. The theorem may now be stated in the following form.

Theorem 6.5: If the function $f(x) \in L^2$ satisfies a Dini-Lipschitz condition of order α at the point ξ with $\alpha > 1$ and if in a neighborhood of ξ both the weight function $\rho(x)$ and the orthogonal system $\{p_n(x)\}$ are bounded, then the expansion

$$f(x) \sim \sum_{n=0}^{\infty} a_n\, p_n(x)$$

converges at the point ξ to $f(\xi)$. If these conditions are satisfied uniformly in the whole subinterval $[a_1, b_1]$ of $[a, b]$ then the convergence to $f(x)$ is uniform in every inner subinterval of $[a_1, b_1]$.

Proof: Consider the difference

$$S_n(\xi) - f(\xi) = \int_a^b \left[f(t) - f(\xi)\right] \rho(t)\, K_n(t, \xi)\, dt.$$

From the Christoffel-Darboux formula

$$S_n(\xi) - f(\xi) = \frac{a_n}{a_{n+1}} \int_a^b \left[f(t) - f(\xi)\right]\rho(t)\, \frac{p_n(\xi)\, p_{n+1}(t) - p_n(t)\, p_{n+1}(\xi)}{t - \xi}\, dt \qquad (6.11)$$

$$= \frac{a_n}{a_{n+1}} \left\{ \int_a^{\xi-\delta} + \int_{\xi-\delta}^{\xi+\delta} + \int_{\xi+\delta}^{b} \right\}$$

where $\delta > 0$ is chosen so small that in $[\xi - \delta, \xi + \delta]$ the functions $|p_n(t)|$, $|p_{n+1}(t)|$ and $\rho(t)$ remain less than some constant C and $f(t)$ satisfies the Dini-Lipschitz condition. The first and third integrals may be estimated as follows:

Define the function $F(t)$ by;

$$F(t) = \begin{cases} \dfrac{f(t) - f(\xi)}{t - \xi} & t \in [a, \xi - \delta) \text{ and } t \in (\xi + \delta, b] \\ 0 & t \in [\xi - \delta, \xi + \delta] \end{cases}$$

By virtue of $f \in L^2$ then $F \in L^2$ and it follows from theorem 6.3 that the Fourier coefficients of F tend to zero; that is, for arbitrary $\epsilon > 0$ there exists an N such that for $n > N$

$$\int_a^b F(t)\, \rho(t)\, p_n(t)\, dt < \epsilon .$$

Thus for $n > N$ the following inequalities hold.

$$\left| \int_a^{\xi-\delta} \right| < \epsilon \left\{ |p_n(\xi)| + |p_{n+1}(\xi)| \right\} \leq 2\epsilon C \qquad (6.12)$$

$$\left| \int_{\xi+\delta}^{b} \right| < \epsilon \left\{ |p_n(\xi)| + |p_{n+1}(\xi)| \right\} \leq 2\epsilon C \qquad (6.13)$$

Consider next the fraction $\dfrac{a_n}{a_{n+1}}$. By equation (6.7)

$$\int_a^b \rho(t)\, p_{n+1}(t)\, t p_n(t)\, dt = \frac{a_n}{a_{n+1}} \int_a^b \rho(t)\, p_{n+1}^2(t)\, dt$$

$$+ \int_a^b \rho(t)\, p_{n+1}(t)\, q_n(t)\, dt$$

where $q_n(t)$ is a polynomial of degree less than $n+1$, hence the last integral vanishes. Therefore by the Cauchy-Schwartz inequality

$$\left|\frac{a_n}{a_{n+1}}\right| \leq \int_a^b \rho(t)\ \left|p_{n+1}(t)\ tp_n(t)\right|\ dt$$

$$\leq \max(|a|, |b|) \left\{\int_a^b \rho(t)\ p_{n+1}^2(t)\ dt \int_a^b \rho(t)\ p_n^2(t)\ dt\right\}^{\frac{1}{2}}$$

$$= \max(|a|, |b|) \equiv A \,. \tag{6.14}$$

From (6.11), (6.12), (6.13), and (6.14) it follows that

$$|S_n(\xi) - f(\xi)| \leq 2AC^3 \int_{\xi-\delta}^{\xi+\delta} \left|\frac{f(t) - f(\xi)}{t - \xi}\right| dt + 4\epsilon AC \,.$$

By the Dini-Lipschitz condition

$$\int_{\xi-\delta}^{\xi+\delta} \left|\frac{f(t) - f(\xi)}{t - \xi}\right| dt \leq K \int_{\xi-\delta}^{\xi+\delta} \frac{dt}{|t - \xi|\ \big|\log|t - \xi|\big|^{\alpha}}$$

$$= \frac{2K}{(\alpha - 1)\ |\log \delta|^{\alpha-1}} \,.$$

Since $\alpha - 1 > 0$ and $\delta > 0$ may be chosen arbitrarily small, the right-hand side can be made less than ϵ. Therefore for $n > N$, $|S_n(\xi) - f(\xi)| < 2A\epsilon(C^3 + 2C)$. The above estimates hold uniformly in every subinterval $[a_1, b_1]$ if the conditions for the theorem are fulfilled uniformly in $[a_1, b_1]$, which completes the proof. It can also be shown that theorem 6.5 is valid if $f(x)$ satisfies in an interval $[a_1, b_1]$ a Lipschitz condition of order α, that is, if in $[a_1, b_1]$

$$|f(x + h) - f(x)| \leq K|h|^{\alpha} \qquad (K = \text{constant})$$

where $0 < \alpha \leq 1$. (Reference 8, page 29.)

Both the Lipschitz and Dini-Lipschitz conditions are more restrictive than continuity. That such a restriction is essential is asserted by the following theorem.

Theorem 6.6: If the Lebesque functions $L_n(x)$ of the orthogonal system $\{p_n(x)\}$ increase without bound for some x_0 that is, $\lim_{n \to \infty} L_n x_0 = \infty$ then there exists a continuous function F whose Fourier series diverges for that x.

Proof: As was seen earlier in this section, the nth partial sum of the generalized Fourier series associated with the function f can be written;

$$S_n(x,f) = \int_a^b \rho(t)\ f(t)\ K_n(x,t)\ dt\ .$$

Then if $|f(t)| \leq 1$

$$|S_n(x,f)| \leq \int_a^b \rho(x)\ |K_n(x,t)|\ dt = L_n(x) \qquad (6.15)$$

If the function f is chosen by:

$$f(x) = \text{sign}\left[K_n(x,t)\right],$$

then $s_n(x,f) = L_n(x)$. This f may be discontinuous, however, a continuous approximation f_n of f can be found such that

$$|f_n(x)| \leq 1 \text{ and } S_n(x_0, f_n) \geq \frac{1}{2} L_n(x_0) \text{ for each } n \qquad (6.16)$$

If the Fourier series of any f_n diverges at x_0, the theorem is proved. Therefore suppose that the Fourier series of each f_n converges to $\gamma_n(x_0)$.

Define the function $F(x)$ by

$$F(x) = \sum_{k=0}^{\infty} \alpha_k f_{n_k}(x) , \tag{6.17}$$

where

$$\alpha_k > 0, \sum_{0}^{\infty} \alpha_k < \infty, \sum_{k+1}^{\infty} \alpha_p \leq \frac{1}{6} \alpha_k . \tag{6.18}$$

For example take $\alpha_k = 7^{-k}$. From the above conditions,

$$\sum_{k=0}^{\infty} \left|\alpha_k f_{n_k}(x)\right| \leq \sum_{0}^{\infty} \alpha_k < \infty .$$

Hence the series

$$\sum_{k=0}^{\infty} \alpha_k f_{n_k}(x)$$

converges. Thus given an $\epsilon > 0$ there exists an N_1 such that

$$\left| F(x) - \sum_{k=0}^{N_1} \alpha_k f_{n_k}(x) \right| < \frac{\epsilon}{2} .$$

From the monotone decreasing property of the α_k it follows that an N_2 can be found such that $\alpha_{N_2} < \epsilon$. Take N equal to the maximum of N_1 and N_2. Then

$$\left| F(x) - \sum_{k=0}^{\infty} \alpha_k f_{n_k} \right| \leq \left| F(x) - \sum_{k=0}^{N} \alpha_k f_{n_k} \right| + \left| \sum_{N+1}^{\infty} \alpha_k \right|$$

$$\leq \frac{\epsilon}{2} + \frac{\epsilon}{6} < \epsilon .$$

Thus the series of continuous functions $\sum \alpha_k f_{n_k}$ converges uniformly and therefore its sum $F(x)$ is continuous and its Fourier series is obtained by addition of the Fourier series of the terms in (6.17). Now chose the n_k by induction as follows. Given $n_1, \cdots, n_{k-1}$, choose n_k such that

$$\alpha_k L_{n_k}(x_0) \to \infty, \quad \sum_0^{k-1} \alpha_p \left| \gamma_{n_p}(x_0) \right| \leq \frac{1}{12} \alpha_k L_{n_k}(x_0).$$

Since

$$S_n\left(x_0, \sum_0^{k-1} \alpha_p f_{n_p}\right) \to \sum_0^{k-1} \alpha_p \gamma_{n_p},$$

it may be supposed that n_k is so large that

$$\left| S_{n_k}\left(x_0, \sum_0^{k-1} \alpha_p f_{n_p}\right) \right| \leq 2 \sum_0^{k-1} \alpha_p \left| \gamma_{n_p} \right| \leq \frac{1}{6} \alpha_k L_{n_k}(x_0). \tag{6.18}$$

Also by (6.15) and (6.18)

$$\left| S_{n_k}\left(x_0, \sum_{k+1}^{\infty} \alpha_p f_{n_p}\right) \right| \leq L_{n_k}(x_0) \sum_{k+1}^{\infty} \alpha_p \leq \frac{1}{6} \alpha_k L_{n_k}(x_0) \tag{6.19}$$

Finally it follows from (6.18) and (6.19) that

$$S_{n_k}(x_0, F) \geq S_{n_k}\left(x_0, \alpha_k f_{n_k}\right) - \frac{1}{6} \alpha_k L_{n_k}(x_0) - \frac{1}{6} \alpha_k L_{n_k}(x_0)$$

$$\geq \frac{1}{6} \alpha_k L_{n_k}(x_0) \to \infty .$$

Hence the Fourier series of F diverges at x_0.

Several examples of the type function whose existence is established by theorem 6.6 have been constructed. (Refs 1 and 11.) However, if the expansion of a specific function is being considered then convergence tests have been developed which consider properties of the orthogonal system as reflected in its Lebesque functions and properties of both the system and the function to be expanded as reflected in the Fourier coefficients. For the general orthogonal system the following theorem holds.

Theorem 6.7: If the relation $\frac{L_n(x)}{\lambda_n} \leq K$ for all n, holds in the set $E \subset [a, b]$ where $\{\lambda_n\}$ is positive and nondecreasing, then

$$\sum_{n=0}^{\infty} c_n^2 \lambda_n < \infty$$

implies the convergence of the orthogonal series $\sum c_n \varphi_n(x)$ almost everywhere in E.

Proof: Reference 8, page 175.

The Lebesque functions for the orthogonal polynomials of this section, under assumptions of boundedness, satisfy the first condition of this theorem. In particular the next theorem has been established.

Theorem 6.8: If $\{p_n(x)\}$ is a system of orthonormal polynomials with respect to the weight function $\rho(x)$ and if the functions $\{p_n(x)\}$ as well as $\rho(x)$ are bounded in $[c, d] \subset (a, b)$ then

$$\frac{L_n(x)}{\log n} < K$$

for all n (K constant) and this condition holds uniformly in every inner subinterval of (c, d).

Proof: Reference 8, page 179.

Thus the next theorem follows as a corollary to theorems 6.7 and 6.8.

<u>Theorem 6.9</u>: If $\{p_n(x)\}$ is an orthonormal system of polynomials with respect to the weight function $\rho(x)$ and if the functions $\{p_n(x)\}$ as well as $\rho(x)$ are bounded in the interval $[c, d] \subset (a, b)$ then the relation

$$\sum_{n=0}^{\infty} c_n^2 \log n < \infty$$

implies the convergence almost everywhere in (c, d) of the orthogonal series

$$\sum_{n=0}^{\infty} c_n p_n(x) .$$

The question of boundedness of the system $\{p_n(x)\}$ has been prevalent in the last few theorems. This question is difficult to answer and even the problem of pointing out sufficient conditions in order that the functions obtained by orthogonalization of a given system of functions should satisfy certain requirements of boundedness has not been satisfactorily solved. However under certain conditions the boundedness of one orthogonal system can be deduced from the boundedness of another system. For example, let $\{p_n(x)\}$ be the orthonormal system of polynomials with respect to the weight function $\rho(x)$ and let $\{q_n(x)\}$ be the orthonormal system with respect to the weight function $\rho(x)\ \sigma(x)$ with $\sigma(x) > 0$ almost everywhere. Then a boundedness relationship can be stated and proven as follows.

<u>Theorem 6.10</u>: Suppose that at a point $x_0 \in (a, b)$ the following conditions are fulfilled:

1*. $p_k(x_0) = O(1)$

2*. $\sigma(x_0) > 0$,

3*. $\left|\dfrac{\sigma(t) - \sigma(x_0)}{t - x_0}\right| = O(1)$ in a neighborhood $x_0 - \epsilon \leq t \leq x_0 + \epsilon$,

4*. $\int_a^b \rho(t)\, \sigma(t)\, p_k^2(t)\, dt = O(1)$,

5*. $\int_a^b \rho(t)\, \dfrac{p_k^2(t)}{\sigma(t)}\, dt = O(1)$

then $q_k(x_0) = O(1)$, where the notation O means that if $\{a_n\}$ is a arbitrary and $\{b_n\}$ a positive sequence of numbers, $a_n = O(b_n)$ signifies that $\dfrac{|a_n|}{b_n}$ is bounded, independent of n.

If the conditions 1* - 3* are uniformly satisfied at the points x of an interval $[c, d] \subset (a, b)$ then $q_n(x) = O(1)$ holds uniformly in $[c, d]$.

Proof: Expanding $q_n(x)$ at the point x_0 in the functions $\{p_n(x)\}$ yields

$$q_n(x_0) = p_n(x_0) \int_a^b \rho(t)\, q_n(t)\, p_n(t)\, dt + \int_a^b \rho(t)\, q_n(t) \sum_{k=0}^{n-1} p_k(t)\, p_k(x_0)\, dt\ .$$

Since $q_n(x)$ is the n^{th} orthonormal polynomial with respect to the weight function $\rho(x)\, \sigma(x)$ it is orthogonal with respect to this weight function to every polynomial of degree less than n. Thus

$$\int_a^b \rho(t)\ \sigma(t)\ q_n(t) \sum_{k=0}^{n-1} p_k(t)\ p_k\left(x_0\right)\ dt = 0\ .$$

Hence because of condition 2* the expansion of q_n can be written

$$q_n\left(x_0\right) = p_n\left(x_0\right) \int_a^b \rho(t)\ q_n(t)\ p_n(t)\ dt$$

$$+ \frac{1}{\sigma\left(x_0\right)} \int_a^b \rho(t) \left[\sigma\left(x_0\right) - \sigma(t)\right] q_n(t) \sum_{k=0}^{n-1} p_k(t)\ p_k\left(x_0\right)\ dt\ .$$

Applying the Cauchy-Schwartz inequality and using condition 5* yields the following estimate for the first integral,

$$\left|\int_a^b\right| \leq \int_a^b \rho(t) \sqrt{\sigma(t)}\ \left|q_n(t)\right| \frac{\left|p_n(t)\right|}{\sqrt{\sigma(t)}}\ dt$$

$$\leq \left\{\int_a^b \rho(t)\ \sigma(t)\ q_n^{\ 2}(t)\ dt \int_a^b \rho(t)\ \frac{p_n^{\ 2}(t)}{\sigma(t)}\ dt\right\}^{\frac{1}{2}} = O(1)\ .$$

Now by considering conditions 1* and 2* the following estimate is obtained

$$\left|q_n\left(x_0\right)\right| = O(1) + O(1) \int_a^b \rho(t) \left|\sigma(t)\right.$$

$$\left. - \sigma\left(x_0\right)\right| \left|q_n(t)\right| \left|\sum_{k=0}^{n-1} p_k(t)\ p_k\left(x_0\right)\right|\ dt\ .$$

Applying the Christoffel-Darboux formula to the sum under the integral sign and using the relationship (6.14) yields

$$\left|q_n(x_0)\right| = O(1) + O(1)\int_a^b \rho(t)\left|\frac{\sigma(t) - \sigma(x_0)}{t - x_0}\right| \left|q_n(t)\right| \left[\left|p_n(t)\right| + \left|p_{n-1}(t)\right|\right] dt \,.$$

Divide the interval of integration into three parts: $[a, x_0 - \epsilon]$, $[x_0 - \epsilon, x_0 + \epsilon]$ and $[x_0 + \epsilon, b]$. On account of 4* and 5* the integral over $[a, x_0 - \epsilon]$ can be estimated as follows:

$$\int_a^{x_0-\epsilon} \leq \frac{1}{\epsilon}\int_a^{x_0-\epsilon} \rho(t)\left[\sigma(x_0) + \sigma(t)\right] \left|q_n(t)\right| \left[\left|p_n(t)\right| + \left|p_{n-1}(t)\right|\right] dt$$

$$= O(1)\int_a^{x_0-\epsilon} \rho(t)\sqrt{\sigma(t)}\left|q_n(t)\right| \left[\frac{\sigma(x_0)}{\sqrt{\sigma(t)}} + \sqrt{\sigma(t)}\right] \left[\left|p_n(t)\right| + \left|p_{n-1}(t)\right|\right] dt$$

$$= O(1)\left\{\int_a^{x_0-\epsilon} \rho(t)\,\sigma(t)\,q_n^2(t)\,dt \int_a^{x_0-\epsilon} \rho(t)\left[\frac{\sigma(x_0)}{\sqrt{\sigma(t)}} + \sqrt{\sigma(t)}\right]^2 \left[\left|p_n(t)\right| + \left|p_{n-1}(t)\right|\right]^2 dt\right.$$

$$= O(1)\left\{\int_a^{x_0-\epsilon} \rho(t)\,\frac{p_n^2(t) + p_{n-1}^2(t)}{\sigma(t)}\,dt + \int_a^{x_0-\epsilon} \rho(t)\,\sigma(t)\left[p_n^2 + p_{n-1}^2(t)\right] dt\right\}^{\frac{1}{2}}$$

$$= O(1) \,.$$

In the same way the estimate $\int_{x_0-\epsilon}^{b} = O(1)$ can be obtained. For the middle integral it is inferred from conditions 3* and 5* that

$$\int_{x_0-\epsilon}^{x_0+\epsilon} = O(1) \int_{x_0-\epsilon}^{x_0+\epsilon} \rho(t) \; |q_n(t)| \; \Big[|p_n(t)| + |p_{n-1}(t)| \Big] \, dt$$

$$= O(1) \left\{ \int_{x_0-\epsilon}^{x_0+\epsilon} \rho(t)\sigma(t) \; q_n^{\,2}(t) \, dt \int_{x_0-\epsilon}^{x_0+\epsilon} \rho(t) \, \frac{p_n^2(t) + \rho_{n-1}^2(t)}{\sigma(t)} \, dt \right\}^{\frac{1}{2}}$$

$$= O(1) \; .$$

Thus the theorem is proven.

Again it seems appropriate to further restrict the class of orthogonal systems under discussion. Thus, in the material that follows interest will be limited to the particular class of orthogonal systems known as Jacobi polynomials.

7. JACOBI POLYNOMIALS

The particular class of orthonormal polynomial systems arising from application of the Gram-Schmidt process to the linearly independent system $\left\{x^n\right\}$ for a weight function $\rho(t) = (b - x)^\alpha (x - a)^\beta$ with $\alpha > -1$, $\beta > -1$ are called Jacobi polynomials. For simplicity the interval $[a, b]$ is usually taken to be $[-1, 1]$ so that the weight function becomes $\rho(x) = (1 - x)^\alpha (1 + x)^\beta$. The general case can be reclaimed from these results by the linear transformation

$$t = -1 + 2\frac{x - a}{b - a} .$$

When $\alpha = \beta$ the corresponding polynomials are called ultraspherical polynomials. Two special cases of particular interest are $\alpha = \beta = 0$ and $\alpha = \beta = -\frac{1}{2}$ which correspond to the Legendre and Chebysheff polynomials, respectively.

Due to the wide range of interest in the Jacobi polynomials they have been treated rather extensively in the literature. Therefore, some of the properties peculiar to these polynomials will be presented without proof, an appropriate reference for the proof being given. First, note that if it is assumed that the Jacobi polynomials are bounded then they become a subset of all previous systems considered in this paper and therefore satisfy the requirements of all preceeding theorems.

Denote by $p_n^{(\alpha,\beta)}(x)$ the nth normed Jacobi polynomial belonging to the weight function $\rho(x) = (1 - x)^\alpha (1 + x)^\beta$ and by $J_n^{(\alpha,\beta)}(x)$ the nth not normed Jacobi polynomial. It has been shown (ref. 8, page 30) that $J_n^{(\alpha,\beta)}(x)$ satisfies the following relationship

$$J_n^{(\alpha,\beta)}(x) = (1 - x)^{-\alpha} (1 + x)^{-\beta} \frac{(-1)^n}{n!2^n} \frac{d^n}{dx^n} \left[(1 - x)^{\alpha+n} (1 + x)^{\beta+n}\right] .$$

(7.1a)

The relationship between $p_n^{(\alpha,\beta)}(x)$ and $J_n^{(\alpha,\beta)}(x)$ is given by (ref. 7, page 172)

$$p_n^{(\alpha,\beta)}(x) = J_n^{(\alpha,\beta)}(x) \Big/ \delta_n^{1/2} \qquad (7.2a)$$

where

$$\delta_n = \frac{2^{\alpha+\beta+1}}{\alpha + \beta + 2n + 1} \frac{\Gamma(\alpha + n + 1)\ \Gamma(\beta + n + 1)}{n!\ \Gamma(\alpha + \beta + n + 1)} \qquad (7.3a)$$

If

$$n = \alpha + \beta + 1 = 0, \quad \delta_0 = \Gamma(\alpha + 1)\ \Gamma(\beta + 1) .$$

The $p_n^{(\alpha,\beta)}(x)$ being orthonormal polynomials in a symmetric interval satisfy the recursion relation developed in section 6:

$$xp_n^{(\alpha,\beta)}(x) = \frac{a_n}{a_{n+1}} p_{n+1}^{(\alpha,\beta)}(x) + \frac{a_{n-1}}{a_n} p_{n-1}(x) . \qquad (7.4a)$$

Evaluation of the leading coefficients yields (ref. 7, page 172)

$$a_n = \frac{1}{2^n\ n!\ \delta_n^{1\ 2}} \frac{\Gamma(\alpha + \beta + 2n + 1)}{\Gamma(\alpha + \beta + n + 1)} . \qquad (7.5a)$$

Hence the recursion relation becomes

$$(\alpha + \beta + 2n)\ (\alpha + \beta + 2n + 1)\ (\alpha + \beta + 2n + 2)\ xJ_n^{(\alpha,\beta)}(x)$$

$$= 2(n + 1)\ (\alpha + \beta + n + 1)\ (\alpha + \beta + 2n)\ J_{n+1}^{(\alpha,\beta)}(x)$$

$$+ 2(\alpha + n)\ (\beta + n)\ (\alpha + \beta + 2n + 2)\ J_{n-1}^{(\alpha,\beta)}(x)$$

(7.6a)

The weight function $\rho(x) = (1 - x)^{\alpha}\ (1 + x)^{\beta}$ satisfies the conditions of theorem 6.5 since

$$\frac{\rho'(x)}{\rho(x)} = \frac{\alpha}{1 - x} + \frac{\beta}{1 + x} = \frac{(\beta - \alpha) - (\alpha + \beta)x}{1 - x^2} \qquad (7.7a)$$

Comparing the coefficients in (7.7a) with those in theorem 6.5 yields

$$A = 1,\ B = 0,\ C = -1,\ D = \beta - \alpha,\ E = -\ (\alpha + \beta).$$

Hence the Jacobi polynomials satisfy the linear homogeneous differential equation

$$\left(1 - x^2\right)\ J_n''^{(\alpha,\beta)}(x) + \left[\beta - \alpha - (\alpha + \beta + 2)x\right]\ J_n'^{(\alpha,\beta)}(x)$$

$$+ n(\alpha + \beta + n + 1)\ J_n^{(\alpha,\beta)}(x) = 0\ .$$

The same equation is obviously satisfied by $p_n^{(\alpha,\beta)}(x)$. From equation (7.4) and the Christoffel-Darboux formula equation (6.10) it follows that the n^{th} kernel of the Jacobi polynomials is given by:

$$K_n(x,t) = \frac{2(n+1)\ \Gamma(\alpha+\beta+n+2)\ \Gamma(\alpha+\beta+2n+1)}{\Gamma(\alpha+\beta+n+1)\ \Gamma(\alpha+\beta+2n+3)} \cdot$$

$$\frac{p_{n+1}{}^{(\alpha,\beta)}(t)\ p_n{}^{(\alpha,\beta)}(x) - p_n{}^{(\alpha,\beta)}(t)\ p_{n+1}{}^{(\alpha,\beta)}(x)}{t - x} \qquad (7.9a)$$

For the Legendre polynomials equations (7.1a), (7.2a), (7.3a), (7.5a), (7.6a), (7.8a), and (7.9a) reduce, respectively, to:

$$P_n(x) = \frac{(-1)^n}{n!\ 2^n} \frac{d^n}{dx^n} \left(1 - x^2\right)^n \qquad (7.1b)$$

$$p_n{}^{(0,0)}(x) = P_n(x) \Big/ \delta_n^{\frac{1}{2}} \qquad (7.2b)$$

where

$$\delta_n = \frac{2}{2n+1} \qquad (7.3b)$$

$$a_n = \frac{1}{2^n\ n!} \frac{\Gamma(2n+1)}{\delta_n^{1/2}\ \Gamma(n+1)} \qquad (7.5b)$$

$$(2n+1)\ xP_n(x) = (n+1)\ P_{n+1}(x) + nP_{n-1}(x) \qquad (7.6b)$$

$$\left(1 - x^2\right) P_n''(x) - 2x\ P_n'(x) + n(n+1)\ P_n(x) = 0 \qquad (7.8b)$$

$$K_n(x,t) = \frac{n+1}{2} \frac{P_{n+1}(t)\ P_n(x) - P_n(t)\ P_{n+1}(x)}{t - x} \qquad (7.9b)$$

For the Chebysheff polynomials these equations become:

$$C_n(x) = \left(1 - x^2\right)^{\frac{1}{2}} \frac{(-1)^n}{n!\ 2^n} \frac{d^n}{dx^n} \left(1 - x^2\right)^{n - \frac{1}{2}} \qquad (7.1c)$$

$$p_n^{\left(-\frac{1}{2},\ -\frac{1}{2}\right)}(x) = C_n(x) \Big/ \delta_n^{\frac{1}{2}} \qquad (7.2c)$$

where

$$\delta_n = \frac{1}{2n} \frac{\left[\Gamma\left(n + \frac{1}{2}\right)\right]^2}{n!\ \Gamma(n)} \qquad (7.3c)$$

$$a_n = \frac{1}{2^n\ n!} \frac{\Gamma(2n)}{\delta_n^{1/2}\ \Gamma(n)} \qquad (7.5c)$$

$$n(2n - 1)\ (2n + 1)\ xC_n(x) = n(n + 1)\ (2n - 1)\ C_{n+1}(x) + \left(n - \frac{1}{2}\right)^2 (2n + 1)\ C_{n-1}(x) \qquad (7.6c)$$

$$\left(1 - x^2\right) C_n''(x) - xC_n'(x) + n^2\ C_n(x) = 0 \qquad (7.8c)$$

$$K_n(x,t) = \frac{4n(n + 1)\ n!\ \Gamma(n + 1)\ \Gamma(2n)}{\Gamma(2n + 2)\ \left[\Gamma\left(n + \frac{1}{2}\right)\right]^2} \frac{C_{n+1}(t)\ C_n(x) - C_n(t)\ C_{n+1}(x)}{t - x} \qquad (7.9c)$$

The normed Chebysheff polynomials are sometimes introduced by the following definition

$$p_0^{\left(-\frac{1}{2},\ -\frac{1}{2}\right)}(x) \equiv \frac{1}{\sqrt{\pi}}$$

$$p_n^{\left(-\frac{1}{2}, -\frac{1}{2}\right)}(x) = \sqrt{\frac{2}{\pi}} \cos (n \text{ arc cos } x), \; n = 1, 2, \cdots \tag{7.10}$$

To show that this definition is valid, introduce the variable $t = \text{arc cos } x$. Then for $m \neq n$

$$\int_{-1}^{1} \frac{1}{\sqrt{1 - x^2}} \cos (m \text{ arc cos } x) \cos (n \text{ arc cos } x) \, dx$$

$$= \int_0^{\pi} \cos mt \cos nt \, dt = 0$$

and for $m = n$

$$\int_{-1}^{1} \frac{1}{\sqrt{1 - x^2}} \cos^2 (n \text{ arc cos } x) \, dx = \int_0^{\pi} \cos^2 nt \, dt$$

$$= \begin{cases} \frac{\pi}{2} \; , & \text{for } n \geq 1 \\ \pi \; , & \text{for } n = 0 \end{cases}$$

Hence the system $\left\{\cos (n \text{ arc cos } x)\right\}$ is orthogonal with respect to the weight function $\rho(x) = \left(1 - x^2\right)^{-1/2}$.

From the relationship

$$\cos nt = 2^{n-1} \cos^n t + \sum_{k=0}^{n-1} \alpha_k \cos^k t \quad (n \geq 1)$$

It follows that $\cos (n \text{ arc cos } x)$ is a polynomial of degree n

having the leading coefficient 2^{n-1}. Since by theorem 6.1 the orthonormal system of polynomials, with respect to the weight function $\rho(x) = \left(1 - x^2\right)^{-1/2}$ with positive leading coefficients in unique it follows that $\left\{\frac{1}{\sqrt{\pi}} ; \sqrt{\frac{2}{\pi}} \cos (n \text{ arc} \cos x)\right\}$ is the system of Chebysheff polynomials.

It can also be seen from the transformation $t = \text{arc} \cos x$ that

$$\left| p_n^{\left(-\frac{1}{2}, -\frac{1}{2}\right)}(x) \right| = \left| \sqrt{\frac{2}{\pi}} \cos nt \right| \leq \sqrt{\frac{2}{\pi}}$$

for $x \in (-1, 1)$. Now let the $\sigma(x)$ of theorem 6.10 be defined

$$\sigma(x) = (1 - x)^{\alpha + \frac{1}{2}} (1 + x)^{\beta + \frac{1}{2}}$$

Then $\sigma(x)$ satisfies conditions 1*, 2*, and 3* of theorem 6.10 and conditions 4* and 5* are fulfilled since the Chebysheff polynomials are uniformly bounded. Hence it can be concluded that the Jacobi polynomials in general are uniformly bounded in (-1, 1).

For one additional system to be considered later in an applied problem, let $\alpha = \beta = 1/2$. Then the above properties reduce to:

$$B_n(x) = \left(1 - x^2\right)^{-\frac{1}{2}} \frac{(-1)^n}{n!\, 2^n} \frac{d^n}{dx^n} \left(1 - x^2\right)^{n + \frac{1}{2}} \tag{7.1d}$$

$$p_n^{\left(\frac{1}{2}, \frac{1}{2}\right)}(x) = B_n(x) \Big/ \delta_n^{\frac{1}{2}} \tag{7.2d}$$

where

$$\delta_n = \frac{2}{(n + 1)\, n!} \frac{\left[\Gamma\left(n + \frac{1}{2}\right)\right]^2}{\Gamma(n + 2)}$$

$$a_n = \frac{1}{2^n \, n!} \frac{\Gamma(2n + 2)}{\delta_n^{1/2} \, \Gamma(n + 2)} \tag{7.5d}$$

$$(2n + 1)\,(n + 1)\,(2n + 3)\, xB_n(x) = (n + 1)\,(n + 2)\,(2n + 1)\, B_{n+1}(x) + \left(n + \frac{1}{2}\right)^2 (2n + 3)\, B_{n-1}(x) \tag{7.6d}$$

$$\left(1 - x^2\right) B_n''(x) - 3x\, B_n'(x) + n(n + 2)\, B_n(x) = 0 \tag{7.8d}$$

$$K_n(x,t) = \frac{(n + 1)^2\, \Gamma(n + 3)\, \Gamma(2n + 3)\, n!}{\Gamma(2n + 4) \left[\Gamma\left(n + \frac{1}{2}\right)\right]^2} \frac{B_{n+1}(t)\, B_n(x) - B_n(t)\, B_{n+1}(x)}{t - x} \tag{7.9d}$$

The remainder of this paper will consist of applying the three particular Jacobi polynomials introduced here to the problem of solving Laplace's equation with boundary conditions.

8. APPLICATION OF JACOBI POLYNOMIALS TO THE PROBLEM OF GRAVITATIONAL FIELD DETERMINATION

Application of Newton's law of gravitation to an inertially fixed bounded distribution of matter M shows that the gravitational attraction at a point P exterior to M can be represented as the gradient of a scalar function U. (ref. 12.) The physical assumption that there exist no gravity sources or sinks exterior to M is the mathematical equivalent of saying U is a harmonic function in the region exterior to M, that is, satisfies Laplace's equation:

$$\nabla^2 U = 0 .$$

Represented in spherical coordinates Laplace's equation has the form

$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial U}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial U}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 U}{\partial\phi^2} = 0 .$$

where r is radius, θ is colatitude, ϕ is longitude. Simplifying the problem by assuming no longitude dependence yields:

$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial U}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta \frac{\partial U}{\partial\theta}\right) = 0 . \qquad (8.1)$$

Separation of variables by assuming $U(r, \theta) = R(r)\,T(\theta)$ leads to a system of differential equations of the following form.

$$r^2 R'' + 2rR' - n(n+1)R = 0 \qquad (8.2)$$

$$T'' + \cot\theta\, T' + n(n+1)T = 0 \qquad (8.3)$$

Transforming equation (8.3) by letting $x = \cos\theta$ yields

$$\left(1 - x^2\right) T'' - 2xT' + n(n + 1) T = 0 \qquad (8.3a)$$

which is Legendre's differential equation. Application of the physical constraints that the gravitational force vanish at infinity and be continuous in the region exterior to M yields the following particular solution. (ref. 7, page 109.)

$$U_n(r, \theta) = \frac{A_n P_n (\cos\theta)}{r^{n+1}} \; ; \; r > 1; \; 0 \leq \theta \leq \pi \qquad (8.4)$$

Suppose that boundary conditions are given at some fixed radius, r_o, by a suitably well behaved function $F\left(f_o, \theta\right)$. The solution to (8.1) which satisfies these boundary conditions is given as follows.

$$U(r, \theta) = \sum_{n=0}^{\infty} \frac{A_n P_n (\cos\theta)}{r^{n+1}} \qquad (8.5)$$

where the coefficients A_n are chosen such that

$$U\left(r_o, \theta\right) = F\left(r_o, \theta\right) \qquad (8.6)$$

In applying the procedure outlined above to the problem of determining the gravitational field of a celestial body, the bounded distribution of mass is defined to be the celestial body of interest. The corresponding gravitational field is defined as the gradient of a potential function of the form (8.5) and is assumed to be sufficiently well behaved so that this series converges uniformly. The boundary conditions are frequently defined through observations of a spacecraft orbiting the body. These observations consist of measurement of some property of the spacecraft's motion and relating them to the gravitational force field constitutes a rather sophisticated problem in the field of dynamics. Let it be assumed that such a relationship has been provided so that for a physically observable property of spacecraft motion, say $f(t)$, there exists a mathematical model which yields a corresponding computed value, $f_{comp}(t)$. The observable $f(t)$ is certainly a function of forces acting on the spacecraft. In the context of this discussion it is only the gravitational force which must be considered. Since this force is defined as the gradient of the function (8.5),

$$f(t) = f\left(A_0, A_1, \cdots, A_n, \cdots, t\right).$$

Suppose that at times t_i, $i = 1, 2, \cdots, K, \cdots$ observations of f are made which are corrupted by errors ϵ_i that is, the observation is $f_{ob}\left(t_i\right) = f\left(t_i\right) + \epsilon_i$. The problem then is to determine the constant coefficients A_n from the discrete observations $f_{ob}\left(t_i\right)$. One of the more common techniques for handling this problem is the differential correction, weighted-least-squares procedure which is discussed in detail in reference 13 and in outline form below.

The differential correction, least-squares technique proceeds as follows. Given the functional form of a measurable event; $f(t) = f(t,\bar{A})$ where $\bar{A}$ is in this case a finite set $\left\{A_i\right\}, i \leq n$ of coefficients in a series approximation to the potential equation; linearize the relationship by setting

$$df(t) = \left(\frac{\partial f(t)}{\partial \bar{A}}\right) d\bar{A} \ . \tag{8.7}$$

Now if K observations, $K > n$, of $f(t)$ have been made at times t_i, $i = 1, \cdots, K$ there exists the set $\left\{f_{ob}\left(t_i, \bar{A}\right)\right\}$. Corresponding to each of these observations a value of $f(t)$ is computed based on some initial estimate $\bar{A}_o$ of the coefficients $\bar{A}$. This generates the set $\left\{f_{comp}\left(t_i, \bar{A}_o\right)\right\}$. Analogous to the differential equation (8.7) the following incremental correspondence is set up at each time t_i:

$$\begin{aligned} \Delta f\left(t_i\right) &= f_{comp}\left(t_i, \bar{A}_o\right) - f_{ob}\left(t_i, \bar{A}\right) \\ &= \left(\left.\frac{\partial f\left(t_i\right)}{\partial \bar{A}}\right|_{\bar{A}=\bar{A}_o}\right) \Delta \bar{A}_o + \epsilon_i \end{aligned} \tag{8.8}$$

where $\Delta \bar{A}_o = \bar{A}_o - \bar{A}$ and ϵ_i is an error arising from such sources as measurement and computational errors, nonlinearities in the relationship and an incomplete mathematical model due to considering only a finite number of terms in the approximation to the potential function.

Simultaneous consideration of all observations yields the following linear matrix equation whose elements are given by equation (8.8).

$$\Delta \bar{f} = B \Delta \bar{A}_o + \bar{\epsilon} \tag{8.9}$$

where

$$\Delta \bar{f} = K \times 1 \text{ known matrix,}$$

$$\Delta \bar{A}_o = n \times 1 \text{ fixed but unknown matrix,}$$

$$\bar{\epsilon} = K \times 1 \text{ unknown matrix,}$$

$$B = K \times n \text{ known matrix of coefficients.}$$

The problem then is to find the "best" estimate $\Delta\hat{A}_o$ of $\Delta\bar{A}_o$ in terms of B and $\Delta\bar{f}$. Denote by $\Delta\hat{f} = B\Delta\hat{A}_o$ the "best" estimate of the true observable and $\Delta\bar{f} - \Delta\hat{f} = \Delta\bar{f} - B\Delta\hat{A}_o = \hat{\epsilon}$ as the "best" estimate of the residual vector. The criterion of "best" as defined in the least-squares procedure is that estimate $\Delta\hat{A}_o$ which minimizes the sum of the squares of the components of the residual vector $\hat{\epsilon}$; that is, minimizes $\hat{\epsilon}^T \hat{\epsilon}$. It is shown in reference 13 that this estimate is given by:

$$\Delta\hat{A}_o = \left[B^T B\right]^{-1} B^T \Delta\bar{f} .$$

Substituting from equation (8.8) yields

$$\Delta\hat{A}_o = \left[\sum_{i=1}^{K}\left(\frac{\partial f(t_i)}{\partial \bar{A}}\Bigg|_{\bar{A}=\bar{A}_o}\right)^T\left(\frac{\partial f(t_i)}{\partial \bar{A}}\Bigg|_{\bar{A}=\bar{A}_o}\right)\right]^{-1}\left[\sum_{i=1}^{K}\left(\frac{\partial f(t_i)}{\partial \bar{A}}\Bigg|_{\bar{A}=\bar{A}_o}\right)\Delta f(t_i)\right] \tag{8.10}$$

where $\Delta\hat{A}_o + \bar{A}_o$ minimizes

$$\sum_{i=1}^{K}\left[f_{ob}\left(t_i, \bar{A}\right) - f_{comp}\left(t_i, \bar{A}_o\right)\right]^2 . \tag{8.11}$$

If observations are made with sufficient frequency to justify substituting integration for summation the general element of the matrix $B^T B$ can be written

$$\left(B^T B\right)_{lm} = \int_{t_i}^{t_k} \left(\frac{\partial f(t)}{\partial A_l} \Bigg|_{\bar{A}=\bar{A}_o} \right) \left(\frac{\partial f(t)}{\partial A_m} \Bigg|_{\bar{A}=\bar{A}_o} \right) dt \; , \tag{8.12}$$

and the function to be minimized can be written

$$\int_{t_i}^{t_k} \left[f_{ob}\,(t, \bar{A}) - f_{comp}\,\left(t, \bar{A}_o\right) \right]^2 dt \; . \tag{8.13}$$

Consider now the following application of the above procedures to a problem of determining the gravitational field of a celestial body. Assume that the gravitational field has no longitude dependence so that its potential function may be given by equation (8.5). For simplicity, assume that the potential can be directly measured at some fixed radius r_o so that the boundary conditions thus provided are in the form of equation (8.6). Since this is a practical application only a finite number of terms can be considered, thus the problem may be stated as follows.

Based upon discrete observations of the boundary conditions $U\left(r_o, \theta\right) = F\left(r_o, \theta\right)$ determine the coefficients A_n in the finite approximation

$$U_m\left(r, \theta, P_n\right) = \sum_{n=0}^{m} \frac{A_n\, P_n\,(\cos\theta)}{r^{n+1}} \tag{8.14}$$

such that

$$\sum_{i=1}^{k} \left[F\left(r_o, \theta_i\right) - U_m\left(r_o, \theta_i, P_n\right)\right]^2 \qquad (8.15)$$

is a minimum. Since it is assumed that r_o is known, let U_m be written in the form

$$U_m\left(r_o, \theta, P_n\right) = \sum_{n=0}^{m} a_n P_n (\cos \theta) \qquad (8.16)$$

where the a_n are the coefficients to be determined.

Applying the least-squares procedures in order to update initial estimates $_o\bar{a}$ of the coefficients $\bar{a}$ yields

$$\Delta_o\bar{a} = \left[B^T B\right]^{-1} B^T \Delta\bar{f}$$

where

$$\left(B^T B\right)_{lj} = \sum_{i=1}^{k} P_l (\cos \theta) P_j (\cos \theta) . \qquad (8.17)$$

In the limiting case of continuous observation the following relationships hold.

$$\left(B^T B\right)_{lj} = \int_{\theta_1}^{\theta_k} P_l (\cos \theta) P_j (\cos \theta) d\theta \qquad (8.18)$$

$$\left(B^T \Delta \bar{f}\right)_l = \int_{\theta_1}^{\theta_k} P_l (\cos \theta) \left[F\left(r_o, \theta\right) - \sum_{n=0}^{m} {}_oa_n P_n (\cos \theta) \right] d\theta$$

$$= \int_{\theta_1}^{\theta_k} F\left(r_o, \theta\right) P_l (\cos \theta) \, d\theta$$

$$- \sum_{n=0}^{m} {}_oa_n \int_{\theta_1}^{\theta_k} P_l (\cos \theta) \, P_n (\cos \theta) \, d\theta \qquad (8.19)$$

and $\Delta_o \bar{a}$ minimizes

$$\int_{\theta_1}^{\theta_k} \left[F\left(r_o, \theta\right) - U_m\left(r_o, \theta\right) \right]^2 d\theta \qquad (8.20)$$

The case just presented is a simplified analogue of the techniques frequently applied in practice in the area of gravitational field determination. This formulation was seen to be quite natural since the separation of variables in Laplace's equation and the transformation from colatitude to cosine of colatitude led to consideration of Legendre's equation, with the independent variable ranging over the interval $[-1,1]$. Due to the completeness of the Legendre polynomials in $[-1,1]$ the requirement to satisfy the boundary condition $F\left(r_o, \theta\right)$, $0 \leq \theta \leq \pi$, did not necessitate considering another form of approximation. However, this particular expansion leaves much to be desired when the techniques of least-squares are applied to determine coefficients. To be more specific, consider the case of continuous observation represented by equations (8.18) - (8.20). It is seen that even if observations span the interval $[0, \pi]$ in colatitude the integral in equation (8.18) is in general not zero, since the orthogonality condition for the Legendre polynomials can be written

$$\int_{-1}^{1} P_l(\cos\theta)\, P_m(\cos\theta)\, d\cos\theta = \int_{0}^{\pi} \sin\theta\, P_l(\cos\theta)\, P_m(\cos\theta)\, d\theta = 0 .$$

It is apparent from both equations (8.18) and (8.19) that the solution would be simplified if the polynomials used in the approximation were chosen to be orthogonal with respect to the unit weight function in the interval $[0, \pi]$. If they had been so chosen then it follows from theorem 5.7 and equation (8.20) that the least squares solution yields the Fourier coefficients of $F(r_o, \theta)$. Also by theorem 6.3, the minimum value for equation (8.20) (among all polynomial approximations) would have been obtained. In addition the matrix inversion required in equation (8.10) becomes a trivial operation.

Although the above features may not be essential, a serious problem in the practical application of the Legendre expansion can be rectified if they are required. This problem arises when in practice consideration must be restricted to a finite term approximation for a function which requires an infinite term expansion. To illustrate this, consider the following situation.

Let $y(\bar{x}) = y(x_1, x_2, \ldots, x_n)$ be a linear function of the n parameters x_i which are to be determined.

Then

$$dy = \left(\frac{\partial y}{\partial \bar{x}}\right) d\bar{x} = B\, d\bar{x} \ . \qquad (8.21)$$

assume that due to practical limitations only $m < n$ of these parameters can be considered at one time. Then the matrix equation (8.21) can be written in partitioned form as follows:

$$dy = \left[B_1 \ \vdots \ B_2 \right] \begin{bmatrix} d\bar{x}_1 \\ --- \\ dx_2 \end{bmatrix} .$$

The least-squares procedures leads to the matrix equation

$$\begin{bmatrix} B_1{}^T B_1 & B_1{}^T B_2 \\ & \\ B_2{}^T B_1 & B_2{}^T B_2 \end{bmatrix} \begin{bmatrix} d\bar{x}_1 \\ \\ d\bar{x}_2 \end{bmatrix} = \begin{bmatrix} B_1{}^T\ dy \\ \\ B_2{}^T\ dy \end{bmatrix} \qquad (8.22)$$

Solving the first of these equations for $d\bar{x}_1$ yields an expression which shows the dependence of $d\bar{x}_1$ on both the magnitude of the parameters in $d\bar{x}_2$ and the matrix $B_1{}^T B_2$:

$$d\bar{x}_1 = \left[B_1{}^T B_1 \right]^{-1} \left\{ B_1{}^T\ dy - B_1{}^T B_2\ d\bar{x}_2 \right\} . \qquad (8.23)$$

However under the restriction that the parameters in $\bar{x}_2$ cannot be considered, the least-squares solution obtained is now

$$d\hat{x}_1 = \left[B_1{}^T B_1 \right]^{-1} B_1{}^T\ dy \ . \qquad (8.24)$$

Subtracting equation (8.23) from (8.24) shows the bias introduced in the solution for $d\bar{x}_1$ to be

$$d\hat{x}_1 - d\bar{x}_1 = \left[B_1^T B_1\right]^{-1} \left[B_1^T B_2\right] d\bar{x}_2 \qquad (8.25)$$

For the Legendre expansion it was shown in equation (8.18) that the elements of $\left[B_1^T B_2\right]$ would in general be nonzero. Thus, the parameter estimates depend upon the parameter set being solved for. This is a highly undesirable situation since quite often in practice the incompleteness of the mathematical model is discovered after the fact. Subsequent extension of the model makes previous results invalid since the inclusion of additional parameters will effect the parameter estimates already obtained.

A second variation of this problem occurs when it is not possible to consider simultaneously all parameters and thus the solution must be sequentially conducted for subsets of the total parameter set. In this situation all estimates obtained will be biased.

Suppose then that a system of polynomials in cosine of colatitude, orthogonal with respect to the unit weight function in $0 \leq \theta \leq \pi$, had been used in equation (8.14). Then the off-diagonal terms of the matrix $B^T B$ are zero, hence the desirable features noted above are realized. The most important advantage is that parameter estimates no longer depend upon the parameter set being solved for. The existence of such an orthogonal system will now be demonstrated.

Consider the Chebysheff polynomials $C_n(x)$ of section 7. The orthogonality property of these polynomials is

$$\int_{-1}^{1} (1 - x^2)^{-1/2} C_n(x)\, C_m(x)\, dx = 0, \; n \neq m.$$

Letting $\cos\theta = x$ transforms this equation to the following form:

$$\int_{-1}^{1} (1 - \cos^2\theta)^{-1/2} C_n(\cos\theta)\, C_m(\cos\theta)\, d\cos\theta$$

$$= \int_{0}^{\pi} C_n(\cos\theta)\, C_m(\cos\theta)\, d\theta = 0\,. \tag{8.26}$$

It remains then to express the function $U_m(r, \theta, P_n)$ of equation (8.16) in terms of Chebysheff polynomials. This is certainly possible since the n^{th} polynomial in any particular Jacobi system is an n^{th} degree polynomial, it can obviously be expressed as a linear combination of the first n polynomials in any other Jacobi system. This transformation from the first few Legendre polynomials to Chebysheff and the B_n polynomials will be given in the next section. Hence, in terms of Chebysheff polynomials the finite term appoximation of $F(r_o, \theta)$ becomes

$$U_m(r_o, \theta, C_n) = \sum_{n=0}^{m} \frac{A_n}{r_o^{\,n+1}} \sum_{k=0}^{n} S_k\, C_k(\cos\theta) \tag{8.27}$$

As in the establishment of equation (8.16), since the radius r_o is assumed constant and known, let equation (8.27) be written in the form

$$U_m(r_o, \theta, C_n) = \sum_{n=0}^{m} b_n\, C_n(\cos\theta) \tag{8.28}$$

where the coefficients b_n are to be determined.

Before leaving this section, consider one additional problem encountered in gravitational field determination based on observations of an artificial satellite. Generally the satellite is placed in an orbit inclined to the equatorial plane of the celestial body. (see Fig. 1). In many instances this inclination is small so that the satellite perturbations due to gravitational forces arising in the polar regions of the celestial body are extremely small, in some cases not measurable. (Ref. 14). This situation has the analogue under the simplifying assumptions of this section, of being able to observe the boundary conditions $F(r_o, \theta)$ only in the restricted region $0 < a \leq \theta \leq b < \pi$. The problem then is to determine the coefficients in a finite term approximation to $F(r_o, \theta)$ based on observations in this restricted region in such a way as to yield the most valid representation of $F(r_o, \theta)$ in the whole interval $[0, \pi]$.

In light of the above discussion, consider the Fourier coefficients of $F(r_o, \theta)$ for the Legendre, Chebysheff and B_n polynomials of section 7, (these coefficients include as a factor the norm of the n^{th} polynomial)

$$b_n = \int_0^\pi F(r_o, \theta)\, C_n(\cos\theta)\, d\theta$$

$$a_n = \int_0^\pi \sin\theta\, F(r_o, \theta)\, P_n(\cos\theta)\, d\theta$$

$$d_n = \int_0^\pi \sin^2\theta\, F(r_o, \theta)\, B_n(\cos\theta)\, d\theta$$

Thus it appears that the Fourier coefficients of the B_n polynomials are least dependent upon information about $F(r_o, \theta)$ in the polar regions and hence these polynomials may be a more appropriate selection for the expansion of $F(r_o, \theta)$ in the case of restricted range of observation.

To check the observations made in this section, the final section of the paper will consist of numerical results obtained by programing for digital computer solution, a least-squares, differential correction, discrete observation procedure for solving equation (8.5) with boundary conditions (8.6) in terms of Legendre formulation (8.14), the Chebysheff formulation (8.28) and the B_n polynomial formulation given below.

$$U_m(r_o, \theta, B_n) = \sum_{n=0}^{m} \frac{A_n}{r_o^{n+1}} \sum_{k=0}^{n} T_k(n)\, B_k(\cos\theta)$$

$$= \sum_{n=0}^{m} d_n B_n(\cos\theta) \qquad (8.29)$$

9. NUMERICAL RESULTS

This section presents numerical results obtained by programing for digital computer solution a least-squares, differential correction, discrete observation procedure for solving equation (8.5) with boundary condition (8.6) in terms of the Legendre expansion (8.14), the Chebysheff expansion (8.28) and the B_n polynomial expansion (8.29). Table I lists the first four polynomials in each of these systems and figure 2 shows plots of these polynomials.

On the strength of the Weierstrass theorem it was assumed that the boundary condition $F\left(r_o, \theta\right)$ was of the form.

$$F\left(r_o, \theta\right) = \sum_{n=0}^{\infty} \frac{e_n}{r_o^{n+1}} \cos^n \theta \qquad (9.1)$$

To simplify the analytic determination of the Fourier coefficients of $F\left(r_o, \theta\right)$, the series was terminated at $n = 4$.

The solution to equation (8.5) which satisfies (9.1) is

$$U_4\left(r, \theta, P_n\right) = \sum_{n=0}^{4} \frac{A_n P_n (\cos \theta)}{r^{n+1}} \quad . \qquad (9.2)$$

Evaluation at $r = r_o$ yields

$$U_4\left(r_o, \theta, P_n\right) = \sum_{n=0}^{4} a_n P_n (\cos \theta) \quad .$$

It can be verified from table I that the following relationships hold between the Legendre, B_n and Chebysheff polynomials.

$$P_o = B_o = C_o$$

$$P_1 = \frac{2}{3} B_1 = 2C_1$$

$$P_2 = \frac{3}{5} B_2 - \frac{1}{8} B_o = 2C_2 + \frac{1}{4} C_o$$

$$P_3 = \frac{4}{7} B_3 - \frac{1}{6} B_1 = 2C_3 + \frac{3}{4} C_1$$

$$P_4 = \frac{5}{9} B_4 - \frac{3}{16} B_2 + \frac{3}{128} B_o = 2C_4 + \frac{5}{6} C_2 + \frac{9}{64} C_o$$

Thus in terms of the B_n polynomials equation (9.2) takes the form

$$U_4\, r,\left(\theta,\ B_n\right) = B_o\left(\frac{A_o}{r} - \frac{A_2}{8r^3} + \frac{3}{128}\frac{A_4}{r^5}\right) + B_1\left(\frac{2A_1}{3r^2} - \frac{A_3}{6r^4}\right) + B_2\left(\frac{2A_2}{5r^3} - \frac{3A_4}{16r^5}\right)$$

$$+ B_3\left(\frac{4A_3}{7r^4}\right) + B_4\left(\frac{5A_4}{9r^5}\right) \quad . \tag{9.3}$$

Evaluation at $r = r_o$ yields

$$U_4\left(r_o,\ \theta,\ B_n\right) = \sum_{n=0}^{4} d_n\, B_n\, (\cos\theta)$$

where

$$A_o = r_o\left[d_o + \frac{5}{24} d_2 + \frac{9}{320} d_4\right]$$

$$A_1 = \frac{3r_o^{\,2}}{2}\left[d_1 + \frac{7}{24} d_3\right]$$

$$A_2 = \frac{5r_o^2}{2} \left[d_2 + \frac{27}{80} d_4 \right]$$

$$A_3 = \frac{7r_o^4}{4} d_3$$

$$A_4 = \frac{9r_o^5}{5} d_4 \; . \tag{9.4}$$

Similarly

$$U_4(r, \theta, C_n) = C_o\left(\frac{A_o}{r} + \frac{A_2}{4r^3} + \frac{9A_4}{64r^5}\right) + C_1\left(\frac{2A_1}{r^2} + \frac{3A_3}{4r^4}\right) + C_2\left(\frac{2A_2}{r^3} + \frac{5A_4}{6r^5}\right)$$

$$+ C_3\left(\frac{2A_3}{r^4}\right) + C_4\left(\frac{2A_4}{r^5}\right) \; . \tag{9.5}$$

Evaluation at $r = r_o$ yields

$$U_4(r_o, \theta, B_n) = \sum_{n=0}^{4} b_n C_n (\cos \theta)$$

where

$$A_o = r_o\left[b_o - \frac{b_2}{8} - \frac{7}{1026} b_4\right]$$

$$A_1 = \frac{r_o^2}{2}\left[b_1 - \frac{3}{8} b_3\right]$$

$$A_2 = \frac{r_o^3}{2}\left[b_2 - \frac{5}{12} b_4\right]$$

$$A_3 = \frac{r_o^{\,4}}{2} b_3$$

$$A_4 = \frac{r_o^{\,5}}{2} b_4 \; . \qquad (9.6)$$

The Fourier coefficients for these three series are presented in table II.

Table III presents the Fourier coefficients evaluated at

$r_o = 1.1$	$e_2 = 0.5$
$e_o = 10.0$	$e_3 = .33$
$e_1 = 1.0$	$e_4 = .25$

and the results of the least squares determination of the coefficients in the Legendre, B_n and Chebysheff expansions of $F(r_o, \theta)$. Five cases are presented for observations of $F(r_o, \theta)$ at 0.5 degree intervals for $0 \leq \theta \leq \pi$. Case 5a corresponds to a solution for the coefficients of a five term approximation and thus represents a complete mathematical model. Cases 4a, 3a, 2a, and 1a correspond respectively to solutions for the coefficients in a 4, 3, 2, and 1 term expansion of $F(r_o, \theta)$. Examination of case 5a on tables IIIa, IIIb, and IIIc shows that in each expansion the least squares solution yields the Fourier coefficients when the mathematical model is complete. Examination of table IIIc shows this to be true for all cases (within an allowable tolerance for numerical error) when the expansion is in terms of Chebysheff polynomials. However, tables IIIa and IIIb shows that the coefficients obtained are not the Fourier coefficients in the case of an incomplete mathematical model due to the bias discussed in section 8 and given by equation (8.25).

Insight into the source of bias can be gained from inspection of the inverse of the normal matrix $\left[B^T B\right]^{-1}$ of equation (8.10), which is rewritten below in a slightly different form.

$$\Delta\hat{A}_0 = \left[DND\right] B^T \Delta\bar{f} \qquad (9.7)$$

where

$$DND = \left[B^T B\right]^{-1}$$

$$D = \text{diag}\left(\frac{1}{\sqrt{B^T B\,{}^{-1}_{ii}}}\right).$$

Note that N is a normalized matrix in the sense that its diagonal elements are unity and its off-diagonal elements are less than or equal to unity in absolute value, since by construction $\left[B^T B\right]^{-1}$ is symmetric and positive semi-definite. Thus, with each parameter a_i in $\Delta\hat{A}$ of equation (9.7) there can be associated a vector which is the i^{th} row of N. Call this vector $\bar{V}_i$. Then a measure of the linear correlation between a_i and a_j can be obtained by $\frac{\bar{V}_i \cdot \bar{V}_j}{\left|\bar{V}_i\right| \left|\bar{V}_j\right|}$, which has the absolute value 1 when a_i and a_j are perfectly correlated and 0 when they are completely independent.

The N matrices corresponding to cases 5a of tables IIIa, IIIb, and IIIc are shown in table IV. It can be inferred from mental approximation of the vector product discussed above that N_{ij} can be used as a measure of the correlation between the i^{th} and j^{th} coefficients being solved for. From table IV it is seen that the highest correlations exist for the B_n expansion and the lowest for the Chebysheff expansion. (The Chebysheff correlations would have been zero for continuous observation of $F(r_o, \theta)$).

Since matrix inversion is involved in the solution process, high correlations indicate an ill-conditioned numerical process. This can be verified from an inspection of the sum of the squares of the residuals resulting from the solutions obtained in cases 5a-1a for each expansion. Since each n^{th} term expansion is an n^{th} degree polynomial, the Legendre, B_n and Chebysheff expansions should yield almost equivalent approximations to $F(r_o, \theta)$ in the event their coefficients can be determined with equal numerical accuracy. The sum of squares for each of the above cases is shown in table V and verifies this assertion. For case 5a the difference in the sum of squares is due to the numerical ill-conditioning arising from high correlations. Reduction of the dimension of the matrix to be inverted in cases 4a-1a improves the numerical conditioning and yields essentially identical approximations. For completeness, residual plots corresponding to cases 5a-1a are shown in figure 3.

Analogous to the cases presented above, a corresponding sequence of cases (5b-1b) were run for a restricted region of observation, $60^\circ \leq \theta \leq 120^\circ$. The resulting solutions are shown in table VI, illustrating the effect of correlations in all cases. The N matrices for case 5b of each expansion is shown in table VII, showing that for this region of observation the B_n polynomials are the least correlated and hence solution in terms of them yields the most valid solution as is borne out by the sum of squares shown in table VIII. Again reduction of the dimension of the matrix improves numerical conditioning and subsequent incomplete expansions yield comparable approximations to $F(r_o, \theta)$. Again for completeness the residuals corresponding to cases 5b-1b are shown in figure 4.

10. CONCLUSIONS BASED ON NUMERICAL RESULTS

The theoretical assertions of section 8 have been verified by the numerical results presented in section 9. It has been shown that for a differential correction least-squares solution based on discrete observations, the greatest numerical accuracy is achieved when the partial derivatives occurring in the normal matrix are orthogonal over the region of observation. This leads to a better approximation to the function being observed. Analogously for a restricted region of observation, $0 < a \leq \theta \leq b < \pi$, it has been shown that the most valid approximation to the observable in the region $0 \leq \theta \leq \pi$ in terms of polynomial systems orthogonal with respect to a weight function in $[0, \pi]$ is obtained when the observable is expanded in terms of the system which is more nearly orthogonal with respect to unity in $[a, b]$.

Finally, for the problem considered it was shown that for an expansion of the observable in terms of polynomial systems orthogonal with respect to a weight function in $[0, \pi]$, completeness of the mathematical model is sufficient for the least-squares procedure to yield the Fourier coefficients, regardless of the weight function. However, for an incomplete mathematical model the Fourier coefficients were obtained only for the expansion in terms of a system orthogonal with respect to the unit weight function.

ACKNOWLEDGMENTS

The author wishes to express his deepest gratitude to Professor W. E. Pace under whose guidance this investigation was conducted. The author is also indebted to Professor L. McFadden, Professor J. W. Layman, and Professor B. C.Horne, Jr. for their participation at the defense of the thesis. He wishes also to acknowledge the use of the facilities at the Langley Research Center where part of the research was undertaken.

Walter Thomas Blackshear

REFERENCES

1. Bary, N. K.: A Treatise on Trignometric Series. Translation by Margaret F. Mullins, The MacMillian Company, vol. 1, 1964.

2. Olmsted, John M. H.: Real Variables. Appleton-Century-Crofts, Inc., 1959.

3. Thielman, Henry P.: Theory of Functions of Real Variables. Prentice-Hall, Inc., 1962.

4. Kolmogorov, A N.; and Fomin, S. V.: Measure, Lebesque Integrals and Hilbert Space. Academic Press, 1961.

5. Vulikh, B. Z.: Introduction to Functional Analysis. Pergamon Press, 1963.

6. Carslaw, H. S.: Theory of Fourier's Series and Integrals. Dover Publications, Inc., 1930.

7. Jackson, Dunham: Fourier Series and Orthogonal Polynomials. Open Court Publishing Company, 1963.

8. Alexita, G.: Convergence Problems of Orthogonal Series. Pergamon Press, 1961.

9. Banach, S.: Sur la divergence des series orthogonalses. Studia Math, 1940, pages 139-154.

10. Rainville, E. D.: Special Functions. The MacMillan Company, 1960.

11. Hardy, G. H.; and Rogosinski, W. W.: Fourier Series. The Cambridge University Press, 1965.

12. Wells, William R.: Notes on Gravitational Potential Theory. LWP-158, Langley Research Center, NASA, 1965.

13. Solloway, C. B.: Elements of the Theory of Orbit Determination. EPD-255, Jet Propulsion Laboratory, 1964.

14. Michael, W. H., Jr.; Tolson, R. H.; and Gapcynski, J. P.: Preliminary Results on the Gravitational Field of the Moon From Analysis of Lunar Orbiter Tracking Data. Paper presented at the American Geophysical Union Annual Meeting, Washington, D.C., April 17-20, 1967.

The vita has been removed from the scanned document

TABLE I

SPECIAL JACOBI POLYNOMIALS

Legendre polynomials

$P_0 (\cos \theta) = 1$

$P_1 (\cos \theta) = \cos \theta$

$P_2 (\cos \theta) = \frac{3}{2} \cos^2\theta - \frac{1}{2}$

$P_3 (\cos \theta) = \frac{5}{2} \cos^3\theta - \frac{3}{2} \cos \theta$

$P_4 (\cos \theta) = \frac{35}{8} \cos^4\theta - \frac{15}{4} \cos^2\theta + \frac{3}{8}$

Chebysheff polynomials

$C_0 (\cos \theta) = 1$

$C_1 (\cos \theta) = \frac{1}{2} \cos \theta$

$C_2 (\cos \theta) = \frac{3}{4} \cos^2\theta - \frac{3}{8}$

$C_3 (\cos \theta) = \frac{5}{2} \cos^3\theta - \frac{15}{8} \cos \theta$

$C_4 (\cos \theta) = \frac{35}{16} \cos^4\theta - \frac{13}{16} \cos^2\theta + \frac{35}{128}$

B_n polynomials

$B_0 (\cos \theta) = 1$

$B_1 (\cos \theta) = \frac{3}{2} \cos \theta$

$B_2 (\cos \theta) = \frac{5}{2} \cos^2\theta - \frac{5}{8}$

$B_3 (\cos \theta) = \frac{35}{8} \cos^3\theta - \frac{35}{16} \cos \theta$

$B_4 (\cos \theta) = \frac{63}{8} \cos^4\theta - \frac{189}{32} \cos^2\theta + \frac{63}{128}$

TABLE II

FOURIER COEFFICIENTS FOR $F(r_o, \theta) = \sum_{n=0}^{4} \frac{e_n}{r_o^{n+1}} \cos^n \theta$

n	a_n	b_n	d_n
0	$e_o + \frac{1}{3} e_2 + \frac{1}{5} e_4$	$e_o + \frac{1}{2} e_2 + \frac{3}{8} e_4$	$e_o + \frac{1}{4} e_2 + \frac{1}{8} e_4$
1	$e_1 + \frac{3}{5} e_3$	$2e_1 + \frac{3}{2} e_3$	$\frac{2}{3} e_1 + \frac{1}{3} e_3$
2	$\frac{2}{3} e_2 + \frac{4}{7} e_4$	$\frac{4}{3} e_2 + \frac{4}{3} e_4$	$\frac{2}{5} e_2 + \frac{3}{10} e_4$
3	$\frac{2}{5} e_3$	$\frac{1}{5} e_3$	$\frac{8}{35} e_3$
4	$\frac{8}{35} e_4$	$\frac{16}{35} e_4$	$\frac{8}{63} e_4$

TABLE III.- COEFFICIENTS FOR SERIES APPROXIMATION OF $F(r_o, \theta)$

(a) COEFFICIENTS FOR LEGENDRE EXPANSION ($0 \leq \theta \leq \pi$)

Fourier		Determined by least squares solution				
n	a_n	Case 5a	Case 4a	Case 3a	Case 2a	Case 1a
0	9.2472	9.2472	9.2485	9.2485	9.3377	9.3377
1	.9617	.9617	.9617	.9958	.9958	
2	.3390	.3391	.3541	.3541		
3	.0902	.0902	.0902			
4	.0355	.0355				

TABLE III.- CONTINUED.

(b) COEFFICIENTS FOR B_n EXPANSION $(0 \leq \theta \leq \pi)$

Fourier		Determined by least squares solution				
n	d_n	Case 5a	Case 4a	Case 3a	Case 2a	Case 1a
0	9.2042	9.2042	9.2042	9.2042	9.3377	9.3377
1	.6261	.6261	.6261	.6639	.6639	
2	.1968	.1968	.2124	.2124		
3	.0515	.0515	.0515			
4	.0197	.0197				

TABLE III.- CONCLUDED.

(c) COEFFICIENTS FOR CHEBYSHEFF EXPANSION ($0 \le \theta \le \pi$)

Fourier		Determined by least squares solution				
n	$\bar{b}_n$	Case 5a	Case 4a	Case 3a	Case 2a	Case 1a
0	9.3369	9.3369	9.3370	9.3370	9.3377	9.3377
1	1.9910	1.9910	1.9910	1.9916	1.9916	
2	.7079	.7079	.7081	.7081		
3	.1803	.1803	.1803			
4	.0710	.0710				

TABLE IV.- CORRELATION MATRICES FOR LEAST-SQUARES SOLUTIONS

(a) N MATRIX FOR LEGENDRE CASE 5a

1.0000	0.0000	-0.3271	-0.0001	-0.0859
.0000	1.0000	.0000	-.5179	.0000
-.3271	.0000	1.0000	.0000	-.4997
-.0001	-.5179	.0000	1.0000	.0000
-.0859	0.0000	-.4997	0.0000	1.0000

TABLE IV.- CONTINUED.

(b) N MATRIX FOR B_n POLYNOMIAL CASE 5a

1.0000	0.0000	-0.4083	0.0000	0.0000
.0000	1.0000	.0000	-.7090	.0000
-.4083	.0000	1.0000	.0000	-.7090
.0000	-.7090	.0000	1.0000	.0000
.0000	.0000	-.7090	.0000	1.0000

TABLE IV.- CONCLUDED.

(c) N MATRIX FOR CHEBYSHEFF CASE 5a

1.0000	0.0000	-0.0039	0.0000	-0.0039
.0000	1.0000	.0000	-.0055	.0000
-.0039	.0000	1.0000	.0000	-.0055
.0000	-.0055	.0000	1.0000	.0000
-.0039	.0000	-.0055	.0000	1.0000

TABLE V

SUM OF SQUARES OF THE RESIDUALS FOR SOLUTIONS BASED ON UNRESTRICTED REGION OF OBSERVATION ($0^\circ \leq \theta \leq 180^\circ$).

Case no.	Legendre	B_n	Chebysheff
5a	0.18×10^{-9}	0.31×10^{-9}	0.09×10^{-9}
4a	.068	.068	.068
3a	.64	.64	.64
2a	13.4	13.4	13.4
1a	192.9	192.9	192.9

TABLE VI.- COEFFICIENTS FOR SERIES APPROXIMATION $F(r_o, \theta)$

(a) COEFFICIENTS FOR LEGENDRE EXPANSION $(60^\circ \leq \theta \leq 120^\circ)$

Fourier		Determined by least squares solution				
n	a_n	Case 5b	Case 4b	Case 3b	Case 2b	Case 1b
0	9.2472	9.2474	9.2266	9.2266	9.1260	9.1260
1	.9617	.9617	.9617	.8613	.8613	
2	.3390	.3397	.2732	.2732		
3	.0902	.0902	.0902			
4	.0355	.0356				

TABLE VI.- CONTINUED.

(b) COEFFICIENTS FOR B_n EXPANSION ($60^\circ \leq \theta \leq 120^\circ$)

Fourier		Determined by least squares solution				
n	d_n	Case 5b	Case 4b	Case 3b	Case 2b	Case 1b
0	9.2042	9.2043	9.1925	9.1925	9.1260	9.1260
1	.6261	.6261	.6261	.5742	.5742	
2	.1968	.1969	.1639	.1639		
3	.0515	.0515	.0515			
4	.0197	.0198				

TABLE VI.- CONCLUDED.

(c) COEFFICIENTS FOR CHEBYSHEFF EXPANSION ($60^\circ \leq \theta \leq 120^\circ$)

Fourier		Determined by least squares solution				
n	b_n	Case 5b	Case 4b	Case 3b	Case 2b	Case 1b
0	9.3369	9.3376	9.2949	9.2949	9.1260	9.1260
1	1.9910	1.9910	1.9910	1.7227	1.7227	
2	.7079	.7104	.5464	.5464		
3	.1803	.1803	.1803			
4	.0710	.0721				

TABLE VII.- CORRELATION MATRICES FOR LEAST-SQUARES SOLUTIONS

(a) N MATRIX FOR LEGENDRE CASE 5b

1.0000	-0.0001	0.9991	-0.0001	0.9922
-.0001	1.0000	-.0001	.9892	-.0001
.9991	-.0001	1.0000	-.0001	.9949
-.0001	.9892	-.0001	1.0000	-.0001
.9922	-.0001	.9949	-.0001	1.0000

TABLE VII.- CONTINUED.

(b) N MATRIX FOR B_n POLYNOMIAL CASE 5b

1.0000	-0.0001	0.9978	-0.0001	0.9885
-.0001	1.0000	-.0001	.9822	-.0001
.9978	-.0001	1.0000	-.0001	.9926
-.0001	.9822	-.0001	1.0000	-.0001
.9885	-.0001	.9926	-.0001	1.0000

TABLE VII.- CONCLUDED.

(c) N MATRIX FOR CHEBYSHEFF CASE 5b

1.0000	-0.0001	0.9997	-0.0001	0.9950
-.0001	1.0000	-.0001	.9939	-.0001
.9997	-.0001	1.0000	-.0001	.9966
-.0001	.9939	-.0001	1.0000	-.0001
.9950	-.0001	.9966	-.0001	1.0000

TABLE VIII

SUM OF SQUARES OF THE RESIDUALS FOR SOLUTIONS BASED ON RESTRICTED REGION OF OBSERVATION ($60^\circ \leq \theta \leq 120^\circ$)

Case no.	Legendre	B_n	Chebysheff
5b	0.11×10^{-3}	0.02×10^{-3}	0.36×10^{-3}
4b	1.38	1.38	1.38
3b	5.2	5.2	5.2
2b	32.9	32.9	32.9
1b	209.1	209.1	209.1

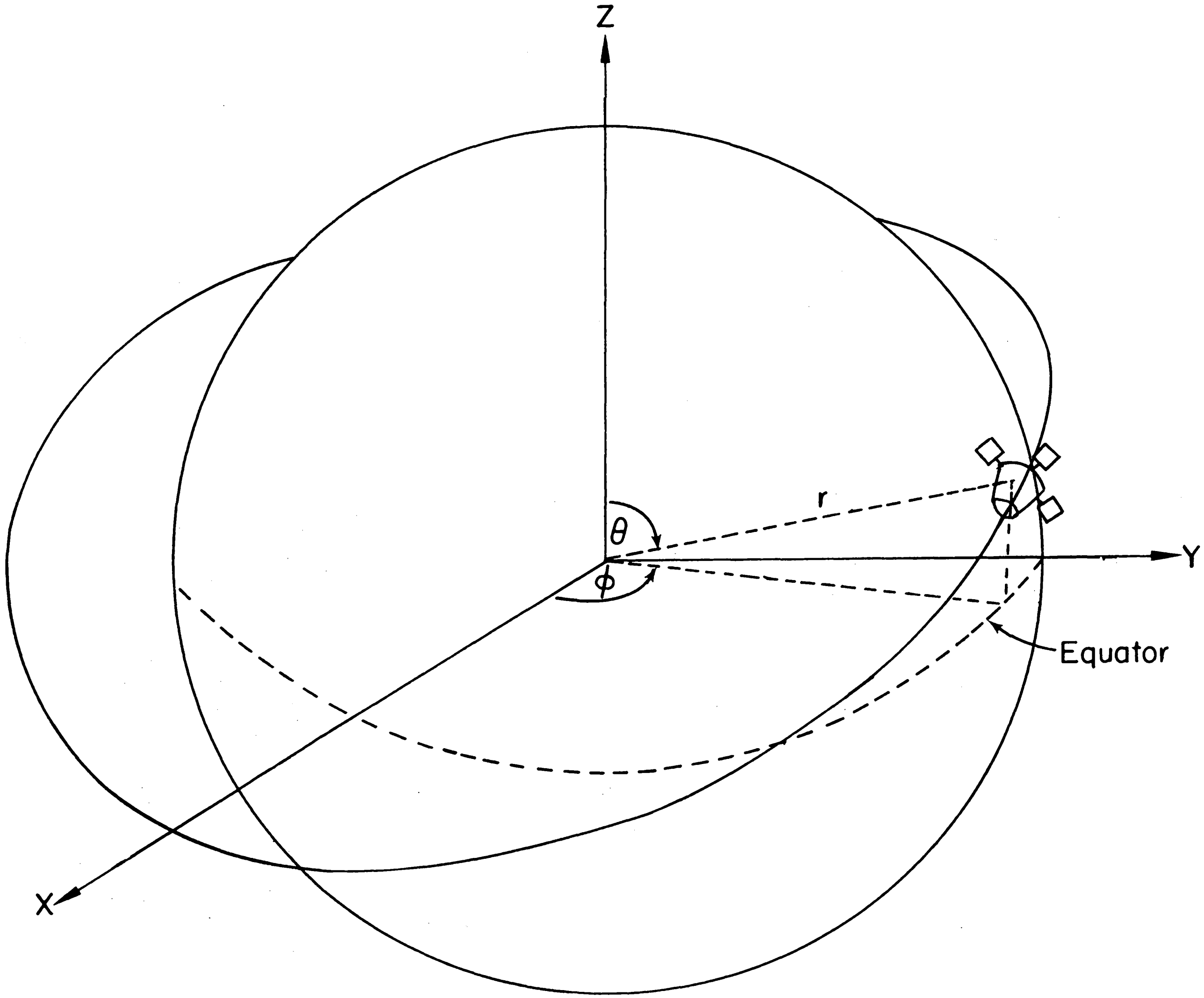
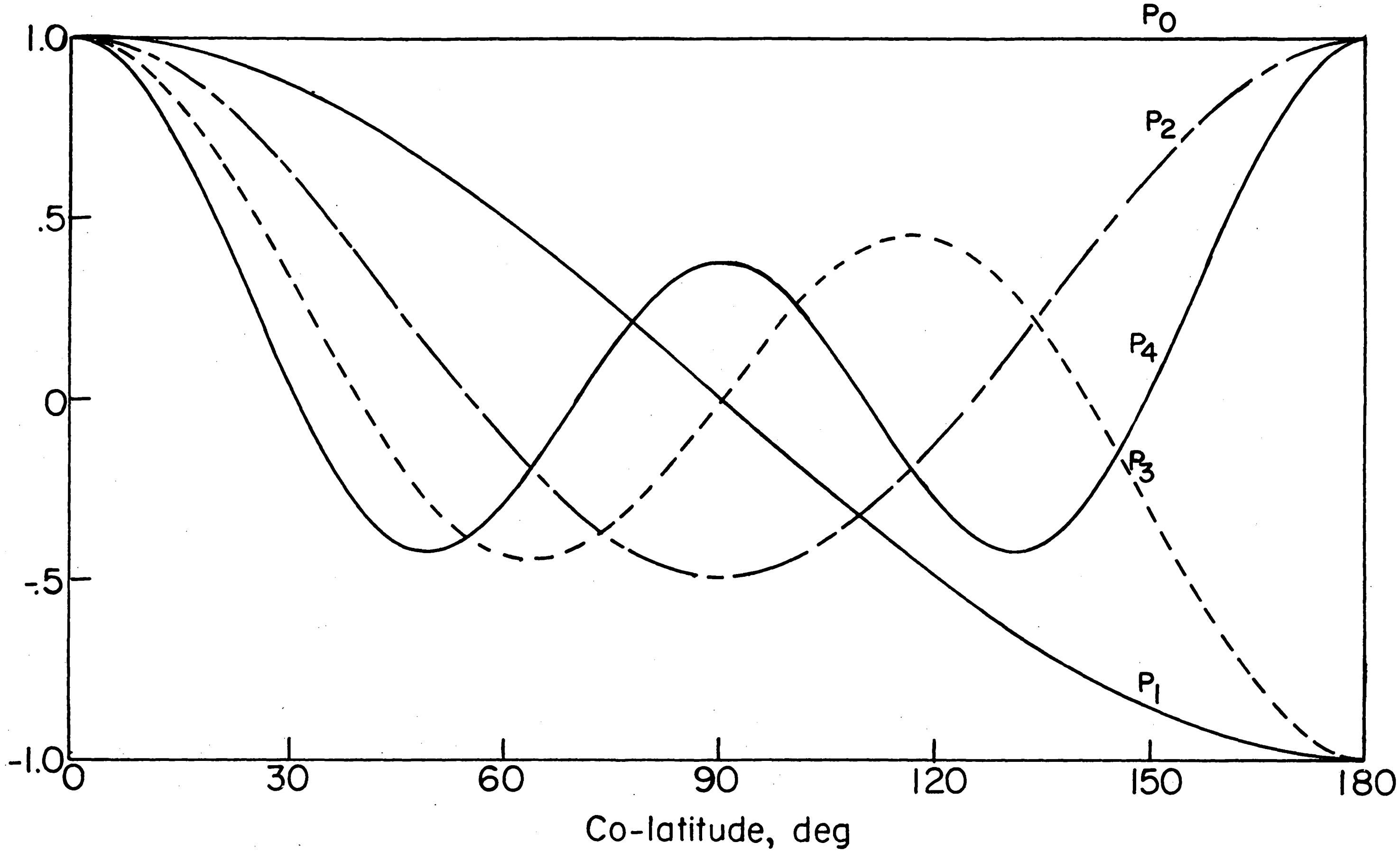
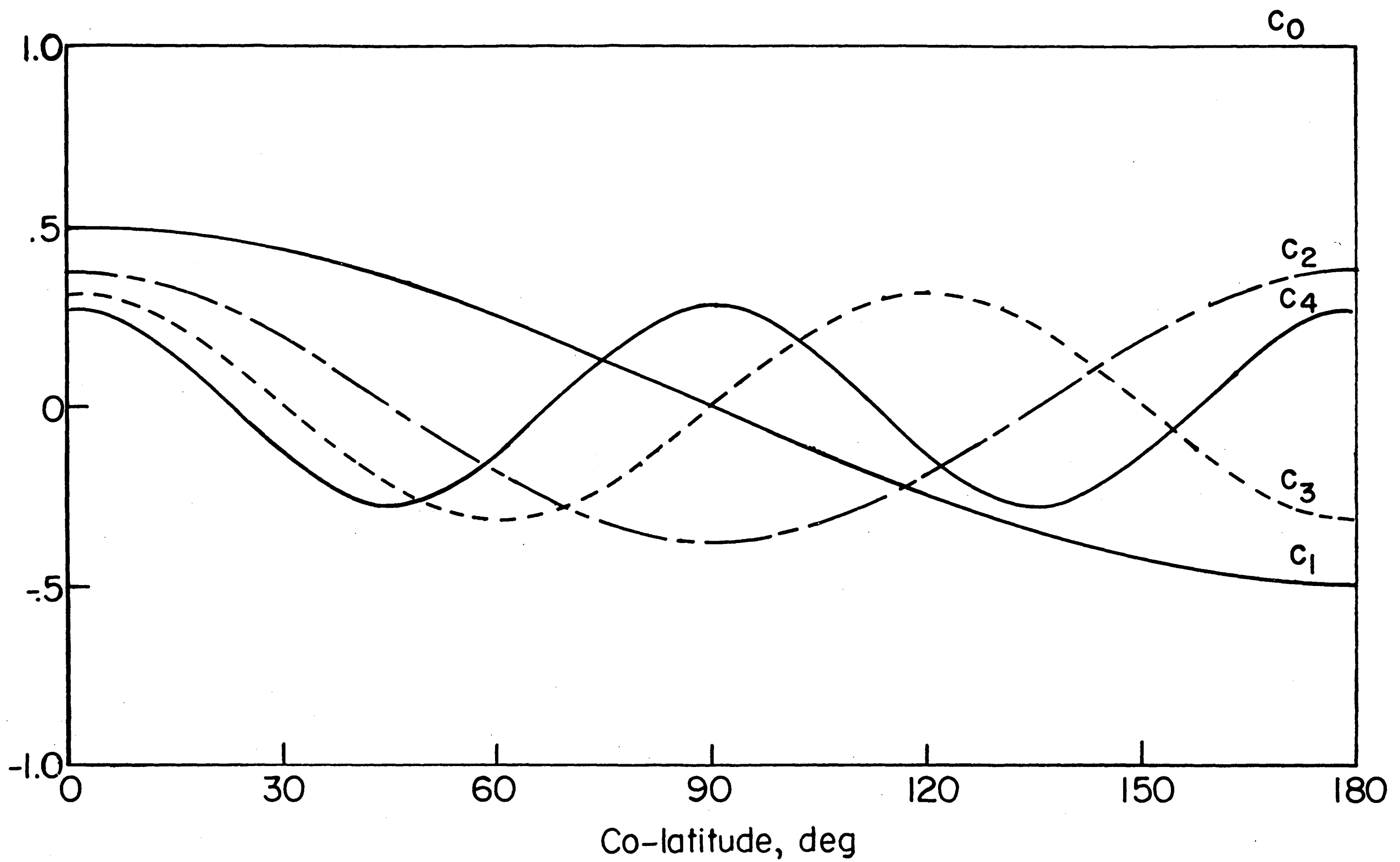


Figure 1.- Coordinate system and orbital orientation.



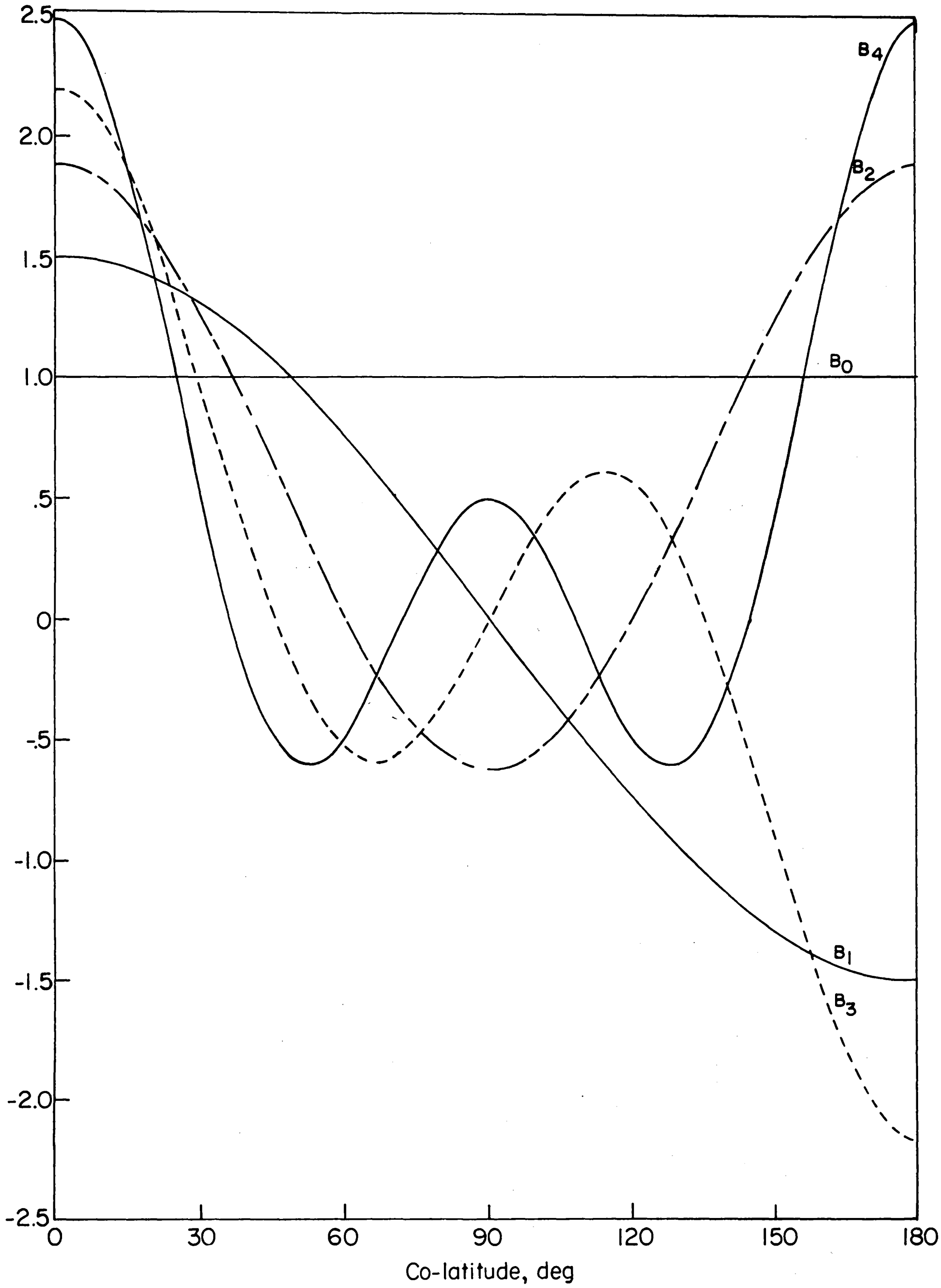
(a) Legendre polynomials.

Figure 2.- Jacobi polynomials.



(b) Chebysheff polynomials

Figure 2.- Continued.



(c) B_n polynomials.

Figure 2.- Concluded.

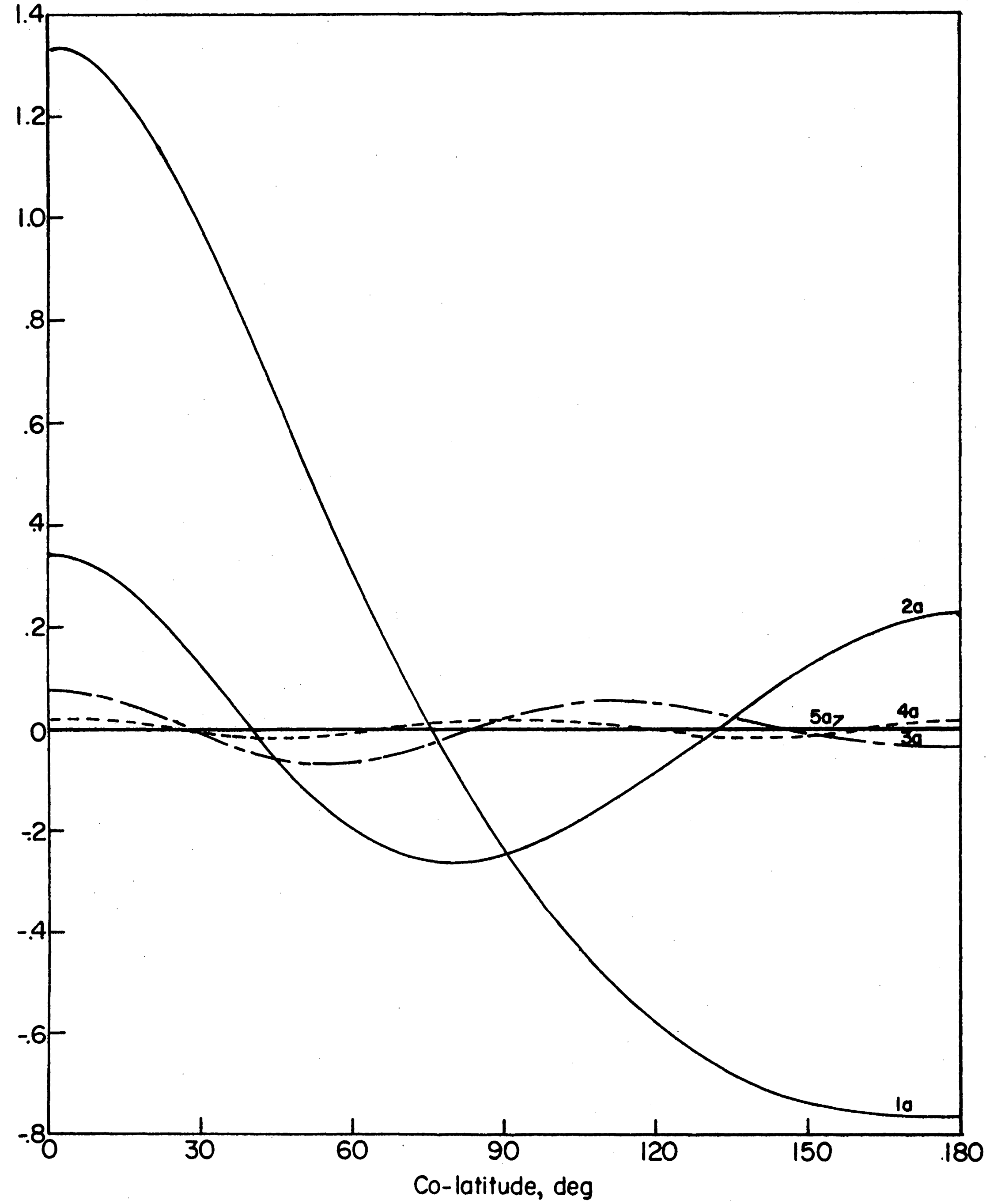


Figure 3.- Residuals for unrestricted region of observation. ($0^\circ \leq \theta \leq 180^\circ$)

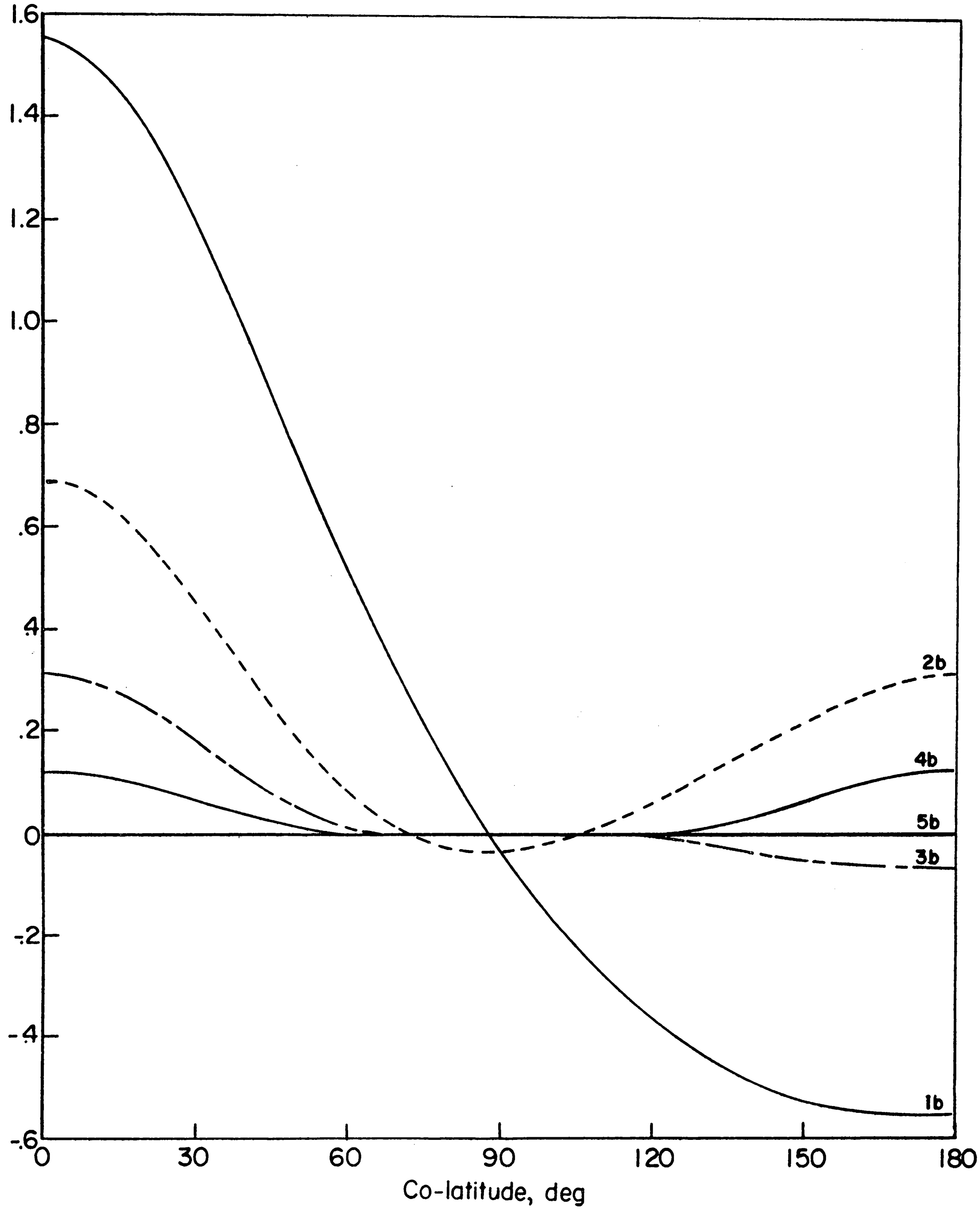


Figure 4.- Residuals for restricted region of observation. ($60^\circ \leq \theta \leq 120^\circ$)

NOTES ON GENERALIZED FOURIER SERIES

WITH APPLICATION TO GRAVITATIONAL FIELD DETERMINATION

by

Walter Thomas Blackshear

ABSTRACT

Let $\{\varphi_n(x)\}$ be an orthonormal system in the set of Lebesgue square integrable functions L^2. Let $f \in L^2$. The generalized Fourier series of f with respect to $\{\varphi_n(x)\}$ is the series

$\sum_{n=0}^{\infty} (f, \varphi_n) \varphi_n(x)$, where (f, φ_n) is the inner product of the functions f and φ_n. The existence of a complete orthonormal system in L^2 is proven. Conditions for convergence of the generalized Fourier series are presented. A discussion of orthogonal polynomials with special emphasis on the Jacobi polynomial systems is presented. A least squares, differential correction, discrete observation procedure is employed to solve the potential equation with boundary conditions in terms of three special Jacobi systems.