AMERICAN MATHEMATICAL SOCIETY
COLLOQUIUM PUBLICATIONS
VOLUME XXIII

ORTHOGONAL POLYNOMIALS

BY

GABOR SZEGÖ

PROFESSOR OF MATHEMATICS
STANFORD UNIVERSITY

1939

PREFACE

Recent years have seen a great deal of progress in the field of orthogonal polynomials, a subject closely related to many important branches of analysis. Orthogonal polynomials are connected with trigonometric, hypergeometric, Bessel, and elliptic functions, are related to the theory of continued fractions and to important problems of interpolation and mechanical quadrature, and are of occasional occurrence in the theories of differential and integral equations. In addition, they furnish comparatively general and instructive illustrations of certain situations in the theory of orthogonal systems. Recently, some of these polynomials have been shown to be of significance in quantum mechanics and in mathematical statistics.

The origins of the subject are to be found in the investigation of a certain type of continued fractions, bearing the name of Stieltjes. Special cases of these fractions were studied by Gauss, Jacobi, Christoffel, and Mehler, among others, while more general aspects of their theory were given by Tchebichef, Heine, Stieltjes, and A. Markoff.

Despite the close relationship between continued fractions and the problem of moments, and notwithstanding recent important advances in this latter subject, continued fractions have been gradually abandoned as a starting point for the theory of orthogonal polynomials. In their place, the orthogonal property itself has been taken as basic, and it is this point of view which has been adopted in the following exposition of the subject. Choosing this same basic property, we discuss certain special orthogonal polynomials, which have been treated in great detail independently of the general theory, and indeed, even before this theory existed at all. In this connection we add the names of Laplace, Legendre, Fourier, Abel, Laguerre, and Hermite to those previously mentioned.

As regards treatises on the subject, we note that the only systematic treatment thus far given is found in J. Shohat's monograph, *Théorie Générale des Polynomes Orthogonaux de Tchebichef*, Mémorial des Sciences Mathématiques, Paris, 1934. Limitations of space have compelled that work to be brief, and consequently, it does not enter into a detailed treatment of many problems which have been especially advanced in recent years. It has therefore seemed desirable to attempt a new and detailed development of the main ideas of this field, devoting, in particular, some space to recent investigations of the distribution of the zeros, of asymptotic representations, of expansion problems, and of certain questions of interpolation and mechanical quadrature.

In what follows, we are concerned partly with the general theory of orthogonal polynomials, and partly with the study of special classes of these polynomials. As might be expected, we have more exhaustive results for these special classes, and we cite as an instance the classical polynomials satisfying linear differential equations of the second order. Also, when the primary importance of these special classes in applications is taken into account, it should not be at all surprising that the present book is mainly devoted to their study. The general theory, however, as developed in Chapters XII and XIII, doubtless represents the most important progress made in recent years.

In the present work, no claim is made for completeness of treatment. On the contrary, the aim has purposely been to make the material suggestive rather than exhaustive. An attempt has been made to indicate the main and characteristic methods and to point out the relation of these to some general ideas in modern analysis. As a rule, preference has been given to those topics to which we were able to make some new, though modest, contributions, or which we could present in a new setting. Thus the book contains a number of results not previously published, some of which originated several years ago. For instance, we have included a discussion of the Cesàro summability of the Jacobi series at the end-points of the orthogonality interval (the method used here is of interest even in the classical case of Legendre series). Further, a new and simpler approach has been given to S. Bernstein's asymptotic formula for orthogonal polynomials. We also refer to certain details of minor importance, such as: simplifications and additions in the asymptotic investigation of Jacobi and Laguerre polynomials and in the discussion of the expansions in terms of these polynomials; the discussion of the cases in which the Jacobi differential equation has only polynomial solutions; the evaluation of the number of zeros of general Jacobi polynomials in the intervals $[-\infty, -1]$, $[-1, +1]$, $[+1, +\infty]$; a new proof of the Heine-Stieltjes theorem on linear differential equations of the second order with polynomial coefficients and polynomial solutions, and so on.

In general, we have preferred to discuss problems which may be stated and treated simply, and which could be presented in a more or less complete form. This was the main reason for devoting no space to the extremely interesting arithmetic and algebraic properties of orthogonal polynomials, such as, for instance, the recent important investigations of I. Schur concerning the irreducibility and related properties of Laguerre and Hermite polynomials. Furthermore, we have attached great importance to the idea of replacing incomplete and overlapping theorems, scattered in the literature, by complete results involving only intrinsic or necessary restrictions. We have also tried to exploit, as far as seemed to be at all possible, definite methods, such as, for instance, Sturm's methods in differential equations (see §§6.3, 6.31, 6.32, 6.83).

A complete treatment of Legendre polynomials was not feasible, and probably not desirable, in the framework of the general theory. Besides, there are already complete treatises on spherical and other harmonics.[1] We have selected and considered only those properties of Legendre polynomials which are the starting points of generalizations to ultraspherical, Jacobi, or to more general polynomials. Another subject which could not be included was Stieltjes'

[1] For instance, E. W. Hobson 1 (see bibliography).

problem of moments, which has been omitted in spite of its great interest; for this subject would have necessitated the development of a complicated apparatus of results and methods. Orthogonal polynomials of more than one variable also have not been treated.[2]

The book is based on a course given at Washington University during the academic year 1935–1936. Acquaintance with the general ideas and methods of the theory of functions of real and complex variables is naturally required. Occasionally, Stieltjes-Lebesgue and Lebesgue integrals are considered. In the greater part of the book, however, these integrals have been avoided, and, except in a very few places, no detailed properties of them were used.

The problems at the end of this book are, with few exceptions, not new, and they are not interconnected as are, for instance, those in Pólya-Szegö's *Aufgaben und Lehrsätze*. They are more or less supplementary in character and serve as illustrations and exercises; they sometimes differ widely from one another both as to subject and method.

The list of references is not complete; it contains only original memoirs, a few text books of primary importance, and monographs to which references are made in the text.

For the suggestion of preparing a book on orthogonal polynomials for the Colloquium Publications, I am indebted to Professor J. D. Tamarkin, who has also participated in the present work by offering a great number of valuable suggestions. It is with the greatest gratitude that I mention his friendly interest.

I have also received valuable advice from my friends and teachers L. Fejér (Budapest), and G. Pólya (Zürich). My colleagues P. Erdös (Manchester), G. Grünwald (Budapest), W. H. Roever (St. Louis), A. Ross (St. Louis), J. Shohat (Philadelphia), and P. Turán (Budapest) gave generously and unstintingly of their time. F. A. Butter, Jr. (at present in Los Angeles) collaborated with me in the preparation of the manuscript. This last aid was made possible through a grant from the Rockefeller Research Fund of Washington University (1936–1937). My student L. H. Kanter also rendered valuable assistance in the preparation of the manuscript.

My gratitude for the encouragement and help of these friends, colleagues, and institutions can hardly be measured by any formal acknowledgment. Lastly, I wish to express to the American Mathematical Society my great appreciation for the inclusion of the present book in its Colloquium Series.

G. Szegö

Washington University, 1938.

[2] Cf. the bibliography in Jackson 8, p. 423.

TABLE OF CONTENTS

		PAGE
PREFACE		v
CHAPTER I.	PRELIMINARIES	1
CHAPTER II.	DEFINITION OF ORTHOGONAL POLYNOMIALS; PRINCIPAL EXAMPLES	22
CHAPTER III.	GENERAL PROPERTIES OF ORTHOGONAL POLYNOMIALS	37
CHAPTER IV.	JACOBI POLYNOMIALS	57
CHAPTER V.	LAGUERRE AND HERMITE POLYNOMIALS	96
CHAPTER VI.	ZEROS OF ORTHOGONAL POLYNOMIALS	107
CHAPTER VII.	INEQUALITIES	154
CHAPTER VIII.	ASYMPTOTIC PROPERTIES OF THE CLASSICAL POLYNOMIALS	185
CHAPTER IX.	EXPANSION PROBLEMS ASSOCIATED WITH THE CLASSICAL POLYNOMIALS	237
CHAPTER X.	REPRESENTATION OF POSITIVE FUNCTIONS	267
CHAPTER XI.	POLYNOMIALS ORTHOGONAL ON THE UNIT CIRCLE	280
CHAPTER XII.	ASYMPTOTIC PROPERTIES OF GENERAL ORTHOGONAL POLYNOMIALS	289
CHAPTER XIII.	EXPANSION PROBLEMS ASSOCIATED WITH GENERAL ORTHOGONAL POLYNOMIALS	306
CHAPTER XIV.	INTERPOLATION	322
CHAPTER XV.	MECHANICAL QUADRATURE	340
CHAPTER XVI.	POLYNOMIALS ORTHOGONAL ON AN ARBITRARY CURVE	355
PROBLEMS AND EXERCISES		368
LIST OF REFERENCES		378
INDEX		395

CHAPTER I

PRELIMINARIES

1.1. Notation

Numbers in bold face type, like **1**, refer to the bibliography at the end of the book. The system of section numbering used is Peano's decimal system, and the numeration of formulas starts anew in each section. Thus, reference to §9.5 and (9.5.2) means section 9.5 in Chapter IX and formula (9.5.2) in the same section, respectively. A similar numeration has been used for the theorems

We use the symbol: $\delta_{nm} = 0$ or 1, according as $n \neq m$, or $n = m$.

The closed real interval $a \leqq x \leqq b$ (a and b finite) will be denoted by $[a, b]$. The same symbol is used if either a or b is infinite or if both are; in this case the equality sign is excluded.

We often write for a real x

$$\text{sgn}\, x = -1, 0, +1, \tag{1.1.1}$$

according as x is negative, zero, or positive; more generally, for arbitrary complex x, $x \neq 0$, we write

$$\text{sgn}\, x = |x|^{-1} x. \tag{1.1.2}$$

The symbol $\bar{x}$ denotes the conjugate complex value, $\Re(x)$ the real part, and $\Im(x)$ the imaginary part of the complex number x.

If two sequences z_n and w_n of complex numbers have the property that $w_n \neq 0$ and $z_n/w_n \to 1$ as $n \to \infty$, we write $z_n \cong w_n$. If z_n and w_n are complex, $w_n \neq 0$, and the sequence $|z_n|/|w_n|$ has finite positive limits of indetermination, we write $z_n \sim w_n$.

Occasionally we make use of the notation

$$z_n = O(a_n), \qquad z_n = o(a_n) \tag{1.1.3}$$

if $a_n > 0$, to state that z_n/a_n is bounded, or tends to 0, respectively, as $n \to \infty$. A similar notation is used for a passage of limit other than $n \to \infty$.

A function $f(x)$ is called increasing (strictly increasing) if $x_1 < x_2$ implies $f(x_1) < f(x_2)$; it is called non-decreasing if $x_1 < x_2$ implies $f(x_1) \leqq f(x_2)$. An analogous terminology will be used for decreasing functions.

Let $p \geqq 1$, and let $\alpha(x)$ be a non decreasing function in $[a, b]$ which is not constant. The class of functions $f(x)$ which are measurable with respect to $\alpha(x)$ and for which the Stielties-Lebesgue integral $\int_a^b |f(x)|^p \, d\alpha(x)$ exists (see §1.4) is called $L_\alpha^p(a, b)$. In case $\alpha(x) = x$ we use the notation $L^p(a, b)$; in case $p = 1$, $\alpha(x)$ arbitrary, the notation $L_\alpha(a, b)$ is used. If $f(x)$ and $g(x)$ belong to the class $L_\alpha^p(a, b)$, the same is true for $f(x) + g(x)$. (Cf. Kaczmarz-Steinhaus **1**, pp. 10–11.)

1.11. Inequalities

(1) *Cauchy's inequality.* Let $\{a_\nu\}$, $\{b_\nu\}$, $\nu = 1, 2, \cdots, n$, be two systems of complex numbers. Then

$$\left|\sum_{\nu=1}^{n} a_\nu b_\nu\right|^2 \leqq \sum_{\nu=1}^{n} |a_\nu|^2 \sum_{\nu=1}^{n} |b_\nu|^2. \tag{1.11.1}$$

The equality sign holds if and only if two numbers λ, μ, not both zero, exist such that $\lambda a_\nu + \mu \bar{b}_\nu = 0$, $\nu = 1, 2, \cdots, n$.

(2) *Schwarz's inequality.* Let $f(x)$ and $g(x)$ be two functions of class $L_\alpha^2(a, b)$. Then $f(x)g(x)$ is of class $L_\alpha(a, b)$, and

$$\left|\int_a^b f(x)g(x)\,d\alpha(x)\right|^2 \leqq \int_a^b |f(x)|^2\,d\alpha(x) \int_a^b |g(x)|^2\,d\alpha(x). \tag{1.11.2}$$

(3) *Inequality for the arithmetic and geometric mean.* If $f(x) > 0$, we have

$$\frac{\int_a^b f(x)\,d\alpha(x)}{\int_a^b d\alpha(x)} \geqq \exp\left\{\frac{\int_a^b \log f(x)\,d\alpha(x)}{\int_a^b d\alpha(x)}\right\}, \tag{1.11.3}$$

provided all integrals exist, and $\int_a^b d\alpha(x) > 0$. (Cf. Hardy-Littlewood-Pólya **1**, pp. 137–138.)

(4) *Abel's transformation and Abel's inequality.* From

$$\begin{aligned} &f_0g_0 + f_1g_1 + \cdots + f_ng_n \\ &\quad = (f_0 - f_1)G_0 + (f_1 - f_2)G_1 + \cdots + (f_{n-1} - f_n)G_{n-1} + f_nG_n, \end{aligned} \tag{1.11.4}$$

where

$$G_\nu = g_0 + g_1 + \cdots + g_\nu, \qquad \nu = 0, 1, 2, \cdots, n, \tag{1.11.5}$$

we obtain, assuming $f_0 \geqq f_1 \geqq \cdots \geqq f_n \geqq 0$, and $|G_\nu| \leqq G$, $\nu = 0, 1, \cdots, n$, the inequality

$$|f_0g_0 + f_1g_1 + \cdots + f_ng_n| \leqq f_0\,G. \tag{1.11.6}$$

(5) *Second mean-value theorem of the integral calculus.* Let $f(x) \geqq 0$ be a non-increasing function, and let $g(x)$ be continuous, $a \leqq x \leqq b$, a and b finite. Then

$$\int_a^b f(x)g(x)\,dx = f(a + 0)\int_a^\xi g(x)\,dx, \qquad a < \xi < b. \tag{1.11.7}$$

1.12. Polynomials and trigonometric polynomials

We shall consider polynomials in x of the form

$$\rho(x) = c_0 + c_1x + c_2x^2 + \cdots + c_mx^m, \tag{1.12.1}$$

with arbitrary complex coefficients $c_0, c_1, c_2, \cdots, c_m$. Here m is called the degree; and if $c_m \neq 0$, the precise degree of $\rho(x)$. In what follows an arbitrary polynomial of degree m will be denoted by π_m. If $\rho_0(x), \rho_1(x), \cdots, \rho_n(x)$ are arbitrary polynomials such that $\rho_m(x)$ has the precise degree m, every π_n can be represented as a linear combination of these polynomials with coefficients which are uniquely determined.

A trigonometric polynomial in θ of degree m has the form

$$g(\theta) = a_0 + a_1 \cos\theta + b_1 \sin\theta + \cdots + a_m \cos m\theta + b_m \sin m\theta, \tag{1.12.2}$$

with arbitrary complex coefficients. Here m is again called the degree of $g(\theta)$; m is the precise degree if $|a_m| + |b_m| > 0$. According as all the b_μ or all the a_μ vanish, $g(\theta)$ is referred to as a cosine or a sine polynomial.

The functions $\cos m\theta$ and $\sin(m+1)\theta/\sin\theta$ are polynomials in $\cos\theta = x$ of the precise degree m and are called *Tchebichef polynomials of the first and second kind*, respectively. These polynomials play a fundamental rôle in subsequent considerations. Setting

$$\cos m\theta = T_m(\cos\theta) = T_m(x), \qquad \frac{\sin(m+1)\theta}{\sin\theta} = U_m(\cos\theta) = U_m(x), \tag{1.12.3}$$

we see that any cosine polynomial of degree m is a polynomial of the same degree in $\cos\theta = x$, and conversely. Any sine polynomial of degree m, divided by $\sin\theta$, furnishes a cosine polynomial of degree $m - 1$. Thus, a sine polynomial can be represented as the product of $\sin\theta = (1 - x^2)^{1/2}$ by a polynomial in $\cos\theta = x$.

The polynomials (1.12.3) are special cases of the so-called Jacobi polynomials (cf. Chapter IV). They contain only even or only odd powers of x according as m is even or odd. Thus $\cos(m + \frac{1}{2})\theta/\cos(\theta/2)$ and $\sin(m + \frac{1}{2})\theta/\sin(\theta/2)$ are cosine polynomials in θ of degree m; they are also connected with the Jacobi polynomials (see (4.1.8)).

We define the "reciprocal" polynomial of (1.12.1) by

$$\rho^*(x) = x^m \bar{\rho}(x^{-1}) = \bar{c}_m + \bar{c}_{m-1}x + \bar{c}_{m-2}x^2 + \cdots + \bar{c}_0 x^m. \tag{1.12.4}$$

If the zeros of $\rho(x)$ are $x_1, x_2, \cdots, x_m$, those of $\rho^*(x)$ are $x_1^*, x_2^*, \cdots, x_m^*$, where $x_\mu^* = \bar{x}_\mu^{-1}$ is the point which is obtained from x_μ by inversion with respect to the unit circle $|x| = 1$ in the complex x-plane. The zeros must be counted according to their multiplicity, and $0^* = \infty$, $\infty^* = 0$; ∞ as a zero of order k means that the coefficients of the k highest powers vanish.

1.2. Fejér's theorem concerning non-negative trigonometric polynomials

THEOREM 1.2.1. *Let $g(\theta)$ be a trigonometric polynomial with real coefficients which is non-negative for all real values of θ. Then there exists a polynomial $\rho(z)$ of the same degree as $g(\theta)$ such that $g(\theta) = |\rho(z)|^2$, where $z = e^{i\theta}$. Conversely, if $z = e^{i\theta}$, the expression $|\rho(z)|^2$ always represents a non-negative trigonometric polynomial in θ of the same degree as the polynomial $\rho(z)$.*

See Fejér 5. The second part of the statement is obvious. The first part is easily derived from (1.12.2) by introducing $z^k + z^{-k}$ for 2 cos $k\theta$ and $z^k - z^{-k}$ for $2i \sin k\theta$. We then find $g(\theta) = z^{-m}G(z)$, where $G(z)$ is a π_{2m} for which $G^*(z) = G(z)$. Now those zeros of $G(z)$ which are different from 0 and ∞, and which do not have the absolute value 1, can be combined in pairs of the form z_μ, z_μ^*, $0 < |z_\mu| < 1$, where z_μ^* has a meaning similar to that in §1.12. Furthermore, every real zero θ_0 of $g(\theta)$ is of even multiplicity, and $e^{i\theta_0}$ is a zero of $G(z)$ of the same multiplicity. Thus

(1.2.1)
$$G(z) = cz^\kappa \prod_{\mu=1}^{\sigma} (z - z_\mu)(z - z_\mu^*) \prod_{\nu=1}^{\tau} (z - \zeta_\nu)^2,$$
$$0 < |z_\mu| < 1, \qquad |\zeta_\nu| = 1; \qquad \kappa + \sigma + \tau = m.$$

Since $g(\theta) = |g(\theta)| = |G(z)|$, $z = e^{i\theta}$, and $|z - z_\mu| = |z_\mu| |z - z_\mu^*|$, $z = e^{i\theta}$, the theorem is established.

The representation in question is, however, not unique. Indeed, if α denotes an arbitrary zero of $\rho(z)$, the polynomial $\rho(z)(1 - \bar{\alpha}z)/(z - \alpha)$ furnishes another representation. Hence assuming $g(\theta) \not\equiv 0$, we can gradually remove all the zeros from $|z| < 1$ and obtain the following theorem:

Theorem 1.2.2. *Let $g(\theta)$ satisfy the condition of Theorem 1.2.1 and $g(\theta) \not\equiv 0$. Then a representation $g(\theta) = |h(e^{i\theta})|^2$ exists such that $h(z)$ is a polynomial of the same degree as $g(\theta)$, with $h(z) \neq 0$ in $|z| < 1$, and $h(0) > 0$. This polynomial is uniquely determined. If $g(\theta)$ is a cosine polynomial, $h(z)$ is a polynomial with real coefficients.*

A generalization of this normalized representation (its extension to a certain class of non-negative functions $g(\theta)$) is of great importance in the discussion of the asymptotic behavior of orthogonal polynomials. (See Chapters X–XIII.)

1.21. Theorem of Lukács concerning non-negative polynomials

(1) Theorem 1.21.1 (Theorem of Lukács). *Let $\rho(x)$ be a π_m non-negative in $[-1, +1]$. Then $\rho(x)$ can be represented in the form*

(1.21.1)
$$\rho(x) = \begin{cases} \{A(x)\}^2 + (1 - x^2)\{B(x)\}^2 & \text{if } m \text{ is even,} \\ (1 + x)\{C(x)\}^2 + (1 - x)\{D(x)\}^2 & \text{if } m \text{ is odd.} \end{cases}$$

Here $A(x)$, $B(x)$, $C(x)$, and $D(x)$ are real polynomials such that the degrees of the single terms on the right-hand side do not exceed m.

The proof can be based on Theorem 1.2.2. We have
$$\rho(\cos\theta) = |h(e^{i\theta})|^2 = |e^{-im\theta/2} h(e^{i\theta})|^2,$$
where $h(z)$ is a π_m with real coefficients. Now the expressions

(1.21.2)
$$\cos m\theta; \quad \frac{\sin(m+1)\theta}{\sin\theta}, \quad \frac{\cos(m+\frac{1}{2})\theta}{\cos(\theta/2)}, \quad \frac{\sin(m+\frac{1}{2})\theta}{\sin(\theta/2)}$$

are all π_m in $\cos\theta$ (see §1.12), so that

$$e^{-im\theta/2}h(e^{i\theta}) = \begin{cases} A(\cos\theta) + i\sin\theta\, B(\cos\theta) & \text{if } m \text{ is even,} \\ 2^{\frac{1}{2}}\cos(\theta/2)\, C(\cos\theta) + i2^{\frac{1}{2}}\sin(\theta/2) D(\cos\theta) & \text{if } m \text{ is odd,} \end{cases}$$

where the degrees of $A(x)$, $B(x)$, $C(x)$, $D(x)$ are, respectively, $m/2$, $m/2 - 1$, $(m-1)/2$, $(m-1)/2$.

(2) The following theorem has a simpler character:

THEOREM 1.21.2. *Every polynomial in x, which is non-negative for all real values of x, can be represented in the form $\{A(x)\}^2 + \{B(x)\}^2$. Every polynomial which is non-negative for $x \geqq 0$, can be represented in the form $\{A(x)\}^2 + \{B(x)\}^2 + x[\{C(x)\}^2 + \{D(x)\}^2]$. Here $A(x)$, $B(x)$, $C(x)$, $D(x)$ are all real polynomials, and the degree of each term does not exceed the degree of the given polynomial.*

These representations can also be written in the form $|P(x)|^2$ and $|P(x)|^2 + x|Q(x)|^2$, respectively, where $P(x)$ and $Q(x)$ are polynomials with complex coefficients; for the degrees the same remark holds as before.

In connection with this section see Pólya-Szegö **1**, vol. 2, pp. 82, 275, 276, problems 44, 45, 47.

1.22. Theorems of S. Bernstein

THEOREM 1.22.1. *If $g(\theta)$ is a trigonometric polynomial of degree m satisfying the condition $|g(\theta)| \leqq 1$, θ arbitrary and real, then $|g'(\theta)| \leqq m$.*

This theorem is due to S. Bernstein. (Cf. M. Riesz **1**.) The upper bound m cannot be replaced by a smaller one as is readily seen by taking $g(\theta) = \cos m\theta$. The following special case is worthy of notice:

THEOREM 1.22.2. *Let $\rho(z)$ be an arbitrary π_m satisfying the condition $|\rho(z)| \leqq 1$, where z is complex, and $|z| \leqq 1$; then $|\rho'(z)| \leqq m$, $|z| \leqq 1$.*

With regard to this theorem see also Szász **1**, pp. 516–517. Finally we mention the following consequence of Theorem 1.22.1:

THEOREM 1.22.3. *Let $\rho(x)$ be a π_m satisfying the condition $|\rho(x)| \leqq 1$ in $-1 \leqq x \leqq +1$. Then*

$$|\rho'(x)| \leqq (1 - x^2)^{-\frac{1}{2}} m.$$

This follows by applying Theorem 1.22.1 to $g(\theta) = \rho(\cos\theta)$.

1.3. Approximation by polynomials

(1) THEOREM 1.3.1 (Theorem of Weierstrass). *A function, continuous in a finite closed interval, can be approximated with a preassigned accuracy by polynomials. A function of a real variable which is continuous and has the period 2π, can be approximated by trigonometric polynomials.*

For information concerning this theorem we refer to Jackson **4**. In the second part of the theorem let the function in question be even (odd); then the approximating trigonometric polynomials can be chosen as cosine (sine) polynomials.

THEOREM 1.3.2. *Let $\omega(\delta)$ be the modulus of continuity of a given function $f(x)$, continuous in the finite interval $[a, b]$,*

$$\omega(\delta) = \max |f(x') - f(x'')| \qquad \text{if } |x' - x''| \leqq \delta. \tag{1.3.1}$$

Then for each m we can find a polynomial $\rho(x)$ of degree m, such that in the given interval of length l we have

$$|f(x) - \rho(x)| < A\omega(l/m). \tag{1.3.2}$$

In the case of a periodic function $f(\theta)$ with period 2π, a trigonometric polynomial $g(\theta)$ of degree m can be found such that

$$|f(\theta) - g(\theta)| < B\omega(2\pi/m). \tag{1.3.3}$$

Here A and B are absolute constants.

In this connection see Jackson **4**, pp. 7, 15.

THEOREM 1.3.3. *Let $f(x)$ have a continuous derivative of order μ in the finite interval $[a, b]$, $\mu \geqq 1$, and let $\omega_\mu(\delta)$ be the modulus of continuity of $f^{(\mu)}(x)$. Then a polynomial $\rho(x)$ of degree $m + \mu$ exists such that*

$$\begin{aligned} &|f(x) - \rho(x)| < C(l/m)^{\mu}\omega_\mu(l/m), \\ &|f'(x) - \rho'(x)| < C(l/m)^{\mu-1}\omega_\mu(l/m), \qquad l = b - a. \end{aligned} \tag{1.3.4}$$

Here C is a constant depending only on μ.

Analogous inequalities can be obtained for all the derivatives $f(x), f'(x), \cdots, f^{(\mu)}(x)$.

For the first inequality see Jackson **4** (p. 18, Theorem VIII). To prove the second inequality we first establish the following lemma:

LEMMA. *Let $f(\theta)$ be a function of period 2π satisfying the Lipschitz condition*

$$|f(\theta_1) - f(\theta_2)| < \lambda |\theta_1 - \theta_2|, \tag{1.3.5}$$

where λ is a positive constant. Then there exist for each m trigonometric polynomials $g(\theta)$ of degree m such that

$$|f(\theta) - g(\theta)| < \frac{D'\lambda}{m}, \qquad |g'(\theta)| < D''\lambda, \tag{1.3.6}$$

where D' and D'' are absolute constants.

For the first inequality (1.3.6) see Jackson **4**, pp. 2–6. When we use his notation and argument, it suffices to show that $|\lambda^{-1}I_m'(\theta)|$ is less than an absolute constant. But

(1.3.7) $$I'_m(\theta) = -\frac{h_m}{2}\int_{-\pi/2}^{+\pi/2} \{f(\theta + 2u) - f(\theta)\}F'_m(u)\,du$$

and

$$\int_0^{\pi/2} u\,|F'_m(u)|\,du = 4\int_0^{\pi/2} u\left|\frac{\sin mu}{m\sin u}\right|^3\left|\frac{d}{du}\frac{\sin mu}{m\sin u}\right|du$$

(1.3.8) $$= O(1)\int_0^{\pi/2} u\left|\frac{\sin mu}{mu}\right|^3\left|\frac{d}{du}\frac{\sin mu}{mu}\right|du + O(1)\int_0^{\pi/2} u\left|\frac{\sin mu}{mu}\right|^3\left|\frac{\sin mu}{mu}\right|du,$$

since $u/\sin u$ is analytic in the closed interval $[0, \pi/2]$. On writing $mu = x$,

$$O(m^{-1})\int_0^{\infty} x\left|\frac{\sin x}{x}\right|^3\left|\frac{d}{dx}\frac{\sin x}{x}\right|dx + O(m^{-2})\int_0^{\infty} x\left|\frac{\sin x}{x}\right|^4 dx = O(m^{-1}).$$

Now we use (cf. loc. cit.) $h_m = O(m)$.

The analogue of the lemma for polynomials can be derived in the usual way. Then in the upper bound of the first inequality of (1.3.6) the factor $b - a = l$ appears. It is convenient to transform the interval $a \leqq x \leqq b$ into $-\frac{1}{2} \leqq y \leqq \frac{1}{2}$ (instead of $-1 \leqq y \leqq 1$, cf. Jackson, loc. cit., p. 14), defining the function in $[-1, -\frac{1}{2}]$ and $[\frac{1}{2}, 1]$ by a constant.

In order to prove Theorem 1.3.3, we apply Theorem VIII of Jackson (loc. cit., p. 18) to $f'(x)$. (For this argument cf. loc. cit., p. 16.) Thus

$$|f'(x) - q(x)| < K(l/m)^{\mu-1}\omega_\mu(l/m),$$

where $q(x)$ is a proper $\pi_{m+\mu-1}$. Applying the lemma to $f(x) - \int_a^x q(t)dt$, which satisfies a Lipschitz condition with

$$\lambda = K(l/m)^{\mu-1}\omega_\mu(l/m),$$

we obtain a π_m, say $\sigma(x)$, such that

$$\left|f(x) - \int_a^x q(t)\,dt - \sigma(x)\right| < K'(l/m)^{\mu}\omega_\mu(l/m), \qquad |\sigma'(x)| < K''(l/m)^{\mu-1}\omega_\mu(l/m).$$

If we write $\int_a^x q(t)\,dt + \sigma(x) = \rho(x)$, the statement is established.

The constants K, K', K'' in the last three inequalities depend only on μ.

(2) **Theorem** 1.3.4 (Theorem of Runge-Walsh). *Let $f(x)$ be an analytic function regular in the interior of a Jordan curve C and continuous in the closed domain bounded by C. Then $f(x)$ can be approximated with an arbitrary accuracy by polynomials.*

See Walsh 1, p. 36. This theorem has been proved by Runge in case $f(x)$ is analytic on C; the general case is due to Walsh.

We need also a supplement to the former theorem, due to Walsh (1, pp. 75–76). Let C be again a Jordan curve in the complex x-plane. Let $x = \phi(z)$ be the map function carrying over the exterior of C into $|z| > 1$ and preserving $x = z = \infty$. Then the circles $|z| = R$, $R > 1$, correspond to certain curves C_R, called *level curves*. We have

THEOREM 1.3.5. *Let $f(x)$ be analytic within and on C, and let C_R be the largest level curve in the interior of which $f(x)$ is regular. Then to an arbitrary r, $0 < r < R$, there corresponds a constant $M > 0$ such that, for each m, a polynomial $\rho_m(x)$ of degree m exists satisfying the inequality*

$$|f(x) - \rho_m(x)| < Mr^{-m}, \qquad x \text{ on } C. \tag{1.3.9}$$

This holds also if C is a Jordan arc, for example, the interval $-1 \leqq x \leqq +1$. In the latter case C_R is an ellipse with foci at ± 1, and R is the sum of the semi-axes (§1.9).

1.4. Orthogonality; weight function; vectors in function spaces

(1) *Let $\alpha(x)$ be a non-decreasing function in $[a, b]$ which is not constant.* If $a = -\infty$ (or $b = +\infty$), we require that $\alpha(-\infty) = \lim_{x\to-\infty} \alpha(x)$ ($\alpha(+\infty) = \lim_{x\to+\infty} \alpha(x)$) should be finite. The *scalar product* of two real functions $f(x)$ and $g(x)$, where x ranges over the real interval $[a, b]$, is defined by the Stieltjes-Lebesgue integral

$$(f, g) = \int_a^b f(x)g(x)\, d\alpha(x), \tag{1.4.1}$$

where we assume that $f(x)g(x)$ is of the class $L_\alpha(a, b)$. This is certainly the case if $f(x)$ and $g(x)$ are both continuous, or both of bounded variation, and $[a, b]$ is a finite interval. For a fixed function $\alpha(x)$ the *orthogonality* with respect to the "distribution" $d\alpha(x)$ may be defined by the relation

$$(f, g) = 0. \tag{1.4.2}$$

We shall also use the expression "$f(x)$ is orthogonal to $g(x)$."

If we permit $f(x)$ and $g(x)$ to be complex functions in general, definition (1.4.1) must be modified to read

$$(f, g) = \int_a^b f(x)\overline{g(x)}\, d\alpha(x). \tag{1.4.3}$$

With this change in the definition of (f, g), we retain (1.4.2) as the definition of orthogonality.

[For the definition of *Stieltjes-Lebesgue integrals* see, for instance, Hildebrandt **1**, pp. 185–194. This definition, given originally for a monotonic $\alpha(x)$, can easily be extended to the case where $\alpha(x)$ is of bounded variation. Hildebrandt **1**, pp. 177–178, may also be consulted for the definition of Riemann-Stieltjes integrals.

In what follows we sometimes need the formula for integration by parts:

$$\int_a^b f(x)\,d\alpha(x) + \int_a^b \alpha(x)\,df(x) = f(b)\alpha(b) - f(a)\alpha(a), \tag{1.4.4}$$

where a and b are finite, $\alpha(x)$ is of bounded variation, and $f(x)$ is continuous. The integrals are taken as Riemann-Stieltjes integrals.

The expression "distribution" used above arises from the classical interpretation of $d\alpha(x)$ as a continuous or discontinuous mass distribution in the interval $[a, b]$, the mass contributed by the interval $[x_1, x_2]$ of $[a, b]$ being $\alpha(x_2) - \alpha(x_1)$.]

(2) If $\alpha(x)$ is absolutely continuous, the scalar product (1.4.1) reduces to

$$(f, g) = \int_a^b f(x)g(x)w(x)\,dx, \tag{1.4.5}$$

where the integral is assumed to exist in Lebesgue's sense. *Here $w(x)$ is a non-negative function measurable in Lebesgue's sense for which $\int_a^b w(x)\,dx > 0$.* We shall call $w(x)$ the *weight function*, referring to a weight function of, or on, the given interval. Instead of "weight function" the term "norm function" is sometimes used in the literature.[3] In the case of a distribution $w(x)\,dx$ the total mass corresponding to the interval $[x_1, x_2]$ is obviously $\int_{x_1}^{x_2} w(x)\,dx$. In what follows we refer to distributions of the form $d\alpha(x)$ as *distributions of Stieltjes type.*

We use the same concept of distribution and weight function on a curve or on an arc in the complex plane, for example, on the unit circle. Then we replace the variable x by the real parameter which is used for the definition of the curve or arc in question. (See Chapters XI and XVI.)

(3) Let $d\alpha(x)$, or $w(x)\,dx$, $a \leqq x \leqq b$, be a fixed distribution, and consider a space of "vectors" defined by the set of real functions $f(x)$ which belong to the class $L^2_\alpha(a, b)$. The scalar product of two vectors (functions) $f(x)$ and $g(x)$ is defined by (1.4.1) and the length (magnitude, norm) of a vector $f(x)$ by $\|f\| = (f, f)^{\frac{1}{2}}$. Vectors (functions) with $\|f\| = 0$ are called zero-vectors (zero-functions); vectors (functions) with $\|f\| = 1$ are said to be normalized. When $f(x)$ is not a zero-function, $\lambda f(x)$ will be normalized provided $\lambda \neq 0$ is a proper constant, uniquely determined save possibly for sign. If the functions $\alpha(x)$ and $w(x)$ satisfy the conditions mentioned in (1) and (2), there exist functions of positive length for both cases. In the second case $f(x)$ is a zero-function if and only if $\{f(x)\}^2 w(x)$, or what amounts to the same thing, $f(x)w(x)$, vanishes everywhere in $[a, b]$ except on a set of measure zero. If $w(x)$ and $f(x)$ are integrable in Riemann's sense, $f(x)$ is a zero-function provided $f(x)w(x)$ vanishes at every point of continuity.

We note the inequality of Schwarz (cf. (1.11.2))

$$\|fg\| \leqq \|f\|\,\|g\|, \tag{1.4.6}$$

[3] Some corresponding German and French terms are: Belegungsfunktion, Gewichtsfunktion, fonction caractéristique (Stekloff), poids (S. Bernstein).

the equality sign holding if and only if $\lambda f(x) + \mu g(x)$ is a zero-function with λ and μ proper constants not both zero.

A finite set of functions $f_0(x), f_1(x), \cdots, f_l(x)$ is said to be linearly independent if the equation

$$\| \lambda_0 f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_l f_l(x) \| = 0$$

can be true only for

$$\lambda_0 = \lambda_1 = \cdots = \lambda_l = 0.$$

Evidently no zero-function can be contained in such a system. An enumerable set of functions ($l = \infty$) is called linearly independent if the preceding condition is satisfied for every finite subset of the given set.

The extension of these considerations to complex vector spaces is not difficult. The scalar product is then defined as in (1.4.3).

Concerning the axiomatic foundation of these concepts see Stone **1**, Chapter I.

1.5. Closure; integral approximations

(1) Definition. *Let $p \geqq 1$, and let $\alpha(x)$ be a non-decreasing function in $[a, b]$ which is not constant.*[4] *Let the functions*

$$f_0(x), f_1(x), f_2(x), \cdots, f_n(x), \cdots \tag{1.5.1}$$

be of the class $L^p_\alpha(a, b)$. The system (1.5.1) is called closed in $L^p_\alpha(a, b)$ if for every $f(x)$ of $L^p_\alpha(a, b)$ and for every $\epsilon > 0$ a function of the form

$$k(x) = c_0 f_0(x) + c_1 f_1(x) + \cdots + c_n f_n(x) \tag{1.5.2}$$

exists such that

$$\int_a^b | f(x) - k(x) |^p \, d\alpha(x) < \epsilon. \tag{1.5.3}$$

With regard to this definition see Kaczmarz-Steinhaus **1**, p. 49. These authors use the term "Abgeschlossenheit" for "closure."

(2) Theorem 1.5.1. *Let p and $\alpha(x)$ have the same meaning as in the previous definition, and let the function $f(x)$ be of the class $L^p_\alpha(a, b)$, a and b finite. Then for every $\epsilon > 0$ a continuous function $F(x)$ can be determined such that*

$$\int_a^b | f(x) - F(x) |^p \, d\alpha(x) < \epsilon. \tag{1.5.4}$$

For a Riemann-integrable function with $\alpha(x) = x$, this follows by a well-known argument from the definition of the integral. In the general case, it is convenient to use the method of W. H. Young of approximating Stieltjes-Lebesgue integrals. (See Hildebrandt **1**, p. 190.)

Applying Weierstrass' theorem, we obtain the following:

[4] See the remark at the beginning of §1.4 (1).

THEOREM 1.5.2. *Let p, a, b, $\alpha(x)$, $f(x)$ satisfy the conditions of Theorem* 1.5.1. *For every $\epsilon > 0$ there exists a polynomial $\rho(x)$ such that*

$$\int_a^b |f(x) - \rho(x)|^p \, d\alpha(x) < \epsilon. \tag{1.5.5}$$

This means the closure of the system

$$\{x^n\}, \qquad n = 0, 1, 2, \cdots, \tag{1.5.6}$$

in the class $L^p_\alpha(a, b)$. In what follows, we shall use in particular the cases $p = 1$ and $p = 2$.

An analogous statement holds for the "mean approximation" of $f(x)$ by trigonometric polynomials, which is equivalent to the property of closure of the system

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \cdots, \cos nx, \sin nx, \cdots \tag{1.5.7}$$

in $L^p_\alpha(-\pi, +\pi)$.

(3) A more precise form of Theorem 1.5.2 is often useful.

THEOREM 1.5.3. *Let p, a, b, $\alpha(x)$, $f(x)$ satisfy the conditions of Theorem* 1.5.1 *and let $f(x)$ be real-valued. Then we can find a polynomial $\rho(x)$ which satisfies* (1.5.5) *and is such that $\rho(x)$ remains between the upper and lower bounds of $f(x)$.*

We refer also to the following property of Riemann-Stieltjes integrals which plays a rôle in Chapter X.

THEOREM 1.5.4. *Let the real-valued function $f(x)$ be bounded in $[a, b]$, a and b finite, $\alpha(x)$ non-decreasing, and let the Riemann-Stieltjes integral $\int_a^b f(x)\, d\alpha(x)$ exist. For every $\epsilon > 0$ there exist polynomials $\rho(x)$ and $\mathrm{P}(x)$ such that*

$$\inf f(x) - \epsilon \leqq \rho(x) \leqq f(x) \leqq \mathrm{P}(x) \leqq \sup f(x) + \epsilon, \tag{1.5.8}$$

and

$$\int_a^b \{\mathrm{P}(x) - \rho(x)\} \, d\alpha(x) < \epsilon. \tag{1.5.9}$$

See (for $\alpha(x) = x$) Pólya-Szegö **1**, vol. 1, pp. 65, 228, problem 137.

Similar statements hold for approximations by trigonometric polynomials. If $f(x)$ is an even function, $-\pi \leqq x \leqq +\pi$, the approximating trigonometric polynomials can be chosen as cosine polynomials.

1.6. Linear functional operations

(1) Let $\mathfrak{U}(f)$ be an operation which makes a number $\mathfrak{U}(f)$ correspond to every function $f(x)$, continuous in the finite interval $[a, b]$. This operation is called *additive* if

$$\mathfrak{U}(c_1 f_1 + c_2 f_2) = c_1 \mathfrak{U}(f_1) + c_2 \mathfrak{U}(f_2) \tag{1.6.1}$$