

POLSKA AKADEMIA NAUK • INSTYTUT MATEMATYCZNY

S. 7133
[26]

ROZPRAWY MATEMATYCZNE

KOMITET REDAKCYJNY

KAROL BORSUK redaktor

ANDRZEJ MOSTOWSKI, MARCELI STARK

STANISŁAW TURSKI

XXVI

S. PASZKOWSKI

The theory of uniform approximation

I. Non-asymptotic theoretical problems

WARSZAWA 1962

PAŃSTWOWE WYDAWNICTWO NAUKOWE

6.7133

COPYRIGHT 1962

by

PAŃSTWOWE WYDAWNICTWO NAUKOWE
WARSZAWA (Poland), ul. Miodowa 10

All Rights Reserved

No part of this book may be translated or reproduced
in any form, by mimeograph or any other means,
without permission in writing from the publishers.



PRINTED IN POLAND

W R O C Ł A W S K A D R U K A R N I A N A U K O W A

SO. 82

14/5

INTRODUCTION

The theory of approximation deals with the properties of approximation of functions from a given space by functions of a more special nature; the last ones form, as a rule, a linear subspace of the original space. Introducing various metrics in our space, hence obtaining various errors of approximation, we get different types of approximation, the uniform one being among them.

In the classical approach to the theory of uniform approximation, the space of approximated functions is the set \mathcal{C}_F of real functions of one variable, continuous on a closed set F , and the subspace \mathcal{W}_n of approximating functions consists of polynomials of the degree not exceeding n . The error of approximation of the function $\xi \in \mathcal{C}_F$ by the polynomial $\omega \in \mathcal{W}_n$ is defined as $\max_{t \in F} |\xi(t) - \omega(t)|$.

The basic concepts of the theory of uniform approximation are:

1. the n -th best polynomial ω_{nF} for the function ξ on the set F ,
2. the n -th error $\varepsilon_n(\xi; F)$ of the best approximation of the function ξ on the set F ,
3. the (n, F) -points of the function ξ .

The polynomial $\omega_{nF} \in \mathcal{W}_n$ is defined as

$$\max_{t \in F} |\xi(t) - \omega_{nF}(t)| = \inf_{\omega \in \mathcal{W}_n} \max_{t \in F} |\xi(t) - \omega(t)|.$$

The error $\varepsilon_n(\xi; F)$ is, by definition, equal to $\max_{t \in F} |\xi(t) - \omega_{nF}(t)|$, and an (n, F) -point of the function ξ is an arbitrary point $u \in F$ for which $|\xi(u) - \omega_{nF}(u)| = \varepsilon_n(\xi; F)$. A large class of theorems is proved only in the special case when F is a closed interval I , or a finite set.

While writing this paper I had in mind mainly the applications of the theory of uniform approximation to the computational practice. These applications appear especially often since electronic computers were invented. For that reason I restrict myself to the description of problems which are more or less closely connected with applications mentioned above. In the subtitle of the first part I called these problems non-asymptotic, in order to distinguish them from problems which, up to the present date, formed the main direction of the development of

the theory of uniform approximation. This branch of the theory deals with relations between the structural properties of function ξ and the behaviour of the terms of sequences $\{\varepsilon_n(\xi; F)\}$ and $\{\omega_{nF}\}$ as $n \rightarrow \infty$.

In an effort to achieve a sort of self-sufficiency of the paper, I have included in it the basic and at the same time classical, theorems of those who created the theory of uniform approximation — Chebyshev, de la Vallée Poussin, Bernstein and others. I have also used the more recent papers of other authors and my own papers dealing with the same topic (the list of references appears at the end of this paper). However, numerous theorems, mainly from Chapters III and IV, are published here for the first time. The known theorems are supplied with the bibliographic references. Sometimes, however, it has not been possible, because it is now difficult to establish who is the author of some of the most obvious theorems, as in §§ 1 and 4.

Now we shall briefly discuss the content of the first part of the present paper. The first chapter deals with the basic properties of the best polynomials; these results are in general, classical. The only exceptions are the theorems of Hornecker from § 2.4 concerning the approximation of certain rational functions, and theorems from § 3 concerning the pairs of best polynomials; the last results have been published in [22].

The second chapter deals with the various estimations of the error of approximation. The theorem of de la Vallée Poussin (§ 5) makes use of the specific properties of the best approximation by polynomials from the class \mathscr{W}_n on the $(n+2)$ -point sets. Some conclusions from this theorem are taken from paper [25] (they have been proved there in a slightly different way). § 6 contains the estimate of the error $\varepsilon_n(\xi; I)$ for functions which are differentiable many times, and § 7 contains the estimates of the error $\varepsilon_{n+1}(\xi; I)$ which depends upon the properties of the n -th best polynomial; the last results are based on papers [26] and [27]. In § 8 are discussed the advantages which follow from the investigation of how the error $\varepsilon_n(\xi; F)$ depends upon the set F .

In the third chapter using two different methods we investigate the distribution of the (n, I) -points in the interval I . The results obtained there may be of certain practical use, which will be explained in the second part of this paper. Except for one theorem of Bernstein, all the results of this chapter seem to be new. In § 9 we discuss in more detail a method which is presented without proof in [23] and which is based upon a certain theorem from the theory of polynomials, proved in [19]. This method gives the estimate of (n, I) -points of function ξ , which depends upon the value of the ratio $\varepsilon_{n+1}(\xi; I)/\varepsilon_n(\xi; I)$. The starting point of the second method, whose particular case has been described in [24], is the theorem of Bernstein, previously mentioned. It gives the

estimate of (n, I) -points of functions which are differentiable many times. This method is described in §§ 11 and 12. § 10 gives the auxiliary lemmas.

The fourth chapter contains the theory of selected methods of computing polynomials which approximate a given function with a sufficiently small error. The most widely used method of computing the best polynomials is so-called second algorithm of Remez (§ 15), modified and enlarged by many authors (§ 16). Since this method is also rather complicated, we often use instead (§ 13) a polynomial, which is not the best one, but does not differ much from it. In § 14 we discuss more special methods. One of them, published here for the first time, has allowed us to find the numerical values of the coefficients of the Zolotarev polynomials. These polynomials are the natural (and important for practice) generalization of the well-known Chebyshev polynomials.

The second part of the paper will appear some months after the first, under the subtitle Computational practice (of the uniform approximation). For the convenience of those readers who are interested in the theory of approximation from the practical point of view only, we shall briefly present in the second part the theorems and concepts from the first part. The description of the practical usage of these notions and theorems (the numerical methods of chapter IV, in particular) will be illustrated by many numerical examples. The second part will also contain numerous tables for using particular theorems or methods, as for instance, the tables of coefficients a_{nm} , f_{nr} and g_{nr} (§ 5), which allow us to compute the errors of the best approximation, and the tables of numbers $t_{n,kg}$ (§ 9) and $c_{pl}(s)$ (§ 12) connected with the estimates of (n, I) -points and the tables of coefficients of Zolotarev polynomials (§ 14). Most of these tables have been computed on the electronic digital computer of the Computational Centre at Moscow University. For the possibility of performing these computations I am greatly indebted to Professor I. S. Berezin, Director of this Centre.

CHAPTER I

BASIC PROPERTIES OF THE BEST POLYNOMIALS

1. Existence, uniqueness and the characteristic properties of the best polynomials. Most of the theorems presented in this paper concern the approximation of continuous functions by polynomials in a closed interval; nevertheless, in the present section we shall consider the problem from a more general view point. We shall need the basic properties of the approximation not only in the closed interval, but in general, in an arbitrary closed set.

1.1. Let \mathcal{C}_F denote the class of all real functions of one variable, continuous on a certain closed set F . Let the norm be defined as

$$(1) \quad \|\xi\|_F = \max_{t \in F} |\xi(t)| \quad (\xi \in \mathcal{C}_F) \quad (1).$$

Let \mathcal{W}_n , where n is a non-negative integer, denote the class of algebraic polynomials with real coefficients, when the order does not exceed n . Finally, let

$$(2) \quad \varepsilon_n(\xi; F) = \inf_{\omega \in \mathcal{W}_n} \|\xi - \omega\|_F.$$

The number $\varepsilon_n(\xi; F)$ will be called the n -th error of approximation of a function ξ on the set F , or, more precisely, the error of approximation of a function ξ by the polynomials of the class \mathcal{W}_n on the set F . The name is justified by the following Theorem 1:1, which states that for the given function ξ , given n , and given set F , there exists a polynomial $\omega_{nF} \in \mathcal{W}_n$ such that $\|\xi - \omega_{nF}\|_F = \varepsilon_n(\xi; F)$. This polynomial, which approximates the function ξ in the best possible way in the sense of the norm (1) from among all polynomials of the class \mathcal{W}_n , will be called the n -th best polynomial for the function ξ on the set F .

1.2. It follows from the definition of the n -th error of the best approximation, and from the definition of n -th best polynomial, that both

(1) Throughout the paper we shall denote the real numbers by small Latin letters; sets of numbers and their systems by capital Latin letters; functions of a real variable by small Greek letters, and the classes of functions by Latin script letters.

of these concepts depend upon: 1. the function ξ which is approximated; 2. the number n which gives the restriction on the degree of the approximating polynomial; 3. the set F on which the function ξ is approximated. Most of the theorems of the theory of approximation deal with the concepts introduced above, and the related concepts for the case of two parameters: the function ξ and the number n . In many cases this allows us to use the symbols $\varepsilon_n(\xi)$ and ω_n instead of $\varepsilon_n(\xi; F)$ and ω_{nF} .

Concerning the relation of $\varepsilon_n(\xi; F)$ and ω_{nF} with the set F we now restrict ourselves to the remark, that if the set F consists of at most $n+1$ points, then any real function defined on this set is identical with a polynomial of the class \mathscr{W}_n , which can be determined from one of the known interpolation formulae. For such sets the problem of the best approximation is a trivial one; the n -th error of the best approximation of any function $\xi \in \mathscr{C}_F$ is equal to zero, and the n -th best polynomial for this function is identically equal to this function on the set F . To exclude this trivial case we shall assume throughout, that the set F , on which we consider our problem of approximation, contains at least $n+2$ points.

1.3. We shall start the systematic investigation of properties of the best polynomials (the content of the first chapter) from the proof of their existence.

THEOREM 1:1 (Borel [5], p. 84). *For any closed set F , any function $\xi \in \mathscr{C}_F$, and any integer n , there exists the n -th best polynomial for the function ξ on the set F .*

Proof. For $n+1$ fixed and distinct points a_0, a_1, \dots, a_n of the set F let us consider the class \mathscr{W}_n^* of polynomials $\omega \in \mathscr{W}_n$, which satisfy the inequality $|\omega(a_k)| \leq \|\xi\|_F + \varepsilon_n(\xi; F) + 1$ ($k = 0, 1, \dots, n$). The class \mathscr{W}_n^* is non-empty, by definition (2), since it contains the set of those polynomials for which $\|\xi - \omega\|_F \leq \varepsilon_n(\xi; F) + 1$. It follows from the last remark, that when computing the lower bound (2) we may restrict our considerations to the polynomials from the class \mathscr{W}_n^* , that is:

$$(3) \quad \varepsilon_n(\xi; F) = \inf_{\omega \in \mathscr{W}_n^*} \|\xi - \omega\|_F.$$

If we interpret the values of the polynomial at the points a_0, a_1, \dots, a_n as the coordinates of a point in the $(n+1)$ -dimensional Euclidean space, we can consider the one-to-one correspondence between the polynomials of the class \mathscr{W}_n^* and the points (x_0, x_1, \dots, x_n) of an $(n+1)$ -dimensional cube such that

$$(4) \quad |x_k| \leq \|\xi\|_F + \varepsilon_n(\xi; F) + 1 \quad (k = 0, 1, \dots, n).$$

Applying the Lagrange interpolation formula we can, for an arbitrary

$\omega \in \mathcal{W}_n^*$, represent the norm $\|\xi - \omega\|_F$ in the form

$$\|\xi - \omega\|_F = \max_{t \in F} |\xi(t) - \omega(t)| = \max_{t \in F} \left| \xi(t) - \sum_{k=0}^n \lambda_k(t) x_k \right|,$$

where

$$\lambda_k = \prod_{i=0, i \neq k}^n \frac{t - a_i}{a_k - a_i}, \quad x_k = \omega(a_k) \quad (k = 0, 1, \dots, n).$$

It is easy to see that this norm is a continuous function of parameters x_k and achieves its minimum in the cube (4). Thus there exists a polynomial $\omega^* \in \mathcal{W}_n^*$ such that

$$\|\xi - \omega^*\|_F = \min_{\omega \in \mathcal{W}_n^*} \|\xi - \omega\|_F.$$

In virtue of (3), it is the n -th best polynomial for the function ξ , whose existence was to be proved.

1.4. The following theorem, which gives the characteristic properties for the best polynomials is the basic tool of the whole theory of the uniform approximation. The formulation of these properties, due to Chebyshev, gave rise to the existence of this theory.

THEOREM 1:2 (Chebyshev, [8], p. 160 and Borel, [5], p. 86-88). *The polynomial $\omega \in \mathcal{W}_n$ is the n -th best polynomial for the function $\xi \in \mathcal{C}_F$ on the set F if and only if there exist $n+2$ points u_0, u_1, \dots, u_{n+1} (where $u_0 < u_1 < \dots < u_{n+1}$, $u_k \in F$, $k = 0, 1, \dots, n+1$) such that*

- (i) $|\xi(u_k) - \omega(u_k)| = \|\xi - \omega\|_F \quad (k = 0, 1, \dots, n+1),$
- (ii) *the numbers $(-1)^k (\xi(u_k) - \omega(u_k))$ have the common sign for $k = 0, 1, \dots, n+1$ ^(*).*

Thus the form of the difference between the polynomial and a given function determines uniquely whether the polynomial is the n -th best polynomial for the given function.

Proof. Let us first consider the case where the function ξ is a polynomial of the class \mathcal{W}_n on the set F . Then this function is the n -th best polynomial for itself, and this polynomial satisfies the conditions (i) and (ii) independently of the points u_0, u_1, \dots, u_{n+1} from the set F . On the other hand, if a polynomial $\omega \in \mathcal{W}_n$ satisfies the conditions (i) and (ii), it is identical with the function ξ . In fact, let us suppose the converse, i.e. that $\|\xi - \omega\|_F > 0$. Let φ denote a polynomial of the class \mathcal{W}_n identical with the function ξ on the set F . By (ii) the difference $\varphi - \omega$,

^(*) That is, they are all positive, all negative, or all equal to zero.

which is also a polynomial from the class \mathscr{W}_n , changes its sign at least $n+1$ times: between the points u_0 and u_1 , between the points u_1 and u_2, \dots , and between the points u_n and u_{n+1} . It follows, however, that the difference $\varphi - \omega$ is identically zero, which contradicts the inequality $\|\xi - \omega\|_F > 0$.

Let us notice that in the case $\xi \in \mathscr{W}_n$ we have in fact proved the uniqueness of n -th best polynomial (generally it will be done in the form of a separate theorem). The n -th best polynomial coincides with this function, or, more precisely, with its polynomial extension over the whole real axis.

Suppose now that $\xi \notin \mathscr{W}_n$. In this case for an arbitrary polynomial $\omega \in \mathscr{W}_n$ we have $\|\xi - \omega\|_F > 0$ and condition (ii) may be formulated in the following way: (ii) the numbers $\xi(u_0) - \omega(u_0), \xi(u_1) - \omega(u_1), \dots, \xi(u_{n+1}) - \omega(u_{n+1})$ are positive and negative alternately.

I. First we shall prove the sufficiency of conditions (i) and (ii). Suppose, conversely, that the polynomial ω satisfying these conditions is not the n -th best polynomial on the set F ; i.e. there exists a polynomial $\omega^* \in \mathscr{W}_n$ such that

$$(5) \quad \|\xi - \omega^*\|_F < \|\xi - \omega\|_F.$$

It follows from this inequality and from condition (i) that

$$|\xi(u_k) - \omega^*(u_k)| \leq \|\xi - \omega^*\|_F < \|\xi - \omega\|_F = |\xi(u_k) - \omega(u_k)|$$

$$(k = 0, 1, \dots, n+1).$$

At the points u_0, u_1, \dots, u_{n+1} the polynomial $\omega^* - \omega = (\xi - \omega) - (\xi - \omega^*)$ from the class \mathscr{W}_n has the same sign as the difference $\xi - \omega$ and hence is positive and negative alternately. This means that the polynomial $\omega^* - \omega$ has at least $n+1$ distinct zeros, thus it is identically zero. This, however, contradicts the inequality (5), which arose from the assumption that conditions (i) and (ii) are not sufficient.

II. The proof of the necessity of conditions (i) and (ii) is slightly more complicated. Let the polynomial ω_{nF} be the n -th best polynomial for the function ξ on the set F . Suppose that there exist $m < n+2$ points $u_0 < u_1 < \dots < u_{m-1}$ of the set F which satisfy conditions (i) and (ii) (at least one such point exists; it is the point at which the function $|\xi - \omega_{nF}|$ achieves its maximum on the set F), and there are no $m+1$ points with these properties. We shall prove that if this assumption were true, there would exist a polynomial ω^* which approximates the function ξ better than the polynomial ω_{nF} , which would lead to a contradiction, in virtue of the definition of the polynomial ω_{nF} .

The assumption expressed above implies that between the points u_i and u_{i+1} one cannot find two points u_i^* and u_{i+1}^* from the set F

($u_i^* < u_{i+1}^*$), at which the function $\xi - \omega_{nF}$ would achieve the same extremal values, as at the points u_{i+1} and u_i . Similarly, on the left from the point u_0 there is no point u_{-1}^* ; on the right from the point u_{m-1} there is no point u_{m-1}^* , at which the function $\xi - \omega_{nF}$ would have the same absolute value, and a different sign, as at the points u_0 and u_{m-1} , respectively. In fact, if such points (denoted by starred u 's) do exist, we could extend the system of points u_0, u_1, \dots, u_{m-1} , adding these points, in contradiction to the assumption that it is impossible. Thus we can find the disjoint closed intervals I_0, I_1, \dots, I_{m-1} which contain in their interiors the points u_0, u_1, \dots, u_{m-1} , and have the properties:

(Ω_1) If at the point u_k the difference $\xi - \omega_{nF}$ achieves its maximal value equal to $\varepsilon_n(\xi; F)$, then at no point of the closed set $F \cap I_k$ does it achieve its minimal value, equal to $-\varepsilon_n(\xi; F)$; if at the point u_k this difference achieves its minimal value equal to $-\varepsilon_n(\xi; F)$, then at no point of the set $F \cap I_k$ does it achieve its maximal value equal to $\varepsilon_n(\xi; F)$;

(Ω_2) At no point of the closed set $F_0 = F - \bigcup_{k=0}^{m-1} \text{Int } I_k$ does the function $|\xi - \omega_{nF}|$ achieve its maximal value $\varepsilon_n(\xi; F)$.

We assume, not restricting the generality, that the difference $\xi - \omega_{nF}$ is positive at the point u_0 . Thus we have

$$(6) \quad \text{sign}(\xi(u_k) - \omega_{nF}(u_k)) = (-1)^k \quad (k = 0, 1, \dots, m-1)$$

and the property (Ω_1) can be formulated in the following way:

(Ω_1) There exists a positive number g such that for $t \in F \cap I_k$ we have the inequalities

$$(7) \quad \begin{aligned} -\varepsilon_n(\xi; F) + g &\leq \xi(t) - \omega_{nF}(t) \leq \varepsilon_n(\xi; F) & (k \text{ even}), \\ -\varepsilon_n(\xi; F) &\leq \xi(t) - \omega_{nF}(t) \leq \varepsilon_n(\xi; F) - g & (k \text{ odd}). \end{aligned}$$

Decreasing the number g if necessary (keeping it positive, however), we can obtain, besides two inequalities (7), also the inequality

$$\|\xi - \omega_{nF}\|_{F_0} \leq \varepsilon_n(\xi; F) - g$$

which follows from the property (Ω_2).

Let $\varrho(t) = 1$ if $m = 1$ and let

$$\varrho(t) = \prod_{k=1}^{m-1} (c_k - t),$$

where c_k is any point between the intervals I_{k-1} and I_k for $m > 1$. Let us also denote $f = g/2\|\varrho\|_F$. We notice easily that the polynomial ϱ belongs to the class \mathcal{W}_n and that its sign, which changes only at the points c_1, c_2, \dots, c_{m-1} , is identical with the sign of the number $\xi(u_k) - \omega_{nF}(u_k)$

in the interval I_k ($k = 0, 1, \dots, m-1$) (assumption (6) is believed to hold).

Let us consider the polynomial $\omega^* = \omega_{nF} + f\varrho$ from the class \mathcal{W}_n . If $t \in F_0$, then

$$|\xi(t) - \omega^*(t)| \leq |\xi(t) - \omega_{nF}(t)| + f|\varrho(t)| \leq \|\xi - \omega_{nF}\|_{F_0} + f\|\varrho\|_{F_0} \leq \varepsilon_n(\xi; F) - \frac{1}{2}g.$$

If $t \in I_k$ and k is even, we have

$$\begin{aligned} \xi(t) - \omega^*(t) &= \xi(t) - \omega_{nF}(t) - f|\varrho(t)| \\ &\leq \|\xi - \omega_{nF}\|_{I_k} - f \min_{t \in I_k} |\varrho(t)| \\ &\leq \varepsilon_n(\xi; F) - f \min_k \min_{t \in I_k} |\varrho(t)| < \varepsilon_n(\xi; F) \end{aligned}$$

and, by (7):

$$\xi(t) - \omega^*(t) \geq -\varepsilon_n(\xi; F) + g - f\|\varrho\|_{I_k} \geq -\varepsilon_n(\xi; F) + \frac{1}{2}g;$$

analogous results may be obtained for odd k .

It follows from all these inequalities that

$$\|\xi - \omega^*\|_F < \varepsilon_n(\xi; F),$$

which contradicts the definition of the number $\varepsilon_n(\xi; F)$ and, therefore, proves the necessity of conditions (i) and (ii).

1.5. We shall complete the exposition of the most important properties of best polynomials, proving their uniqueness.

THEOREM 1:3 (Borel, [5], p. 85). *For any closed set F , any function $\xi \in \mathcal{C}_F$ and for any non-negative integer n there is exactly one n -th best polynomial for the function ξ on the set F .*

Proof ([34], p. 77-78). We have already proved this theorem for functions ξ such that $\varepsilon_n(\xi; F) = 0$. Let us suppose that $\varepsilon_n(\xi; F) > 0$ and suppose that ω_{nF} and ω_{nF}^* are two best polynomials for the function ξ . The same property is true for their arithmetical mean since by the triangle inequality for the norm (1) we get

$$\|\xi - \frac{1}{2}(\omega_{nF} + \omega_{nF}^*)\|_F = \frac{1}{2}\|(\xi - \omega_{nF}) + (\xi - \omega_{nF}^*)\|_F \leq \varepsilon_n(\xi; F).$$

From Theorem 1:2 applied to n -th best polynomial $\frac{1}{2}(\omega_{nF} + \omega_{nF}^*)$ follows the existence of points u_0, u_1, \dots, u_{n+1} (with $u_0 < u_1 < \dots < u_{n+1}$, $u_k \in F$, $k = 0, 1, \dots, n+1$) such that

$$(8) \quad \left| \xi(u_k) - \frac{1}{2}(\omega_{nF}(u_k) + \omega_{nF}^*(u_k)) \right| = \varepsilon_n(\xi; F) \quad (k = 0, 1, \dots, n+1).$$

To prove Theorem 1:3 it suffices to show that for $k = 0, 1, \dots, n+1$ we have

$$(9) \quad \begin{aligned} \xi(u_k) - \omega_{nF}(u_k) &= \xi(u_k) - \omega_{nF}^*(u_k) = \\ &= \xi(u_k) - \frac{1}{2}(\omega_{nF}(u_k) + \omega_{nF}^*(u_k)) = \pm \varepsilon_n(\xi; F), \end{aligned}$$

since it would then follow that $\omega_{nF} = \omega_{nF}^*$.

Suppose the contrary of formula (9); i.e. that we have the inequality $|\xi(u_k) - \omega_{nF}(u_k)| < \varepsilon_n(\xi; F)$. It leads immediately to the contradiction:

$$\begin{aligned} \varepsilon_n(\xi; F) &= \frac{1}{2} |(\xi(u_k) - \omega_{nF}(u_k)) + (\xi(u_k) - \omega_{nF}^*(u_k))| \\ &< \frac{1}{2} \varepsilon_n(\xi; F) + \frac{1}{2} \varepsilon_n(\xi; F). \end{aligned}$$

The inequality $|\xi(u_k) - \omega_{nF}^*(u_k)| < \varepsilon_n(\xi; F)$ cannot hold either; we have therefore proved that the differences $\xi(u_k) - \omega_{nF}(u_k)$ and $\xi(u_k) - \omega_{nF}^*(u_k)$ have a common absolute value, equal to $\varepsilon_n(\xi; F)$. Now it suffices to note that these differences have also a common sign. Otherwise the left-hand side of (8) would have been zero, which would contradict the assumption $\varepsilon_n(\xi; F) > 0$.

1.6. Having proved that the n -th best polynomial for the function ξ on the set F exists and is uniquely determined by three parameters (ξ , n and F) we may introduce some new concepts and notations. We shall denote the relation between the function ξ and its n -th best polynomial ω_{nF} by the symbol $[\xi, n, F | \omega_{nF}]$.

Any point of the set F at which the function $|\xi - \omega_{nF}|$ achieves its maximal value $\varepsilon_n(\xi; F)$ will be called the (n, F) -point of the function ξ . We shall specify this definition further, according to the sign of the difference at this point. If $\xi(u) - \omega_{nF}(u) = \varepsilon_n(\xi; F)$, we shall call the (n, F) -point u the $(n, +, F)$ -point of function ξ , and if $\xi(u) - \omega_{nF}(u) = -\varepsilon_n(\xi; F)$ we shall call it the $(n, -, F)$ -point of function ξ .

In all notations introduced we shall consequently omit the symbol F , except the cases for which it may be misleading.

Applying the notation of (n, F) -points, we may reformulate the "necessity" part of Theorem 1:2 as follows:

THEOREM 1:4. *For any closed set F , any function $\xi \in \mathcal{C}_F$ and any non-negative integer n there exist at least $n+2$ (n, F) -points of the function ξ , such that when we arrange them according to their magnitude, they become alternately the $(n, +, F)$ -points and $(n, -, F)$ -points.*

We stress the fact that the number of (n, F) -points is restricted by Theorem 1:4 only from below. If there are more than $n+2$ (n, F) -points satisfying the assertion of the theorem, we have the following important consequences concerning the n -th best polynomial of a given function. Namely, we have

THEOREM 1:5. *Let $[\xi, p, F | \omega_{pF}]$ for $p = n, n+1, \dots$. The equation $\omega_{nF} = \omega_{n+1,F} = \dots = \omega_{n+m,F}$ where $m \geq 1$ holds if and only if there exist $n+m+2$ (n, F) -points of function ξ which are (after being arranged according to their magnitude) alternately the $(n, +, F)$ -points and $(n, -, F)$ -points.*

Proof. The polynomial ω_{nF} belongs not only to the class \mathscr{W}_n , but also to the classes $\mathscr{W}_{n+1}, \dots, \mathscr{W}_{n+m}$. If such points exist, then by Theorem 1:2 it is also the $(n+1)$ -st, \dots , $(n+m)$ -th best polynomial for function ξ , and from Theorem 1:3 it follows that $\omega_{nF} = \omega_{n+1,F} = \dots = \omega_{n+m,F}$. If, conversely, the last equation holds, the $n+m+2$ $(n+m, F)$ -points, which exist by Theorem 1:4 are, at the same time, the (n, F) -points of function ξ , which was to be proved.

Until now we have investigated the properties of the best polynomial for a fixed, though arbitrary, function. It is natural to ask how such a polynomial changes when we change the function which determines it. This problem has not yet been solved completely, and our knowledge of the solution cannot be stated in one theorem. In this, introductory chapter, we shall mention only the simplest facts concerning the case.

It is known that if we change the metric which determines our best polynomial into the mean square metric \mathscr{L}^2 , then the n -th best polynomial would depend linearly upon the function approximated. In the theory of uniform approximation this fact, important for applications, is no longer true, and generally, from the relations $[\xi, n, F | \omega_{nF}]$ and $[\eta, n, F | \psi_{nF}]$ it does not follow that $[\xi + \eta, n, F | \omega_{nF} + \psi_{nF}]$. Nevertheless, one can notice some particular features of the linear dependence of the best polynomial from the function approximated in the theory considered in here as may be seen from Theorems 1:6 and 1:7.

THEOREM 1:6. *If $[\xi, n, F | \omega_{nF}]$ then $[c\xi, n, F | c\omega_{nF}]$ for any number c .*

THEOREM 1:7. *If $[\xi, n, F | \omega_{nF}]$ and $\varphi \in \mathscr{W}_n$, then $[\xi + \varphi, n, F | \omega_{nF} + \varphi]$.*

Both theorems follow directly from the characteristic properties of best polynomials as presented in Theorem 1:2. Thus, for instance, Theorem 1:7 follows from the fact that $(\xi + \varphi) - (\omega_{nF} + \varphi) = \xi - \omega_{nF}$, i.e. that the difference between function $\xi + \varphi$ and the polynomial $\omega_{nF} + \varphi$ from the class \mathscr{W}_n satisfies all those conditions of Theorem 1:2 which determine the form of the difference between the function and its n -th best polynomial.

THEOREM 1:8. *If the set F^* can be transformed into F by formula $t = at^* + b$ ($t \in F$), if for some function $\xi \in \mathscr{C}_F$ we have $[\xi, n, F | \omega_{nF}]$, and if $\xi^*(t^*) = \xi(at^* + b)$ then $[\xi^*, n, F^* | \omega_{nF^*}^*]$, where $\omega_{nF^*}^*(t^*) = \omega_{nF}(at^* + b)$.*

This theorem becomes quite obvious if we notice that $\xi^*(t^*) - \omega_{nF^*}^*(t^*) = \xi(at^* + b) - \omega_{nF}(at^* + b) = \xi(t) - \omega_{nF}(t)$ and that the properties of the

difference $\xi - \omega_{nF}$ on the set F which concern its extremal values hold also for the difference $\xi^* - \omega_{nF}^*$ on the set F^* .

The importance of this theorem follows from the fact that most of the theorems of the theory of approximation have their simplest form for certain particular sets, for instance the interval $\langle -1, 1 \rangle$.

The last theorem of § 1 will be proved only for the case of sets F which are symmetric with respect to a certain point a (which implies that for every t the relations $a + t \in F$ and $a - t \in F$ are equivalent), hence it is particularly so for all closed intervals and all discrete sets which have the same distance between every pair of consecutive points.

THEOREM 1:9. *If the set F is symmetric with respect to the point a and if the function $\xi \in \mathcal{C}_F$ is an even function (or an odd function, respectively) with respect to the point a , i.e. if*

$$\xi(a+t) = \xi(a-t) \quad \text{for } a+t \in F$$

(or $\xi(a+t) = -\xi(a-t)$ for $a+t \in F$, respectively), then for an even (or odd, respectively) number n we have the equation

$$\omega_{nF} = \omega_{n+1,F}, \quad \text{where } [\xi, n, F | \omega_{nF}], [\xi, n+1, F | \omega_{n+1,F}]$$

and the polynomial ω_{nF} is even (or odd, respectively) with respect to the point a .

From Theorem 1:9 it follows, in particular, that for $a = 0$ and $F = \langle -b, b \rangle$ the polynomial $\omega_{n, \langle -b, b \rangle}$, which is the best polynomial for an even (or odd) function in the usual sense of this word, contains only the even (odd) powers of the variable t .

We shall prove the theorem for the case when the function ξ is even. It follows from the definition of the best polynomial $\omega_{n+1,F}$ that

$$|\xi(a+t) - \omega_{n+1,F}(a+t)| \leq \varepsilon_{n+1}(\xi; F) \quad (a+t \in F).$$

The last inequality may be also written in the form:

$$|\xi(a-t) - \omega_{n+1,F}(a-t)| \leq \varepsilon_{n+1}(\xi; F) \quad (a-t \in F).$$

If we replace $\xi(a-t)$ by $\xi(a+t)$ in the last inequality, and then add both inequalities, we get

$$|\xi(a+t) - \frac{1}{2}(\omega_{n+1,F}(a+t) + \omega_{n+1,F}(a-t))| \leq \varepsilon_{n+1}(\xi; F) \quad (a+t \in F).$$

It means that the polynomial $\frac{1}{2}(\omega_{n+1,F}(a+t) + \omega_{n+1,F}(a-t))$ is the $(n+1)$ -st best polynomial for the function $\xi(a+t)$. By Theorem 1:3 this polynomial is identical with the polynomial $\omega_{n+1,F}(a+t)$, i.e. $\omega_{n+1,F}(a+t) = \omega_{n+1,F}(a-t)$ and the polynomial $\omega_{n+1,F}$ is even with respect to the point a . If the number n is also even, the polynomial $\omega_{n+1,F}$ cannot con-

tain the term with $(n+1)$ -st power of t , hence it belongs to the class \mathscr{W}_n . Thus it is also the n -th best polynomial for the function ξ , which was to be proved.

2. The direct application of the theorem concerning the (n, F) -points to computing the best polynomials. The Chebyshev polynomials.

2.1. The exact analytical expressions for the best polynomials are known only in few cases. It is easiest to find the n -th best polynomial, if the set F consists of exactly $n+2$ points. In fact, let $F = \{t_0, t_1, \dots, t_{n+1}\}$, where $t_0 < t_1 < \dots < t_{n+1}$ and let $\xi \in \mathscr{C}_F$. By Theorem 1:4 all these points t_k are alternately the $(n, +, F)$ -points and $(n, -, F)$ -points of the approximated function ξ , or — according to their definition — the difference $\xi - \omega_{nF}$ (where $[\xi, n, F | \omega_{nF}]$) takes on at the points t_0, t_1, \dots, t_{n+1} the values equal to $\varepsilon_n(\xi; F)$ and $-\varepsilon_n(\xi; F)$ alternately.

We have

THEOREM 2:1 (Vallée Poussin, [34], p. 81). *If $F = \{t_0, t_1, \dots, t_{n+1}\}$, then the coefficients of the n -th best polynomial ω_{nF} for the function ξ on the set F and the number e_n with the absolute value equal to $\varepsilon_n(\xi; F)$ and the suitably chosen sign, satisfy the system of $n+2$ equations:*

$$(1) \quad \xi(t_k) - \omega_{nF}(t_k) = (-1)^k e_n \quad (k = 0, 1, \dots, n+1).$$

The existence and the uniqueness of the solution of this system follows from Theorems 1:1 and 1:3, and since the system is linear with respect to all $n+2$ unknowns, the effective solution can be relatively easy to find. Further remarks concerning the system (1) are given in § 5, where there the formulas are given for the n -th error of the best approximation $\varepsilon_n(\xi; F)$, and in the second part of the paper, where we explain how to compute $\varepsilon_n(\xi; F)$ and the coefficients of polynomial ω_{nF} in practice.

2.2. Let us now take as the set F the closed interval $\langle -1, 1 \rangle$. We believe that it will be easiest to investigate the properties of the approximation of relatively simple functions, namely the polynomials, but of the degree higher than the approximating ones. This suggestion is, in general, valid, but the results easily applied in practice concern only the case when the degree of the approximated polynomial exceeds the degree of approximating polynomials by 1. We shall investigate this case in more detail, as we shall introduce at the same time the so-called Chebyshev polynomials which are of great importance for the theory of approximation.

Let us assume that we want to find the n -th best polynomial ω_n for the polynomial $\eta = a_0 + a_1 t + \dots + a_n t^n + a_{n+1} t^{n+1}$ in the interval $\langle -1, 1 \rangle$.

Applying Theorems 1:6 and 1:7 we obtain

$$\omega_n = a_0 + a_1 t + \dots + a_n t^n + a_{n+1} \sigma_n,$$

where σ_n is the n -th best polynomial for the function t^{n+1} in this interval. Thus we have succeeded in reducing the problem to that of finding the polynomial σ_n , which does not already depend upon the coefficients of the polynomial η .

To find the polynomial σ_n let us consider for a non-negative integer p the function

$$(2) \quad \tau_p = \cos(p \arccos t)$$

defined in the interval $\langle -1, 1 \rangle$. For $p = 0$ we have $\tau_p = 1$ and for $p > 0$ this function is a polynomial of the degree p with the coefficient 2^{p-1} at t^p , since from the formula

$$\begin{aligned} \cos px &= 2^{p-1} \cos^p x + p \sum_{j=1}^{[p/2]} \frac{(-1)^j}{2j} \binom{p-j-1}{j-1} (2\cos x)^{p-2j} \\ &= p \sum_{j=0}^{[p/2]} \frac{(-1)^j}{2(p-j)} \binom{p-j}{j} (2\cos x)^{p-2j} \quad (p > 0) \end{aligned}$$

([32], p. 41, formula 1.331.3) it follows, after putting $x = \arccos t$, that

$$(3) \quad \tau_p = p \sum_{j=0}^{[p/2]} \frac{(-1)^j}{2(p-j)} \binom{p-j}{j} (2t)^{p-2j} \quad (p > 0).$$

The polynomial (2), whose definition can be extended by (3) to the whole real axis, will be called the p -th Chebyshev polynomial (of the first kind).

It follows easily from (2) that at $p+1$ points

$$t_{pk} = -\cos \frac{k\pi}{p} \quad (k = 0, 1, \dots, p)$$

the polynomial τ_p achieves its extremal values equal to $+1$, and -1 alternately. In fact,

$$\tau_p(t_{pk}) = \cos(p \arccos(-\cos k\pi/p)) = \cos(p-k)\pi = (-1)^{p-k}$$

and $|\tau_p(t)| \leq 1$ for all $t \in \langle -1, 1 \rangle$. It is also easy to note that all zeros of the polynomial τ_p are the numbers $-\cos(2k-1)\pi/2p$ for $k = 1, 2, \dots, p$. Finally, it follows from (3) that the polynomial τ_p is an even function for even p and an odd function for odd p . We shall make use of these facts on many occasions.

Among many interesting properties of the Chebyshev polynomials we shall mention in this chapter only those which will be of use to us in the sequel. The first theorem gives the solution of the problem formulated at the beginning of § 2.2.

THEOREM 2:2 (Chebyshev, [8], p. 173). *The n -th best polynomial σ_n for the function t^{n+1} in the interval $\langle -1, 1 \rangle$ and the n -th error of the approximation of this function are equal respectively to*

$$\sigma_n = t^{n+1} - 2^{-n}\tau_{n+1}, \quad \varepsilon_n(t^{n+1}; \langle -1, 1 \rangle) = 2^{-n}$$

and the $(n, \langle -1, 1 \rangle)$ -points of the function t^{n+1} are

$$t_{n+1,k} = -\cos \frac{k\pi}{n+1} \quad (k = 0, 1, \dots, n+1).$$

Proof. The polynomial $\sigma_n = t^{n+1} - 2^{-n}\tau_{n+1}$ belongs to the class \mathscr{W}_n since, as we have already noticed, the coefficient of t^{n+1} in the polynomial τ_{n+1} of the degree $n+1$ is equal to 2^n . The difference $t^{n+1} - \sigma_n$ is equal to $2^{-n}\tau_{n+1}$, i. e. its norm in the interval $\langle -1, 1 \rangle$ equals $2^{-n}\|\tau_{n+1}\|_{\langle -1, 1 \rangle} = 2^{-n}$, and at $n+2$ points $t_{n+1,0}, t_{n+1,1}, \dots, t_{n+1,n+1}$ this difference takes on the extremal values equal alternately to 2^{-n} and -2^{-n} . Thus Theorem 2:2 is proved.

2.3. The following three theorems specify the relation between the behaviour of the polynomial inside and outside a given interval.

THEOREM 2:3 (Chebyshev, [9]). *An arbitrary polynomial $\varphi \in \mathscr{W}_n$ satisfies the inequality*

$$(4) \quad |\varphi(t)| \leq \|\varphi\|_{\langle -1, 1 \rangle} |\tau_n(t)| \quad (|t| \geq 1).$$

Proof. It suffices to prove (4) for $|t| > 1$ since for $|t| = 1$ it follows directly from the equality $|\tau_n(-1)| = |\tau_n(1)| = 1$. Thus, let us suppose that

$$(5) \quad |\varphi(t^*)| > \|\varphi\|_{\langle -1, 1 \rangle} |\tau_n(t^*)|$$

for $|t^*| > 1$ and let us consider the polynomial from the class \mathscr{W}_n

$$(6) \quad \varrho = \varphi(t^*)\tau_n - \tau_n(t^*)\varphi.$$

By the definition of the Chebyshev polynomial τ_n we have

$$\varrho(t_{nk}) = (-1)^{n-k}\varphi(t^*) - \tau_n(t^*)\varphi(t_{nk})$$

and, in view of (5), we have the relations

$$\text{sign } \varrho(t_{nk}) = (-1)^{n-k} \text{sign } \varphi(t^*) \neq 0 \quad (k = 0, 1, \dots, n).$$

Thus the polynomial ϱ has at least one zero in each of n intervals $(t_{n0}, t_{n1}), (t_{n1}, t_{n2}), \dots, (t_{n,n-1}, t_{nn})$. The number t^* which does not belong to any of these intervals is also a zero of this polynomial. It follows that the polynomial ϱ is identically zero, that is, for all t

$$(7) \quad \varphi(t^*)\tau_n(t) = \tau_n(t^*)\varphi(t).$$

This, however, contradicts (5) for $t = 1$ since $\tau_n(1) = \cos n \arccos 1 = 1$ and $|\varphi(1)| \leq \|\varphi\|_{\langle -1, 1 \rangle}$. Thus the inequality (4) is proved.

THEOREM 2:4 (Paszkowski, [22], p. 47). *If the numbers $-\cos \pi/2n$ and $\cos \pi/2n$ are the zeros of a polynomial φ from the class \mathscr{W}_n , then we have the inequality*

$$(8) \quad |\varphi(t)| \leq \|\varphi\|_C |\tau_n(t)| \quad (|t| \geq \cos \pi/2n),$$

where $C = \langle -\cos \pi/2n, \cos \pi/2n \rangle$.

To understand the meaning of this theorem it is worth-while to remember that the numbers $-\cos \pi/2n$ and $\cos \pi/2n$ are the zeros of the polynomial τ_n . Thus the meaning of Theorems 2:3 and 2:4 is, roughly speaking, the following: in some class of polynomials (determined by the assumptions of the theorems), the values of the polynomial τ_n , which belongs to this class, tend to $\pm \infty$ in the most immediate way. Theorem 2:5, which will be proved later, shows that the Chebyshev polynomials may serve also for estimating the values of other polynomials from below. It should be added that the inequality (8) is stronger than (4) (for $C \subset \langle -1, 1 \rangle$, hence $\|\varphi\|_C \leq \|\varphi\|_{\langle -1, 1 \rangle}$) and it is valid on a larger set of points t ; it is not, however, satisfied for all polynomials from the class \mathscr{W}_n . The estimate (8) is, in fact, interesting only for those polynomials φ , which have no real zeros outside the interval C , since in the opposite case this estimate could be essentially sharpened, at least for some values of t .

Proof. Let us suppose, similarly as in the preceding theorem that

$$(9) \quad |\varphi(t^*)| > \|\varphi\|_C |\tau_n(t^*)|$$

for $|t^*| > \cos \pi/2n$. Under this assumption the polynomial (6) has its zeros in $n-2$ open intervals $(t_{n1}, t_{n2}), \dots, (t_{n,n-2}, t_{n,n-1})$, at the point $-\cos \pi/2n$ (which lies on the left from t_{n1}), at the point $\cos \pi/2n$ (which lies on the right from $t_{n,n-1}$) and at the point t^* . Thus we have the identity (7), but for $t = t_{n1} \in C$ it contradicts the inequality (9).

THEOREM 2:5 (Paszkowski, [22], p. 49). *If the zeros z_1, z_2, \dots, z_n of a polynomial φ of the degree n are real and such that*

$$-\cos \pi/2n = z_1 < z_2 < \dots < z_n = \cos \pi/2n,$$

then the inequality

$$|\varphi(t)| \geq \min_{1 \leq k \leq n-1} \|\varphi\|_{\langle z_k, z_{k+1} \rangle} |\tau_n(t)| \quad (|t| \geq \cos \pi/2n)$$

holds.

Proof. Suppose that

$$(10) \quad |\varphi(t^*)| < \min_{1 \leq k \leq n-1} \|\varphi\|_{\langle z_k, z_{k+1} \rangle} |\tau_n(t^*)|$$

for $|t^*| > \cos \pi/2n$ and consider the polynomial (6). According to the definition of norm $\|\varphi\|$, in each of $n-1$ intervals (z_i, z_{i+1}) there exists a point v_i such that

$$|\varphi(v_i)| = \|\varphi\|_{\langle z_i, z_{i+1} \rangle} \geq \min_{1 \leq k \leq n-1} \|\varphi\|_{\langle z_k, z_{k+1} \rangle}$$

and the signs of the values $\varphi(v_1), \varphi(v_2), \dots, \varphi(v_{n-1})$ are alternately positive and negative. Thus

$$|\tau_n(t^*)\varphi(v_i)| \geq \min_{1 \leq k \leq n-1} \|\varphi\|_{\langle z_k, z_{k+1} \rangle} |\tau_n(t^*)| > |\varphi(t^*)| \geq |\varphi(t^*)\tau_n(v_i)|,$$

$$\text{sign } \varrho(v_i) = -\text{sign } \tau_n(t^*) \text{sign } \varphi(v_i)$$

and the polynomial $\varrho = \varphi(t^*)\tau_n - \tau_n(t^*)\varphi$ has $n+1$ zeros (at the points $-\cos \pi/2n, \cos \pi/2n, t^*$ and in each of the intervals $(v_1, v_2), \dots, (v_{n-2}, v_{n-1})$), which contradicts the inequality (10).

The theorem 2:5 will still be true, probably, for the polynomials φ with multiple zeros, if we replace the Chebyshev polynomials by their proper generalization. However we shall not deal with this generalization, since it is not necessary for our purposes of approximation. Even Theorem 2:4 will be used only for the polynomials φ which satisfy the assumptions of Theorem 2:5.

2.4. The properties of Chebyshev polynomials are due to the periodicity of trigonometric functions, and the relations between some trigonometric and algebraic expressions. The method, by which Hornecker in paper [15] has found the best polynomials for functions $1/(t-c)$ and $1/(t^2-c)$, is based upon the same facts. His results (in a slightly more general form) are given in Theorems 2:7 and 2:8. The same method also provides the best polynomials for function $t/(t^2-c)$ (Theorem 2:9). All formulas are effective, however one usually cannot obtain exactly the (n) -points of approximated functions from them. The situation is still worse with the function $t^{n+2} + at^{n+1}$ (a constant), if we approximate it by the polynomials from the class \mathscr{W}_n . The formula for n -th best polynomial for this function is known (see, for example, [1], p. 265-268), but it is complicated. It is even difficult to think of its use in numerical

approximation. These few remarks should suffice to show the difficulties which we meet when trying to solve effectively the problem of approximation of seemingly simple functions. Therefore, it increases the practical importance of approximate formulas, which we shall discuss in next chapters.

Hornecker's method, stated above, makes use of the following lemma, which we shall presently prove.

THEOREM 2:6. *If $|p| < 1$, m and n are positive integers, and*

$$(11) \quad \varrho_{nm}(x) = \frac{\cos(n+m)x - 2p \cos nx + p^2 \cos(n-m)x}{1 - 2p \cos mx + p^2},$$

then $\|\varrho_{nm}\|_{\langle 0, \pi \rangle} = 1$ and there exist points x_0, x_1, \dots, x_{n+m} with $0 = x_0 < x_1 < \dots < x_{n+m} = \pi$, such that

$$\varrho_{nm}(x_k) = (-1)^k \quad (k = 0, 1, \dots, n+m).$$

Proof. Let us notice at first, that the function ϱ_{nm} is well defined and continuous for every x , which is implied by the relation $1 - 2p \cos mx + p^2 \geq (1 - |p|)^2 > 0$. The inequality $|\varrho_{nm}(x)| \leq 1$ follows from the fact that

$$\begin{aligned} \varrho_{nm}(x) + 1 &= \frac{1 + \cos(n+m)x - 2p(\cos nx + \cos mx) + p^2(1 + \cos(n-m)x)}{1 - 2p \cos mx + p^2} \\ &= \frac{2(\cos^2 \frac{1}{2}(n+m)x - 2p \cos \frac{1}{2}(n+m)x \cos \frac{1}{2}(n-m)x + p^2 \cos^2 \frac{1}{2}(n-m)x)}{1 - 2p \cos mx + p^2} \\ &= \frac{2\gamma^2(x)}{1 - 2p \cos mx + p^2}, \end{aligned}$$

where $\gamma(x) = \cos \frac{1}{2}(n+m)x - p \cos \frac{1}{2}(n-m)x$ and also from

$$\begin{aligned} \varrho_{nm}(x) - 1 &= \frac{\cos(n+m)x - 1 - 2p(\cos nx - \cos mx) + p^2(\cos(n-m)x - 1)}{1 - 2p \cos mx + p^2} \\ &= \frac{2\sigma^2(x)}{1 - 2p \cos mx + p^2}, \end{aligned}$$

where $\sigma(x) = \sin \frac{1}{2}(n+m)x - p \sin \frac{1}{2}(n-m)x$.

The functions $\gamma(x)$ and $\sigma(x)$ are, respectively, the real and imaginary part of the complex function

$$(12) \quad e^{(n+m)xi/2} - pe^{(n-m)xi/2} = e^{(n+m)xi/2}(1 - pe^{-mxi}).$$

As x increases from 0 to π , the argument $\frac{1}{2}(n+m)x$ of the first term of this function increases from 0 to $\frac{1}{2}(n+m)\pi$. The graph of the second

term on the complex plane is the circle with centre at the point 1 and radius $|p| < 1$. The argument of this term oscillates between $-\arcsin|p|$ and $\arcsin|p|$, and for $x = 0$ and $x = \pi$ it is equal to 0. Thus for $x \in \langle 0, \pi \rangle$ the argument of expression (12) varies in a continuous way from 0 to $\frac{1}{2}(n+m)\pi$, and, at some points x_0, x_1, \dots, x_{n+m} with $0 = x_0 < x_1 < \dots < x_{n+m} = \pi$ it takes on the values $0, \frac{1}{2}\pi, \dots, \frac{1}{2}(n+m)\pi$.

It is easily seen that at the points $x_0, x_2, \dots, x_{2[(n+m)/2]}$ the imaginary part of expression (12) vanishes, i.e. $\sigma(x) = 0$ and $\varrho_{nm}(x) = 1$. On the other hand, at the points $x_1, x_3, \dots, x_{2[(n+m-1)/2]+1}$ the real part of expression (12) vanishes, i.e. $\gamma(x) = 0$ and $\varrho_{nm}(x) = -1$. In view of the inequality $|\varrho_{nm}(x)| \leq 1$ this completes the proof of Theorem 2:6.

Now we shall approximate in the interval $I = \langle -1, 1 \rangle$ the three rational functions, mentioned at the beginning of § 2.4.

THEOREM 2:7. *If $|c| > 1$, then for every positive integer n we have*

$$(13) \quad \left[\frac{1}{t-c}, n, I \right] - \frac{4p}{1-p^2} \left(\frac{1}{2} + \sum_{k=1}^{n-1} p^k \tau_k(t) + \frac{p^n \tau_n(t)}{1-p^2} \right)$$

and the equation

$$(14) \quad \varepsilon_n \left(\frac{1}{t-c}; I \right) = \frac{|p|^n}{c^2-1},$$

where $p = c - (\text{sign } c)\sqrt{c^2-1}$.

Proof. From the formula given in [32] (p. 45, formula 1.353.3), it follows that

$$(15) \quad \sum_{k=1}^{n-1} p^k \cos kx = \frac{-p^2 + p \cos x - p^n \cos nx + p^{n+1} \cos(n-1)x}{1 - 2p \cos x + p^2}$$

(in the quoted formula there has been a misprint: $p^n \cos x$ instead of $p^n \cos nx$). If $x = \arccost$, then $\cos kx = \tau_k(t)$ ($k = 0, 1, \dots$) and, in particular, $\cos x = t$ and:

$$\frac{1}{2} + \sum_{k=1}^{n-1} p^k \tau_k(t) = \frac{\frac{1}{2}(1-p^2) - p^n \tau_n(t) + p^{n+1} \tau_{n-1}(t)}{1 - 2pt + p^2}.$$

Let us notice next, that $1+p^2 = 1+2\sigma^2-1-2|\sigma|\sqrt{c^2-1} = 2c \times (c - (\text{sign } \sigma)\sqrt{c^2-1}) = 2cp$, hence

$$\frac{2p}{1-2pt+p^2} = \frac{1}{t-c}.$$

Thus, the difference $\delta(t)$ between the function $1/(t-c)$ and the polynomial, which forms the last term in relation (13), is equal to

$$\begin{aligned}\delta(t) &= -\frac{4p}{1-p^2} \left(\frac{p^n \tau_n(t) - p^{n+1} \tau_{n-1}(t)}{1-2pt+p^2} - \frac{p^n \tau_n(t)}{1-p^2} \right) \\ &= -\frac{4p^{n+1}}{(1-p^2)^2} \cdot \frac{(1-p^2)\tau_n(t) - p(1-p^2)\tau_{n-1}(t) - (1+p^2)\tau_n(t) + 2pt\tau_n(t)}{1-2pt+p^2} \\ &= -\frac{4p^{n+2}}{(1-p^2)^2} \cdot \frac{2t\tau_n(t) - \tau_{n-1}(t) - 2p\tau_n(t) + p^2\tau_{n-1}(t)}{1-2pt+p^2}.\end{aligned}$$

Now let us come back to the variable $x = \arccost$. Since

$$2\cos x \cos nx - \cos(n-1)x = \cos(n+1)x,$$

we have

$$\delta(\cos x) = -\frac{4p^{n+2}}{(1-p^2)^2} \varrho_{n1}(x),$$

where the function $\varrho_{n1}(x)$ is defined by (11). Moreover, $|p| = (|c| + \sqrt{c^2 - 1})^{-1} < 1$. It follows from Theorem 2:6, that $\|\delta(t)\|_{\langle -1, 1 \rangle} = 4|p|^{n+2}/(1-p^2)^2$ and that there exist points $u_0 = \cos x_{n+1}$, $u_1 = \cos x_n$, ..., $u_{n+1} = \cos x_0$ such that $-1 = u_0 < u_1 < \dots < u_{n+1} = 1$ and

$$\delta(u_k) = \frac{4(-1)^{n-k} p^{n+2}}{(1-p^2)^2} \quad (k = 0, 1, \dots, n+1).$$

It means that the difference $\delta(t)$ has the characteristic properties—according to Theorem 1:2—of the difference between the function and its n -th best polynomial. Hence the polynomial, which appears in relation (13) is, in fact, the n -th best polynomial for the function $1/(t-c)$. At the same time we have proved that

$$\varepsilon_n \left(\frac{1}{t-c}; I \right) = \|\delta\|_{\langle -1, 1 \rangle} = \frac{4|p|^{n+2}}{(1-p^2)^2}.$$

Since

$$\begin{aligned}1-p^2 &= 1-2c^2+1+2|c|\sqrt{c^2-1} \\ &= 2\sqrt{c^2-1}(|c|-\sqrt{c^2-1}) \\ &= 2p(\operatorname{sign} c)\sqrt{c^2-1},\end{aligned}$$

the formula (14) follows.

THEOREM 2:8. *If $c < 0$ or $c > 1$, then for every positive even integer n we have*

$$(16) \quad \left[\frac{1}{t^2 - c}, n, I \right] - \frac{8p}{1 - p^2} \left(\frac{1}{2} + \sum_{k=1}^{n/2-1} p^k \tau_{2k}(t) + \frac{p^{n/2} \tau_n(t)}{1 - p^2} \right),$$

$$(17) \quad \varepsilon_n \left(\frac{1}{t^2 - c}; I \right) = \frac{|p|^{n/2}}{2(c^2 - c)},$$

where $p = 2c - 1 - 2(\text{sign } c)\sqrt{c^2 - c}$.

Proof. Let us replace n by $n/2$ and x by $2x$ in identity (15). We get

$$\sum_{k=1}^{n/2-1} p^k \cos 2kx = \frac{-p^2 + p \cos 2x - p^{n/2} \cos nx + p^{n/2+1} \cos(n-2)x}{1 - 2p \cos 2x + p^2}.$$

After putting $x = \arccos t$ it follows that

$$\frac{1}{2} + \sum_{k=1}^{n/2-1} p^k \tau_{2k}(t) = \frac{\frac{1}{2}(1 - p^2) - p^{n/2} \tau_n(t) + p^{n/2+1} \tau_{n-2}(t)}{1 - 2p \tau_2(t) + p^2}.$$

Since $\tau_2(t) = 2t^2 - 1$ and

$$(1 + p)^2 = 4(c - (\text{sign } c)\sqrt{c^2 - c})^2 = 4c(2c - 1 - 2(\text{sign } c)\sqrt{c^2 - c}) = 4cp,$$

we have

$$\frac{4p}{1 - 2p \tau_2(t) + p^2} = \frac{4p}{(1 + p)^2 - 4pt^2} = \frac{1}{t^2 - c}.$$

If we denote by $\delta(t)$ the difference between the function $1/(t^2 - c)$ and the polynomial, which forms the last term in relation (16), we get

$$\begin{aligned} \delta(t) &= -\frac{8p}{1 - p^2} \left(\frac{p^{n/2} \tau_n(t) - p^{n/2+1} \tau_{n-2}(t)}{1 - 2p \tau_2(t) + p^2} - \frac{p^{n/2}}{1 - p^2} \tau_n(t) \right) \\ &= -\frac{8p^{n/2+2}}{(1 - p^2)^2} \cdot \frac{2\tau_2(t) \tau_n(t) - \tau_{n-2}(t) - 2p \tau_n(t) + p^2 \tau_{n-2}(t)}{1 - 2p \tau_2(t) + p^2}. \end{aligned}$$

Since $2\cos 2x \cos nx - \cos(n-2)x = \cos(n+2)x$, we have

$$\begin{aligned} \delta(t) &= -\frac{8p^{n/2+2}}{(1 - p^2)^2} \cdot \frac{\cos(n+2)x - 2p \cos nx + p^2 \cos(n-2)x}{1 - 2p \cos 2x + p^2} \\ &= -\frac{8p^{n/2+2}}{(1 - p^2)^2} Q_{n/2}(x). \end{aligned}$$

It follows from Theorem 2:6 that $\|\delta\|_I = 8|p|^{n/2+2}/(1-p^2)^2$ and there are $n+3$ points of the interval I at which the function $\delta(t)$ takes on the values $\|\delta\|_I$ and $-\|\delta\|_I$ alternately. Thus this polynomial is the n -th (and, at the same time, the $(n+1)$ -st) best polynomial for the function $1/(t^2-c)$. The n -th error of best approximation is equal to

$$\varepsilon_n \left(\frac{1}{t^2-c}; I \right) = \frac{8|p|^{n/2+2}}{(1-p^2)^2}.$$

From the last equation and from the relation

$$\begin{aligned} 1-p^2 &= 1-4c^2+4c-1-4c^2+4c+4(\operatorname{sign} c)(2c-1)\sqrt{c^2-c} \\ &= 4(\operatorname{sign} c)\sqrt{c^2-c}(2c-1-2(\operatorname{sign} c)\sqrt{c^2-c}) = 4p(\operatorname{sign} c)\sqrt{c^2-c} \end{aligned}$$

follows formula (17).

THEOREM 2:9. *If $c < 0$ or $c > 1$, then for every positive odd integer n we have*

$$(18) \quad \left[\frac{t}{t^2-c}, n, I \right] = \frac{4}{1-p} \left(\sum_{k=1}^{(n-1)/2} p^k \tau_{2k-1}(t) + \frac{p^{(n+1)/2} \tau_n(t)}{1-p^2} \right),$$

$$(19) \quad \varepsilon_n \left(\frac{t}{t^2-c}; I \right) = \frac{|p|^{(n-1)/2} (1+p)}{4(c^2-c)},$$

where $p = 2c-1-2(\operatorname{sign} c)\sqrt{c^2-c}$.

Proof. Let us first consider the sum

$$\sum_{k=1}^{n-1} p^k \cos(2k-1)x = \cos x \sum_{k=1}^{n-1} p^k \cos 2kx + \sin x \sum_{k=1}^{n-1} p^k \sin 2kx.$$

Applying identity (15) and an analogous identity for sine (see [32], p. 45, formula 1.353.1) — we get, after changing x into $2x$:

$$\begin{aligned} &\sum_{k=1}^{n-1} p^k \cos(2k-1)x \\ &= [\cos x(-p^2 + p \cos 2x - p^n \cos 2nx + p^{n+1} \cos 2(n-1)x) + \\ &\quad + \sin x(p \sin 2x - p^n \sin 2nx + p^{n+1} \sin 2(n-1)x)] / (1 - 2p \cos 2x + p^2) \\ &= \frac{(p-p^2) \cos x - p^n \cos(2n-1)x + p^{n+1} \cos(2n-3)x}{1 - 2p \cos 2x + p^2}. \end{aligned}$$

Let us put into the last formula $x = \arccost$ and replace n by $(n+1)/2$:

$$\sum_{k=1}^{(n-1)/2} p^k \tau_{2k-1}(t) = \frac{(p-p^2)t - p^{(n+1)/2} \tau_n(t) + p^{(n+3)/2} \tau_{n-2}(t)}{1 - 2p\tau_2(t) + p^2}.$$

Let $\delta(t)$ be the difference between the function $t/(t^2-c)$ and the polynomial which appears in relation (18). We check easily that

$$-\frac{4}{1-p} \cdot \frac{(p-p^2)t}{1-2p\tau_2(t)+p^2} = -\frac{4pt}{(1+p)^2-4pt^2} = \frac{t}{t^2-c}$$

and then we can write equation

$$\begin{aligned} \delta(t) &= -\frac{4}{1-p} \left(\frac{p^{(n+1)/2} \tau_n(t) - p^{(n+3)/2} \tau_{n-2}(t)}{1-2p\tau_2(t)+p^2} - \frac{p^{(n+1)/2} \tau_n(t)}{1-p^2} \right) \\ &= -\frac{4p^{(n+3)/2}}{(1-p)(1-p^2)} \varrho_{n2}(x). \end{aligned}$$

By Theorem 2:6 the mentioned polynomial is the n -th (and, at the same time, the $(n+1)$ -st) best polynomial for the function $t/(t^2-c)$. We have also derived the formula

$$\varepsilon_n \left(\frac{t}{t^2-c}; I \right) = \frac{4|p|^{(n+3)/2}}{(1-p)(1-p^2)} = \frac{4|p|^{(n+3)/2} (1+p)}{(1-p^2)^2}.$$

Since, as we already know, $1-p^2 = 4p(\text{sign } c)\sqrt{c^2-c}$, we can transform the last formula to the form (19).

We should mention that for each of the functions $1/(t-c)$, $1/(t^2-c)$ and $t/(t^2-c)$ its n -th best polynomial differs from the corresponding partial sum of the expansion of this function into a series with respect to the Chebyshev polynomials only by the coefficient with τ_n . In fact, this series is the following:

$$\left. \begin{aligned} \frac{1}{t-c} &= -\frac{4p}{1-p^2} \left(\frac{1}{2} + \sum_{k=1}^{\infty} p^k \tau_k(t) \right) & (p = c - (\text{sign } c)\sqrt{c^2-1}), \\ \frac{1}{t^2-c} &= -\frac{8p}{1-p^2} \left(\frac{1}{2} + \sum_{k=1}^{\infty} p^k \tau_{2k}(t) \right) \\ \frac{t}{t^2-c} &= -\frac{4}{1-p} \sum_{k=1}^{\infty} p^k \tau_{2k-1}(t) \end{aligned} \right\} (p = 2c - 1 - 2(\text{sign } c)\sqrt{c^2-c}).$$

It is also worth-while to compare relations (13), (16) and (18) (the last ones for $c > 1$) with the more complicated Bernstein's formulas ([4], p. 81-83).

3. The characteristic properties of pairs of successive best polynomials. In § 1 we have given the characteristic properties of the best polynomials. It is quite obvious that an arbitrary polynomial from the class \mathcal{W}_n is the n -th best polynomial on the set F for some function from \mathcal{C}_F and for any arbitrarily given approximation error (to obtain this function it suffices to add to the given polynomial the auxiliary function, which would take on at $n+2$ points of the set F the maximal and the minimal values alternately, with modulus equal to the given n -th approximation error). Thus we can say, that the class \mathcal{W}_n coincides with the class of all possible n -th best polynomials. If, however, we consider two polynomials $\psi_n \in \mathcal{W}_n$ and $\psi_{n+1} \in \mathcal{W}_{n+1}$ then it turns out that such a function $\xi \in \mathcal{C}_F$ for which these polynomials are the n -th and the $(n+1)$ -st best polynomials on the set F does not always exist.

3.1. In this section we shall discuss the properties of pairs of successive best polynomials which correspond to a given function, but we shall deal only with the case when F is a closed interval. This case has important practical consequences. All theorems of § 3 have been proved in [22].

THEOREM 3:1. *If the polynomials ω_n and ω_{n+1} are the n -th and the $(n+1)$ -st best polynomials for the function $\xi \in \mathcal{C}_I$ where I is a closed interval (a, b) and if $\omega_n \neq \omega_{n+1}$, then all the zeros z_1, z_2, \dots, z_{n+1} of the difference $\omega_n - \omega_{n+1}$ are real, single, and satisfy the inequality*

$$u_0 < z_1 < u_1 < z_2 < \dots < z_{n+1} < u_{n+1},$$

where u_0, u_1, \dots, u_{n+1} are, alternately, the $(n, +)$ -points and $(n, -)$ -points of the function ξ .

Proof. The assumption that the polynomials ω_n and ω_{n+1} are not identical is equivalent to the inequality

$$(1) \quad \|\xi - \omega_n\|_I > \|\xi - \omega_{n+1}\|_I.$$

The difference $\xi - \omega_n$ is, at the points u_k , positive and negative alternately, and it has the absolute value equal to $\|\xi - \omega_n\|_I$. Then, it follows from inequality (1) that the difference $\omega_n - \omega_{n+1} = (\xi - \omega_{n+1}) - (\xi - \omega_n)$ has, at the points u_k , the sign opposite to the sign of the difference $\xi - \omega_n$ at these points, i.e. positive and negative alternately. It follows at once that in each of $n+1$ intervals $(u_0, u_1), (u_1, u_2), \dots, (u_n, u_{n+1})$ the polynomial $\omega_n - \omega_{n+1}$ has at least one zero. If we consider the fact that the degree of this polynomial does not exceed $n+1$, we conclude that all

the mentioned zeros are single and that there are no other zeros of this polynomial.

We shall keep the notation z_1, z_2, \dots, z_{n+1} for the zeros of the polynomial $\omega_n - \omega_{n+1}$ in the next theorem, and we shall, in addition, assume (in order to have the uniformity of notations) that $z_0 = a, z_{n+2} = b$.

THEOREM 3:2. *If $\omega_n \neq \omega_{n+1}$, then*

$$(2) \quad \varepsilon_n(\xi) + \varepsilon_{n+1}(\xi) \geq \|\omega_n - \omega_{n+1}\|_I,$$

$$(3) \quad \varepsilon_n(\xi) - \varepsilon_{n+1}(\xi) \leq \min_{0 \leq k \leq n+1} \|\omega_n - \omega_{n+1}\|_{I_k},$$

where $I_k = \langle z_k, z_{k+1} \rangle$ ($k = 0, 1, \dots, n+1$).

Proof. Inequality (2) follows directly from the equations $\varepsilon_n(\xi) = \|\xi - \omega_n\|_I$, $\varepsilon_{n+1}(\xi) = \|\xi - \omega_{n+1}\|_I$, which define the polynomials ω_n and ω_{n+1} , if we consider the triangle inequality for the norm $\|\cdot\|_I$. To prove (3) let us notice that in view of the inequality

$$z_0 \leq u_0 < z_1 < u_1 < \dots < z_{n+1} < u_{n+1} \leq z_{n+2}$$

in any interval $\langle z_k, z_{k+1} \rangle$ ($k = 0, 1, \dots, n+1$) we have an (n) -point u_k at which

$$\begin{aligned} |\xi(u_k) - \omega_n(u_k)| &= \varepsilon_n(\xi), \\ |\xi(u_k) - \omega_{n+1}(u_k)| &\leq \varepsilon_{n+1}(\xi). \end{aligned}$$

It follows that

$$\varepsilon_n(\xi) - \varepsilon_{n+1}(\xi) \leq |\omega_n(u_k) - \omega_{n+1}(u_k)| \leq \|\omega_n - \omega_{n+1}\|_{I_k},$$

which was to be proved.

3.2. The next theorem shows that the conditions formulated in Theorems 3:1 and 3:2, satisfied by the best polynomials ω_n and ω_{n+1} , errors of approximation $\varepsilon_n(\xi)$, $\varepsilon_{n+1}(\xi)$ and the interval I are characteristic for these notions.

THEOREM 3:3. *We are given: the polynomial ω_n of the degree not exceeding n and ω_{n+1} of the degree $n+1$, such that their difference $\omega_n - \omega_{n+1}$ has real single zeros $z_1 < z_2 < \dots < z_{n+1}$ in the open interval (a, b) . Let the numbers e_n and e_{n+1} satisfy the inequalities $0 \leq e_{n+1} < e_n$,*

$$(4) \quad e_n + e_{n+1} \geq \|\omega_n - \omega_{n+1}\|_I,$$

$$(5) \quad e_n - e_{n+1} \leq \min_{0 \leq k \leq n+1} \|\omega_n - \omega_{n+1}\|_{I_k},$$

where $I = \langle a, b \rangle$, $I_k = \langle z_k, z_{k+1} \rangle$, $z_0 = a, z_{n+2} = b$.

If the above conditions are satisfied, then there exists a function $\xi \in \mathcal{E}_I$ such that $[\xi, n, I | \omega_n]$, $[\xi, n+1, I | \omega_{n+1}]$, $\varepsilon_n(\xi; I) = e_n$, $\varepsilon_{n+1}(\xi; I) = e_{n+1}$.

Proof. Let

$$A^+ = \{t: t \in I, |\omega_n(t) - \omega_{n+1}(t) - e_n| \leq e_{n+1}\},$$

$$A^- = \{t: t \in I, |\omega_n(t) - \omega_{n+1}(t) + e_n| \leq e_{n+1}\}.$$

Since $e_{n+1} < e_n$, the difference $\omega_n - \omega_{n+1}$ must be positive in the set A^+ and negative in the set A^- . It follows at once that these sets are disjoint. Their sum A consists of the points t at which

$$-e_{n+1} + e_n \leq \omega_n(t) - \omega_{n+1}(t) \leq e_{n+1} + e_n$$

or

$$-e_{n+1} - e_n \leq \omega_n(t) - \omega_{n+1}(t) \leq e_{n+1} - e_n,$$

hence

$$A = A^+ \cup A^- = \{t: t \in I, e_n - e_{n+1} \leq |\omega_n(t) - \omega_{n+1}(t)| \leq e_n + e_{n+1}\}.$$

Since all zeros of the difference $\omega_n - \omega_{n+1}$ are real and single, in each of the intervals I_1, I_2, \dots, I_n , whose end points are just these zeros, this difference has only one extremal point. For the same reason the difference $\omega_n - \omega_{n+1}$ is monotone on the semi-axes $(-\infty, z_1)$ and $(z_{n+1}, +\infty)$ and more so on the intervals $I_0 = \langle a, z_1 \rangle$ and $I_{n+1} = \langle z_{n+1}, b \rangle$. From assumptions (4) and (5) it follows that the set A has points in each of the intervals I_k and, by these remarks we have

$$A = \bigcup_{i=0}^{n+1} J_i,$$

where

$$J_0 = \langle a, b \rangle \subset \langle a, z_1 \rangle,$$

$$J_k = \langle a_k, b_k \rangle \subset (z_k, z_{k+1}) \quad (k = 1, 2, \dots, n),$$

$$J_{n+1} = \langle a_{n+1}, b \rangle \subset (z_{n+1}, b).$$

Moreover,

$$(6) \quad A^+ = J_0 \cup J_2 \cup \dots \quad \text{and} \quad A^- = J_1 \cup J_3 \cup \dots$$

or

$$(7) \quad A^+ = J_1 \cup J_3 \cup \dots \quad \text{and} \quad A^- = J_0 \cup J_2 \cup \dots,$$

since the difference $\omega_n - \omega_{n+1}$ is alternately positive and negative in the intervals J_0, J_1, \dots, J_{n+1} .

On the other hand, we know that at end points a_k and b_k of the intervals J_k we have the equation $|\omega_n(t) - \omega_{n+1}(t)| = e_n - e_{n+1}$.

Let us define the function ξ in the interval I by the equations

$$\xi(t) = \begin{cases} (8) & \left. \begin{aligned} & \omega_n(t) - e_n && (t \in A^+), \\ (9) & \omega_n(t) + e_n && (t \in A^-), \\ (10) & \omega_{n+1}(t) - \frac{6e_{n+1}s}{a_1 - b_0} \left(\frac{5b_0 + a_1}{6} - t \right) && (t \in \langle b_0, \frac{2b_0 + a_1}{3} \rangle), \\ (11) & \omega_{n+1}(t) - \frac{6e_{n+1}s}{a_1 - b_0} \left(t - \frac{b_0 + a_1}{2} \right) && (t \in \langle \frac{2b_0 + a_1}{3}, \frac{b_0 + 2a_1}{3} \rangle), \\ (12) & \omega_{n+1}(t) - \frac{6e_{n+1}s}{a_1 - b_0} \left(\frac{b_0 + 5a_1}{6} - t \right) && (t \in \langle \frac{b_0 + 2a_1}{3}, a_1 \rangle), \\ (13) & \omega_{n+1}(t) - \frac{2(-1)^k e_{n+1}s}{a_{k+1} - b_k} \left(\frac{b_k + a_{k+1}}{2} - t \right) && (t \in \langle b_k, a_{k+1} \rangle, \\ & && k = 1, 2, \dots, n), \end{aligned} \right\} \end{cases}$$

where $s = \text{sign}(\omega_n(a) - \omega_{n+1}(a))$.

Since the interval I is the sum of the intervals

$$J_0 = \langle a, b_0 \rangle, \langle b_0, a_1 \rangle, J_1 = \langle a_1, b_1 \rangle, \dots, \langle b_n, a_{n+1} \rangle, J_{n+1} = \langle a_{n+1}, b \rangle,$$

the above equations define the function ξ in the whole interval I . To check whether the function ξ is continuous in this interval it suffices to find out if, at the points

$$b_0, a_1, b_1, \dots, b_n, a_{n+1}, \frac{2b_0 + a_1}{3}, \frac{b_0 + 2a_1}{3}$$

the particular parts of the definition do not contradict each other, for each of them separately defines the function ξ in a continuous way. For the last two points the consistency follows at once from (10)-(12). In the case of (6) (the case (7) is analogous):

$$\omega_n(a_k) - \omega_{n+1}(a_k) = \omega_n(b_k) - \omega_{n+1}(b_k) = (-1)^k (e_n - e_{n+1})$$

(we remember that in the sets A^+ , A^- the difference $\omega_n - \omega_{n+1}$ is positive and negative, respectively). In this case, since $a \in J_0 \subset A^+$ — we have $s = 1$ and from definitions (8) and (13) it follows that if k is even, then

$$\xi(a_k) = \omega_n(a_k) - e_n = \omega_{n+1}(a_k) - e_{n+1}$$

and the last equation holds. In a similar way we check that definitions (8), (10) and (13) are identical for $b_k \in A^+$, definitions (9), (12) and (13)

are identical for $a_k \in A^-$; finally that definitions (9) and (13) are identical for $b_k \in A^-$.

We have already proved that $\xi \in \mathcal{C}_I$. Now we shall show that the function ξ has all the properties mentioned in Theorem 3:3. We start from the proof of the inequalities

$$\|\xi - \omega_n\|_I \leq e_n, \quad \|\xi - \omega_{n+1}\|_I \leq e_{n+1}.$$

The inequality

$$(14) \quad |\xi(t) - \omega_n(t)| \leq e_n$$

is obvious in the set A (see (8) and (9)). In the set $I - A$ we have the inequality

$$(15) \quad |\xi(t) - \omega_{n+1}(t)| \leq e_{n+1},$$

since in each of the subintervals of which this set consists, the difference $\xi - \omega_{n+1}$ is linear, and takes on the values e_{n+1} or $-e_{n+1}$ at the end points of these intervals. At the same time, on the set $I - A$ we have the inequality $|\omega_n(t) - \omega_{n+1}(t)| \leq e_n - e_{n+1}$ which, added to (15), gives the required inequality (14). It remains to show that the inequality (15) is valid also for $t \in A$. It is so by the definitions (8) and (9) and the definition of sets A^+ and A^- , since for example, for $t \in A^+$ we have $\xi(t) = \omega_n(t) - e_n$ and $|\omega_n(t) - e_n - \omega_{n+1}(t)| \leq e_{n+1}$.

Now it remains only to show that: 1. at $n+2$ points $a, a_1, a_2, \dots, a_{n-1}$ the difference $\xi - \omega_n$ takes on the values equal to e_n and $-e_n$ alternately, 2. at $n+3$ points $b_0, (2b_0 + a_1)/3, (b_0 + 2a_1)/3, a_1, a_2, \dots, a_n$ the difference $\xi - \omega_{n+1}$ takes on the values equal to e_{n+1} and $-e_{n+1}$ alternately. The first part of this assertion follows from definition (8) and (9) and from the remark that the points $a, a_1, a_2, \dots, a_{n+1}$ belong to the sets A^+ and A^- alternately. The second part follows from definitions (10)-(13).

Thus Theorem 3:3 is completely proved. In connection with this it is worthy of mention that if we change in the suitable way the definition of the function ξ in the interval $\langle b_0, a_1 \rangle$ we can—for any fixed integer m —achieve the polynomial ω_{n+1} is for ξ not only the $(n+1)$ -st, but also the $(n+m)$ -th best polynomial. In fact, to do this it suffices to separate $\langle b_0, a_1 \rangle$ into the number of subintervals greater than before and change definitions (10)-(12) in such a way, that the difference $\xi - \omega_{n+1}$ would change linearly in each of these intervals. It follows that between two successive (n) -points we may have many $(n+1)$ -points (in fact, not only in the case when the last ones are at the same time $(n+m)$ -points with $m > 1$). In general, if we know about the function ξ only that $\varepsilon_n(\xi) > \varepsilon_{n+1}(\xi)$, then it is difficult to say anything which could have any practical importance about the (n) -points and $(n+1)$ -points.

THEOREM 3:4. *Let $[\xi, n | \omega_n]$, $[\xi, n+1, \omega_{n+1}]$ and $\omega_n \neq \omega_{n+1}$. If u_0, u_1, \dots, u_{n+1} (where $u_0 < u_1 < \dots < u_{n+1}$) are, alternately, the $(n, +)$ -points and $(n, -)$ -points of the function ξ , and v_0, v_1, \dots, v_{n+2} (where $v_0 < v_1 < \dots < v_{n+2}$) are, alternately, the $(n+1, +)$ -points and $(n+1, -)$ -points of the function ξ , then we have at least one of the inequalities*

$$(16) \quad v_0 \geq u_0, u_{n+1} \geq v_{n+2}.$$

Proof. Suppose, on the contrary, that we have the inequalities $v_0 < u_0, u_{n+1} < v_{n+2}$. We shall prove that if $v_0 < u_0$ then

$$(17) \quad \text{sign}(\xi(u_0) - \omega_n(u_0)) = -\text{sign}(\xi(v_0) - \omega_{n+1}(v_0)).$$

Let, e.g., $\xi(u_0) - \omega_n(u_0) = \varepsilon_n(\xi)$. Since $\omega_n - \omega_{n+1} = (\xi - \omega_{n+1}) - (\xi - \omega_n)$, we have

$$\omega_n(u_0) - \omega_{n+1}(u_0) \leq \varepsilon_{n+1}(\xi) - \varepsilon_n(\xi) < 0.$$

We know that all the zeros of the difference $\omega_n - \omega_{n+1}$ are real, and the smallest among them lies on the right from u_0 (Theorem 3:1). Since $v_0 < u_0$, we have

$$\omega_n(v_0) - \omega_{n+1}(v_0) < \omega_n(u_0) - \omega_{n+1}(u_0) \leq \varepsilon_{n+1}(\xi) - \varepsilon_n(\xi).$$

From two equations

$$\xi(v_0) - \omega_{n+1}(v_0) = \varepsilon_{n+1}(\xi), \quad \xi(v_0) - \omega_{n+1}(v_0) = -\varepsilon_{n+1}(\xi),$$

which define the $(n+1)$ -point v_0 the first one contradicting the above inequality (it would follow, that $\xi(v_0) - \omega_n(v_0) > \varepsilon_n(\xi)$), hence the second one must hold, which proves (17).

In a similar way (assuming that $v_{n+2} > u_{n+1}$) we show that

$$(18) \quad \text{sign}(\xi(u_{n+1}) - \omega_n(u_{n+1})) = -\text{sign}(\xi(v_{n+2}) - \omega_{n+1}(v_{n+2})).$$

Since, however, the values of the difference $\xi - \omega_n$ at the points u_0, u_1, \dots, u_{n+1} are alternately positive and negative, and the values of the difference $\xi - \omega_{n+1}$ at the points v_0, v_1, \dots, v_{n+2} are also alternately positive and negative, we have

$$\text{sign}(\xi(u_{n+1}) - \omega_n(u_{n+1})) = (-1)^{n+1} \text{sign}(\xi(u_0) - \omega_n(u_0)),$$

$$\text{sign}(\xi(v_{n+2}) - \omega_{n+1}(v_{n+2})) = (-1)^{n+2} \text{sign}(\xi(v_0) - \omega_{n+1}(v_0))$$

and equations (17) and (18), which follow from the negation of the theorem's assertion, are contradictory.

It follows in particular from Theorem 3:4, that the systems of (n) -points and $(n+1)$ -points never have the "separating property":

$$v_0 < u_0 < v_1 < u_1 < \dots < v_{n+1} < u_{n+1} < v_{n+2}.$$

The "separating property" is the feature, for instance, of the zeros of orthogonal polynomials. Under rather severe restrictions on the function ξ (see § 11), we have, however, the inequality which is not much weaker than the preceding one.

3.3. In § 3 we have investigated the characteristic properties of pairs of successive best polynomials. The properties of longer sequences (of three, four, ...) successive best polynomials are not known. Theorems 3:1 and 3:2 give only the necessary conditions under which some polynomials $\omega_n, \omega_{n+1}, \dots, \omega_{n+m}$ are the n -th, $(n+1)$ -st, ..., $(n+m)$ -th best polynomial for the function ξ . It is not obvious how to generalize Theorem 3:3 for the case of an arbitrary closed set F . In any case, the construction of the function ξ given in the proof of the mentioned theorem cannot be carried out in the general case.

CHAPTER II

ESTIMATION OF ERROR OF THE BEST APPROXIMATION

4. Basic theorems. In this section we shall give the most elementary and the most frequently applied estimations concerning the error of the best approximation $\varepsilon_n(\xi; F)$, which we shall denote by $\varepsilon_n(\xi)$, unless it may lead to misunderstanding. We remember that

$$(1) \quad \varepsilon_n(\xi) = \min_{\omega \in \mathcal{W}_n} \|\xi - \omega\|_F = \|\xi - \omega_n\|_F,$$

where \mathcal{W}_n is the class of algebraic polynomials of the degree not exceeding n , F is a closed set on which the function ξ is defined and continuous, and finally, ω_n is the n -th best polynomial for the function ξ on the set F .

4.1. Without the basic theorem 4:1 the whole theory of uniform approximation would be of no practical application. Hence we have

THEOREM 4:1 (Weierstrass, [37]). *For any closed set F , and any function $\xi \in \mathcal{C}_F$, the sequence $\{\varepsilon_n(\xi)\}$ is non-increasing and tends to 0.*

The practical meaning of this theorem is the following: any function ξ continuous on the closed set F can be approximated with an arbitrary degree of accuracy in the sense of norm $\|\cdot\|_F$ by algebraic polynomials.

The proof presented here is due to Lebesgue [16]. The proof based upon the properties of Bernstein polynomials (see Bernstein, [3]) is more frequently used.

The inequality $\varepsilon_n(\xi) \geq \varepsilon_{n+1}(\xi)$ follows directly from definition (1) and from the fact that the class \mathcal{W}_{n+1} in which we take the minimum in the formula for $\varepsilon_{n+1}(\xi)$ contains the class \mathcal{W}_n . It remains to show that for a given continuous function ξ and for an arbitrary positive number d there exists a polynomial ψ_d such that $\|\xi - \psi_d\|_F \leq d$. It would then follow, that if n_d is the degree of the polynomial ψ_d , then $\varepsilon_{n_d}(\xi) \leq d$ and $\lim_{n \rightarrow \infty} \varepsilon_n(\xi) = 0$.

Let us extend the function ξ in a continuous way to the whole closed interval $I = \langle a, b \rangle$, which contains the set F , and put

$$t_{mi} = a + \frac{(b-a)i}{m} \quad (i = 0, 1, \dots, m).$$

In the interval I we define the function λ_m which satisfies the interpolation conditions

$$(2) \quad \lambda_m(t_{mi}) = \xi(t_{mi}) \quad (i = 0, 1, \dots, m).$$

Its graph is a polygonal line which vertices have the ordinates t_{mi} . Thus the function λ_m is of the form

$$\lambda_m = \sum_{i=0}^m c_{mi} |t - t_{mi}|,$$

where the coefficients c_{mi} are defined by (2).

In each of the subintervals $\langle t_{mi}, t_{m,i+1} \rangle$ the function λ_m is linear:

$$\lambda_m(t) = \xi(t_{mi}) + \frac{\xi(t_{m,i+1}) - \xi(t_{mi})}{t_{m,i+1} - t_{mi}} (t - t_{mi});$$

thus for $t \in \langle t_{mi}, t_{m,i+1} \rangle$ we have

$$\begin{aligned} |\lambda_m(t) - \xi(t)| &\leq |\xi(t_{mi}) - \xi(t)| + \frac{|\xi(t_{m,i+1}) - \xi(t_{mi})|}{t_{m,i+1} - t_{mi}} (t - t_{mi}) \\ &\leq |\xi(t_{mi}) - \xi(t)| + |\xi(t_{m,i+1}) - \xi(t_{mi})|. \end{aligned}$$

Since the function ξ is continuous in the interval I , there exists a number m such that if $t, u \in I$, $|t - u| \leq (b - a)/m$, then $|\xi(t) - \xi(u)| \leq d/4$. For this m we have the inequality $|\lambda_m(t) - \xi(t)| \leq d/2$ for $t \in \langle t_{mi}, t_{m,i+1} \rangle$, $i = 0, 1, \dots, m - 1$, i.e. the inequality

$$(3) \quad \|\lambda_m - \xi\|_I \leq d/2.$$

Now we shall use the well-known expansion

$$\sqrt{1 - u^2} = 1 - \frac{1}{2}u^2 - \frac{1}{2 \cdot 4}u^4 - \frac{1 \cdot 3}{2 \cdot 4 \cdot 6}u^6 - \dots$$

valid for $|u| \leq 1$. Introducing the new variable $|t| = \sqrt{1 - u^2}$ we write this expansion in the form

$$|t| = 1 - \frac{1}{2}(1 - t^2) - \frac{1}{2 \cdot 4}(1 - t^2)^2 - \frac{1 \cdot 3}{2 \cdot 4 \cdot 6}(1 - t^2)^3 - \dots$$

for $|t| \leq 1$. From the last equation, and by a suitable linear transformation of the variable t , we can put the expansion of the function $|t - t_{mi}|$ into a polynomial series, uniformly convergent in the interval I . Thus there exists a polynomial ψ_{di} equal to a sufficiently large sum of this series, and such that

$$||t - t_{mi}| - \psi_{di}| \leq \frac{d}{2(m+1)(|c_{mi}| + 1)} \quad (t \in I; i = 0, 1, \dots, m).$$

It follows that for $t \in I$

$$\begin{aligned} \left| \lambda_m - \sum_{i=0}^m c_{mi} \psi_{di} \right| &\leq \sum_{i=0}^m |c_{mi}| \cdot |t - t_{mi} - \psi_{di}| \\ &\leq \frac{d}{2(m+1)} \sum_{i=0}^m \frac{|c_{mi}|}{|c_{mi}| + 1} \leq \frac{d}{2}. \end{aligned}$$

Thus, for $\psi_d = \sum_{i=0}^m c_{mi} \psi_{di}$ we have the inequality $\|\lambda_m - \psi_d\|_I \leq d/2$, which combined with (3), gives the inequality $\|\xi - \psi_d\|_I \leq d$, hence also the required inequality $\|\xi - \psi_d\|_F \leq d$.

In connection with the last theorem we shall make two remarks:

I. We have proved that $\varepsilon_n(\xi) \geq \varepsilon_{n+1}(\xi)$. The equation $\varepsilon_n(\xi) = \varepsilon_{n+1}(\xi) = \dots = \varepsilon_{n+m}(\xi)$ is equivalent to the fact that the n -th best polynomial for the function ξ is, at the same time, the $(n+1)$ -st, ..., $(n+m)$ -th best polynomial for the same function. For that reason Theorem 1:5 determines when such an equation holds.

II. If the set F is a closed interval, then the sequence $\{\varepsilon_n(\xi; F)\}$ is completely determined by the properties mentioned in Theorem 4:1. More precisely, for any sequence $\{e_n\}$ of numbers, satisfying the condition $e_0 \geq e_1 \geq e_2 \geq \dots$, $\lim_{n \rightarrow \infty} e_n = 0$ and for every closed interval I with positive length there exists a function $\xi \in \mathcal{C}_I$ for which $\varepsilon_n(\xi; I) = e_n$ ($n = 0, 1, \dots$). This has been proved by Bernstein, [2]. The assumption that F is an interval is not very essential however, the fact mentioned is no longer valid if F is a finite set.

4.2. Now we shall present a theorem which follows directly from Theorems 1:6-1:9.

THEOREM 4:2. For any function $\xi \in \mathcal{C}_F$ and any number c we have the relation

$$\varepsilon_n(c\xi) = |c| \varepsilon_n(\xi).$$

THEOREM 4:3. For any function $\xi \in \mathcal{C}_F$ and any polynomial $q \in \mathcal{H}_n$ we have the relation

$$\varepsilon_n(\xi + q) = \varepsilon_n(\xi).$$

THEOREM 4:4. If the set F^* can be transformed into the set F by the formula $t = at^* + b$ ($t \in F$) and if $\xi^*(t^*) = \xi(at^* + b)$, where $\xi \in \mathcal{C}_F$, then we have

$$\varepsilon_n(\xi^*; F^*) = \varepsilon_n(\xi; F).$$

THEOREM 4:5. *If the set F is symmetric with respect to the point a and if the function $\xi \in \mathcal{C}_F$ is even (or odd, respectively) with respect to the point a , then for any even (odd, respectively) number n we have the relation*

$$\varepsilon_n(\xi) = \varepsilon_{n+1}(\xi).$$

These theorems will be used mainly in the case when F is a closed interval. Theorem 4:5 may be slightly generalized by using Theorem 4:3; namely it is valid also in the case when there exists a polynomial $\varphi \in \mathcal{W}_n$ such that the function $\xi + \varphi$ is even (or odd) with respect to the point a .

THEOREM 4:6. *For any functions $\xi, \eta \in \mathcal{C}_F$ we have the inequality*

$$(4) \quad |\varepsilon_n(\xi) - \varepsilon_n(\eta)| \leq \varepsilon_n(\xi + \eta) \leq \varepsilon_n(\xi) + \varepsilon_n(\eta).$$

Proof. Suppose we are given $[\xi, n, F | \omega_n], [\eta, n, F | \psi_n]$. From the definition of the n -th approximation error it follows that

$$\begin{aligned} \varepsilon_n(\xi + \eta) &= \min_{\nu \in \mathcal{W}_n} \|(\xi + \eta) - \nu\|_F \leq \|(\xi + \eta) - (\omega_n + \psi_n)\|_F \\ &\leq \|\xi - \omega_n\|_F + \|\eta - \psi_n\|_F = \varepsilon_n(\xi) + \varepsilon_n(\eta). \end{aligned}$$

If in the last inequality we replace the functions ξ, η and $\xi + \eta$ by $-\eta, \xi + \eta$ and $(-\eta) + (\xi + \eta) = \xi$ respectively, we get

$$\begin{aligned} \varepsilon_n(\xi) &\leq \varepsilon_n(-\eta) + \varepsilon_n(\xi + \eta) = \varepsilon_n(\eta) + \varepsilon_n(\xi + \eta), \\ \varepsilon_n(\xi) - \varepsilon_n(\eta) &\leq \varepsilon_n(\xi + \eta). \end{aligned}$$

In the same way $\varepsilon_n(\eta) - \varepsilon_n(\xi) \leq \varepsilon_n(\xi + \eta)$, which completes the proof of Theorem 4:6.

If we consider arbitrary linearly dependent functions ξ and η , we easily see that Theorem 4:6 cannot be improved. More exactly, this problem is explained by the following

THEOREM 4:7. *The equation $\varepsilon_n(\xi + \eta) = \varepsilon_n(\xi) + \varepsilon_n(\eta)$ holds if and only if, in the set F there exist points $u_0 < u_1 < \dots < u_{n+1}$ such that*

$$(5) \quad \begin{aligned} \xi(u_k) - \omega_n(u_k) &= (-1)^k c \varepsilon_n(\xi), & \eta(u_k) - \psi_n(u_k) &= (-1)^k c \varepsilon_n(\eta) \\ & & (k = 0, 1, \dots, n+1), \end{aligned}$$

where $[\xi, n, F | \omega_n], [\eta, n, F | \psi_n], |c| = 1$.

Proof. If equation (5) are satisfied, then

$$(6) \quad \begin{aligned} (\xi(u_k) + \eta(u_k)) - (\omega_n(u_k) + \psi_n(u_k)) &= (-1)^k c (\varepsilon_n(\xi) + \varepsilon_n(\eta)) \\ & \quad (k = 0, 1, \dots, n+1) \end{aligned}$$

and in the whole set F , as we have observed in the preceding theorem, we have

$$\|(\xi + \eta) - (\omega_n + \psi_n)\| \leq \varepsilon_n(\xi) + \varepsilon_n(\eta).$$

It follows that the points u_0, u_1, \dots, u_{n+1} are the (n) -points of the function $\xi + \eta$, and the n -th error of the best approximation of this function is the number $\varepsilon_n(\xi) + \varepsilon_n(\eta)$.

Suppose now that $\varepsilon_n(\xi + \eta) = \varepsilon_n(\xi) + \varepsilon_n(\eta)$. In Theorem 4:6 we have proved that the error of approximation of the sum $\xi + \eta$ by the polynomial $\omega_n + \psi_n$ does not exceed $\varepsilon_n(\xi) + \varepsilon_n(\eta)$, i. e., in view of the above assumption, the number $\varepsilon_n(\xi + \eta)$. Thus the polynomial $\omega_n + \psi_n$ is the n -th best polynomial for the function $\xi + \eta$ and there exist the (n) -points of this function u_0, u_1, \dots, u_{n+1} satisfying equations (6).

Suppose that one of the equations of group (5) is not satisfied, i. e. that $\xi(u_k) - \omega_n(u_k) = (-1)^k c(\varepsilon_n(\xi) - d)$ for some k and a positive d . Then $\eta(u_k) - \psi_n(u_k) = (-1)^k c(\varepsilon_n(\eta) + d)$, which is not consistent with the definition of the polynomial ψ_n and number $\varepsilon_n(\eta)$. Thus the first group of equations (5) holds, and in view of (6) the second group must also hold.

It follows easily from Theorem 4:7 that the second limiting case in inequality (4): $\varepsilon_n(\xi + \eta) = |\varepsilon_n(\xi) - \varepsilon_n(\eta)|$ holds if and only if there exist points $u_0 < u_1 < \dots < u_{n+1}$ such that

$$\begin{aligned} \xi(u_k) - \omega_n(u_k) &= (-1)^k c \varepsilon_n(\xi), & \eta(u_k) - \psi_n(u_k) &= -(-1)^k c \varepsilon_n(\eta) \\ & & & (k = 0, 1, \dots, n+1). \end{aligned}$$

4.3. The next theorem shows for which functions ξ and η we may sharpen the left-hand side of inequality (4).

THEOREM 4:8. *If the set F is symmetric with respect to the point a and if the functions $\xi, \eta \in \mathcal{C}_F$ are even and odd, respectively, with respect to the point a , then*

$$(7) \quad \varepsilon_n(\xi + \eta) \geq \max\{\varepsilon_n(\xi), \varepsilon_n(\eta)\}.$$

Proof. By the assumptions of the functions ξ and η we have the relations $\xi(a-t) = \xi(a+t)$, $\eta(a-t) = -\eta(a+t)$ for $a+t \in F$. Let now $[\xi + \eta, n, F | v_n]$. Then

$$|\xi(a+t) + \eta(a+t) - v_n(a+t)| \leq \varepsilon_n(\xi + \eta) \quad (a+t \in F),$$

hence also

$$|\xi(a-t) + \eta(a-t) - v_n(a-t)| \leq \varepsilon_n(\xi + \eta) \quad (a-t \in F).$$

Adding these inequalities we get

$$|\xi(a+t) + \xi(a-t) + \eta(a+t) + \eta(a-t) - v_n(a+t) - v_n(a-t)| \leq 2\varepsilon_n(\xi + \eta),$$

which can be reduced to

$$|\xi(a+t) - \frac{1}{2}(v_n(a+t) + v_n(a-t))| \leq \varepsilon_n(\xi + \eta) \quad (a+t \in F)$$

by the use of assumptions concerning functions ξ and η .

It follows that there exists a polynomial from the class \mathcal{W}_n which approximates the function ξ with the error not exceeding $\varepsilon_n(\xi + \eta)$,

i.e. $\varepsilon_n(\xi) \leq \varepsilon_n(\xi + \eta)$. Similarly we prove that $\varepsilon_n(\eta) \leq \varepsilon_n(\xi + \eta)$, which completes the proof of (7).

Theorem 4:8 may be sharpened in a similar way as was Theorem 4:5; inequality (7) holds in the case when polynomials $\varphi, \chi \in \mathcal{W}_n$ exist such that the function $\xi + \varphi$ is even, and the function $\eta + \chi$ is odd with respect to the point a , for in this case:

$$\begin{aligned} \varepsilon_n(\xi + \eta) &= \varepsilon_n((\xi + \varphi) + (\eta + \chi)) \geq \max\{\varepsilon_n(\xi + \varphi), \varepsilon_n(\eta + \chi)\} \\ &= \max\{\varepsilon_n(\xi), \varepsilon_n(\eta)\}. \end{aligned}$$

5. Vallée Poussin type estimations.

5.1. By this title we mean the group of estimations based upon the exact expression existing for the n -th error of the best approximation on the set consisting of $n + 2$ points. Basic estimations of this kind have been given by Vallée Poussin.

THEOREM 5:1 (Vallée Poussin, [34], p. 81). *If $F^* = \{t_0, t_1, \dots, t_{n+1}\}$, where $t_0 < t_1 < \dots < t_{n+1}$ then*

$$\varepsilon_n(\xi; F^*) = \frac{\left| \sum_{k=0}^{n+1} (-1)^k \xi(t_k) / w_k \right|}{\sum_{k=0}^{n+1} 1/w_k},$$

where

$$w_k = \prod_{i=0, i \neq k}^{n+1} |t_k - t_i|.$$

Proof. Theorem 2:1 states that the coefficients of the n -th best polynomial ω_{nF^*} for the function ξ on an $(n + 2)$ -point set F^* , and the number e_n with the absolute value equal to $\varepsilon_n(\xi; F^*)$, satisfy the system of $n + 2$ linear equations

$$(1) \quad \omega_{nF^*}(t_k) + (-1)^k e_n = \xi(t_k) \quad (k = 0, 1, \dots, n + 1).$$

Thus, by the Cramer formulas we have for the last unknown the equation

$$e_n = \frac{\begin{vmatrix} 1 & t_0 & \dots & t_0^n & \xi(t_0) \\ 1 & t_1 & \dots & t_1^n & \xi(t_1) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & t_{n+1} & \dots & t_{n+1}^n & \xi(t_{n+1}) \end{vmatrix}}{\begin{vmatrix} 1 & t_0 & \dots & t_0^n & 1 \\ 1 & t_1 & \dots & t_1^n & -1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & t_{n+1} & \dots & t_{n+1}^n & (-1)^{n+1} \end{vmatrix}} = \frac{\sum_{k=0}^{n+1} (-1)^k d_k \xi(t_k)}{\sum_{k=0}^{n+1} d_k},$$

where d_k is the Vandermonde determinant of the form

$$\begin{vmatrix} 1 & t_0 & \dots & t_0^n \\ \dots & \dots & \dots & \dots \\ 1 & t_{k-1} & \dots & t_{k-1}^n \\ 1 & t_{k+1} & \dots & t_{k+1}^n \\ \dots & \dots & \dots & \dots \\ 1 & t_{n+1} & \dots & t_{n+1}^n \end{vmatrix}.$$

It follows from the well-known formula that

$$\begin{aligned} d_k &= \prod_{0 \leq i < j \leq n+1; i, j \neq k} (t_j - t_i) \\ &= \frac{\prod_{0 \leq i < j \leq n+1} (t_j - t_i)}{\prod_{i=0}^{k-1} (t_k - t_i) \prod_{j=k+1}^{n+1} (t_j - t_k)} = \frac{\prod_{0 \leq i < j \leq n+1} (t_j - t_i)}{w_k}, \end{aligned}$$

hence

$$(2) \quad e_n = \frac{\sum_{k=0}^{n+1} (-1)^k \xi(t_k) / w_k}{\sum_{k=0}^{n+1} 1 / w_k},$$

which was to be proved.

If $F^* \subset F$ where F is a closed set, and if $\xi \in \mathcal{C}_{F^*}$, then $\varepsilon_n(\xi; F^*) \leq \varepsilon_n(\xi; F)$, since for an arbitrary polynomial ω we have $\|\xi - \omega\|_{F^*} \leq \|\xi - \omega\|_F$. On the other hand, if we consider the $n+2$ points $u_0 < u_1 < \dots < u_{n+1}$ which are alternately the $(n, +, F)$ - and $(n, -, F)$ -points of the function ξ , then the n -th best polynomial for the function ξ on the set F satisfies, by definition, the equations (1) for $t_k = u_k$ and $|e_n| = \varepsilon_n(\xi; F)$. Thus we get

THEOREM 5:2 (Vallée Poussin, [34], p. 86). *For any function $\xi \in \mathcal{C}_F$ we have*

$$(3) \quad \varepsilon_n(\xi; F) = \max_{\{t_0, t_1, \dots, t_{n+1}\} \subset F} \varepsilon_n(\xi; \{t_0, t_1, \dots, t_{n+1}\}),$$

where the maximum is taken for all systems of points $t_k \in F$ such that $t_0 < t_1 < \dots < t_{n+1}$.

Theorem 5:2 may serve for estimating from below the n -th error of the best approximation on the closed set which contains more than $n+2$ points by the means of such errors which correspond to $(n+2)$ -points subset. If we succeed in selecting the proper subset, this estima-

tion may be quite satisfactory. Usually, however, we apply this theorem in a slightly modified form. We can replace the function ξ on both sides of equation (3) by the function $\xi - \omega$ where ω is a suitably chosen polynomial from the class \mathscr{W}_n , and then, applying Theorem 4:3, it follows that $\varepsilon_n(\xi - \omega) = \varepsilon_n(\xi)$.

THEOREM 5:3 (Vallée Poussin, [34], p. 85). *If, for the function $\xi \in \mathscr{C}_F$, the polynomial $\omega \in \mathscr{W}_n$ has the property that the numbers*

$$(4) \quad \xi(t_0) - \omega(t_0), \xi(t_1) - \omega(t_1), \dots, \xi(t_{n+1}) - \omega(t_{n+1}) \\ (t_0 < t_1 < \dots < t_{n+1}, t_i \in F \quad \text{for } i = 0, 1, \dots, n+1)$$

are alternately non-negative and non-positive, then

$$\varepsilon_n(\xi; F) \geq \frac{\sum_{k=0}^{n+1} |\xi(t_k) - \omega(t_k)| / w_k}{\sum_{k=0}^{n+1} 1/w_k}, \quad \text{where } w_k = \prod_{j=0, j \neq k}^{n+1} |t_k - t_j|.$$

This inequality follows directly from Theorems 5:1, 5:2, equation $\varepsilon_n(\xi - \omega) = \varepsilon_n(\xi)$ and the assumption concerning the sign of the difference $\xi - \omega$, which allows us to eliminate the factor $(-1)^k$ in the sum on the right-hand side of this inequality.

5.2. If the degree n of the approximating polynomial is large, the computation of the expressions w_k may be cumbersome. We may apply a theorem of Remez (see [31], p. 37), which is somewhat weaker than the preceding one.

THEOREM 5:4. *If for the function $\xi \in \mathscr{C}_F$ the polynomial $\omega \in \mathscr{W}_n$ has the property that the numbers (4) are alternately non-negative and non-positive, then*

$$\varepsilon_n(\xi; F) \geq \frac{1}{2} \min_{0 \leq k \leq n} (|\xi(t_k) - \omega(t_k)| + |\xi(t_{k+1}) - \omega(t_{k+1})|).$$

Proof. In equation (2) we get

$$w_k = \prod_{i=0, i \neq k}^{n+1} |t_k - t_i| = (-1)^{n+1-k} \prod_{i=0, i \neq k}^{n+1} (t_k - t_i) = (-1)^{n+1-k} \pi'(t_k),$$

where $\pi(t) = (t - t_0)(t - t_1) \dots (t - t_{n+1})$. Thus, replacing in (2) the function ξ by the function $\delta = \xi - \omega$ (the equation (2) is still valid, as stated by Theorem 4:3), we obtain

$$e_n = \frac{\sum_{k=0}^{n+1} \delta(t_k) / \pi'(t_k)}{\sum_{k=0}^{n+1} (-1)^k / \pi'(t_k)}.$$

We know from the theory of finite differences that the numerator of this expression is equal to the $(n+1)$ -st divided difference of the function δ , computed for points t_0, t_1, \dots, t_{n+1} ([18], p. 9). Let us denote this difference by $\delta[t_0, t_1, \dots, t_{n+1}]$. If we introduce the auxiliary function σ defined at the points t_k by equations $\sigma(t_k) = (-1)^k$ ($k = 0, 1, \dots, n+1$) then the denominator of the above expression for e_n will become the $(n+1)$ -st divided difference $\sigma[t_0, t_1, \dots, t_{n+1}]$ and

$$e_n = \frac{\delta[t_0, t_1, \dots, t_{n+1}]}{\sigma[t_0, t_1, \dots, t_{n+1}]}.$$

Now we remember that the $(n+1)$ -st divided difference is defined by n -th differences with equations

$$\delta[t_0, t_1, \dots, t_{n+1}] = \frac{-\delta[t_0, t_1, \dots, t_n] + \delta[t_1, t_2, \dots, t_{n+1}]}{t_{n+1} - t_0}$$

([18], p. 1). Applying the analogous formulas for the differences of lower and lower orders, we finally get the equations

$$\begin{aligned} & \delta[t_0, t_1, \dots, t_{n+1}] \\ &= \frac{1}{t_{n+1} - t_0} \left(\frac{-\delta[t_0, t_1, \dots, t_{n-1}] + \delta[t_1, t_2, \dots, t_n]}{t_n - t_0} + \right. \\ & \quad \left. + \frac{-\delta[t_1, t_2, \dots, t_n] + \delta[t_2, t_3, \dots, t_{n+1}]}{t_{n+1} - t_1} \right) \\ &= \frac{1}{(t_{n+1} - t_0)(t_n - t_0)} \delta[t_0, t_1, \dots, t_{n-1}] - \\ & \quad - \frac{1}{t_{n+1} - t_0} \left(\frac{1}{t_n - t_0} + \frac{1}{t_{n+1} - t_1} \right) \delta[t_1, t_2, \dots, t_n] + \\ & \quad \frac{1}{(t_{n+1} - t_0)(t_{n+1} - t_1)} \delta[t_2, t_3, \dots, t_{n+1}] = \dots \\ &= \sum_{k=0}^n (-1)^{n-k} c_k \delta[t_k, t_{k+1}] = \sum_{k=0}^n (-1)^{n-k} c_k \frac{\delta(t_{k+1}) - \delta(t_k)}{t_{k+1} - t_k}, \end{aligned}$$

where the coefficients c_k are positive. The same equation holds for the

function σ , hence

$$e_n = \frac{\sum_{k=0}^n (-1)^{n-k} c_k \frac{\delta(t_{k+1}) - \delta(t_k)}{t_{k+1} - t_k}}{\sum_{k=0}^n (-1)^{n-k} c_k \frac{(-1)^{k+1} - (-1)^k}{t_{k+1} - t_k}} = \frac{\sum_{k=0}^n (-1)^{k+1} \frac{2c_k}{t_{k+1} - t_k} \frac{\delta(t_{k+1}) - \delta(t_k)}{2}}{\sum_{k=0}^n \frac{2c_k}{t_{k+1} - t_k}}.$$

Considering equation $\delta = \xi - \omega$ and the assumption about the sign of the difference $\xi - \omega$ we can say that

$$|e_n| = \frac{\sum_{k=0}^n \frac{2c_k}{t_{k+1} - t_k} \left(|\xi(t_{k+1}) - \omega(t_{k+1})| + |\xi(t_k) - \omega(t_k)| \right)}{2 \sum_{k=0}^n \frac{2c_k}{t_{k+1} - t_k}}.$$

Thus the number $|e_n|$ is the weighted mean, with the weights $2c_k/(t_{k+1} - t_k)$, of numbers $\frac{1}{2} (|\xi(t_{k+1}) - \omega(t_{k+1})| + |\xi(t_k) - \omega(t_k)|)$, hence the assertion follows.

- 5.3.** The applications of the Vallée Poussin type estimations can be simplified by another method, if we fix in advance a certain special way of selecting the points t_0, t_1, \dots, t_{n+1} . Now we assume that the set F is the interval $I = \langle -1, 1 \rangle$. From the theorems which will be presented in Chapter III it follows that for functions which are sufficiently regular in this interval, their (n, I) -points lie close to the points

$$t_{n+1,k} = -\cos \frac{k\pi}{n+1} \quad (k = 0, 1, \dots, n+1).$$

Thus we may suppose that the n -th error of the best approximation $\varepsilon_n(\xi; F^*)$ gives for $F^* = \{t_{n+1,0}, t_{n+1,1}, \dots, t_{n+1,n+1}\}$ a sufficiently exact estimation of $\varepsilon_n(\xi; I)$. This supposition is supported by many numerical examples. Now we shall consider in some detail the conclusions which may be drawn from Theorems 5:1 and 5:2 in the case when $t_k = t_{n+1,k}$.

THEOREM 5:5 (Paszkowski, [25]). *For any function $\xi \in \mathcal{C}_{\langle -1, 1 \rangle}$ we have the inequality*

$$(5) \quad \varepsilon_n(\xi; \langle -1, 1 \rangle) \geq \frac{1}{n+1} \left| \frac{1}{2} \xi(t_{n+1,0}) - \xi(t_{n+1,1}) + \dots + (-1)^n \xi(t_{n+1,n}) + \left| \frac{(-1)^{n+1}}{2} \xi(t_{n+1,n+1}) \right| \right|.$$

Proof. We shall show that the number (2) for $t_k = t_{n+1,k}$ is equal to

$$\frac{1}{n+1} \left(\frac{1}{2} \xi(t_0) - \xi(t_1) + \dots + (-1)^n \xi(t_n) + \frac{(-1)^{n+1}}{2} \xi(t_{n+1}) \right)$$

(here and in the sequel we shall write for simplicity t_k instead of $t_{n+1,k}$). Then inequality (5) will follow from Theorem 5:1 for $F^* = \{t_0, t_1, \dots, t_{n+1}\}$ and from Theorem 5:2 for $F = \langle -1, 1 \rangle$.

First we shall compute the numbers w_k . To do this consider the $(n+1)$ -st Chebyshev polynomial $\tau_{n+1} = \cos((n+1)\arccost)$ defined in § 2.2. We already know that it is a polynomial of the degree $n+1$ with the coefficient 2^n of t^{n+1} . Thus the coefficient of t^n in its derivative

$$(6) \quad \tau'_{n+1} = \frac{n+1}{\sqrt{1-t^2}} \sin((n+1)\arccost)$$

is equal to $2^n(n+1)$. All the zeros of the polynomial τ'_{n+1} are numbers t_1, t_2, \dots, t_n , for

$$\begin{aligned} \sin((n+1)\arccost_k) &= \sin\left((n+1)\arccos\left(\cos\left(\frac{(n+1-k)\pi}{n+1}\right)\right)\right) \\ &= \sin(n+1-k)\pi = 0 \end{aligned}$$

and $1-t_k^2 \neq 0$, hence

$$\prod_{j=1}^n (t-t_j) = \frac{1}{2^n(n+1)} \tau'_{n+1} = \frac{1}{2^n \sqrt{1-t^2}} \sin((n+1)\arccost).$$

We have $t_0 = -\cos 0 = -1$, $t_{n+1} = -\cos \pi = 1$, thus

$$\left| \prod_{j=0}^{n+1} (t-t_j) \right| = \frac{1}{2^n} \sqrt{1-t^2} \sin((n+1)\arccost) = \frac{1}{2^n} \sqrt{(1-t^2)(1-\tau_{n+1}^2)}.$$

Now we can compute the expression

$$w_0 = |(t_0-t_1)(t_0-t_2)\dots(t_0-t_{n+1})| = |(-1-t_1)(-1-t_2)\dots(-1-t_{n+1})|.$$

In view of the last equation we have

$$\begin{aligned} (7) \quad w_0 &= \left| \prod_{j=1}^{n+1} (-1-t_j) \right| = \lim_{t \rightarrow -1} \frac{\left| \prod_{j=0}^{n+1} (t-t_j) \right|}{|t+1|} \\ &= \frac{1}{2^n} \lim_{t \rightarrow -1} \sqrt{\frac{(1-t^2)(1-\tau_{n+1}^2)}{(1+t)^2}} = \frac{1}{2^n} \sqrt{2 \lim_{t \rightarrow -1} \frac{1-\tau_{n+1}^2}{1+t}}. \end{aligned}$$

From the definition of polynomial τ_{n+1} it follows that $\tau_{n+1}(-1) = (-1)^{n+1}$, and identity (8), derived in the proof, gives for $t = -1$ the equation $\tau'_{n+1}(-1) = (-1)^n(n+1)^2$. Thus, applying the de l'Hôpital rule to the last limit in the sequence of equations (7) we obtain

$$w_0 = \frac{1}{2^n} \sqrt{2 \lim_{t \rightarrow -1} (-2\tau_{n+1}\tau'_{n+1})} = \frac{n+1}{2^{n-1}}.$$

In a similar way we get

$$w_{n+1} = \frac{n+1}{2^{n-1}}.$$

The remaining numbers w_k , for $0 < k < n+1$, we compute in the following way:

$$\begin{aligned} w_k &= |(t_k - t_0) \dots (t_k - t_{k-1})(t_k - t_{k+1}) \dots (t_k - t_{n+1})| = \lim_{t \rightarrow t_k} \frac{\prod_{j=0}^{n+1} (t - t_j)}{t - t_k} \\ &= \frac{1}{2^n} \sqrt{(1 - t_k^2) \lim_{t \rightarrow t_k} \frac{1 - \tau_{n+1}^2}{(t - t_k)^2}} = \frac{1}{2^n} \sqrt{(1 - t_k^2) \lim_{t \rightarrow t_k} \frac{-2\tau_{n+1}\tau'_{n+1}}{2(t - t_k)}} \\ &= \frac{1}{2^n} \sqrt{(1 - t_k^2)(-1)^{n-k} \lim_{t \rightarrow t_k} \frac{\tau'_{n+1}}{t - t_k}} = \frac{1}{2^n} \sqrt{(1 - t_k^2) |\tau''_{n+1}(t_k)|}. \end{aligned}$$

It follows from formula (6) that $(1 - t^2)\tau''_{n+1} = (n+1)^2(1 - \tau_{n+1}^2)$. Differentiating this equation and dividing the obtained equation by $2\tau'_{n+1}$ we get

$$(8) \quad (1 - t^2)\tau''_{n+1} - t\tau'_{n+1} = -(n+1)^2\tau_{n+1}.$$

This polynomial-type identity holds also in the case when $\tau'_{n+1}(t) = 0$. Since $\tau'_{n+1}(t_k) = 0$ for $0 < k < n+1$, we have

$$|\tau''_{n+1}(t_k)| = \frac{(n+1)^2}{1 - t_k^2}, \quad w_k = \frac{n+1}{2^n} \quad (0 < k < n+1).$$

From the formulas for w_0, w_1, \dots, w_{n+1} it follows that

$$\begin{aligned} (9) \quad e_n &= \frac{\frac{2^{n-1}}{n+1} \xi(t_0) + \sum_{k=1}^n (-1)^k \frac{2^n}{n+1} \xi(t_k) + (-1)^{n+1} \frac{2^{n-1}}{n+1} \xi(t_{n+1})}{\frac{2^{n-1}}{n+1} + n \frac{2^n}{n+1} + \frac{2^{n-1}}{n+1}} \\ &= \frac{1}{n+1} \left(\frac{1}{2} \xi(t_0) + \sum_{k=1}^n (-1)^k \xi(t_k) + \frac{(-1)^{n+1}}{2} \xi(t_{n+1}) \right), \end{aligned}$$

which was to be proved.

When applying estimation (5) we should compute the values of the function ξ at the points $t_{n+1,k}$ with great accuracy, so that the right-hand side of (5) will have least two or three significant digits. Theorems 5:6, 5:8 and 5:9 which are conclusions from the preceding theorem, are free of this disadvantage. We may use them for function ξ in a particular form.

5.4. For our further investigation it is essential that the Chebyshev polynomials τ_0, τ_1, \dots are orthogonal in the interval $\langle -1, 1 \rangle$ with the weight function $(1-t^2)^{-1/2}$. This follows from the equation

$$\int_{-1}^1 \tau_j(t) \tau_k(t) (1-t^2)^{-1/2} dt = \begin{cases} 0 & (j \neq k), \\ \pi & (j = k = 0), \\ \pi/2 & (j = k \neq 0). \end{cases}$$

For every function $\xi \in \mathcal{C}_{\langle -1, 1 \rangle}$ we may construct an orthogonal series

$$(10) \quad \sum_{i=0}^{\infty} a_i[\xi] \tau_i,$$

where

$$(11) \quad a_l[\xi] = \begin{cases} \frac{1}{\pi} \int_{-1}^1 \xi(t) (1-t^2)^{-1/2} dt & (l = 0), \\ \frac{2}{\pi} \int_{-1}^1 \xi(t) \tau_l(t) (1-t^2)^{-1/2} dt & (l > 0). \end{cases}$$

If the function ξ has a bounded first derivative in the interval $\langle -1, 1 \rangle$, then this series is uniformly convergent to the function ξ in this interval ([20], p. 245).

In the next theorem we shall use the method proposed by Hornecker in [14].

THEOREM 5:6. *If*

$$\xi = \sum_{i=0}^{\infty} a_i \tau_i \quad (t \in \langle -1, 1 \rangle),$$

then

$$(12) \quad \varepsilon_n(\xi; \langle -1, 1 \rangle) \geq \left| \sum_{i=0}^{\infty} a_{(2i+1)(n+1)} \right|.$$

Proof. First, let us compute the values of an arbitrary Chebyshev polynomial at the points t_{pk} (p positive integer, $0 \leq k \leq p$). Since

$$t_{pk} = -\cos \frac{k\pi}{p} = \cos \frac{(p-k)\pi}{p},$$

then if $l = ip + j$ (i —non-negative integer, $0 \leq j \leq p-1$) we get

$$\begin{aligned}\tau_l(t_{pk}) &= \cos \left(l \arccos \left(\cos \frac{(p-k)\pi}{p} \right) \right) \\ &= \cos \left(\frac{(ip+j)(p-k)\pi}{p} \right) = \cos \left(i(p-k)\pi + \frac{j(p-k)\pi}{p} \right) \\ &= \cos i(p-k)\pi \cos \frac{j(p-k)\pi}{p} = (-1)^{i(p-k)} \tau_j(t_{pk}).\end{aligned}$$

Now we can compute the value of the function ξ for $t = t_{pk}$:

$$\xi(t_{pk}) = \sum_{l=0}^{\infty} a_l \tau_l(t_{pk}) = \sum_{i=0}^{\infty} \sum_{j=0}^{p-1} (-1)^{i(p-k)} a_{ip+j} \tau_j(t_{pk}).$$

Let $p = n+1$, $q_k = \frac{1}{2}$ for $k=0$ and $k=p$, $q_k = 1$ for $0 < k < p$. Using these notations we write (9) in the form

$$e_n = \frac{1}{p} \sum_{k=0}^{\mu} (-1)^k q_k \xi(t_{pk})$$

and we compute that

$$e_n = \frac{1}{p} \sum_{i=0}^{\infty} \sum_{j=0}^{p-1} a_{2ip+j} \sum_{k=0}^{\mu} (-1)^k (-1)^{i(p-k)} q_k \tau_j(t_{pk}).$$

If i is even (odd, respectively), then the exponent $k+i(p-k)$ can be replaced by the exponent k (or p , respectively). Thus we split the above sum into two parts as follows with respect to the index i :

$$\begin{aligned}e_n &= \frac{1}{p} \sum_{i=0}^{\infty} \sum_{j=0}^{p-1} a_{2ip+j} \sum_{k=0}^{\mu} (-1)^k q_k \tau_j(t_{pk}) + \\ &\quad + \frac{(-1)^{\mu}}{p} \sum_{i=0}^{\infty} \sum_{j=0}^{p-1} a_{(2i+1)p+j} \sum_{k=0}^{\mu} q_k \tau_j(t_{pk}).\end{aligned}$$

The expression

$$\frac{1}{p} \sum_{k=0}^{\mu} (-1)^k q_k \tau_j(t_{pk})$$

is analogous as is e_n , but instead of function ξ we have here the polynomial τ_j . Thus for $j = 0, 1, \dots, p-1$ the absolute value of this expression equals the n -th error of the best approximation of a certain polynomial from the class \mathscr{W}_n on the set $\{t_{p0}, t_{p1}, \dots, t_{pp}\}$, and this error is obviously 0.

Thus

$$e_n = \frac{(-1)^n}{p} \sum_{i=0}^{\infty} \sum_{j=0}^{p-1} a_{(2i+1)p} \sum_{k=0}^p q_k \tau_j(t_{pk}).$$

Now we notice that

$$\begin{aligned} \tau_{p-j}(t_{pk}) &= \cos \frac{(p-j)(p-k)\pi}{p} = \cos \left((p-k)\pi - \frac{j(p-k)\pi}{p} \right) \\ &= \cos(p-k)\pi \cos \frac{j(p-k)\pi}{p} = (-1)^{p-k} \tau_j(t_{pk}). \end{aligned}$$

It follows that

$$\sum_{k=0}^p q_k \tau_j(t_{pk}) = (-1)^p \sum_{k=0}^p (-1)^k q_k \tau_{p-j}(t_{pk})$$

and by the reasons given above this sum is equal to zero for $p-j = 0, 1, \dots, p-1$; that is, for $j = 1, 2, \dots, p$. On the other hand, for $j = 0$ we have $\tau_j = 1$ and

$$\sum_{k=0}^p q_k \tau_j(t_{pk}) = \sum_{k=0}^p q_k = p.$$

Thus, finally

$$(13) \quad e_n = (-1)^n \sum_{i=0}^{\infty} a_{(2i+1)p} = (-1)^{n+1} \sum_{i=0}^{\infty} a_{(2i+1)(n+1)}.$$

Since $|e_n| = \varepsilon_n(\xi; \{t_{n+1,0}, t_{n+1,1}, \dots, t_{n+1,n+1}\})$, inequality (12) follows from Theorem 5:2.

Now we shall prove an auxiliary theorem which expresses the coefficients of the series (10) by the coefficients of the expansion of the function into power series.

THEOREM 5:7 (Linskiy, [17]). *If*

$$\xi = \sum_{l=0}^{\infty} x_l t^l \quad (t \in \langle -1, 1 \rangle),$$

then

$$a_l[\xi] = \begin{cases} \sum_{j=0}^{\infty} \frac{1}{2^{2j}} \binom{2j}{j} x_{2j} & (l=0), \\ \sum_{j=0}^{\infty} \frac{1}{2^{2j+l-1}} \binom{2j+l}{j} x_{2j+l} & (l=1, 2, \dots). \end{cases}$$

Proof. For $l = 0, 1, \dots$ we have

$$a_l[\xi] = \sum_{k=0}^{\infty} x_k a_l[t^k],$$

thus we should compute $a_l[t^k]$. First we shall do it for $l = 0$. From definition (11) it follows that

$$a_0[t^k] = \frac{1}{\pi} \int_{-1}^1 t^k (1-t^2)^{-1/2} dt.$$

For an odd k the function $t^k(1-t^2)^{-1/2}$ is odd and this integral equals zero. Thus we put $k = 2j$ and $t = \cos u$, which leads to the equation

$$a_0[\xi] = \sum_{j=0}^{\infty} x_{2j} a_0[t^{2j}] = \frac{1}{\pi} \sum_{j=0}^{\infty} x_{2j} \int_0^{\pi} \cos^{2j} u du.$$

Since

$$\cos^{2j} u = \frac{1}{2^{2j}} \left\{ \sum_{i=0}^{j-1} 2 \binom{2j}{i} \cos 2(j-i)u + \binom{2j}{j} \right\}$$

(see [32], p. 39, formula 1.320.5), and

$$\int_0^{\pi} \cos mu du = 0$$

for $m = 1, 2, \dots$, we have

$$a_0[\xi] = \frac{1}{\pi} \sum_{j=0}^{\infty} x_{2j} \int_0^{\pi} \frac{1}{2^{2j}} \binom{2j}{j} du,$$

which was to be proved.

Now we assume that $l > 0$. Then

$$a_l[t^k] = \frac{2}{\pi} \int_{-1}^1 t^k \tau_l(t) (1-t^2)^{-1/2} dt.$$

The polynomial τ_l is orthogonal with the weight function $(1-t^2)^{-1/2}$ to the polynomials $\tau_0, \tau_1, \dots, \tau_{l-1}$. The functions $1, t, \dots, t^{l-1}$ are linear combinations of these polynomials, thus the polynomial τ_l is orthogonal to these functions, i.e. $a_l[t^k] = 0$ for $k < l$. The same equation holds also for $k \geq l$, if the difference $k-l$ is odd. It follows from fact that by (3) from § 2, the polynomial τ_l is an even function for an even l and an odd function for an odd l . Thus we may restrict our considerations to

the values of indices k , equal to $2j+l$ for non-negative integer j . Putting $t = \cos u$ we get

$$a_l[t^k] = \frac{2}{\pi} \int_0^\pi \cos^k u \cos lu \, du.$$

For $k \geq 2$ we transform this integral, integrating twice by parts as follows:

$$\begin{aligned} \int_0^\pi \cos^k u \cos lu \, du &= \frac{k}{l} \int_0^\pi \sin u \cos^{k-1} u \sin lu \, du \\ &= \frac{k}{l^2} \int_0^\pi (\cos^k u - (k-1) \sin^2 u \cos^{k-2} u) \cos lu \, du \\ &= \frac{k^2}{l^2} \int_0^\pi \cos^k u \cos lu \, du - \frac{k(k-1)}{l^2} \int_0^\pi \cos^{k-2} u \cos lu \, du. \end{aligned}$$

Thus if $k \neq l$ we get

$$a_l[t^k] = \frac{k(k-1)}{(k-l)(k+l)} a_l[t^{k-2}].$$

Applying this equation for $k = 2j+l, 2j+l-2, \dots, l+2$ we get

$$\begin{aligned} a_l[t^{2j+l}] &= \frac{(2j+l)(2j+l-1)}{2j(2j+2l)} a_l[t^{2j+l-2}] = \dots \\ &= \frac{(2j+l)(2j+l-1)\dots(2+l)(1+l)}{(2j(2j-2)\dots 2)((2j+2l)(2j+2l-2)\dots(2l+2))} a_l[t^l] \\ &= \frac{1}{2^{2j-1}\pi} \binom{2j+l}{j} \int_0^\pi \cos^l u \cos lu \, du = \frac{1}{2^{2j-1}\pi} \binom{2j+l}{j} \int_{-1}^1 t^l \tau_l(t) (1-t^2)^{-1/2} dt. \end{aligned}$$

The function t^l is the sum of the term $2^{-l+1}\tau_l$ and of the linear combination of the Chebyshev polynomials $\tau_{l-2}, \tau_{l-4}, \dots$ which follows from formula (3), § 2. Thus

$$\int_{-1}^1 t^l \tau_l(t) (1-t^2)^{-1/2} dt = 2^{-l}\pi,$$

$$a_l[t^{2j+l}] = \frac{1}{2^{2j+l-1}} \binom{2j+l}{j}.$$

We have already established that the coefficients $a_l[t^k]$ for $k = 0, 1, \dots, l-1, l+1, l+3, \dots$ are equal to 0, thus for $l > 0$

$$a_l[\xi] = \sum_{j=0}^{\infty} x_{2j+l} a_l[t^{2j+l}] = \sum_{j=0}^{\infty} \frac{1}{2^{2j+l-1}} \binom{2j+l}{j} x_{2j+l},$$

which completes the proof of Theorem 5:7.

THEOREM 5:8 (Paszkowski, [25]). *If*

$$\xi = \sum_{l=0}^{\infty} x_l t^l \quad (t \in \langle -1, 1 \rangle),$$

then

$$\varepsilon_n(\xi; \langle -1, 1 \rangle) \geq \sum_{m=0}^{\infty} a_{nm} x_{n+2m+1},$$

where

$$a_{nm} = \frac{1}{2^{n+2m}} \sum_{i=0}^{[m/(n+1)]} \binom{n+2m+1}{m-i(n+1)} \quad (m = 0, 1, \dots).$$

Proof. We can apply Theorem 5:6 to the function ξ , changing in inequality (12) the coefficients a_l by the expressions, obtained from Theorem 5:7:

$$\begin{aligned} \sum_{i=0}^{\infty} a_{(2i+1)(n+1)} &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{2^{2j+(2i+1)(n+1)-1}} \binom{2j+(2i+1)(n+1)}{j} x_{2j+(2i+1)(n+1)} \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{1}{2^{n+2m}} \binom{n+2m+1}{j} x_{n+2m+1}, \end{aligned}$$

where $m = j + i(n+1)$. Instead of adding with respect to j we add with respect to m , also from 0 to ∞ . Then—since $i \geq 0$ and $m \geq i(n+1)$ —the index i will change from 0 to $m/(n+1)$, or, more precisely, to the greatest integer not exceeding this number, and

$$\begin{aligned} (14) \quad \sum_{i=0}^{\infty} a_{(2i+1)(n+1)} &= \sum_{m=0}^{\infty} \sum_{i=0}^{[m/(n+1)]} \frac{1}{2^{n+2m}} \binom{n+2m+1}{m-i(n+1)} x_{n+2m+1} \\ &= \sum_{m=0}^{\infty} a_{nm} x_{n+2m+1}. \end{aligned}$$

Combining the last result with (12) we get the assertion of Theorem 5:8.

THEOREM 5:9 (Paszkowski, [25]). Let $\xi = \sum_{l=0}^{r-1} x_l t^l + \xi_r t^r$, where $r \geq n+1$ and $\xi_r \in \mathcal{C}_{(-1,1)}$, and let

$$x_r^{\min} \leq \xi_r(t) + (-1)^{n+r+1} \xi_r(-t) \leq x_r^{\max} \quad (0 \leq t \leq 1),$$

$$f_{nr} = \frac{1}{n+1} \left(\sum_{0 \leq j < (n+1)/4} t_{n+1, n-2j+1}^r - \frac{1}{2} \right),$$

$$g_{nr} = \frac{1}{n+1} \sum_{0 \leq j < (n-1)/4} t_{n+1, n-2j}^r,$$

$$a_{nr}^{\min} = x_r^{\min} f_{nr} - x_r^{\max} g_{nr},$$

$$a_{nr}^{\max} = x_r^{\max} f_{nr} - x_r^{\min} g_{nr}.$$

If the numbers

$$e_{nr}^{\min} = \sum_{m=0}^{[(r-n)/2]-1} a_{nm} x_{n+2m+1} + a_{nr}^{\min}, \quad e_{nr}^{\max} = \sum_{m=0}^{[(r-n)/2]-1} a_{nm} x_{n+2m+1} + a_{nr}^{\max}$$

have a common sign, then

$$\varepsilon_n(\xi; \langle -1, 1 \rangle) \geq \min\{|e_{nr}^{\min}|, |e_{nr}^{\max}|\}.$$

Proof. Let $\xi^* = \sum_{l=0}^{\infty} x_l^* t^l$, where $x_l^* = x_l$ for $l \leq r-1$, $x_l^* = 0$ for the remaining l . Let us denote by e_n^* the expression analogous to (9) which corresponds to the function ξ^* . Applying formulas (13) and (14) to this function we get

$$e_n^* = (-1)^{n+1} \sum_{m=0}^{\infty} a_{nm} x_{n+2m+1}^*.$$

By the definition of coefficients x_l^* we may restrict the summation to the non-negative values of m , such that $n+2m+1 \leq r-1$, i.e. $0 \leq m \leq (r-n)/2-1$. Hence

$$(-1)^{n+1} e_n^* = \sum_{m=0}^{[(r-n)/2]-1} a_{nm} x_{n+2m+1}^*.$$

Let $\xi^{**} = \xi_r t^r$. Let us denote by e_n^{**} the expression analogous to (9) which corresponds to the function ξ^{**} . Thus we have

$$\begin{aligned} & (-1)^{n+1} e_n^{**} \\ &= \frac{1}{n+1} \left(\frac{(-1)^{n+1}}{2} \xi_r(t_0) t_0^r + (-1)^n \xi_r(t_1) t_1^r + \dots - \xi_r(t_n) t_n^r + \frac{1}{2} \xi_r(t_{n+1}) t_{n+1}^r \right), \end{aligned}$$

where $t_k \equiv t_{n+1,k}$. We shall evaluate this expression. If $t_k < 0$, then $t_{n+1-k} = -t_k > 0$ and

$$\begin{aligned} & (-1)^{n+1-k} \xi_r(t_k) t_k^r + (-1)^k \xi_r(t_{n+1-k}) t_{n+1-k}^r \\ & = (-1)^k (\xi_r(t_{n+1-k}) + (-1)^{n+r+1} \xi_r(-t_{n+1-k})) t_{n+1-k}^r. \end{aligned}$$

For an even value of k this expression may be estimated from below and from above by $x_r^{\min} t_{n+1-k}^r$ and $x_r^{\max} t_{n+1-k}^r$, respectively. For an odd k it is estimated by $-x_r^{\max} t_{n+1-k}^r$ and $-x_r^{\min} t_{n+1-k}^r$, respectively. It follows that the number $(-1)^{n+1} e_n^{**}$ may be estimated from below and from above by numbers:

$$\begin{aligned} & \frac{1}{n+1} x_r^{\min} (\frac{1}{2} t_{n+1}^r + t_{n-1}^r + \dots) - \frac{1}{n+1} x_r^{\max} (t_n^r + t_{n-2}^r + \dots), \\ & \frac{1}{n+1} x_r^{\max} (\frac{1}{2} t_{n+1}^r + t_{n-1}^r + \dots) - \frac{1}{n+1} x_r^{\min} (t_n^r + t_{n-2}^r + \dots), \end{aligned}$$

where only the powers of positive numbers t_l are added, i.e. the numbers with the indices $l > (n+1)/2$. Since $t_{n+1} = 1$, the sums of powers of these numbers may be written in the form

$$\frac{1}{2} t_{n+1}^r + t_{n-1}^r + \dots = \sum_{0 \leq j < (n+1)/4} t_{n-2j+1}^r - \frac{1}{2}, \quad t_n^r + t_{n-2}^r + \dots = \sum_{0 \leq j < (n-1)/4} t_{n-2j}^r.$$

Thus we have proved that

$$a_{nr}^{\min} = x_r^{\min} f_{nr} - x_r^{\max} g_{nr} \leq (-1)^{n+1} e_n^{**} \leq x_r^{\max} f_{nr} - x_r^{\min} g_{nr} = a_{nr}^{\max}.$$

Since $\xi = \xi^* + \xi^{**}$, and expression (9) is linear with respect to the function ξ , then $e_n = e_n^* + e_n^{**}$. From the formula for $(-1)^{n+1} e_n^*$ and from the estimate for $(-1)^{n+1} e_n^{**}$ obtained above, it follows that

$$e_{nr}^{\min} \leq (-1)^{n+1} e_n \leq e_{nr}^{\max}.$$

If the numbers e_{nr}^{\min} and e_{nr}^{\max} have common signs, then this inequality gives us a non-trivial estimation from below for $|e_n|$ by $\min\{|e_{nr}^{\min}|, |e_{nr}^{\max}|\}$ and such an estimation from below for the error $\varepsilon_n(\xi; \langle -1, 1 \rangle)$.

Theorems 5:5, 5:6, 5:8 and 5:9 give the estimation of the n -th error of the best approximation of the function ξ in the interval $\langle -1, 1 \rangle$. Using Theorem 4:4 it is easy to obtain the analogous results for the functions $\xi \in \mathcal{C}_{\langle a, b \rangle}$. In fact, if we take $F = \langle a, b \rangle$, $F^* = \langle -1, 1 \rangle$ and we consider the fact that the interval F^* may be transformed into F by the formula $t = \frac{1}{2}((b-a)t^* + a + b)$ ($t^* \in F^*$), we get

$$\varepsilon_n(\xi; \langle a, b \rangle) = \varepsilon_n(\xi^*; \langle -1, 1 \rangle) = \varepsilon_n\left(\xi\left(\frac{(b-a)t + a + b}{2}\right); \langle -1, 1 \rangle\right),$$

thus reducing the problem of estimating the error of the best approximation to the "standard" case of the interval $\langle -1, 1 \rangle$.

6. Estimation of the error of the best approximation for functions which are differentiable several times.

6.1. In this section we shall estimate the n -th error of the best approximation of the function in the interval $I = \langle a, b \rangle$, assuming that the function is defined and differentiable $n+1$ times in this interval.

THEOREM 6:1 (Bernstein, [4], p. 46). *If the functions ξ and η have the $(n+1)$ -st derivatives in the interval I , and if*

$$\xi^{(n+1)}(t) \geq |\eta^{(n+1)}(t)| \quad (t \in I) \quad \text{OR} \quad -\xi^{(n+1)}(t) \geq |\eta^{(n+1)}(t)| \quad (t \in I),$$

then we have the inequality

$$(1) \quad \varepsilon_n(\xi) \geq \varepsilon_n(\eta).$$

Proof. We notice easily that we may restrict our consideration to the case when $\xi^{(n+1)}(t) \geq |\eta^{(n+1)}(t)|$. Otherwise, if the function $\xi^{(n+1)}$ were not positive, we could have considered the function $-\xi$ instead of ξ , and use the equation $\varepsilon_n(-\xi) = \varepsilon_n(\xi)$, which follows from Theorem 4:2.

It is also easy to notice that it suffices to prove inequality (1) in the case when $\xi^{(n+1)}(t) > |\eta^{(n+1)}(t)|$. In fact, if $\xi^{(n+1)}(t) \geq |\eta^{(n+1)}(t)|$ then $\xi^{(n+1)}(t) + h > |\eta^{(n+1)}(t)|$ for an arbitrary $h > 0$ and it would follow from the assertion which we are being to prove, that

$$(2) \quad \varepsilon_n \left(\xi + \frac{ht^{n+1}}{(n+1)!} \right) \geq \varepsilon_n(\eta).$$

However, by Theorems 4:6 and 4:2 we have

$$\varepsilon_n \left(\xi + \frac{ht^{n+1}}{(n+1)!} \right) \leq \varepsilon_n(\xi) + \varepsilon_n \left(\frac{ht^{n+1}}{(n+1)!} \right) = \varepsilon_n(\xi) + \frac{h}{(n+1)!} \varepsilon_n(t^{n+1}),$$

thus, passing to the limit as $h \rightarrow 0$ in inequality (2), we get the inequality (1), assuming the condition $\xi^{(n+1)}(t) \geq |\eta^{(n+1)}(t)|$.

Let us assume, then, that $\xi^{(n+1)}(t) > |\eta^{(n+1)}(t)|$ for $t \in I$, and let us suppose further, that $\varepsilon_n(\xi) < \varepsilon_n(\eta)$. Let $[\xi, n | \omega_n]$, $[\eta, n | \psi_n]$, i.e. let the polynomials ω_n and ψ_n be the n -th best polynomials for the functions ξ and η , respectively. By the definition of these polynomials we have $\|\xi - \omega_n\|_I = \varepsilon_n(\xi)$, $\|\eta - \psi_n\|_I = \varepsilon_n(\eta)$.

Let us denote by v_0, v_1, \dots, v_{n+1} (where $a \leq v_0 < v_1 < \dots < v_{n+1} \leq b$) the points which are alternately the $(n, +)$ -points and $(n, -)$ -points of the function η , i.e. such points that

$$\eta(v_k) - \psi_n(v_k) = (-1)^k c \varepsilon_n(\eta) \quad (k = 0, 1, \dots, n+1),$$

where $|c| = 1$. Further let $\delta_- = (\eta - \psi_n) - (\xi - \omega_n)$, $\delta_+ = (\eta - \psi_n) + (\xi - \omega_n)$. The functions δ_- and δ_+ have, for $\varepsilon_n(\xi) < \varepsilon_n(\eta)$, the same sign in each of the points v_k as the difference $\eta - \psi_n$:

$$(3) \quad \text{sign } \delta_-(v_k) = \text{sign } \delta_+(v_k) = (-1)^k c \quad (k = 0, 1, \dots, n+1).$$

It follows that each of the functions δ_- , δ_+ has one zero in each of the $n+1$ intervals $(v_0, v_1), (v_1, v_2), \dots, (v_n, v_{n+1})$. On the other hand, we know that $\delta_-^{(n+1)} = \eta^{(n+1)} - \xi^{(n+1)}$, $\delta_+^{(n+1)} = \eta^{(n+1)} + \xi^{(n+1)}$ and in view of the assumption $\xi^{(n+1)}(t) > |\eta^{(n+1)}(t)|$ the function $\delta_-^{(n+1)}$ is negative, and the function $\delta_+^{(n+1)}$ is positive in the interval I . From Rolle's theorem applied to the functions δ_- and δ_+ and their derivatives, it follows that δ_- and δ_+ have exactly one zero in each of the intervals mentioned, and have no zeros outside.

Let us denote the greatest zero of the function δ_- by $z^{(0)}$ and the greatest zero of its k -th derivative by $z^{(k)}$ ($k = 1, 2, \dots, n$). The derivative $\delta_-^{(k)}$ has only $n+1-k$ zeros, which separate the zeros of the derivative of the lower order. The existence of these zeros follows from Rolle's theorem, i. e. from the existence of $n+1$ single zeros of the function δ_- . Thus $z^{(n)} < \dots < z^{(1)} < z^{(0)} < v_{n+1}$. In the interval $(z^{(0)}, b)$ the function δ_- , together with its first $n+1$ derivatives is negative. Indeed, since $\delta_-^{(n+1)}(t) < 0$, the function $\delta_-^{(n)}$ decreases in the interval I . It has the single zero $z^{(n)}$ and must be negative to the right of it. The function $\delta_-^{(n-1)}$ decreases in this domain, thus it is negative to the right of its greatest zero.

By arguments similar to those used previously in this section, the statements in the preceding paragraph may be proved. In the same way, starting from the inequality $\delta_+^{(n+1)}(t) > 0$, we prove that the function δ_+ is positive in the interval which contains the point v_{n+1} . However, it follows from (3) that the functions δ_- and δ_+ have the same sign at the point v_{n+1} . The contradiction obtained shows that the assumption $\varepsilon_n(\xi) < \varepsilon_n(\eta)$ is false.

In practice, we apply Theorem 6:1 to one of the functions ξ , η only when it is easy to compute the n -th error of the best approximation for the other one, and when the $(n+1)$ -st derivative of this function has a relatively simple form. As an example, let us consider the function $\xi = pt^{n+1}/(n+1)!$ in the interval $I = \langle a, b \rangle$. Putting $F = I$, $F^* = \langle -1, 1 \rangle$ and considering the fact that the interval F^* may be trans-

formed into F by the formula $t = \frac{1}{2}((b-a)t^* + a + b)$, we get from Theorem 4:4:

$$\begin{aligned} \varepsilon_n(\zeta; I) &= \varepsilon_n(\zeta; F) = \varepsilon_n\left(\zeta\left(\frac{(b-a)t^* + a + b}{2}\right); F^*\right) \\ &= \varepsilon_n\left(\frac{p}{(n+1)!} \left(\frac{(b-a)t + a + b}{2}\right)^{n+1}; \langle -1, 1 \rangle\right). \end{aligned}$$

The function

$$\frac{p}{(n+1)!} \left(\frac{(b-a)t + a + b}{2}\right)^{n+1}$$

is a sum of the expression

$$\frac{p(b-a)^{n+1}t^{n+1}}{2^{n+1}(n+1)!}$$

and some polynomial φ of the degree n , thus it follows from Theorems 4:3, 4:2 and 2:2, applied successively, that:

$$\begin{aligned} \varepsilon_n(\zeta; I) &= \varepsilon_n\left(\frac{p(b-a)^{n+1}t^{n+1}}{2^{n+1}(n+1)!}; \langle -1, 1 \rangle\right) \\ &= \frac{p(b-a)^{n+1}}{2^{n+1}(n+1)!} \varepsilon_n(t^{n+1}; \langle -1, 1 \rangle) = \frac{p(b-a)^{n+1}}{2^{2n+1}(n+1)!}. \end{aligned}$$

Since the $(n+1)$ -st derivative of the function ζ is constant, equal to p , and we have succeeded in computing its n -th error of the best approximation we may use the functions of this type in Theorem 6:1 for estimating the errors of approximation for other functions. Thus we get:

THEOREM 6:2 (Bernstein, [4], p. 47). *If the $(n+1)$ -st derivative of the function ξ satisfies in the interval $I = \langle a, b \rangle$ the inequality*

$$0 \leq p \leq \xi^{(n+1)}(t) \leq q \quad \text{or} \quad 0 \leq p \leq -\xi^{(n+1)}(t) \leq q,$$

where p and q are constants, then

$$(4) \quad \frac{p(b-a)^{n+1}}{2^{2n+1}(n+1)!} \leq \varepsilon_n(\xi) \leq \frac{q(b-a)^{n+1}}{2^{2n+1}(n+1)!}.$$

The left-hand side of the inequality (4) follows from Theorem 6:1 for $\eta = pt^{n+1}/(n+1)!$, and the right-hand side follows from the same theorem, after changing ξ and η , for $\eta = qt^{n+1}/(n+1)!$.

6.2. If the $(n+1)$ -st derivative of the function ξ equals zero even at one point the interval I , then the estimate from below of the number

$\varepsilon_n(\xi)$ which follow from Theorem 6:2 gives only the obvious inequality $\varepsilon_n(\xi) \geq 0$. In this case it is sometimes possible to use the fact that $\varepsilon_n(\xi; I) \geq \varepsilon_n(\xi; J)$ (see (1), § 8) for such an interval $J \subset I$ in which the function $\xi^{(n+1)}$ does not vanish, and apply Theorem 6:2 to the interval J . However, in this manner we usually obtain rather weak estimations. Sometimes we can use the following generalization of Theorem 6:1:

THEOREM 6:3. *Let the functions ξ and η have the $(n+1)$ -st derivatives in the interval I . Let m be a positive integer, and let $c_1, c_2, \dots, c_m, d_1, d_2, \dots, d_m, p_1, p_2, \dots, p_m$ be the numbers such that*

$$(5) \quad c_i t + d_i \in I \quad (t \in I, i = 1, 2, \dots, m),$$

$$(6) \quad \sum_{i=1}^m |p_i| = 1.$$

If one of the following two inequalities

$$(7) \quad \begin{aligned} \sum_{i=1}^m p_i c_i^{n+1} \xi^{(n+1)}(c_i t + d_i) &\geq |\eta^{(n+1)}(t)| \quad (t \in I), \\ - \sum_{i=1}^m p_i c_i^{n+1} \xi^{(n+1)}(c_i t + d_i) &\geq |\eta^{(n+1)}(t)| \quad (t \in I) \end{aligned}$$

is satisfied, then $\varepsilon_n(\xi) \geq \varepsilon_n(\eta)$.

The proof is similar to that of Theorem 6:1. It is reduced to showing that if

$$(8) \quad \sum_{i=1}^m p_i c_i^{n+1} \xi^{(n+1)}(c_i t + d_i) > |\eta^{(n+1)}(t)| \quad (t \in I),$$

then the assumption $\varepsilon_n(\xi) < \varepsilon_n(\eta)$ leads to a contradiction.

Let $[\xi, n | \omega_n]$. It follows from the definition of the polynomial ω_n that $|\xi(t) - \omega_n(t)| \leq \varepsilon_n(\xi)$ for $t \in I$. According to (5) we have also:

$$|\xi(c_i t + d_i) - \omega_n(c_i t + d_i)| \leq \varepsilon_n(\xi) \quad (t \in I, i = 1, 2, \dots, m).$$

If we multiply both sides of the i -th of these inequalities by $|p_i|$ and add all these inequalities using (6), we get

$$(9) \quad \left| \sum_{i=1}^m p_i \xi(c_i t + d_i) - \sum_{i=1}^m p_i \omega_n(c_i t + d_i) \right| \leq \varepsilon_n(\xi).$$

Let

$$\xi^* = \sum_{i=1}^m p_i \xi(c_i t + d_i), \quad \omega_n^* = \sum_{i=1}^m p_i \omega_n(c_i t + d_i),$$

where, of course, $\omega_n^* \in \mathscr{W}_n$. Then it would follow from (9) that $\|\xi^* - \omega_n^*\|_I \leq \varepsilon_n(\xi) < \varepsilon_n(\eta)$. Now we use inequality (8), whose left-hand side equals

the $(n+1)$ -st derivative of the function $\xi^* - \omega_n^*$ for considering the functions $\delta_- = (\eta - \psi_n) - (\xi^* - \omega_n^*)$, $\delta_+ = (\eta - \psi_n) + (\xi^* - \omega_n^*)$, where $[\eta, n | \psi_n]$. The contradiction at which we were to arrive consists of the fact that the functions δ_- and δ_+ would have, at the last (n) -point v_{n+1} of the function η , the same and different signs, simultaneously.

Now we present one of the particular cases of Theorem 6:3. Let $c_1 = 1$, $d_1 = 0$, $c_2 = -1$, $d_2 = a + b$, $p_1 = p_2 = \frac{1}{2}$. Then the left-hand side of the first of inequalities (7) is the $(n+1)$ -st derivative of the function

$$\xi_{\text{even}}(t) = \frac{1}{2}(\xi(t) + \xi(a+b-t)),$$

i.e. of the even (with respect to the centre of the interval $I = \langle a, b \rangle$) part of the function ξ . If, however, $p_2 = -\frac{1}{2}$, then, with the remaining parameters being the same as above, the left-hand side of inequality (7) is the $(n+1)$ -st derivative of the function

$$\xi_{\text{odd}}(t) = \frac{1}{2}(\xi(t) - \xi(a+b-t)),$$

i.e. the odd part of the function ξ with respect to the centre of the interval I . Thus the estimation $\varepsilon_n(\xi) \geq \varepsilon_n(\eta)$ may be applied to the function ξ not only in the case when its $(n+1)$ -st derivative has the constant sign in the interval I , but under a weaker assumption: if it has a constant sign for the even or odd part of the considered function.

7. Estimation of the $(n+1)$ -st error of the best approximation based upon the knowledge of the n -th best polynomial.

7.1. Now we shall show how the knowledge of the n -th best polynomial of the function ξ in the interval I , hence the knowledge of the n -th error of approximation $\varepsilon_n(\xi)$, permits, in certain special cases, an estimation from below the ratio $\varepsilon_{n+1}(\xi)/\varepsilon_n(\xi)$; i.e., estimate the number $\varepsilon_{n+1}(\xi)$ from below. The corresponding results, which incidentally are of more theoretical than practical interest, are given in [27]; here we shall present them in slightly simplified form. They are based on the estimation of the measure of the set

$$A_h = \{t: t \in I, |\xi(t) - \omega_n(t)| \geq h\varepsilon_n(\xi)\} \quad (0 \leq h \leq 1),$$

which can be computed, provided we know the n -th best polynomial ω_n for the function ξ in the interval I . In § 7 we shall assume throughout that $\varepsilon_n(\xi) > 0$, for otherwise we would have $\varepsilon_{n+1}(\xi) = 0$ also, and the estimation would not have been necessary.

THEOREM 7:1. *If m is the smallest integer such that $\varepsilon_{n+m}(\xi) < \varepsilon_n(\xi)$, then*

$$(1) \quad |A_h| \leq (b-a) \sqrt{\frac{1-h+2\delta_n(\xi)}{1+\delta_n(\xi)}} \quad (h \geq \delta_n(\xi)),$$

where $\delta_n(\xi) = \varepsilon_{n+m}(\xi)/\varepsilon_n(\xi)$.

The existence of the number n , mentioned in the theorem follows from the assumption $\varepsilon_n(\xi) > 0$ and from Theorem 4:1 of Weierstrass.

Proof. By the definition, the $(n+m)$ -th best polynomial ω_{n+m} for the function ξ satisfies the inequality $|\xi(t) - \omega_{n+m}(t)| \leq \varepsilon_{n+m}(\xi)$ for $t \in I$. Thus, it follows that:

$$\varepsilon_{n+m}(\xi) \geq |\xi(t) - \omega_n(t)| - |\omega_n(t) - \omega_{n+m}(t)| \quad (t \in I).$$

If we replaced the expression $|\xi(t) - \omega_n(t)|$ by the sum $|\omega_n(t) - \omega_{n+m}(t)| + \varepsilon_{n+m}(\xi)$ in the definition of the set A_h , the set so obtained would contain the set A_h , i.e.

$$|A_h| \leq \{t: t \in I, |\omega_n(t) - \omega_{n+m}(t)| \geq h\varepsilon_n(\xi) - \varepsilon_{n+m}(\xi)\}.$$

We assume that $\varepsilon_n(\xi) = \varepsilon_{n+1}(\xi) = \dots = \varepsilon_{n+m-1}(\xi)$. This means that the polynomials $\omega_n, \omega_{n+1}, \dots, \omega_{n+m-1}$ that is, the best polynomials for the function ξ are identical; hence in particular $\omega_n - \omega_{n+m} = \omega_{n+m-1} - \omega_{n+m}$. The properties of the difference of pairs of successive best polynomials have been investigated in § 3.1. We know from these investigations that the polynomial $\omega_{n+m-1} - \omega_{n+m}$ of the $(n+m)$ -th degree has only the real and single zeros $z_1 < z_2 < \dots < z_{n+m}$ which lie inside the interval $I = \langle a, b \rangle$ (Theorem 3:1 with n changed to $n+m-1$). It has also been shown in § 3.1 that

$$(2) \quad \varepsilon_n(\xi) - \varepsilon_{n+m}(\xi) \leq \|\omega_n - \omega_{n+m}\|_{\langle z_k, z_{k+1} \rangle} \leq \varepsilon_n(\xi) + \varepsilon_{n+m}(\xi) \\ (k = 0, 1, \dots, n+m),$$

where $z_0 = a, z_{n+m+1} = b$ (Theorem 3:2). The left-hand side of the above inequality follows from (3), § 3, and the right-hand side—in view of the relation $\langle z_k, z_{k+1} \rangle \subset I$ —from (2), § 3.

Considering the mentioned conclusions from § 3 we may write that

$$|A_h| \leq \sum_{k=0}^{n+m} \left\{ t: t \in \langle z_k, z_{k+1} \rangle, |\omega_n(t) - \omega_{n+m}(t)| \geq h\varepsilon_n(\xi) - \varepsilon_{n+m}(\xi) \right\} \\ \leq \sum_{k=0}^{n+m} \left\{ t: t \in \langle z_k, z_{k+1} \rangle, \right. \\ \left. |\omega_n(t) - \omega_{n+m}(t)| \geq \frac{h\varepsilon_n(\xi) - \varepsilon_{n+m}(\xi)}{\varepsilon_n(\xi) + \varepsilon_{n+m}(\xi)} \|\omega_n - \omega_{n+m}\|_{\langle z_k, z_{k+1} \rangle} \right\} \\ = \sum_{k=0}^{n+m} \left\{ t: t \in \langle z_k, z_{k+1} \rangle, \right. \\ \left. |\omega_n(t) - \omega_{n+m}(t)| \geq \frac{h - \delta_n(\xi)}{1 + \delta_n(\xi)} \|\omega_n - \omega_{n+m}\|_{\langle z_k, z_{k+1} \rangle} \right\}.$$

This equation is interesting only in the case when $h \geq \delta_n(\xi)$, since otherwise the sum on the right-hand side equals $b - a$, and the estimation $|A_h| \leq b - a$ is valid for every h .

In paper [13] Erdős has proved (see also [26]) that any polynomial χ with only real zeros satisfies the inequality

$$(3) \quad \{t: t \in \langle c, d \rangle, |\chi(t)| \geq h \|\chi\|_{\langle c, d \rangle}\} \leq (d - c) \sqrt{1 - h} \quad (0 \leq h \leq 1),$$

where c and d are arbitrary successive zeros of this polynomial. The same inequality is also true when c is the greatest zero of the polynomial χ and d is an arbitrary number greater than c . In fact, the function $|\chi(t)|$ is convex for $t \geq c$, i.e. the straight line which passes through points $(c, 0)$ and $(d, |\chi(d)|)$ lies above the graph of this function in the interval $\langle c, d \rangle$. Thus, if we replace in (3) the symbol χ by the symbol λ of a linear function, with the graph passing through the mentioned points, then the measure of the set which appears in (3) will increase. This new measure, connected with the function λ equals simply $(d - c)(1 - h)$, i.e. it does not exceed $(d - c)\sqrt{1 - h}$. We notice similarly that inequality (3) is valid also in the case when d is the smallest zero of the polynomial χ and c is an arbitrary number smaller than d .

All these remarks show that we can apply inequality (3) to the polynomial $\chi = \omega_n - \omega_{n+m}$ and to the intervals $\langle c, d \rangle = \langle z_k, z_{k+1} \rangle$ for $k = 0, 1, \dots, n + m$. Thus we get

$$\left\{ t: t \in \langle z_k, z_{k+1} \rangle, |\omega_n(t) - \omega_{n+m}(t)| \geq \frac{h - \delta_n(\xi)}{1 + \delta_n(\xi)} \|\omega_n - \omega_{n+m}\|_{\langle z_k, z_{k+1} \rangle} \right\} \\ \leq (z_{k+1} - z_k) \sqrt{1 - \frac{h - \delta_n(\xi)}{1 + \delta_n(\xi)}} = (z_{k+1} - z_k) \sqrt{\frac{1 - h + 2\delta_n(\xi)}{1 + \delta_n(\xi)}} \\ (h \geq \delta_n(\xi); k = 0, 1, \dots, n + m),$$

$$|A_h| \leq \sum_{k=0}^{n+m} (z_{k+1} - z_k) \sqrt{\frac{1 - h + 2\delta_n(\xi)}{1 + \delta_n(\xi)}} \quad (h \geq \delta_n(\xi)),$$

which was to be proved.

In Theorem 7:1 we estimated the measure of the set A_h from above. The estimation of this measure from below is more complicated; we shall deal with it in the next theorem. The symbol $\kappa_p(h)$ for $h \geq 0$ will denote the greatest root of the equation $|\tau_p(t)| = h$, where τ_p is the p -th Chebyshev polynomial. The greatest zero of this polynomial equals $\cos \pi/2p$, hence $\kappa_p(h) \geq \cos \pi/2p$. In addition, let

$$\mu_p(h) = \{t: t \in \langle 0, 1 \rangle, t^{p-1}(1-t) \geq h(p-1)^{p-1} p^{-p}\} \quad (0 \leq h \leq 1).$$

Since $(p-1)^{p-1}p^{-p} = \|t^{p-1}(1-t)\|_{\langle 0,1 \rangle}$, the function μ_p is a measure similar to the measure of the set A_h , except that it is connected with a different function and a different interval. The function μ_p decreases from 1 to 0 as h increases from 0 to 1.

THEOREM 7:2. *If $\delta_n(\xi) \leq \frac{1}{2}$ then we have the inequality*

$$(4) \quad |A_h| \geq (b-a) \frac{\mu_{n+m} \left(\frac{h + \delta_n(\xi)}{1 - \delta_n(\xi)} \right) \cos \frac{\pi}{2(n+m)}}{\chi_{n+m} \left(\frac{h + \delta_n(\xi)}{1 - \delta_n(\xi)} \right)} \quad (h \leq 1 - 2\delta_n(\xi)),$$

where the symbols $\delta_n(\xi)$ and m have the same meaning as in Theorem 7:1.

Proof. From the inequality $|\xi(t) - \omega_{n+m}(t)| \leq \varepsilon_{n+m}(\xi)$ it follows that

$$\varepsilon_{n+m}(\xi) \geq |\omega_n(t) - \omega_{n+m}(t)| - |\xi(t) - \omega_n(t)| \quad (t \in I).$$

Thus, if we replace $|\xi(t) - \omega_n(t)|$ by the difference $|\omega_n(t) - \omega_{n+m}(t)| - \varepsilon_{n+m}(\xi)$ in the definition of the set A_h , the set obtained would be contained in A_h , i.e.:

$$\begin{aligned} |A_h| &\geq |\{t: t \in I, |\omega_n(t) - \omega_{n+m}(t)| \geq h\varepsilon_n(\xi) + \varepsilon_{n+m}(\xi)\}| \\ &= \sum_{k=0}^{n+m} |\{t: t \in \langle z_k, z_{k+1} \rangle, |\omega_n(t) - \omega_{n+m}(t)| \geq h\varepsilon_n(\xi) + \varepsilon_{n+m}(\xi)\}|. \end{aligned}$$

If we use now the left-hand side of inequality (2), we can write that

$$(5) \quad |A_h| \geq \sum_{k=1}^{n+m-1} |\{t: t \in \langle z_k, z_{k+1} \rangle, |\omega_n(t) - \omega_{n+m}(t)| \geq g \|\omega_n - \omega_{n+m}\|_{\langle z_k, z_{k+1} \rangle}\}| + |\{t: t \in \langle a, z_1 \rangle, |\omega_n(t) - \omega_{n+m}(t)| \geq g \min_{1 \leq k \leq n+m-1} \|\omega_n - \omega_{n+m}\|_{\langle z_k, z_{k+1} \rangle}\}| + |\{t: t \in \langle z_{n+m}, b \rangle, |\omega_n(t) - \omega_{n+m}(t)| \geq g \min_{1 \leq k \leq n+m-1} \|\omega_n - \omega_{n+m}\|_{\langle z_k, z_{k+1} \rangle}\}|,$$

where

$$g = \frac{h\varepsilon_n(\xi) + \varepsilon_{n+m}(\xi)}{\varepsilon_n(\xi) - \varepsilon_{n+m}(\xi)} = \frac{h + \delta_n(\xi)}{1 - \delta_n(\xi)}.$$

It has been proved in [26] that

$$|\{t: t \in \langle c, d \rangle, |\chi(t)| \geq h \|\chi\|_{\langle c, d \rangle}\}| \geq (d-c) \mu_p(h) \quad (0 \leq h \leq 1),$$

where χ is an arbitrary polynomial of the degree p with only real zeros, and c and d are its arbitrary successive zeros. We can apply this theorem to the polynomial $\chi = \omega_n - \omega_{n+m}$ of the degree $p = n+m$ and to the

intervals $\langle c, d \rangle = \langle z_k, z_{k+1} \rangle$ ($k = 1, 2, \dots, n+m-1$). Adding all these inequalities with respect to k we get

$$(6) \quad \sum_{k=1}^{n+m-1} \{t: t \in \langle z_k, z_{k+1} \rangle, |\omega_n(t) - \omega_{n+m}(t)| \geq g \|\omega_n - \omega_{n+m}\|_{\langle z_k, z_{k+1} \rangle}\} \\ \geq (z_{n+m} - z_1) \mu_{n+m}(g).$$

To estimate two other terms of the right-hand side of inequality (5) we shall use a different method. Let us for simplicity denote $\chi = \omega_n - \omega_{n+m}$ and let us introduce a new polynomial

$$\varphi(t) = \chi \left(\frac{z_{n+m} - z_1}{2 \cos \pi/2(n+m)} t + \frac{z_1 + z_{n+m}}{2} \right).$$

Since all the zeros of the polynomial $\chi = \omega_n - \omega_{n+m}$ are real numbers $z_1 < z_2 < \dots < z_{n+m}$, all zeros of the polynomial φ are also real numbers w_k , such that $-\cos \pi/2(n+m) = w_1 < w_2 < \dots < w_{n+m} = \cos \pi/2(n+m)$, and we may apply Theorem 2:5 to the polynomial φ , after changing the number n into $n+m$. Thus we get the inequality

$$|\varphi(t)| \geq \min_{1 \leq k \leq n+m-1} \|\varphi\|_{\langle w_k, w_{k+1} \rangle} |\tau_{n+m}(t)| \quad (|t| \geq \cos \pi/2(n+m)),$$

which, in view of the definition of the polynomial φ , may be written in the form

$$(7) \quad |\chi(t)| \geq \min_{1 \leq k \leq n+m-1} \|\chi\|_{\langle z_k, z_{k+1} \rangle} \tau_{n+m} \left(\frac{2t - z_1 - z_{n+m}}{z_{n+m} - z_1} \cos \frac{\pi}{2(n+m)} \right) \\ (t \leq z_1, t \geq z_{n+m}).$$

Let us consider the set

$$\{t: t \in \langle z_{n+m}, b \rangle, |\chi(t)| \geq g \min_{1 \leq k \leq n+m-1} \|\chi\|_{\langle z_k, z_{k+1} \rangle}\},$$

the measure of which equals the last term of the right-hand side of inequality (5). This set will not increase if we replace in its definition the function $|\chi(t)|$ by the right-hand side of inequality (7). It means that

$$(8) \quad \{t: t \in \langle z_{n+m}, b \rangle, |\chi(t)| \geq g \min_{1 \leq k \leq n+m-1} \|\chi\|_{\langle z_k, z_{k+1} \rangle}\} \\ \geq \left\{ t: t \in \langle z_{n+m}, b \rangle, \tau_{n+m} \left(\frac{2t - z_1 - z_{n+m}}{z_{n+m} - z_1} \cos \frac{\pi}{2(n+m)} \right) \geq g \right\}.$$

The number $\cos \pi/2(n+m)$ is the greatest zero of the polynomial τ_{n+m} , and this polynomial, which has only the real zeros, increases from

0 to $+\infty$ for $t \geq \cos \pi/2(n+m)$. Thus the polynomial

$$(9) \quad \tau_{n+m} \left(\frac{2t - z_1 - z_{n+m}}{z_{n+m} - z_1} \cos \frac{\pi}{2(n+m)} \right)$$

increases for $t \geq z_{n+m}$ from 0 to $+\infty$ and according to the definition of the function \varkappa_{n+m} , it achieves the value g at the point t_0 such that

$$\frac{2t_0 - z_1 - z_{n+m}}{z_{n+m} - z_1} \cos \frac{\pi}{2(n+m)} = \varkappa_{n+m}(g),$$

i.e. such that

$$t_0 = \frac{(z_{n+m} - z_1) \varkappa_{n+m}(g)}{2 \cos \pi/2(n+m)} + \frac{z_1 + z_{n+m}}{2}.$$

If $t_0 \leq b$, i.e. if the polynomial (9) achieves the value g in the interval $\langle z_{n+m}, b \rangle$, then the set, whose measure appears on the right-hand side of (8), equals the interval $\langle t_0, b \rangle$. If, however, $t_0 > b$, then the set is empty. Thus

$$(10) \quad \left\{ t: t \in \langle z_{n+m}, b \rangle, |\chi(t)| \geq g \min_{1 \leq k \leq n+m-1} \|\chi\|_{\langle z_k, z_{k+1} \rangle} \right\} \\ \geq \max \left\{ 0, b - \frac{(z_{n+m} - z_1) \varkappa_{n+m}(g)}{2 \cos \pi/2(n+m)} - \frac{z_1 + z_{n+m}}{2} \right\}.$$

In a similar way we prove that

$$(11) \quad \left\{ t: t \in \langle a, z_1 \rangle, |\chi(t)| \geq g \min_{1 \leq k \leq n+m-1} \|\chi\|_{\langle z_k, z_{k+1} \rangle} \right\} \\ \geq \max \left\{ 0, -a - \frac{(z_{n+m} - z_1) \varkappa_{n+m}(g)}{2 \cos \pi/2(n+m)} + \frac{z_1 + z_{n+m}}{2} \right\}.$$

If we consider inequalities (5), (6), (10) and (11) in which the notation $\chi = \omega_n - \omega_{n+m}$ is used, we get the following estimate for $|A_h|$:

$$|A_h| \geq (z_{n+m} - z_1) \mu_{n+m}(g) + \\ + \max \left\{ 0, b - \frac{(z_{n+m} - z_1) \varkappa_{n+m}(g)}{2 \cos \pi/2(n+m)} - \frac{z_1 + z_{n+m}}{2} \right\} + \\ + \max \left\{ 0, -a - \frac{(z_{n+m} - z_1) \varkappa_{n+m}(g)}{2 \cos \pi/2(n+m)} + \frac{z_1 + z_{n+m}}{2} \right\}.$$

For arbitrary numbers x and y we have the inequality $\max\{0, x\} + \max\{0, y\} \geq \max\{0, x+y\}$, hence

$$|A_h| \geq (z_{n+m} - z_1) \mu_{n+m}(g) + \max \left\{ 0, b - a - \frac{(z_{n+m} - z_1) \varkappa_{n+m}(g)}{\cos \pi/2(n+m)} \right\}.$$

To obtain the desired ultimate estimation for $|A_h|$ from the last result we shall treat the right-hand side of the obtained inequality as a function of the difference $z_{n+m} - z_1$ and we compute the minimum of this function. It changes linearly on the left and on the right from the point

$$(12) \quad \frac{\cos \pi/2(n+m)}{\kappa_{n+m}(g)} (b-a),$$

and on the left from this point the coefficient of $z_{n+m} - z_1$ equals

$$(13) \quad \mu_{n+m}(g) - \frac{\kappa_{n+m}(g)}{\cos \pi/2(n+m)},$$

and on the right it equals

$$(14) \quad \mu_{n+m}(g).$$

From the definition of the functions μ_{n+m} and κ_{n+m} it follows that the first of them is non-negative and does not exceed 1, the other one is no smaller than $\cos \pi/2(n+m)$. Thus coefficient (13) is non-positive, and (14)—non-negative and the right-hand side of the considered inequality achieves its minimum at the point (12). Thus inequality (4), which we were to prove, follows at once. Moreover, since the function μ_{n+m} is defined only for arguments from the interval $\langle 0, 1 \rangle$, and the function κ_{n+m} only for non-negative arguments, we must have

$$\frac{h + \delta_n(\xi)}{1 - \delta_n(\xi)} \leq 1, \quad \text{i.e.} \quad h \leq 1 - 2\delta_n(\xi).$$

Thus the values of the parameter h for which inequality (4) is satisfied exist only in the case when $\delta_n(\xi) \leq \frac{1}{2}$.

7.2. Now we shall apply the estimations of the measure of the set A_h from below and from above, obtained in Theorems 7:1 and 7:2 to the estimation of the ratio $\delta_n(\xi) = \varepsilon_{n+m}(\xi)/\varepsilon_n(\xi)$.

THEOREM 7:3. *If m is the smallest integer such that $\varepsilon_{n+m}(\xi) < \varepsilon_n(\xi)$, and if $\delta_n(\xi) = \varepsilon_{n+m}(\xi)/\varepsilon_n(\xi)$, then for all $h \in \langle 0, 1 \rangle$ we have the inequality*

$$(15) \quad \delta_n(\xi) \geq \frac{h + \left(\frac{|A_h|}{b-a}\right)^2 - 1}{2 - \left(\frac{|A_h|}{b-a}\right)^2}.$$

Proof. For $h \geq \delta_n(\xi)$ we may transform inequality (1) to the form (15). Inequality (15) is also true, however, for $\delta_n(\xi) > h$. It follows from

the inequality

$$h \geq \frac{h + \left(\frac{|A_h|}{b-a}\right)^2 - 1}{2 - \left(\frac{|A_h|}{b-a}\right)^2},$$

which is equivalent to the evident inequality

$$\left(1 - \left(\frac{|A_h|}{b-a}\right)^2\right)(1+h) \geq 0.$$

THEOREM 7:4. *If the symbols $\delta_n(\xi)$ and m have the same meaning as in Theorem 7:3, then for all $h \in \langle 0, 1 \rangle$ we have*

$$(16) \quad \delta_n(\xi) \geq \frac{\iota_{n+m}\left(\frac{|A_h|}{b-a}\right) - h}{\iota_{n+m}\left(\frac{|A_h|}{b-a}\right) + 1},$$

where the function ι_{n+m} is the function inverse to the function of variable h :

$$(17) \quad \frac{\mu_{n+m}(h) \cos \pi/2(n+m)}{\kappa_{n+m}(h)}.$$

Proof. Since the function μ_{n+m} decreases from 1 to 0, and κ_{n+m} increases from $\cos \pi/2(n+m)$ to 1 as h increases from 0 to 1, the function (17) decreases from 1 to 0 and the inverse function exists. The latter is defined in the interval $\langle 0, 1 \rangle$ and decreases from 1 to 0 in this interval. Thus, inequality (4) for $\delta_n(\xi) \leq \frac{1}{2}$ and $h \leq 1 - 2\delta_n(\xi)$ may be written in the form

$$\iota_{n+m}\left(\frac{|A_h|}{b-a}\right) \leq \frac{h + \delta_n(\xi)}{1 - \delta_n(\xi)},$$

i.e. we have inequality (16) under the above restrictions. The same inequality is also true for $h > 1 - 2\delta_n(\xi)$, i.e. for $\delta_n(\xi) > \frac{1}{2}(1-h)$, since

$$\frac{1}{2}(1-h) \geq \frac{\iota_{n+m}\left(\frac{|A_h|}{b-a}\right) - h}{\iota_{n+m}\left(\frac{|A_h|}{b-a}\right) + 1}, \quad \text{i.e.} \quad (1+h) \left(1 - \iota_{n+m}\left(\frac{|A_h|}{b-a}\right)\right) \geq 0$$

and, the more it is true for $\delta_n(\xi) > \frac{1}{2}$.

Theorems 7:3 and 7:4 give the positive estimates from below for the ratio $\delta_n(\xi)$ only if the inequalities

$$h > 1 - \left(\frac{|A_h|}{b-a}\right)^2 \quad \text{and} \quad h < \iota_{n+m} \left(\frac{|A_h|}{b-a}\right)$$

respectively, are satisfied. In the opposite case, i.e. if for all $h \in \langle 0, 1 \rangle$

$$\begin{aligned} \iota_{n+m} \left(\frac{|A_h|}{b-a}\right) \leq h \leq 1 - \left(\frac{|A_h|}{b-a}\right)^2, \\ \text{i.e.} \quad \frac{\mu_{n+m}(h) \cos \pi/2(n+m)}{\kappa_{n+m}(h)} \leq \frac{|A_h|}{b-a} \leq \sqrt{1-h} \end{aligned}$$

nothing follows from Theorems 7:3 and 7:4 except the fact that $\delta_n(\xi)$ is non-negative. Probably the effective application of these theorems might be considerably extended through the sharpening of the estimations for the measures of the type

$$(18) \quad \{t: t \in I, |\chi(t)| \geq h \|\chi\|_I\},$$

where χ is a polynomial with real zeros. These estimations were obtained by finding analogous estimations for subintervals of I such that their end points are the successive zeros of the polynomial χ . The latter estimates cannot be sharpened, but one can sharpen their conclusions which concern measure (18).

7.3. We shall add the remark that by investigating the difference $\xi - \omega_n$, where ξ is a function continuous in the interval I , and ω_n is its n -th best polynomial on this interval, we may obtain estimates for the ratio $\delta_n(\xi) = \varepsilon_{n+m}(\xi)/\varepsilon_n(\xi)$ by a different method; this will be shown in § 9.5. In this way we shall obtain estimates from below for $\delta_n(\xi)$ by a positive number, except in the improbable case when the (n) -points $u_0 < u_1 < \dots < u_{n+m}$ of the function ξ are identical with the points $\frac{1}{2}(a+b) - \frac{1}{2}(b-a) \cos k\pi/(n+m)$ for $k = 0, 1, \dots, n+m$, respectively.

8. The relation between the error of the best approximation and the interval of approximation. The error $\varepsilon_n(\xi; I)$ of the best approximation depends upon the approximated function ξ , the degree n of approximating polynomial, and the interval I over which we perform our approximation. The remarks given in § 8.1 concerning the last relation characterize this relation rather weakly. Nevertheless, they seem to show the direction in which we should proceed in order to get the interesting results. We shall show in § 8.2 that these results might be useful for the estimation of the error $\varepsilon_n(\xi; I)$.

8.1. We shall start from the most obvious inequalities. If $J \subset I$, where I and J are closed intervals, then

$$(1) \quad \varepsilon_n(\xi; J) \leq \varepsilon_n(\xi; I);$$

the more general inequality $\varepsilon_n(\xi; G) \leq \varepsilon_n(\xi; F)$, where G and F are closed sets, is also true; we have used this inequality several times in proofs of previous theorems. If the function ξ satisfies some additional assumptions in the interval I (for instance, if its $(n+1)$ -st derivative exists and is of a constant sign) and if $I \neq J$, then we can replace the symbol \leq by the symbol $<$ in inequality (1).

It is also obvious that if the interval J is not contained in the interval K , neither K is contained in J , then the errors of approximation $\varepsilon_n(\xi; J)$ and $\varepsilon_n(\xi; K)$ are independent in the sense that for arbitrary non-negative numbers e_J, e_K there exists a function $\xi \in \mathcal{C}_{J \cup K}$ which satisfies the equations $\varepsilon_n(\xi; J) = e_J, \varepsilon_n(\xi; K) = e_K$. The function ξ may be, for instance, the function such that: 1. it is identically 0 in the interval $J \cap K$ (provided this set is not empty); 2. in the interval $J - K$ its graph is a polygonal line with $n+4$ vertices with ordinates successively equal to 0, $e_J, -e_J, \dots, (-1)^{n+1}e_J, 0$; 3. in the interval $K - J$ its graph is a polygonal line with $n+4$ vertices with ordinates successively equal to 0, $e_K, -e_K, \dots, (-1)^{n+1}e_K, 0$. From this description, and from the fundamental Theorem 1:2 it follows that the n -th best polynomial for this function in the interval J as well as in the interval K is identically zero.

If the numbers $\varepsilon_n(\xi; J)$ and $\varepsilon_n(\xi; K)$ are known, then the n -th error of the best approximation of the function ξ in the set $J \cup K$ is no longer arbitrary. From inequality (1) generalized to the case arbitrary closed sets, we get the following estimation of this error from below:

$$\varepsilon_n(\xi; J \cup K) \geq \max\{\varepsilon_n(\xi; J), \varepsilon_n(\xi; K)\},$$

and in the case when the intervals J and K have a common part with positive length, we can also give an estimation from above; this will be done in the next theorem. If, however, the set $J \cap K$ consists of a single point, or is an empty set, then the number $\varepsilon_n(\xi; J \cup K)$ may be arbitrarily large.

THEOREM 8:1. For any function ξ continuous in the interval $I = \langle a, b \rangle$ any numbers c and d such that $a \leq c < d \leq b$ and any non-negative integer n we have the inequality

$$(2) \quad \varepsilon_n(\xi; I) \leq \frac{t_J(1+t_K)\varepsilon_n(\xi; J) + t_K(1+t_J)\varepsilon_n(\xi; K)}{t_J + t_K},$$

where $J = \langle a, d \rangle, K = \langle c, b \rangle, t_J = \tau_n \left(\frac{2b-c-d}{d-c} \right), t_K = \tau_n \left(\frac{c+d-2a}{d-c} \right)$.

Proof. If ω_{nJ} and ω_{nK} are the n -th best polynomials for the function ξ in the intervals J and K , respectively, then

$$|\xi(t) - \omega_{nJ}(t)| \leq \varepsilon_n(\xi; J) \quad (t \in J),$$

$$|\xi(t) - \omega_{nK}(t)| \leq \varepsilon_n(\xi; K) \quad (t \in K).$$

Both of these inequalities are satisfied for $t \in J \cap K = \langle c, d \rangle$, hence it follows that

$$|\omega_{nJ}(t) - \omega_{nK}(t)| \leq \varepsilon_n(\xi; J) + \varepsilon_n(\xi; K) \quad (t \in J \cap K).$$

Thus, if we transform linearly the variable t , we see that the norm of the polynomial

$$\omega_{nJ}\left(\frac{(d-c)u + c + d}{2}\right) - \omega_{nK}\left(\frac{(d-c)u + c + d}{2}\right)$$

in the interval $\langle -1, 1 \rangle$ of the variable u does not exceed $\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K)$. Applying Theorem 2:3 to this polynomial we can write the inequality

$$\begin{aligned} \left| \omega_{nJ}\left(\frac{(d-c)u + c + d}{2}\right) - \omega_{nK}\left(\frac{(d-c)u + c + d}{2}\right) \right| \\ \leq (\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K)) |\tau_n(u)| \quad (|u| \geq 1), \end{aligned}$$

i.e., after coming back to the variable t —the inequality

$$|\omega_{nJ}(t) - \omega_{nK}(t)| \leq (\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K)) \left| \tau_n\left(\frac{2t - c - d}{d - c}\right) \right| \quad (t \leq c, t \geq d).$$

Let us now consider the polynomial

$$\omega = h\omega_{nJ} + (1-h)\omega_{nK} = \omega_{nJ} - (1-h)(\omega_{nJ} - \omega_{nK}) = \omega_{nK} + h(\omega_{nJ} - \omega_{nK}),$$

where, at this moment, h is an arbitrary number from the interval $\langle 0, 1 \rangle$. For $t \in \langle c, d \rangle$ we have

$$\begin{aligned} |\xi(t) - \omega(t)| &\leq h|\xi(t) - \omega_{nJ}(t)| + (1-h)|\xi(t) - \omega_{nK}(t)| \\ &\leq h\varepsilon_n(\xi; J) + (1-h)\varepsilon_n(\xi; K). \end{aligned}$$

For $t \in \langle a, c \rangle$ we have the inequality

$$\begin{aligned} |\xi(t) - \omega(t)| &\leq |\xi(t) - \omega_{nJ}(t)| + (1-h)|\omega_{nJ}(t) - \omega_{nK}(t)| \\ &\leq \varepsilon_n(\xi; J) + (1-h)(\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K)) \left| \tau_n\left(\frac{2t - c - d}{d - c}\right) \right|, \end{aligned}$$

and since the Chebyshev polynomial τ_n (which has only the real zeros

in the interval $\langle -1, 1 \rangle$ is monotone outside this interval and satisfies the equation $|\tau_n(-t)| = \tau_n(t)$ for $t \geq 1$, we get

$$|\xi(t) - \omega(t)| \leq \varepsilon_n(\xi; J) + (1-h)(\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K)) \tau_n \left(\frac{c+d-2a}{d-c} \right) \\ (t \in \langle a, c \rangle).$$

Finally, for $t \in \langle d, b \rangle$ we have the inequality

$$|\xi(t) - \omega(t)| \leq |\xi(t) - \omega_{nK}(t)| + h |\omega_{nJ}(t) - \omega_{nK}(t)| \\ \leq \varepsilon_n(\xi; K) + h(\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K)) \tau_n \left(\frac{2b-c-d}{d-c} \right).$$

Thus

$$(3) \quad \varepsilon_n(\xi; I) \leq \|\xi - \omega\|_{J \cup K} \leq \max \{ h\varepsilon_n(\xi; J) + (1-h)\varepsilon_n(\xi; K), \\ \varepsilon_n(\xi; J) + (1-h)(\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K))t_K, \varepsilon_n(\xi; K) + h(\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K))t_J \}.$$

Since the numbers t_J and t_K are the values of the Chebyshev polynomial for the values of argument greater than 1, we have $t_J > 1$ and $t_K > 1$. It is easy to see that the first of the three numbers which appear on the right-hand side of (3) is smaller than the remaining ones for an arbitrary $h \in \langle 0, 1 \rangle$. The second of these three numbers is a decreasing function of h , and the third one is an increasing function of h . Thus the number which estimates from above the n -th error of approximation $\varepsilon_n(\xi; I)$ in inequality (3) will be smallest for h such that

$$\varepsilon_n(\xi; J) + (1-h)(\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K))t_K \\ = \varepsilon_n(\xi; K) + h(\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K))t_J,$$

i.e. for

$$h = \frac{\varepsilon_n(\xi; J) - \varepsilon_n(\xi; K) + (\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K))t_K}{(\varepsilon_n(\xi; J) + \varepsilon_n(\xi; K))(t_J + t_K)}.$$

For such a value of h , which, as it may be seen, belongs to the interval $\langle 0, 1 \rangle$, inequality (3) becomes inequality (2) which was to be proved.

8.2. Theorem 8:1, despite all its generality, is very weak, as may be seen from the following remark. If the numbers a and b are fixed, c tends to a and d tends to b , then from (2) we obtain in the limit the estimate $\varepsilon_n(\xi; I) \leq \varepsilon_n(\xi; J) + \varepsilon_n(\xi; K)$, in spite of the fact that the approximation errors $\varepsilon_n(\xi; J)$ and $\varepsilon_n(\xi; K)$ tend to the common limit $\varepsilon_n(\xi; I)$. It seems to be difficult to obtain sharper estimations, but it may be worthwhile to indicate the advantages, which could follow from such estimate.

Suppose that the function ξ has in the interval I the $(n+1)$ -st derivative, which is near to zero in the whole interval, except a neighbourhood of the end point b , where it can assume large values. The error $\varepsilon_n(\xi; I)$ may be estimated from above by using Theorem 6:2, and the result to this estimation will be a large number, since the maximum of $|\xi^{(n+1)}(t)|$ in the interval I is large. If, however, we know sufficiently exact estimates of the type

$$(4) \quad \varepsilon_n(\xi; I) \leq s_J \varepsilon_n(\xi; J) + s_K \varepsilon_n(\xi; K)$$

(at least for sufficiently regular continuous functions), where s_J, s_K are positive numbers independent of ξ , we could have represented the interval I as a sum of intersecting intervals $J = \langle a, d \rangle$ and $K = \langle c, b \rangle$ such that: 1. in the interval J the maximum of $|\xi^{(n+1)}(t)|$ would be considerably smaller than in the interval I ; 2. the interval K would be considerably shorter than I . Then applying inequality (4), and estimating the approximation errors $\varepsilon_n(\xi; J), \varepsilon_n(\xi; K)$ by Theorem 6:2, we should obtain the results better than before, since the value of the $(n+1)$ -st derivative of the function ξ and the length of interval I have a very strong influence on the upper bound in the estimate of the error of the best approximation (see (4), § 6).

CHAPTER III

THE DISTRIBUTION OF (n) -POINTS IN THE INTERVAL OF APPROXIMATION

9. The estimates dependent upon the value of the ratio $\varepsilon_{n+1}(\xi)/\varepsilon_n(\xi)$ [23].

9.1. In the whole section we shall assume that the interval of approximation is the interval $\langle -1, 1 \rangle$, and that we have the inequality

$$(1) \quad \varepsilon_n(\xi) > \varepsilon_{n+1}(\xi).$$

The results obtained for the case of interval $\langle -1, 1 \rangle$ can easily be generalized for the case of any other interval by means of Theorem 1:8.

It follows from inequality (1) that $\omega_n \neq \omega_{n+1}$, where, as usual, ω_n and ω_{n+1} are the n -th and the $(n+1)$ -st best polynomials for the function ξ in the interval $\langle -1, 1 \rangle$. It also follows from this inequality that we can find only $n+2$ points u_0, u_1, \dots, u_{n+1} (where $-1 \leq u_0 < u_1 < \dots < u_{n+1} \leq 1$), which are, alternately, the $(n, +)$ -points and $(n, -)$ -points of the function ξ .

In Theorem 3:1 we have proved under the assumption (1) that the difference $\omega_n - \omega_{n+1}$ is a polynomial of the degree $n+1$ with real single zeros which lie in the interval $(-1, 1)$. These zeros, arranged in increasing order, will be denoted by the symbols z_1, z_2, \dots, z_{n+1} ; in addition, we denote $z_0 = -1, z_{n+2} = 1$.

THEOREM 9:1. *If $\xi \in \mathcal{C}_{\langle -1, 1 \rangle}$, then the (n) -point u_k ($k = 0, 1, \dots, n+1$) of the function ξ belongs to the set*

$$\langle z_k, z_{k+1} \rangle \cap \left\{ t: \frac{|\omega_n(t) - \omega_{n+1}(t)|}{\|\omega_n - \omega_{n+1}\|_{\langle -1, 1 \rangle}} \geq \frac{\varepsilon_n(\xi) - \varepsilon_{n+1}(\xi)}{\varepsilon_n(\xi) + \varepsilon_{n+1}(\xi)} \right\}.$$

Proof. The fact that u_k belongs to the interval $\langle z_k, z_{k+1} \rangle$ has already been proved in Theorem 3:1. If we subtract the inequalities

$$|\xi(u_k) - \omega_{n+1}(u_k)| \leq \varepsilon_{n+1}(\xi),$$

which follow from the definition of the polynomial ω_{n+1} , from the corresponding equations

$$|\xi(u_k) - \omega_n(u_k)| = \varepsilon_n(\xi),$$

which follow from the definition of the (n) -points, we get

$$|\omega_n(u_k) - \omega_{n+1}(u_k)| \geq \varepsilon_n(\xi) - \varepsilon_{n+1}(\xi) \quad (k = 0, 1, \dots, n+1).$$

Since $\|\omega_n - \omega_{n+1}\|_{\langle -1, 1 \rangle} \leq \varepsilon_n(\xi) + \varepsilon_{n+1}(\xi)$ (inequality (2) from § 3 for $I = \langle -1, 1 \rangle$) we have

$$(2) \quad \frac{|\omega_n(u_k) - \omega_{n+1}(u_k)|}{\|\omega_n - \omega_{n+1}\|_{\langle -1, 1 \rangle}} \geq \frac{\varepsilon_n(\xi) - \varepsilon_{n+1}(\xi)}{\varepsilon_n(\xi) + \varepsilon_{n+1}(\xi)} \quad (k = 0, 1, \dots, n+1)$$

and the assertion of Theorem 9:1 follows.

Let τ be an arbitrary polynomial of the degree $n+1$, with real zeros $z_1^\tau, z_2^\tau, \dots, z_{n+1}^\tau$ and such that $-1 < z_1^\tau < z_2^\tau < \dots < z_{n+1}^\tau < 1$. For every such polynomial we shall also denote $z_0^\tau = -1, z_{n+2}^\tau = 1$. For any number $g \in (0, 1)$ we shall denote by $\mathscr{W}_{n+1, g}$ the class of polynomials τ which satisfy the condition

$$(3) \quad \min_{0 \leq j \leq n+1} \|\tau\|_{\langle z_j^\tau, z_{j+1}^\tau \rangle} \geq g \|\tau\|_{\langle -1, 1 \rangle}.$$

THEOREM 9:2. *If $\xi \in \mathscr{C}_{\langle -1, 1 \rangle}$ then for an arbitrary system of points u_0, u_1, \dots, u_{n+1} arranged according to their magnitude, and which are alternately the $(n, +)$ -points and $(n, -)$ -points of the function ξ , we have*

$$(4) \quad \left\{ \begin{array}{l} -1 \leq u_0 \leq \sup_{\tau \in \mathscr{W}_{n+1, g}} s_0^\tau, \\ \inf_{\tau \in \mathscr{W}_{n+1, g}} r_k^\tau \leq u_k \leq \sup_{\tau \in \mathscr{W}_{n+1, g}} s_k^\tau \quad (k = 1, 2, \dots, n), \\ \inf_{\tau \in \mathscr{W}_{n+1, g}} r_{n+1}^\tau \leq u_{n+1} \leq 1, \end{array} \right.$$

where

$$(5) \quad g = \frac{\varepsilon_n(\xi) - \varepsilon_{n+1}(\xi)}{\varepsilon_n(\xi) + \varepsilon_{n+1}(\xi)},$$

$$(6) \quad \langle r_k^\tau, s_k^\tau \rangle = \langle z_k^\tau, z_{k+1}^\tau \rangle \cap \{t: |\tau(t)| \geq g \|\tau\|_{\langle -1, 1 \rangle}\} \quad (k = 0, 1, \dots, n+1).$$

Proof. Let us notice first that obviously the set on the right-hand side of equation (6) is an interval. For $k = 1, 2, \dots, n$ this set consists of points which lie between the successive zeros of the polynomial τ , and for which the absolute value of this polynomial is not smaller than $g \|\tau\|_{\langle -1, 1 \rangle}$. In intervals $\langle -1, z_1^\tau \rangle = \langle z_0^\tau, z_1^\tau \rangle$ and $\langle z_{n+1}^\tau, 1 \rangle = \langle z_{n+1}^\tau, z_{n+2}^\tau \rangle$ the polynomial τ is a monotone function, hence formulas (6) are meaningful also for $k = 0$ and $k = n+1$; we have here $r_0^\tau = z_0^\tau = -1, s_{n+1}^\tau = z_{n+2}^\tau = 1$ for an arbitrary polynomial $\tau \in \mathscr{W}_{n+1, g}$.

The polynomial $\omega_n - \omega_{n+1}$ belongs to the class $\mathscr{W}_{n+1, g}$, since inequality (3) follows from (2) and from $u_k \in \langle z_k, z_{k+1} \rangle$ for this polynomial. In

view of Theorem 9:1 the (n) -point u_k of the function ξ belongs to interval (6), which corresponds to the polynomial $\tau^* = \omega_n - \omega_{n+1}$. Thus we have the inequalities

$$-1 \leq u_0 \leq s_0^{\tau^*}, \quad r_k^{\tau^*} \leq u_k \leq s_k^{\tau^*} \quad (k = 1, 2, \dots, n), \quad r_{n+1}^{\tau^*} \leq u_{n+1} \leq 1$$

and, in addition, inequalities (4).

9.2. Theorem 9:2 gives the estimation of the (n) -points of the function ξ , which depends only upon the value of the ratio (5); i. e., it depends only upon the ratio value of the successive errors of the best approximation:

$$\frac{\varepsilon_{n+1}(\xi)}{\varepsilon_n(\xi)} = \frac{1-g}{1+g}.$$

In § 9.3 we shall find $\inf_{\tau} r_k^{\tau}$, $\sup_{\tau} s_k^{\tau}$, which will allow us to apply inequality (4). But first we shall prove an auxiliary theorem in which we establish the conditions for the existence of a solution of a certain system of linear inequalities.

THEOREM 9:3. *Suppose we are given $2n+3$ distinct numbers $y_0, y_1, \dots, y_{n+1}, z_1, z_2, \dots, z_{n+1}$ and the numbers s_0, s_1, \dots, s_{n+1} different from 0. The system of linear inequalities*

$$(7) \quad \sum_{j=1}^{n+1} \frac{s_i x_j}{y_i - z_j} < 0 \quad (i = 0, 1, \dots, n+1)$$

does not have a solution if and only if the numbers

$$(8) \quad \frac{\prod_{j=1}^{n+1} (z_j - y_i)}{s_i \prod_{j=0, j \neq i}^{n+1} (y_j - y_i)} \quad (i = 0, 1, \dots, n+1)$$

have a common sign.

Proof. System (7) does not have a solution if and only if the system

$$(9) \quad \sum_{j=1}^{n+1} \frac{s_i x_j}{y_i - z_j} \leq q_i \quad (i = 0, 1, \dots, n+1)$$

does not have a solution for any system of negative numbers q_0, q_1, \dots, q_{n+1} . Now we shall deal briefly with a more general system of linear inequalities:

$$(10) \quad \sum_{j=1}^{n+1} a_{ij} x_j \leq a_i \quad (i = 0, 1, \dots, n+1).$$

Let

$$\bar{d}_i = \begin{vmatrix} a_{01} & a_{02} & \dots & a_{0,n+1} \\ \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,n+1} \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,n+1} \\ \dots & \dots & \dots & \dots \\ a_{n+1,1} & a_{n+1,2} & \dots & a_{n+1,n+1} \end{vmatrix} \quad (i = 0, 1, \dots, n+1).$$

It is well known that if system (10) has a matrix of the order $n+1$, then a necessary and sufficient condition for the existence of a solution for this system is the existence of a determinant \bar{d}_i , different from zero, and such that

$$\frac{1}{\bar{d}_i} \begin{vmatrix} a_{01} & a_{02} & \dots & a_{0,n+1} & a_0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{i-1,1} & a_{i-1,2} & \dots & a_{i-1,n+1} & a_{i-1} \\ a_{i+1,1} & a_{i+1,2} & \dots & a_{i+1,n+1} & a_{i+1} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n+1,1} & a_{n+1,2} & \dots & a_{n+1,n+1} & a_{n+1} \\ a_{i1} & a_{i2} & \dots & a_{i,n+1} & a_i \end{vmatrix} \geq 0,$$

or

$$\frac{(-1)^i}{\bar{d}_i} \begin{vmatrix} a_0 & a_{01} & \dots & a_{0,n+1} \\ a_1 & a_{11} & \dots & a_{1,n+1} \\ \dots & \dots & \dots & \dots \\ a_{n+1} & a_{n+1,1} & \dots & a_{n+1,n+1} \end{vmatrix} = \frac{(-1)^i}{\bar{d}_i} \sum_{j=0}^{n+1} (-1)^j a_j \bar{d}_j \geq 0$$

(see [10], p. 27). In other words, system (10) does not have a solution if and only if: (i) the numbers $(-1)^i \bar{d}_i$, different from zero have a common sign; (ii) this sign is different than the sign of the non-vanishing expression

$$\sum_{j=0}^{n+1} (-1)^j a_j \bar{d}_j.$$

We may apply this theorem for the investigation of system (9), since it will soon turn out that the matrix of this system is of the order $n+1$. In this case the absolute terms $a_0 = q_0, a_1 = q_1, \dots, a_{n+1} = q_{n+1}$ are negative, and we may omit condition (ii) in the criterion formulated above, since it follows from condition (i).

Let us introduce the symbols

$$\text{cau}(a_1, a_2, \dots, a_p; b_1, b_2, \dots, b_p) = \begin{vmatrix} \frac{1}{a_1 - b_1} & \frac{1}{a_1 - b_2} & \dots & \frac{1}{a_1 - b_p} \\ \frac{1}{a_2 - b_1} & \frac{1}{a_2 - b_2} & \dots & \frac{1}{a_2 - b_p} \\ \dots & \dots & \dots & \dots \\ \frac{1}{a_p - b_1} & \frac{1}{a_p - b_2} & \dots & \frac{1}{a_p - b_p} \end{vmatrix}$$

($a_j \neq b_k$ for $j = 1, 2, \dots, p$ and $k = 1, 2, \dots, p$) for the Cauchy determinant, and

$$\text{van}(c_1, c_2, \dots, c_p) = \begin{vmatrix} 1 & c_1 & c_1^2 & \dots & c_1^{p-1} \\ 1 & c_2 & c_2^2 & \dots & c_2^{p-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & c_p & c_p^2 & \dots & c_p^{p-1} \end{vmatrix}$$

for the Vandermonde determinant.

The determinant d_i for system (9) is of the form

$$d_i = \begin{vmatrix} \frac{s_0}{y_0 - z_1} & \frac{s_0}{y_0 - z_2} & \dots & \frac{s_0}{y_0 - z_{n+1}} \\ \dots & \dots & \dots & \dots \\ \frac{s_{i-1}}{y_{i-1} - z_1} & \frac{s_{i-1}}{y_{i-1} - z_2} & \dots & \frac{s_{i-1}}{y_{i-1} - z_{n+1}} \\ \frac{s_{i+1}}{y_{i+1} - z_1} & \frac{s_{i+1}}{y_{i+1} - z_2} & \dots & \frac{s_{i+1}}{y_{i+1} - z_{n+1}} \\ \dots & \dots & \dots & \dots \\ \frac{s_{n+1}}{y_{n+1} - z_1} & \frac{s_{n+1}}{y_{n+1} - z_2} & \dots & \frac{s_{n+1}}{y_{n+1} - z_{n+1}} \end{vmatrix}$$

$$= \frac{1}{s_i} \text{cau}(y_0, \dots, y_{i-1}, y_{i+1}, \dots, y_{n+1}; z_1, z_2, \dots, z_{n+1}) \prod_{j=0}^{n+1} s_j$$

and, in view of the well-known formula

$$\text{cau}(a_1, a_2, \dots, a_p; b_1, b_2, \dots, b_p) = (-1)^{p(p-1)/2} \frac{\text{van}(a_1, a_2, \dots, a_p) \text{van}(b_1, b_2, \dots, b_p)}{\prod_{j=1}^p \prod_{k=1}^p (a_j - b_k)},$$

we can write that

$$d_i = (-1)^{n(n+1)/2} \frac{\text{van}(y_0, \dots, y_{i-1}, y_{i+1}, \dots, y_{n+1}) \text{van}(z_1, z_2, \dots, z_{n+1})}{s_i \prod_{j=0, j \neq i}^{n+1} \prod_{k=1}^{n+1} (y_j - z_k)} \prod_{j=0}^{n+1} s_j.$$

Since

$$\text{van}(c_1, c_2, \dots, c_p) = (c_2 - c_1)(c_3 - c_1) \dots (c_p - c_1)(c_3 - c_2) \dots (c_p - c_2) \dots (c_p - c_{p-1}),$$

we have

$$\begin{aligned} & \text{van}(y_0, \dots, y_{i-1}, y_{i+1}, \dots, y_{n+1}) \\ &= \frac{\text{van}(y_0, y_1, \dots, y_{n+1})}{(y_i - y_0) \dots (y_i - y_{i-1})(y_{i+1} - y_i) \dots (y_{n+1} - y_i)} = \frac{\text{van}(y_0, y_1, \dots, y_{n+1})}{(-1)^i \prod_{j=0, j \neq i}^{n+1} (y_j - y_i)}, \end{aligned}$$

$$(-1)^i d_i = (-1)^{n(n+1)/2} \times$$

$$\begin{aligned} & \times \frac{\text{van}(y_0, y_1, \dots, y_{n+1}) \text{van}(z_1, z_2, \dots, z_{n+1}) \prod_{j=1}^{n+1} (y_i - z_j)}{s_i \prod_{j=0}^{n+1} \prod_{k=1}^{n+1} (y_j - z_k) \prod_{j=0, j \neq i}^{n+1} (y_j - y_i)} \prod_{j=0}^{n+1} s_j \\ &= \frac{\prod_{j=1}^{n+1} (z_j - y_i)}{s_i \prod_{j=0, j \neq i}^{n+1} (y_j - y_i)} d, \end{aligned}$$

where

$$d = (-1)^{(n+1)(n-2)/2} \frac{\text{van}(y_0, y_1, \dots, y_{n+1}) \text{van}(z_1, z_2, \dots, z_{n+1})}{\prod_{j=0}^{n+1} \prod_{k=1}^{n+1} (y_j - z_k)} \prod_{j=0}^{n+1} s_j.$$

Now, the assertion of the theorem follows directly from the formula for $(-1)^i d_i$ and from condition (i), which is equivalent to the impossibility of the solution of systems (9) and (7).

9.3. Let the numbers $v_1^r, v_2^r, \dots, v_n^r$ (where $z_1^r < v_1^r < z_2^r < v_2^r < \dots < z_n^r < v_n^r < z_{n+1}^r$) be the zeros of the derivative of the polynomial τ and let $v_0^r = -1, v_{n+1}^r = 1$. Then $v_0^r, v_1^r, \dots, v_{n+1}^r$ are all the points at which the polynomial τ achieves its extremal values in the interval $\langle -1, 1 \rangle$.

We shall denote by the symbol $\tau_{nk\theta}$ the polynomial τ , which satisfies the conditions formulated earlier, and such that for some $g \in (0, 1)$ and some integer k ($1 \leq k \leq n+1$) we have

$$\tau(v_j^r) = \begin{cases} (-1)^{n+1-j}g & (j = 0, 1, \dots, k-1), \\ (-1)^{n+1-j} & (j = k, k+1, \dots, n+1) \end{cases}$$

(we remember that $n+1$ is the degree of this polynomial). The existence and uniqueness of polynomials $\tau_{nk\theta}$ for arbitrary values of parameters n , k , and g have been proved in paper [19]. We easily see that for any k the polynomial τ_{nk1} is identical with the $(n+1)$ -st Chebyshev polynomial τ_{n+1} .

By definition (6) the numbers r_k^r and s_k^r (where $r_k^r \leq s_k^r$) are the roots of the equation $|\tau(t)| = g \|\tau\|_{\langle -1, 1 \rangle}$, which belong to the interval $\langle z_k^r, z_{k+1}^r \rangle$; the end points of this interval are the successive zeros of the polynomial τ or numbers $-1, 1$. The points r_k^r and s_k^r are separated by the point v_k^r , at which the polynomial τ achieves its extremal value. As $g \leq 1$, we have $\|\tau_{nk\theta}\|_{\langle -1, 1 \rangle} = 1$ and the point $r_k^{r_{nk\theta}}$ is the only root of the equation $|\tau_{nk\theta}(t)| = g$ which lies in the interval $\langle z_k^{r_{nk\theta}}, v_k^{r_{nk\theta}} \rangle$, where the function $|\tau_{nk\theta}|$ increases from 0 to 1. We shall use the symbol $t_{nk\theta} = r_k^{r_{nk\theta}}$.

THEOREM 9:4. *If $g \in (0, 1)$, then*

$$(11) \quad \min_{\tau \in \mathcal{W}_{n+1, g}} r_k^r = t_{nk\theta} \quad (k = 1, 2, \dots, n+1),$$

$$(12) \quad \max_{\tau \in \mathcal{W}_{n+1, g}} s_k^r = -t_{n, n+1-k, \theta} \quad (k = 0, 1, \dots, n).$$

Proof. If $g = 1$, then the class $\mathcal{W}_{n+1, g}$ consists of only one polynomial $\tau_{nk1} = \tau_{n+1}$ (up to a constant), since this is the only polynomial which satisfies the conditions

$$\|\tau\|_{\langle z_j^r, z_{j+1}^r \rangle} = \|\tau\|_{\langle -1, 1 \rangle} \quad (j = 0, 1, \dots, n+1)$$

which follow from (3). Thus, in this case, equation (11) is a consequence of the definition of the point t_{nk1} . For the polynomial $\tau = \tau_{n+1}$ the interval $\langle r_k^r, s_k^r \rangle$ reduces to the single point v_k^r at which the polynomial achieves its extremal value. According to formulas given in § 2.2 we have

$$r_k^{r_{n+1}} = s_k^{r_{n+1}} = v_k^{r_{n+1}} = -\cos k\pi/(n+1) \quad (k = 0, 1, \dots, n+1)$$

and equation (12) follows from the fact that

$$s_k^{r_{n+1}} = -\cos k\pi/(n+1) = -r_{n+1-k}^{r_{n+1}} = -t_{n, n+1-k, 1}.$$

Let us now assume that $0 < g < 1$. We have to verify that for every polynomial τ satisfying conditions (3) we have $r_k^r \geq r_k^{r_{nk\theta}} = t_{nk\theta}$. It suffices to consider the polynomials τ , for which $\|\tau\|_{\langle -1, 1 \rangle} = 1$ and $\tau(t) > 0$

for $t \geq 1$, as the multiplication of a polynomial τ by a number different from zero does not change the points r_k^τ and s_k^τ . Thus we may assume that

$$\tau = c \prod_{j=1}^{n+1} (t - z_j),$$

where $-1 < z_1 < z_2 < \dots < z_{n+1} < 1$. The upper indices in the symbols z_i^τ , r_i^τ , v_i^τ will be omitted from now on. Here the coefficient c is a positive number such that at some point v_m (where $0 \leq m \leq n+1$) we have the equation

$$(13) \quad |\tau(v_m)| = 1,$$

and at all points v_i we have the inequalities

$$(14) \quad g \leq |\tau(v_i)| \leq 1 \quad (i = 0, 1, \dots, n+1).$$

The polynomial τ is a function of parameters $c, z_1, z_2, \dots, z_{n+1}$ which determine it. We shall change these parameters in such a way that condition (13) for a fixed m and condition (14) will not be disturbed.

Let

$$w_i = \tau(v_i) = c \prod_{j=1}^{n+1} (v_i - z_j) \quad (i = 0, 1, \dots, n+1).$$

Taking the logarithms of both sides of this equation we compute the complete differential at the extremal value w_i (divided by w_i):

$$\frac{dw_i}{w_i} = \frac{dc}{c} + \sum_{j=1}^{n+1} \frac{dv_i - dz_j}{v_i - z_j} = \frac{dc}{c} + dv_i \sum_{j=1}^{n+1} \frac{1}{v_i - z_j} - \sum_{j=1}^{n+1} \frac{dz_j}{v_i - z_j}.$$

From the formula for the logarithmic derivative of the polynomial τ :

$$(15) \quad \frac{\tau'(t)}{\tau(t)} = \sum_{j=1}^{n+1} \frac{1}{t - z_j}$$

and from the definition of the points v_i it follows that

$$\sum_{j=1}^{n+1} \frac{1}{v_i - z_j} = 0 \quad (i = 1, 2, \dots, n).$$

Besides, $v_0 = -1, v_{n+1} = 1$ for any polynomial τ , i.e., $dv_0 = dv_{n+1} = 0$. Thus

$$(16) \quad \frac{dw_i}{w_i} = \frac{dc}{c} - \sum_{j=1}^{n+1} \frac{dz_j}{v_i - z_j}.$$

The value $w_m = \tau(v_m)$ is constant by condition (13), hence its complete differential is equal to zero and

$$(17) \quad \frac{dc}{c} = \sum_{j=1}^{n+1} \frac{dz_j}{v_m - z_j}.$$

Putting this value to equation (16) we get

$$(18) \quad \frac{dw_i}{w_i} = (v_i - v_m) \sum_{j=1}^{n+1} \frac{dz_j}{(v_i - z_j)(v_m - z_j)} \quad (i = 0, 1, \dots, n+1).$$

Let us now compute the complete differential of the root r_k of the equation $|\tau(t)| = g$. Since the value $\tau(r_k)$ is constant, we have $d\tau(r_k) = 0$ and

$$\frac{dc}{c} + dr_k \sum_{j=1}^{n+1} \frac{1}{r_k - z_j} - \sum_{j=1}^{n+1} \frac{dz_j}{r_k - z_j} = 0.$$

In view of (15) and (17) we have the equation

$$(19) \quad dr_k = (v_m - r_k) \frac{\tau(r_k)}{\tau'(r_k)} \sum_{j=1}^{n+1} \frac{dz_j}{(r_k - z_j)(v_m - z_j)},$$

provided that $\tau'(r_k) \neq 0$. The case $\tau'(r_k) = 0$ will be discussed later.

Let us split the system of numbers v_0, v_1, \dots, v_{n+1} into three subsets F , G , and H as follows:

$$v_i \in \begin{cases} F, & \text{if } |\tau(v_i)| = 1 \quad (\text{in particular if } i = m), \\ G, & \text{if } |\tau(v_i)| = g, \\ H, & \text{if } g < |\tau(v_i)| < 1. \end{cases}$$

We want to prove that if the polynomials τ and $\tau_{nk\varphi}$ are not identical, then there exists a polynomial φ such that $r_k^\varphi < r_k = r_k^{\tau}$. It follows from the previous considerations that such a polynomial may be obtained from τ by changing the parameters defining it, if: (i) there exist increments $dc, dz_1, dz_2, \dots, dz_{n+1}$ such that $dr_k < 0$; i.e.,

$$(20) \quad (v_m - r_k) \frac{\tau(r_k)}{\tau'(r_k)} \sum_{j=1}^{n+1} \frac{dz_j}{(r_k - z_j)(v_m - z_j)} < 0,$$

(ii) these increments determine the polynomial $\tau(t) + d\tau(t)$ which satisfies the conditions analogous to (13) and (14).

According to the definition of the sets F , G and H and formula (18), condition (ii) will be satisfied if

$$(21) \quad (v_i - v_m) \sum_{j=1}^{n+1} \frac{dz_j}{(v_i - z_j)(v_m - z_j)} \begin{cases} < 0 & (v_i \in F, i \neq m), \\ > 0 & (v_i \in G). \end{cases}$$

Here we do not have the inequalities which would correspond to the points of the subset H , since for sufficiently small increments the inequality $g < |\tau(v_i)| < 1$ implies the inequality $g \leq |\tau(v_i) + d\tau(v_i)| \leq 1$. Nor do we have the condition which would correspond to the point v_m , since the equation $|\tau(v_m) + d\tau(v_m)| = 1$ has already been considered in the derivation of formulas (18) and (19).

Let us also notice that in inequality (20) we can omit the factor $\tau(r_k)/\tau'(r_k)$. Indeed, if $\tau(r_k) > 0$, then $\tau'(r_k) > 0$, since the polynomial τ increases in the interval (z_k, v_k) (which contains the point r_k) from 0 to $\tau(v_k)$. Similarly, if $\tau(r_k) < 0$, then $\tau'(r_k) < 0$. If we introduce the notations:

$$(22) \quad x_j = \frac{dz_j}{v_m - z_j} \quad (j = 1, 2, \dots, n+1),$$

$$s_i = \begin{cases} -1 & (v_i \in F, i < m \text{ and } v_i \in G, i > m), \\ 1 & (v_i \in F, i > m \text{ and } v_i \in G, i < m), \\ \text{sign}(v_m - r_k) & (i = m), \end{cases}$$

$$y_i = \begin{cases} v_i & (i \neq m), \\ r_k & (i = m), \end{cases}$$

we can write the system of inequalities (20) and (21) in the form:

$$(23) \quad \sum_{j=1}^{n+1} \frac{s_i x_j}{y_i - z_j} < 0 \quad (v_i \in F \cup G).$$

One of the inequalities of this system, the one which concerns the point $y_m = r_k$, guarantees the condition $dr_k < 0$. We have derived it under the assumption $\tau'(r_k) \neq 0$. Now we shall show that if $\tau'(r_k) = 0$, then the condition $dr_k < 0$ is implied by another inequality from system (23); i. e., that in this case the system contains one inequality less.

If $\tau'(r_k) = 0$, then $r_k = v_k$, since the points r_k and v_k lie in the interval (z_k, z_{k+1}) , in which the derivative of the polynomial τ has only one zero. At the point $r_k = v_k$ we have $|\tau(r_k)| = g$, hence this point belongs to the subset G . In system (21), introduced to (23), we have the inequality which corresponds to this point:

$$(v_k - v_m) \sum_{j=1}^{n+1} \frac{dz_j}{(v_k - z_j)(v_m - z_j)} > 0.$$

If we compare this inequality with the formula for the complete differential of the polynomial τ at the fixed point t :

$$\frac{d\tau(t)}{\tau(t)} = \frac{dc}{c} - \sum_{j=1}^{n+1} \frac{dz_j}{t - z_j} = (t - v_m) \sum_{j=1}^{n+1} \frac{dz_j}{(t - z_j)(v_m - z_j)},$$

we see that $d\tau(t)/\tau(t) > 0$ for $t = r_k$. This means, for instance, if $\tau(v_k) > 0$, then at the point r_k the polynomial $\tau(t) + d\tau(t)$ has the value greater than $\tau(v_k) = g$. Thus the point $r_k + dr_k$, at which this polynomial achieves the value g lies on the left from the point r_k and $dr_k < 0$.

System (23) has as many inequalities, as many points v_0, v_1, \dots, v_{n+1} contain the subsets F and G . We already know that in the case $\tau'(r_k) = 0$, system (23) has one inequality less, and hence consists of at most $n+1$ inequalities. The determinants of matrices of this reduced system with $n+1$ unknowns x_1, x_2, \dots, x_{n+1} are different from zero (they were computed in the proof of Theorem 9:3), thus the system has a solution. For the same reasons system (23) has a solution, if $\tau'(r_k) \neq 0$, and, at the same time, not all the points v_0, v_1, \dots, v_{n+1} belong either to F or to G .

Let us now assume that $\tau'(r_k) \neq 0$ and $F \cup G = \{v_0, v_1, \dots, v_{n+1}\}$, i.e., that system (23) may be written in the form

$$\sum_{j=1}^{n+1} \frac{s_i x_j}{y_i - z_j} < 0 \quad (i = 0, 1, \dots, n+1).$$

Theorem 9:3 states that the system obtained does not have a solution if and only if the numbers (8) have a common sign. Let us now investigate the signs of these numbers, considering the notations in (22) and the inequality $v_0 < z_1 < v_1 < \dots < z_k < r_k < v_k < \dots < z_{n+1} < v_{n+1}$.

If $i = m$, then

$$\text{sign} \prod_{j=1}^{n+1} (z_j - y_i) = \text{sign} \prod_{j=1}^{n+1} (z_j - r_k) = (-1)^k,$$

$$\text{sign} s_i \prod_{j=0, j \neq i}^{n+1} (y_j - y_i) = s_m \prod_{j=0, j \neq m}^{n+1} (v_j - r_k) = \text{sign} \prod_{j=0}^{n+1} (v_j - r_k) = (-1)^k,$$

hence

$$(24) \quad \frac{\prod_{j=1}^{n+1} (z_j - y_m)}{s_m \prod_{j=0, j \neq m}^{n+1} (y_j - y_m)} > 0.$$

On the other hand, if $i \neq m$, then

$$\begin{aligned} \operatorname{sign} \prod_{j=1}^{n+1} (z_j - y_i) &= \operatorname{sign} \prod_{j=1}^{n+1} (z_j - v_i) = (-1)^i, \\ \operatorname{sign} s_i \prod_{j=0, j \neq i}^{n+1} (y_j - y_i) &= s_i \operatorname{sign}(r_k - v_i) \operatorname{sign} \prod_{j=0, j \neq i, m}^{n+1} (v_j - v_i) \\ &= s_i \operatorname{sign}(r_k - v_i) (v_m - v_i) \operatorname{sign} \prod_{j=0, j \neq i}^{n+1} (v_j - v_i) \\ &= (-1)^i s_i \operatorname{sign}(r_k - v_i) (v_m - v_i) \\ &= \begin{cases} (-1)^{i+1} \operatorname{sign}(r_k - v_i) & (v_i \in F), \\ (-1)^i \operatorname{sign}(r_k - v_i) & (v_i \in G), \end{cases} \\ \operatorname{sign} \frac{\prod_{j=1}^{n+1} (z_j - y_i)}{s_i \prod_{j=0, j \neq i}^{n+1} (y_j - y_i)} &= \begin{cases} -\operatorname{sign}(r_k - v_i) & (v_i \in F), \\ \operatorname{sign}(r_k - v_i) & (v_i \in G). \end{cases} \end{aligned}$$

By (24) the numbers (8) have a common sign if and only if $r_k - v_i < 0$ for $v_i \in F$ and $r_k - v_i > 0$ for $v_i \in G$; i.e.—according to the definition of the sets F and G —if

$$\begin{aligned} |\tau(v_0)| &= |\tau(v_1)| = \dots = |\tau(v_{k-1})| = g, \\ |\tau(v_k)| &= |\tau(v_{k+1})| = \dots = |\tau(v_{n+1})| = 1, \end{aligned}$$

and the last equations are just the ones that define the polynomial $\tau_{nk\theta}$. Thus the root r_k^τ of the equation $|\tau(t)| = g \|\tau\|_{\langle -1, 1 \rangle}$ achieves its minimal value for $\tau = \tau_{nk\theta}$. Thus we have proved equation (11).

To prove equation (12) let us consider the polynomial $\sigma(t) = \tau(-t)$, whose zeros are $z_1^\sigma = -z_{n+1}^\tau < z_2^\sigma = -z_n^\tau < \dots < z_{n+1}^\sigma = -z_1^\tau$. The zeros of the derivative of the polynomial σ are $v_1^\sigma = -v_n^\tau < v_2^\sigma = -v_{n-1}^\tau < \dots < v_n^\sigma = -v_1^\tau$. The number s_k^τ has been defined as the root of the equation $|\tau(t)| = g \|\tau\|_{\langle -1, 1 \rangle}$, which belongs to the interval $\langle v_k^\tau, z_{k+1}^\tau \rangle$. At the same time $-s_k^\tau$ is the root of the equation $|\sigma(t)| = g \|\sigma\|_{\langle -1, 1 \rangle}$, which belongs to the interval $\langle -z_{k+1}^\tau, -v_k^\tau \rangle = \langle z_{n+1-k}^\sigma, v_{n+1-k}^\sigma \rangle$, hence it follows that $s_k^\tau = -r_{n+1-k}^\sigma$. Since the polynomials $\tau(-t)$ run over the same class $\mathscr{W}_{n+1, g}$ as the polynomials $\tau(t)$, we obtain, according to equation (11) proved above, that

$$\max_{\tau \in \mathscr{W}_{n+1, g}} s_k^\tau = - \min_{\sigma \in \mathscr{W}_{n+1, g}} r_{n+1-k}^\sigma = -t_{n, n+1-k, g}.$$

9.4. Before we present the estimation of the (n) -points of the function ξ in its final form, we shall prove an auxiliary theorem which specifies the dependence t_{nkg} on g .

THEOREM 9:5. For any values of parameters n and k the quantity t_{nkg} is an increasing function of g .

Proof. The class $\mathcal{W}_{n+1, g}$ consists of the polynomials of the degree $n+1$ which satisfy inequality (3). If $g' > g$, then $\mathcal{W}_{n+1, g'} \subset \mathcal{W}_{n+1, g}$, hence $\min_{\tau \in \mathcal{W}_{n+1, g'}} r_k^i \geq \min_{\tau \in \mathcal{W}_{n+1, g}} r_k^i$, or, in view of equation (11), $t_{nkg'} \geq t_{nkg}$. Thus we know already that t_{nkg} is a not-decreasing function of g .

Suppose now, contrary to the assertion of our theorem, that $t_{nkg} = t_{nkg'}$ for some numbers g and $g' > g$ from the interval $(0, 1)$. Let us consider the polynomial from the class \mathcal{W}_{n+1} :

$$(25) \quad \delta = g' \tau_{nkg} - g'' \tau_{nkg'},$$

where $g < g'' < g'$.

From the definition of the points $v_j^{inkg'}$ and $t_{nkg'}$ it follows that

$$\begin{aligned} \tau_{nkg'}(v_j^{inkg'}) &= (-1)^{n+1-j} g' \quad (j = 0, 1, \dots, k-1), \\ \tau_{nkg'}(t_{nkg'}) &= (-1)^{n+1-k} g'. \end{aligned}$$

On the other hand, in the whole interval $\langle -1, t_{nkg} \rangle = \langle -1, t_{nkg'} \rangle$ we have the inequality $|\tau_{nkg}(t)| \leq g$. Thus at the points $v_j^{inkg'}$ ($j = 0, 1, \dots, k-1$) and $t_{nkg'}$ the absolute value of the subtrahend in the difference (25) equals $g'g''$, and the absolute value of the minuend is not greater than $g'g < g'g''$, hence

$$\begin{aligned} \text{sign } \delta(v_j^{inkg'}) &= -\text{sign } \tau_{nkg'}(v_j^{inkg'}) = (-1)^{n-j} \quad (j = 0, 1, \dots, k-1), \\ \text{sign } \delta(t_{nkg'}) &= (-1)^{n-k}. \end{aligned}$$

From the definition of the points v_j^{inkg} it follows that

$$\tau_{nkg}(v_j^{inkg}) = (-1)^{n+1-j} \quad (j = k, k+1, \dots, n+1).$$

In the whole interval $\langle t_{nkg'}, 1 \rangle$ we have the inequality $|\tau_{nkg}(t)| \leq 1$. Thus at the points v_j^{inkg} ($j = k, k+1, \dots, n+1$) the absolute value of the minuend of difference (25) equals g' , and the absolute value of the subtrahend of this difference does not exceed $g'' < g'$, hence

$$\text{sign } \delta(v_j^{inkg}) = \text{sign } \tau_{nkg}(v_j^{inkg}) = (-1)^{n+1-j} \quad (j = k, k+1, \dots, n+1).$$

In this way we have found $n+3$ points $v_0^{inkg'} < v_1^{inkg'} < \dots < v_{k-1}^{inkg'} < t_{nkg'} = t_{nkg} < v_k^{inkg} < v_{k+1}^{inkg} < \dots < v_{n+1}^{inkg}$ at which the polynomial $\delta \in \mathcal{W}_{n+1}$ has alternately positive and negative values, which is impossible. This contradiction has arisen from the assumption that $t_{nkg} = t_{nkg'}$ for $g < g'$.

THEOREM 9:6. *If, for a given number $h \in (0, 1)$, the function $\xi \in \mathcal{C}_{(-1, 1)}$ satisfies the inequality*

$$\frac{\varepsilon_{n+1}(\xi)}{\varepsilon_n(\xi)} \leq h,$$

and the points u_0, u_1, \dots, u_{n+1} , where $-1 \leq u_0 < u_1 < \dots < u_{n+1} \leq 1$, are alternately the $(n, +)$ -points and the $(n, -)$ -points of this function, then

$$(26) \quad \begin{cases} -1 \leq u_0 \leq -t_{n, n+1, (1-h)/(1+h)}, \\ t_{n, k, (1-h)/(1+h)} \leq u_k \leq -t_{n, n+1-k, (1-h)/(1+h)} \quad (k = 1, 2, \dots, n), \\ t_{n, n+1, (1-h)/(1+h)} \leq u_{n+1} \leq 1. \end{cases}$$

Proof. From Theorems 9:2 and 9:4 we have the inequalities

$$\begin{aligned} -1 \leq u_0 &\leq -t_{n, n+1, g}, \\ t_{n, kg} \leq u_k &\leq -t_{n, n+1-k, g} \quad (k = 1, 2, \dots, n), \end{aligned}$$

$$t_{n, n+1, g} \leq u_{n+1} \leq 1,$$

where

$$g = \frac{\varepsilon_n(\xi) - \varepsilon_{n+1}(\xi)}{\varepsilon_n(\xi) + \varepsilon_{n+1}(\xi)}.$$

Since $\varepsilon_{n+1}(\xi)/\varepsilon_n(\xi) \leq h$, we have $g \geq (1-h)/(1+h)$ and Theorem 9:5 provides us directly with inequalities (26).

Theorem 9:6 in the above form cannot be applied if $\varepsilon_n(\xi) = \varepsilon_{n+1}(\xi)$. However in this case—provided that $\varepsilon_n(\xi) > 0$ (the case $\varepsilon_n(\xi) = 0$ is not of interest here)—there exists an integer $m \geq 2$ for which $\varepsilon_n(\xi) = \varepsilon_{n+1}(\xi) = \dots = \varepsilon_{n+m-1}(\xi) > \varepsilon_{n+m}(\xi)$. Then all the (n) -points of the function ξ , are, at the same time, its $(n+m-1)$ -points. The estimates for these points can be obtained from inequalities (26), if we replace the number n by $n+m-1$, and

$$\frac{\varepsilon_{n+m}(\xi)}{\varepsilon_n(\xi)} = \frac{\varepsilon_{n+m}(\xi)}{\varepsilon_{n+m-1}(\xi)} \leq h < 1.$$

Let us note also that Theorem 9:6 gives an estimate of the (n) -points of the function ξ , which depends only upon the value of the ratio $\varepsilon_{n+1}(\xi)/\varepsilon_n(\xi)$, hence is not connected with the analytical properties of this function.

9.5. It was stated in § 7.3 that the theorems of the present section can provide us with some new estimates of the error of the best approximation. To obtain these estimates, let us consider once more the value t_{nkg} . We already know from Theorem 9:5, that it is an increasing func-

tion of g . Besides, it is easy to see that $t_{nk0} = -1$ (or, more precisely, $\lim_{g \rightarrow 0^+} t_{nk0} = -1$, since we have defined t_{nk0} only for $g > 0$) and that $t_{nk1} = -\cos k\pi/(n+1)$. The first equation follows from the fact that for $g \rightarrow 0$ the polynomial τ_{nk0} tends to the polynomial whose zeros z_1, z_2, \dots, z_k are equal to -1 , and t_{nk0} is, by definition, the root of the equation $|\tau_{nk0}(t)| = g$ which belongs to the interval (z_k, v_k) . The second equation was mentioned earlier, in the beginning of the proof of Theorem 9:4.

From the above remarks it follows that as h increases from 0 to 1, the value of $t_{n,k,(1-h)/(1+h)}$ decreases from $-\cos k\pi/(n+1)$ to -1 , and the value of $-t_{n,n+1-k,(1-h)/(1+h)}$ increases from $-\cos k\pi/(n+1)$ to 1. Thus, to an arbitrary number $u \in \langle -1, 1 \rangle$ and arbitrary $k = 1, 2, \dots, n$ there corresponds one and only one number $h_{ku} \in \langle 0, 1 \rangle$ such that

$$(27) \quad u = t_{n,k,(1-h_{ku})/(1+h_{ku})} \quad \text{OR} \quad u = -t_{n,n+1-k,(1-h_{ku})/(1+h_{ku})}.$$

As we see, these equations may be satisfied simultaneously only if $u = -\cos k\pi/(n+1)$ for $h_{ku} = 0$.

Similarly, using the equations

$$(28) \quad u = -t_{n,n+1,(1-h_{0u})/(1+h_{0u})}, \quad u = t_{n,n+1,(1-h_{n+1,u})/(1+h_{n+1,u})}$$

we can uniquely determine the numbers h_{0u} and $h_{n+1,u}$.

THEOREM 9:7. *If u_0, u_1, \dots, u_{n+1} , where $u_0 < u_1 < \dots < u_{n+1}$, are alternately the $(n, +)$ -points and $(n, -)$ -points of the function ξ , and if the numbers $h_{0u_0}, h_{1u_1}, \dots, h_{n+1,u_{n+1}}$ are defined by (27) and (28), then we have the inequality*

$$\varepsilon_{n+1}(\xi) \geq \varepsilon_n(\xi) \max_{0 \leq k \leq n+1} h_{ku_k}.$$

In fact, suppose that for some k we have

$$\varepsilon_{n+1}(\xi)/\varepsilon_n(\xi) = h^* < h_{ku_k}.$$

Then, putting h^* in place of h in inequalities (26), which correspond to this particular k , we arrive at the contradiction with the definition of the number h_{ku_k} .

10. Auxiliary theorems.

10.1. Now we shall prove some theorems which will be applied later in § 11. We shall derive some estimates of the (n) -points which depend upon the properties of the derivatives of the approximated functions.

THEOREM 10:1. *If a sequence, which consists of l zeros and ones contains j pairs of neighbouring ones, then it contains at least $-[(j-l+1)/2]$ zeros⁽³⁾.*

⁽³⁾ $[x]$ denotes the integral part of the number x .

Proof. Let us denote by z the number of zeros in the sequence under consideration, the number of ones preceding the first zero by y_0 , the number of ones following the i -th and preceding the $(i+1)$ -st zero by y_i ($i = 1, 2, \dots, z-1$), and finally, the number of ones which follow the last zero by y_z . This means that $y_0 + y_1 + \dots + y_z = l - z$. Since y_i neighbouring ones form at least $y_i - 1$ pairs of neighbouring ones (more precisely, they form $y_i - 1$ pairs if $y_i \geq 2$ and 0 pairs if $y_i < 2$), we have

$$j \geq (y_0 - 1) + (y_1 - 1) + \dots + (y_z - 1) = l - z - (z + 1) = l - 2z - 1.$$

It follows that $-z \leq [(j - l + 1)/2]$, as asserted.

Theorem 10:1 remains true, if we replace in it the ones by any numbers different from zero.

In the next two theorems, and also in § 11, we shall estimate the number of zeros of some functions. If the function under investigation is differentiable p times we can introduce the notion of its zero of the order k ($1 \leq k \leq p + 1$). By definition, the point t_0 is a zero of the order k ($1 \leq k \leq p$) of the function η if $\eta(t_0) = \eta'(t_0) = \dots = \eta^{(k-1)}(t_0) = 0$, $\eta^{(k)}(t_0) \neq 0$ and the zero of the order $p + 1$ is a point t_0 such that $\eta(t_0) = \eta'(t_0) = \dots = \eta^{(p)}(t_0) = 0$. The expression "the number of zeros with their orders" means the sum of all orders of zeros of the function η .

THEOREM 10:2. *If: (i) the functions β and γ are differentiable in the interval $I = \langle -1, 1 \rangle$; (ii) $\|\beta\|_I = \|\gamma\|_I$; (iii) there exist points $-1 = c_0 < c_1 < \dots < c_p = 1$ such that $\gamma(c_k) = (-1)^k \|\gamma\|_I$ for $k = 0, 1, \dots, p$ or $\gamma(c_k) = (-1)^{k+1} \|\gamma\|_I$ for the same k , then each of the functions $\delta_- = \beta - \gamma$ and $\delta_+ = \beta + \gamma$ has, in the intervals $\langle -1, c_l \rangle$ and $\langle c_l, 1 \rangle$ for $0 \leq l \leq p$, at least l and $p - l$ zeros with their orders, respectively.*

Proof. Let us notice first that if we succeed in proving that the function δ_- has at least l zeros (with their orders) in the interval $\langle -1, c_l \rangle$, the rest of the assertion will follow. Indeed, let $\beta^*(t) = \beta(-t)$, $\gamma^*(t) = \gamma(-t)$, $c_k^* = -c_{p-k}$ ($k = 0, 1, \dots, p$). The functions β^* , γ^* and the points c_k^* satisfy all the assumptions of Theorems 10:2. From this part of its assertion which concerns the function δ_- in the interval $\langle -1, c_l \rangle$ it would then follow for $l^* = p - l$, that the function $\delta^* = \beta^* - \gamma^*$ has at least l^* zeros in the interval $\langle -1, c_{l^*}^* \rangle = \langle -1, c_{p-l}^* \rangle = \langle -1, -c_l \rangle$. This, in turn would imply that the function $\delta_-(t) = \delta^*(-t)$ has at least $p - l$ zeros in the interval $\langle c_l, 1 \rangle$, which coincides with the assertion of our theorem.

The pair of functions β and $-\gamma$ also satisfies the assumption of the theorem. Since $\delta_+ = \beta - (-\gamma)$, the part of assertion which concerns the zeros of the function δ_+ can be obtained directly from the part which concerns the function δ_- .

Let us now investigate the properties of the function δ_- in the interval $\langle -1, c_l \rangle$, where $0 < l \leq p$ (for $l = 0$ the theorem is obvious). We notice easily that if $\delta_-(c_i) = 0$ for some $c_i \in (-1, 1)$, then the point c_i is a double zero of the function δ_- ; i.e., $\delta'_-(c_i) = 0$. In fact, by (iii) we have $|\gamma(c_i)| = \|\gamma\|_I$. Hence, from the equation $\delta_-(c_i) = 0$ it follows that $|\beta(c_i)| = |\gamma(c_i)| = \|\gamma\|_I = \|\beta\|_I$. Thus, at the point c_i both functions β, γ have their extremum, and since this point lies inside the interval I , we have $\beta'(c_i) = \gamma'(c_i) = 0$.

If $\|\gamma\|_I > 0$ (and only this case is of any interest for us), then from (iii) it follows that

$$(1) \quad \gamma(c_k)\gamma(c_{k+1}) < 0 \quad (k = 0, 1, \dots, p-1).$$

Suppose that in the sequence $\delta_-(c_0), \delta_-(c_1), \dots, \delta_-(c_p)$ two neighbouring numbers, e.g. $\delta_-(c_i)$ and $\delta_-(c_{i+1})$, are different from zero. The inequality $\delta_-(c_i) \neq 0$, or $\beta(c_i) \neq \gamma(c_i)$, is equivalent to the fact that $|\beta(c_i)| < |\gamma(c_i)|$ or $\beta(c_i) = -\gamma(c_i)$. In both cases the difference $\delta_-(c_i) = \beta(c_i) - \gamma(c_i)$ and the value $\gamma(c_i)$ have opposite signs. The signs of $\delta_-(c_{i+1})$ and $\gamma(c_{i+1})$ are also different. Thus from (1) for $k = i$ it follows that $\delta_-(c_i)\delta_-(c_{i+1}) < 0$ and the continuous function δ_- has a zero in the interval (c_i, c_{i+1}) , which corresponds to the numbers $\delta_-(c_i)$ and $\delta_-(c_{i+1})$.

Case I. $\delta_-(c_0) = 0$. Let us denote by j the number of neighbouring numbers different from zero in the sequence

$$(2) \quad \delta_-(c_1), \delta_-(c_2), \dots, \delta_-(c_l),$$

which consists of l numbers. We already know that for an arbitrary pair of numbers $\delta_-(c_i) \neq 0, \delta_-(c_{i+1}) \neq 0$ there exists a zero of the function δ_- which belongs to the interval (c_i, c_{i+1}) . By Theorem 10:1, sequence (2) contains at least $-[(j-l+1)/2]$ zeros.

If $l < p$ or $l = p$ and $\delta_-(c_p) \neq 0$, then to the number $\delta_-(c_k) = 0$ in sequence (2) there corresponds the point c_k which lies inside the interval I . We have previously proved that every such point is a double zero of the function δ_- . Finally, from the assumption of case I the point $c_0 = -1$ is also a zero of the function δ_- . Thus we know that this function has at least

$$(3) \quad j - 2[(j-l+1)/2] + 1$$

zeros together with their orders in the interval $\langle -1, c_l \rangle$.

If $l = p$ and $\delta_-(c_p) = 0$, then in the sequence $\delta_-(c_1), \delta_-(c_2), \dots, \delta_-(c_{p-1})$ there exists j non-vanishing pairs, and at least $-[(j-p+1+1)/2]$ zeros. The numbers c_0 and c_p are the zeros of the function δ_- and in the

whole interval $\langle -1, c_p \rangle = \langle -1, 1 \rangle$ it has at least

$$(4) \quad j - 2[(j - p + 2)/2] + 2 = j - 2[(j - p)/2]$$

zeros. Since the number (3) equals either l or $l + 1$ and the number (4), which concerns the case when $l = p$, equals either p or $p + 1$, Theorem 10:2 is true in case I.

Case II. $\delta_-(c_0) \neq 0$. Let us consider the sequence $\delta_-(c_0), \delta_-(c_1), \dots, \delta_-(c_l)$ consisting of $l + 1$ number. Let us denote by j the number of pairs of neighbouring non-vanishing numbers in this sequence, and let us proceed similarly as in case I. We prove the existence of

$$j - 2[(j - l)/2] \quad (l < p \text{ or } l = p \text{ and } \delta_-(c_p) \neq 0),$$

$$j - 2[(j - p + 1)/2] + 1 \quad (l = p \text{ and } \delta_-(c_p) = 0)$$

zeros together with their orders of the function δ_- . Since these numbers are not smaller than l and p , respectively, we completed the proof of Theorem 10:2.

THEOREM 10:3 ([24], p. 47). *If: (i) the functions β and γ are continuous in the interval $\langle a, b \rangle$, (ii) $\|\beta\|_{\langle a, b \rangle} = \|\gamma\|_{\langle a, b \rangle} = w > 0$, (iii) there exist points $a = a_{\beta 0} < a_{\beta 1} < \dots < a_{\beta k} = b$, $a = a_{\gamma 0} < a_{\gamma 1} < \dots < a_{\gamma l} = b$ such that $\beta(a_{\beta i}) = (-1)^i w$ for $i = 0, 1, \dots, k$ and $\gamma(a_{\gamma j}) = (-1)^j w$ for $j = 0, 1, \dots, l$, (iv) $a_{\beta i} \neq a_{\gamma j}$ for an even difference $i - j$ and for $0 < i < k$, $0 < j < l$, (v) $|\beta(t)| < w$ for $t \neq a_{\beta i}$, and $|\gamma(t)| < w$ for $t \neq a_{\gamma j}$, then the function $\beta - \gamma$ has at least*

$$(5) \quad \max\{k, l\} - 1 - \text{sign}|k - l| \quad (k - l \text{ even}),$$

$$(6) \quad \max\{k, l\} - 1 \quad (k - l \text{ odd})$$

distinct zeros in the interval (a, b) .

Proof. By assumptions (iv) and (v) none of the points $a_{\beta 1}, a_{\beta 2}, \dots, a_{\beta, k-1}, a_{\gamma 1}, a_{\gamma 2}, \dots, a_{\gamma, l-1}$ are a zero of the function $\beta - \gamma$. In fact, if they were, e.g., $\beta(a_{\beta i}) - \gamma(a_{\beta i}) = 0$, then $\gamma(a_{\beta i}) = \beta(a_{\beta i}) = (-1)^i w$, and $a_{\beta i} = a_{\gamma j}$ for j such that $(-1)^i = (-1)^j$; i.e., for j such that the difference $i - j$ is even. Thus, at each point $a_{\beta i}$ the function $\beta - \gamma$ has a sign consistent with the sign of the function β , i.e., equal to $(-1)^i$, and at each point $a_{\gamma j}$ it has the opposite sign than the function γ , i.e. it is equal to $(-1)^{j+1}$.

Case I. $k = l$. We must prove the existence of $\max\{k, l\} - 1 - \text{sign}|k - l| = k - 1$ zeros of the function $\beta - \gamma$.

Case I. 1. k is even. Let us arrange according to the magnitude, in one sequence consisting of k points, the points $a_{\beta 1}, a_{\beta 3}, \dots, a_{\beta, k-1}$ (at which $\beta(t) - \gamma(t) < 0$) and the points $a_{\gamma 1}, a_{\gamma 3}, \dots, a_{\gamma, l-1} = a_{\gamma, k-1}$ (at which $\beta(t) - \gamma(t) > 0$). It suffices to prove that among every pair of successive points of this sequence lies a zero of the function $\beta - \gamma$. It is

true if one point belongs to the system $\{a_{\beta i}\}$ and the other belongs to the system $\{a_{\gamma j}\}$, since in this case it would follow from the remark made at the beginning, that the difference $\beta - \gamma$ has the values of opposite signs at these points. If these successive points belong to one system, and are equal, e.g., $a_{\beta 1}$ and $a_{\beta 3}$, then the difference $\beta - \gamma$ changes its sign twice in the interval $(a_{\beta 1}, a_{\beta 3})$: between the points $a_{\beta 1}$ and $a_{\beta 2}$, and between the points $a_{\beta 2}$ and $a_{\beta 3}$.

Case I.2. k is odd. Let us arrange the $k-1$ points $a_{\beta 1}, a_{\beta 3}, \dots, a_{\beta, k-2}, a_{\gamma 1}, a_{\gamma 3}, \dots, a_{\gamma, k-2}$ in a sequence whose elements will be denoted by d_1, d_2, \dots, d_{k-1} , where $d_1 < d_2 < \dots < d_{k-1}$. The elements of an analogous sequence, obtained by arranging the points $a_{\beta 2}, a_{\beta 4}, \dots, a_{\beta, k-1}, a_{\gamma 2}, a_{\gamma 4}, \dots, a_{\gamma, k-1}$ will be denoted by e_1, e_2, \dots, e_{k-1} . The reasoning in case I.1 shows that in each of the intervals

$$(7) \quad (d_1, d_2), (d_2, d_3), \dots, (d_{k-2}, d_{k-1}),$$

$$(8) \quad (e_1, e_2), (e_2, e_3), \dots, (e_{k-2}, e_{k-1})$$

there exists a zero of the function $\beta - \gamma$. If, however, the end points of one of these intervals belong to the same system $\{a_{\beta i}\}$ or $\{a_{\gamma j}\}$, then such an interval would contain two zeros of the function $\beta - \gamma$. Thus the existence of $k-1$ distinct zeros of the function $\beta - \gamma$ becomes obvious if: I.2.1. *there exists an interval from group (7) disjoint from the interval (e_1, e_{k-1}) , or there exists an interval from group (8) disjoint from the interval (d_1, d_{k-1}) , which, in view of definition of the points d_i and e_i is equivalent to the condition*

$$d_2 \leq e_1 \quad \text{or} \quad d_{k-1} \leq e_{k-2}.$$

The same assertion is true if: I.2.2. *end points of some of the intervals (7) or (8) belong to the same system $\{a_{\beta i}\}$ or $\{a_{\gamma j}\}$.*

Let us now assume that neither of the possibilities I.2.1, I.2.2 holds. We know that $d_1 = \min\{a_{\beta 1}, a_{\gamma 1}\}$, $e_1 = \min\{a_{\beta 2}, a_{\gamma 2}\}$. Let, for instance, $a_{\beta 1} < a_{\gamma 1}$. Then $d_1 = a_{\beta 1}$ and d_2 is equal to $a_{\gamma 1}$ or $a_{\beta 3}$. From the negation of condition I.2.2 it follows that $d_2 = a_{\gamma 1}$; i.e., $a_{\gamma 1} < a_{\beta 3}$. From the definition of the number e_1 and from the inequality $e_1 < d_2$ (i.e. the negation of I.2.1) it follows also, that $e_1 = a_{\beta 2}$. Thus the function $\beta - \gamma$ has, besides the $k-2$ zeros in the intervals (8), one more zero in the interval $(d_1, e_1) = (a_{\beta 1}, a_{\beta 2})$ disjoint with those intervals.

Case II. $k-l$ is even, $k \neq l$. Since the numbers k and l play a symmetric role in Theorem 10:2, we may assume that $k > l$. Then $\max\{k, l\} - 1 - \text{sign}|k-l| = k-2$. Theorem 10:3 in case II follows from the fact that the function $\beta - \gamma$ has a zero in each of the $k-2$ intervals $(a_{\beta 1}, a_{\beta 2}), \dots, (a_{\beta, k-2}, a_{\beta, k-1})$.

Case III. $k - l$ is odd. Assuming, as before, that $k > l$, we get the equation $\max\{k, l\} - 1 = k - 1$. The function $\beta - \gamma$ has a zero in each of the $k - 2$ intervals $(a_{\beta 1}, a_{\beta 2}), \dots, (a_{\beta, k-2}, a_{\beta, k-1})$ and also in the interval $(a_{\beta, k-1}, a_{\beta k}) = (a_{\beta, k-1}, b)$ (only for an odd $k - l$), since its values at each of the end points of these intervals are of opposite signs.

It is easy to see that Theorem 10:3 remains true if we replace assumptions (iii) and (iv) by the equations $\beta(a_{\beta i}) = (-1)^{k-i}w, \gamma(a_{\gamma j}) = (-1)^{l-j}w$ and the condition $a_{\beta i} \neq a_{\gamma j}$ for even values of the difference $(k - i) - (l - j)$ and for $0 < i < k, 0 < j < l$. The proof of this modification of Theorem 10:3 is based upon the introduction of the functions $\beta^*(t) = \beta(a + b - t), \gamma^*(t) = \gamma(a + b - t)$ and the points $a_{\beta i}^* = a + b - a_{\beta, k-i}, a_{\gamma j}^* = a + b - a_{\gamma, l-j}$. They satisfy all the assumptions of Theorem 10:3 in its original form.

10.2. In the next theorem we shall determine the sign of a functional determinant. But first, we shall present some necessary facts from the theory of interpolation. It is well known that the coefficient with t^p in the interpolation polynomial for the function η and for the nodes t_0, t_1, \dots, t_p is equal to the p -th divided difference $\eta[t_0, t_1, \dots, t_p]$ (see [18], p. 3; from the definition:

$$\eta[t_i] = \eta(t_i), \quad \eta[t_i, t_{i+1}, \dots, t_{i+k}] = \frac{\eta[t_{i+1}, \dots, t_{i+k}] - \eta[t_i, \dots, t_{i+k-1}]}{t_{i+k} - t_i}.$$

The same coefficient can be computed by using directly the equations which determine the interpolation polynomial. Thus we obtain the equation:

$$(9) \quad \eta[t_0, t_1, \dots, t_p] = \begin{vmatrix} 1 & t_0 & \dots & t_0^{p-1} & \eta(t_0) & 1 & t_0 & \dots & t_0^{p-1} & t_0^p \\ 1 & t_1 & \dots & t_1^{p-1} & \eta(t_1) & 1 & t_1 & \dots & t_1^{p-1} & t_1^p \\ \dots & \dots \\ 1 & t_p & \dots & t_p^{p-1} & \eta(t_p) & 1 & t_p & \dots & t_p^{p-1} & t_p^p \end{vmatrix} :$$

The denominator of the right-hand side equals the Vandermonde determinant, and is positive if $t_0 < t_1 < \dots < t_p$.

The class of functions, which have the continuous p -th derivative in the interval $I = \langle a, b \rangle$ will be denoted by \mathcal{C}_I^p . If the function η belongs to this class and $a \leq t_0 < t_1 < \dots < t_p \leq b$, then

$$\frac{1}{p!} \min_{t \in I} \eta^{(p)}(t) \leq \eta[t_0, t_1, \dots, t_p] \leq \frac{1}{p!} \max_{t \in I} \eta^{(p)}(t)$$

(see [18], p. 6). Thus, if $\eta \in \mathcal{C}_I^p$ and the p -th derivative of the function η has a constant sign in the interval I (and is everywhere positive or every-

where negative), then from the above inequality and from (9) it follows that

$$(10) \quad \text{sign} \begin{vmatrix} 1 & t_0 & \dots & t_0^{\nu-1} & \eta(t_0) \\ 1 & t_1 & \dots & t_1^{\nu-1} & \eta(t_1) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & t_p & \dots & t_p^{\nu-1} & \eta(t_p) \end{vmatrix} = \text{sign } \eta^{(\nu)}(t).$$

THEOREM 10:4. *If $I = \langle a, b \rangle$, $\xi \in \mathcal{C}_I^{\nu+1}$, $\gamma \in \mathcal{C}_I^{\nu+1}$, and if the following assumptions are true:*

- (Ξ_1) *the functions $\xi^{(\nu+1)}$ and $\gamma^{(\nu)}$ have a constant sign in the interval I ,*
- (Ξ_2) *$|\xi^{(\nu+1)}(t)| \geq |\gamma^{(\nu+1)}(t)|$ for $t \in I$,*
- (Ξ_3) *$|\xi^{(\nu)}(t)| < |\gamma^{(\nu)}(t)|$ for $t \in I$,*

or the assumptions

- (Γ_1) *the functions $\xi^{(\nu)}$ and $\gamma^{(\nu+1)}$ have a constant sign in the interval I ,*
- (Γ_2) *$|\xi^{(\nu+1)}(t)| \leq |\gamma^{(\nu+1)}(t)|$ for $t \in I$,*
- (Γ_3) *$|\xi^{(\nu)}(t)| > |\gamma^{(\nu)}(t)|$ for $t \in I$,*

then for an arbitrary system of points t_0, t_1, \dots, t_{p+1} , such that $a \leq t_0 < t_1 < \dots < t_{p+1} \leq b$, we have the equation

$$(11) \quad \text{sign} \begin{vmatrix} 1 & t_0 & \dots & t_0^{\nu-1} & \gamma(t_0) & \xi(t_0) \\ 1 & t_1 & \dots & t_1^{\nu-1} & \gamma(t_1) & \xi(t_1) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & t_{p+1} & \dots & t_{p+1}^{\nu-1} & \gamma(t_{p+1}) & \xi(t_{p+1}) \end{vmatrix} = \begin{cases} \text{sign } \gamma^{(\nu)}(t) \xi^{(\nu+1)}(t) & \text{in the case } (\Xi_{123}), \\ -\text{sign } \gamma^{(\nu+1)}(t) \xi^{(\nu)}(t) & \text{in the case } (\Gamma_{123}). \end{cases}$$

Proof. We shall restrict the proof to the case of the assumptions (Ξ_{123}), since the assumptions (Γ_{123}) may be obtained from (Ξ_{123}) by changing the role of functions ξ and γ . At the same time in the determinant, the sign of which is defined by equation (11), we have to interchange the two last columns, which accounts for the minus sign in the right-hand side of the equation in the case (Γ_{123}). Let

$$\begin{aligned} \varphi(u) &= \begin{vmatrix} 1 & t_0 & \dots & t_0^{\nu-1} & \gamma(t_0) & \xi(t_0) \\ 1 & t_1 & \dots & t_1^{\nu-1} & \gamma(t_1) & \xi(t_1) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & t_p & \dots & t_p^{\nu-1} & \gamma(t_p) & \xi(t_p) \\ 1 & u & \dots & u^{\nu-1} & \gamma(u) & \xi(u) \end{vmatrix} \\ &= \begin{vmatrix} 1 & t_0 & \dots & t_0^{\nu-1} & \gamma(t_0) \\ 1 & t_1 & \dots & t_1^{\nu-1} & \gamma(t_1) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & t_p & \dots & t_p^{\nu-1} & \gamma(t_p) \end{vmatrix} \xi(u) - \begin{vmatrix} 1 & t_0 & \dots & t_0^{\nu-1} & \xi(t_0) \\ 1 & t_1 & \dots & t_1^{\nu-1} & \xi(t_1) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & t_p & \dots & t_p^{\nu-1} & \xi(t_p) \end{vmatrix} \gamma(u) + c_0 + c_1 u + \dots \\ & \qquad \qquad \qquad + c_{\nu-1} u^{\nu-1}, \end{aligned}$$

where c_0, c_1, \dots, c_{p-1} are some constants. Let us differentiate this equation $p+1$ times, writing these two determinants as one:

$$\varphi^{(p+1)}(u) = \begin{vmatrix} 1 & t_0 & \dots & t_0^{p-1} & \gamma(t_0) \xi^{(p+1)}(u) - \xi(t_0) \gamma^{(p+1)}(u) \\ 1 & t_1 & \dots & t_1^{p-1} & \gamma(t_1) \xi^{(p+1)}(u) - \xi(t_1) \gamma^{(p+1)}(u) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & t_p & \dots & t_p^{p-1} & \gamma(t_p) \xi^{(p+1)}(u) - \xi(t_p) \gamma^{(p+1)}(u) \end{vmatrix}.$$

If we fix u and put $\eta(t) = \gamma(t) \xi^{(p+1)}(u) - \xi(t) \gamma^{(p+1)}(u)$, the determinant so obtained will be equal to the determinant (10). By the assumptions (Ξ_{123}) the sign $\eta^{(p)}(t) = \gamma^{(p)}(t) \xi^{(p+1)}(u) - \xi^{(p)}(t) \gamma^{(p+1)}(u)$ is constant in the interval I and for any u it is identical with the sign of the function $\gamma^{(p)}(t) \xi^{(p+1)}(u)$. From this fact and from (10) it follows that $\text{sign} \varphi^{(p+1)}(u) = \text{sign} \eta^{(p)}(t) = \text{sign} \gamma^{(p)}(t) \xi^{(p+1)}(u) = \text{const}$. Thus the function $\varphi(u)$ has at most $p+1$ zeros, which are equal to the numbers t_0, t_1, \dots, t_p . It is easy to show by means of considerations already applied in the proof of Theorem 6:1, that in the left neighbourhood of the point b , containing t_{p+1} , the signs of the functions $\varphi^{(p+1)}, \varphi^{(p)}, \dots, \varphi', \varphi$ are identical. Thus in the case (Ξ_{123}) we have $\text{sign} \varphi(t_{p+1}) = \text{sign} \varphi^{(p+1)}(t_{p+1}) = \text{sign} \gamma^{(p)}(t) \xi^{(p+1)}(u)$, which was to be proved.

THEOREM 10:5 ([4], p. 85). *If the $(n+1)$ -st derivative of the function ξ exists and is of a constant sign in the interval $I = \langle a, b \rangle$, then there exist exactly $n+2$ (n, I) -points u_0, u_1, \dots, u_{n+1} of the function ξ and they satisfy the conditions $a = u_0 < u_1 < \dots < u_{n+1} = b$ and*

$$(12) \quad \xi(u_k) - \omega_{nI}(u_k) = (-1)^{n+1-k} \varepsilon_n(\xi; I) \text{sign} \xi^{(n+1)}(t) \quad (k = 0, 1, \dots, n+1),$$

where $[\xi, n, I \mid \omega_{nI}]$.

This theorem directly concerns the theory of approximation; however, it is only a fragment of the Bernstein theorem which will be given in its entirety in § 12. We shall use this part of it in § 11.

Proof. All the (n, I) -points of the function ξ which lie inside the interval I are the zeros of the first derivative of the difference $\xi - \omega_{nI}$, since this difference achieves its extremum at these points. The function $((\xi - \omega_{nI})')^{(n)} = \xi^{(n+1)}$ has a constant sign, hence there is at most n of these (n, I) -points. Thus, by Theorem 1:2 the end points a and b of the interval I should also be the (n, I) -points. We have to show that the signs of the difference $\xi(u_{n+1}) - \omega_{nI}(u_{n+1}) = \xi(b) - \omega_{nI}(b)$ and the function $\xi^{(n+1)}(t)$ coincide. This can be proved in the same manner as Theorem 6:1, and the rest of equations (12) follows from it, since the points u_0, u_1, \dots, u_{n+1} are, alternately, the $(n, +, I)$ -points and $(n, -, I)$ -points of the function.

11. The estimates dependent upon the properties of the derivatives of the function ξ . General theorems.

11.1. The estimates of the (n) -points of the function ξ given in this section, form a continuation of some results of Bernstein (see [4], p. 85-87). These estimates are also generalizations of theorems contained in [24]. They are valid for those functions which have continuous derivatives of a sufficiently high order and they satisfy some additional conditions. For instance, such a condition would be the constancy of sign of a certain derivative of this function. The subclass of the class \mathcal{C}_I^p , consisting of functions whose p -th derivatives have a constant (positive or negative) sign in the interval I will be denoted by the symbol \mathcal{D}_I^p .

In the whole section the interval of approximation will be $I = \langle -1, 1 \rangle$. All the theorems can easily be generalized in the case of an arbitrary interval using Theorem 1:8.

Let us introduce an auxiliary function γ_p (p positive integer). The form of this function will be made precise later, in order to obtain the required properties. Now we shall assume only that

(Γ_1^p) $\gamma_p = \beta_p - \psi_{p-1}$, where $\beta_p \in \mathcal{C}_I^{p+1}$ and where ψ_{p-1} is the $(p-1)$ -st best polynomial for the function β_p ;

(Γ_2^p) there exist exactly $p+1$ $(p-1, I)$ -points $c_{p0} < c_{p1} < \dots < c_{pp}$ of the function β_p , where $c_{p0} = -1$, $c_{pp} = 1$.

Concerning the function ξ we shall assume that

(Ξ) there exist exactly $n+2$ (n, I) -points $u_0 < u_1 < \dots < u_{n+1}$ of the function ξ , where $u_0 = -1$, $u_{n+1} = 1$.

We know that the assumption (Ξ) is satisfied if the $(n+1)$ -st derivative of the function ξ exists and has a constant sign in the interval I (Theorem 10:5). For the same reason the assumption (Γ_2^p) is satisfied, if the p -th derivative of the function β_p has a constant sign in the interval I . Then we have

$$(1) \quad \gamma_p(c_{pl}) = (-1)^{p-l} \varepsilon_{p-1}(\beta_p) \operatorname{sign} \beta_p^{(p)}(t) = (-1)^{p-l} \|\gamma_p\|_I \operatorname{sign} \gamma_p^{(p)}(t) \\ (l = 0, 1, \dots, p)$$

by analogy to (12), § 10. On the other hand, it always follows from (Γ_2^p) that

$$(2) \quad |\gamma_p(t)| < \|\gamma_p\|_I \quad (t \neq c_{pl}, l = 0, 1, \dots, p).$$

In all theorems of this section the assumptions (Γ_2^p) and (Ξ) will be assumed to hold. Sometimes they will follow from other assumptions, which determine the signs of the functions $\xi^{(n+1)}$ and $\gamma_p^{(p)}$.

Now we shall formulate some assumptions which will appear in various combinations in theorems from § 11:

- $(\Xi_1^{nm}) \quad \xi \in \mathcal{D}_I^{n+m}, \gamma_{n+m-1} \in \mathcal{C}_I^{n+m},$
- $(\Xi_2^{nm}) \quad \varepsilon_n(\xi) < \|\gamma_{n+m-1}\|_I,$
- $(\Xi_3^{nm}) \quad |\xi^{(n+m)}(t)| \geq |\gamma_{n+m-1}^{(n+m)}(t)| \text{ for } t \in I,$
- $(\Xi_4^{nm}) \quad \xi \in \mathcal{D}_I^{n+1}, \gamma_{n+m-1} \in \mathcal{D}_I^{n+m-1},$
- $(\Xi_5^{nm}) \quad |\xi^{(n+m-1)}(t)| < |\gamma_{n+m-1}^{(n+m-1)}(t)| \text{ for } t \in I,$
- $(\Gamma_1^{nm}) \quad \xi \in \mathcal{C}_I^{n+m}, \gamma_{n+m-1} \in \mathcal{D}_I^{n+m},$
- $(\Gamma_2^{nm}) \quad \varepsilon_n(\xi) > \|\gamma_{n+m-1}\|_I,$
- $(\Gamma_3^{nm}) \quad |\xi^{(n+m)}(t)| \leq |\gamma_{n+m-1}^{(n+m)}(t)| \text{ for } t \in I,$
- $(\Gamma_4^{nm}) \quad \xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+m-1}, \gamma_{n+m-1} \in \mathcal{D}_I^{n+m-1},$
- $(\Gamma_5^{nm}) \quad |\xi^{(n+m-1)}(t)| > |\gamma_{n+m-1}^{(n+m-1)}(t)| \text{ for } t \in I$

(m is a positive integer).

THEOREM 11 : 1. *Let*

$$(3) \quad s_\xi = \text{sign}(\xi(-1) - \omega_n(-1)), \quad s_\gamma = \text{sign} \gamma_{n+m-1}(-1),$$

$$\delta_- = \|\gamma_{n+m-1}\|_I (\xi - \omega_n) - \varepsilon_n(\xi) \gamma_{n+m-1},$$

$$\delta_+ = \|\gamma_{n+m-1}\|_I (\xi - \omega_n) + \varepsilon_n(\xi) \gamma_{n+m-1},$$

where ω_n is the n -th best polynomial for the function ξ in the interval I . Let z_- and z_+ denote the number of zeros (with their orders) of the functions δ_- and δ_+ , respectively. If the assumptions (Ξ_{123}^{nm}) or (Γ_{123}^{nm}) are satisfied, then

$$z_- \leq \begin{cases} n+m-1 & (s_\xi = -s_\gamma \text{ and } m \text{ is even}), \\ n+m & (\text{otherwise}), \end{cases}$$

$$z_+ \leq \begin{cases} n+m-1 & (s_\xi = s_\gamma \text{ and } m \text{ is even}), \\ n+m & (\text{otherwise}). \end{cases}$$

Proof. The degree of the polynomial ω_n does not exceed n and, after differentiating $n+m$ times the functions δ_- and δ_+ we get

$$\delta_-^{(n+m)} = \|\gamma_{n+m-1}\|_I \xi^{(n+m)} - \varepsilon_n(\xi) \gamma_{n+m-1}^{(n+m)},$$

$$\delta_+^{(n+m)} = \|\gamma_{n+m-1}\|_I \xi^{(n+m)} + \varepsilon_n(\xi) \gamma_{n+m-1}^{(n+m)}.$$

From the assumptions (Ξ_2^{nm}) and (Ξ_3^{nm}) it follows that

$$\|\gamma_{n+m-1}\|_I |\xi^{(n+m)}(t)| > \varepsilon_n(\xi) |\gamma_{n+m-1}^{(n+m)}(t)| \quad (t \in I).$$

This means that the functions $\delta_-^{(n+m)}$ and $\delta_+^{(n+m)}$ have, for every $t \in I$, the same sign as the function $\xi^{(n+m)}$, hence, by the assumption (Ξ_1^{nm}) ,

they have a constant sign. Similarly, if the assumptions (Γ_{123}^{nm}) are satisfied, then

$$\|\gamma_{n+m-1}\|_I |\xi^{(n+m)}(t)| < \varepsilon_n(\xi) |\gamma_{n+m-1}^{(n+m)}(t)| \quad (t \in I).$$

Then, at every point $t \in I$ the sign of the function $\delta_+^{(n+m)}$ coincides with the sign of $\gamma_{n+m-1}^{(n+m)}$, and the sign of $\delta_-^{(n+m)}$ is opposite to the sign of $\gamma_{n+m-1}^{(n+m)}$. Thus, in this case also, the functions $\delta_+^{(n+m)}$ and $\delta_-^{(n+m)}$ have a constant sign. It follows from Rolle's theorem that each of the functions δ_- , δ_+ has at most $n+m$ zeros together with their orders; i.e., $z_- \leq n+m$, $z_+ \leq n+m$.

Let us now consider the function δ_- in the case when $s_\xi = -s_\gamma$ and m is even. Since

$$\begin{aligned} \delta_-(-1) &= \|\gamma_{n+m-1}\|_I (\xi(-1) - \omega_n(-1)) - \varepsilon_n(\xi) \gamma_{n+m-1}(-1) \\ &= (\|\gamma_{n+m-1}\|_I |\xi(-1) - \omega_n(-1)| + \varepsilon_n(\xi) |\gamma_{n+m-1}(-1)|) s_\xi, \end{aligned}$$

we have $\text{sign } \delta_-(-1) = s_\xi$. By the assumption (Ξ) it follows that the points u_0, u_1, \dots, u_{n+1} are alternately the $(n, +)$ -points and the $(n, -)$ -points of the function ξ . By the assumption (Γ_2^{n+m-1}) we see that the points $c_0, c_1, \dots, c_{n+m-1}$ (for simplicity we omit the first index in the symbols $c_{n+m-1, l}$) are alternately the $(n+m-2, +)$ -points and the $(n+m-2, -)$ -points of the function β_{n+m-1} . Thus

$$\begin{aligned} \xi(1) - \omega_n(1) &= \xi(u_{n+1}) - \omega_n(u_{n+1}) \\ &= (-1)^{n+1} (\xi(u_0) - \omega_n(u_0)) = (-1)^{n+1} s_\xi \varepsilon_n(\xi), \end{aligned}$$

$$\begin{aligned} \gamma_{n+m-1}(1) &= \gamma_{n+m-1}(c_{n+m-1}) \\ &= (-1)^{n+m-1} \gamma_{n+m-1}(c_0) = (-1)^{n+m} s_\xi \|\gamma_{n+m-1}\|_I \end{aligned}$$

and for an even m

$$(4) \quad \text{sign } \delta_-(1) = (-1)^{n+1} s_\xi = (-1)^{n+1} \text{sign } \delta_-(-1).$$

The function δ_- is differentiable $n+m$ times, and we may define its zeros of the order k for $k \leq n+m+1$. From what we have proved above, it follows that the order of any zero of the function δ_- does not exceed $n+m$. At a zero of order k for $k \leq n+m$ the function δ_- changes its sign if k is odd and does not change its sign if k is even. If we assume that the total multiplicity of the zeros of the function ξ in the interval $I = \langle -1, 1 \rangle$ equals $n+m$, then values of this function at the points -1 and 1 will have opposite signs if $n+m$ is odd, and the same signs if $n+m$ is even. In other words $\text{sign } \delta_-(1) = (-1)^{n+m} \text{sign } \delta_-(-1)$. For an even m this equation contradicts equation (4) obtained above. It follows that the function δ_- , in this case, can have at most $n+m-1$ zeros together with their orders.

In a similar way we prove that $z_{\pm} \leq n + m - 1$, if $s_{\xi} = s_{\gamma}$ and m is even.

11.2. Theorem 11:2 which we shall now prove, contains some estimate of the (n) -points u_0, u_1, \dots, u_{n+1} but it is still in an inconvenient form. In the next theorem (11:3) we shall obtain stronger estimates, which will allow some applications.

THEOREM 11:2. *If the assumptions (Ξ_{123}^{nm}) or (Γ_{123}^{nm}) are satisfied, then in each of the intervals $\langle u_0, u_1 \rangle, \langle u_1, u_2 \rangle, \dots, \langle u_n, u_{n+1} \rangle$ lies at least one number from among the numbers $c_{n+m-1,0}, c_{n+m-1,1}, \dots, c_{n+m-1,n+m-1}$.*

Proof. We have to show that for an arbitrary $k = 0, 1, \dots, n$ there exists a l such that $c_l \in \langle u_k, u_{k+1} \rangle$, where $c_l \equiv c_{n+m-1,l}$. We already know that such a l exists for $k = 0$ and $k = n$, since from the assumptions (Ξ) and (Γ_2^{n+m-1}) we get $c_0 = u_0 = -1, c_{n+m-1} = u_{n+1} = 1$. Thus the negation of the assertion is equivalent to the fact that for some k from among $1, 2, \dots, n-1$ all the points $c_0, c_1, \dots, c_{n+m-1}$ lie outside the interval $\langle u_k, u_{k+1} \rangle$; i.e., that for this k and some l we have the inequality

$$(5) \quad c_l < u_k < u_{k+1} < c_{l+1}.$$

We shall prove that the negation of the assertion implies the existence of at least $n + m + 1$ zeros of the function δ_- or δ_+ , which contradicts Theorem 11:1.

Assuming that inequality (5) is true, we shall prove that $z_{\pm} \geq n + m + 1$, if

$$(6) \quad \gamma_{n+m-1}(c_l) (\xi(u_k) - \omega_n(u_k)) > 0$$

and at the same time

$$(7) \quad \gamma_{n+m-1}(c_{l+1}) (\xi(u_{k+1}) - \omega_n(u_{k+1})) > 0.$$

In the case of opposite inequalities we may show in an analogous way that $z_{\pm} \geq n + m + 1$.

Part I. The proof of the existence of $l+1$ (with their orders) zeros of the function δ_- in the interval $\langle -1, u_k \rangle$. Let us notice first that the points c_l and the functions $\beta = \|\gamma_{n+m-1}\|_I (\xi - \omega_n), \gamma = \varepsilon_n(\xi) \gamma_{n+m-1}$ whose difference equals δ_- , satisfy the assumptions of Theorem 10:2 for $p = n + m - 1$. These functions have, in particular, the same norm, equal to $\|\gamma_{n+m-1}\|_I \|\xi - \omega_n\|_I = \|\gamma_{n+m-1}\|_I \varepsilon_n(\xi)$. Thus, using the considerations employed already in the proof of Theorem 10:2 one can show that the function δ_- has the following properties:

(Δ_1) if $\delta_-(c_i) = 0$ for $0 < i < n + m - 1$, then c_i is at least a double zero of the function δ_- ,

(Δ_2) if $\delta_-(c_l) \neq 0$, then in the interval (c_l, u_k) there exists a zero of the function δ_- ,

(Δ_3) if $\delta_-(c_{l+1}) \neq 0$, then in the interval (u_{k+1}, c_{l+1}) there exists a zero of the function δ_- ,

(Δ_4) in the interval (u_k, u_{k+1}) there exists a zero of the function δ_- .

In fact, the property (Δ_1) is given in the proof of Theorem 10:2. We have also noticed in this proof, that if $\delta_-(c_l) \neq 0$ then the numbers $\delta_-(c_l)$ and $\gamma(c_l) = \varepsilon_n(\xi)\gamma_{n+m-1}(c_l)$ have opposite signs:

$$(8) \quad \delta_-(c_l)\gamma_{n+m-1}(c_l) < 0.$$

In view of assumption (5), inequality (2) is true for $t = u_k$; i.e., $|\gamma_{n+m-1}(u_k)| < \|\gamma_{n+m-1}\|_I$. From this fact and from the equation $|\xi(u_k) - \omega_n(u_k)| = \varepsilon_n(\xi)$ it follows that the numbers $\delta_-(u_k)$ and $\xi(u_k) - \omega_n(u_k)$ have the same signs:

$$(9) \quad \delta_-(u_k)(\xi(u_k) - \omega_n(u_k)) > 0.$$

Inequalities (6), (8) and (9) show that the values of the function δ_- at the end points of the interval (c_l, u_k) have opposite signs, and in this interval there exists a zero of the function δ_- (property (Δ_2)). In the same way we verify the property (Δ_3). Finally, from inequality (9), and from an analogous inequality concerning the point u_{k+1} :

$$\delta_-(u_{k+1})(\xi(u_{k+1}) - \omega_n(u_{k+1})) > 0$$

and the inequality

$$(\xi(u_k) - \omega_n(u_k))(\xi(u_{k+1}) - \omega_n(u_{k+1})) < 0$$

it follows that $\delta_-(u_k)\delta_-(u_{k+1}) < 0$ (property (Δ_4)).

Case I.1. $\delta_-(c_l) \neq 0$. Applying Theorem 10:2 to the functions $\beta = \|\gamma_{n+m-1}\|_I(\xi - \omega_n)$ and $\gamma = \varepsilon_n(\xi)\gamma_{n+m-1}$ for $p = n + m - 1$, we see that the function δ_- has at least l zeros in the interval $\langle -1, c_l \rangle$. In view of the assumption $\delta_-(c_l) \neq 0$ the property (Δ_2) shows the existence of the $(l+1)$ -st zero in the interval (c_l, u_k) .

Case I.2. $\delta_-(c_l) = 0$. The assertion of part I for $l = 0$ follows directly from this assumption. If $l > 0$ (and, at the same time $l < n + m - 1$, since the point c_{l+1} appears in (5)), then, by Theorem 10:2, the function δ_- has $l-1$ zeros in the interval $\langle -1, c_{l-1} \rangle$. Besides, by property (Δ_1) for $i = l$ the point c_l is at least a double zero of the function δ_- .

Part II. Proof of the existence of $n + m - l - 1$ zeros of the function δ_- in the interval $(u_{k+1}, 1)$ proceeds in an analogous way. We should use the property (Δ_3) instead of (Δ_2) and inequality (7) instead of (6).

Collecting the results of parts I and II, and using the property (Δ_4) , we get the inequality $z_- \geq (l+1) + (n+m-l-1) + 1 = n+m+1$. This inequality contradicts Theorem 11:1, which proves our theorem.

THEOREM 11:3. *Let*

$$(10) \quad r = \min \left\{ \left[\frac{m-1}{2} \right], \left[\frac{2n+m-1}{4} \right] \right\}.$$

If the assumptions (Ξ_{123}^{nm}) or (Γ_{123}^{nm}) are satisfied, then the equation

$$(11) \quad u_k = c_{n+m-1, l}$$

has, for $0 < l < n+m-1$, at most r solutions in numbers k and l such that the difference $k-l$ is even, and at most r solutions in numbers k and l such that the difference $k-l$ is odd.

Proof. We shall use the numbers s_ξ and s_γ , defined by (3) and equal to either -1 or 1 , and also the known functions $\delta_- = \beta - \gamma$, $\delta_+ = \beta + \gamma$ where $\beta = \|\gamma_{n+m-1}\|_I(\xi - \omega_n)$, $\gamma = \varepsilon_n(\xi)\gamma_{n+m-1}$.

The values of the function β at the points u_0, u_1, \dots, u_{n+1} have alternating signs:

$$\text{sign } \beta(u_k) = (-1)^k s_\xi \quad (k = 0, 1, \dots, n+1)$$

(the conclusion from property (Ξ) and the definition of the sign s_ξ). Similarly, the values of the function γ at the points $c_0, c_1, \dots, c_{n+m-1}$, where $c_l = c_{n+m-1, l}$, have alternating signs:

$$\text{sign } \gamma(c_l) = (-1)^l s_\gamma \quad (l = 0, 1, \dots, n+m-1)$$

(this is a conclusion from the property (Γ_2^{n+m-1}) and the definition of the sign s_γ). We also know that $|\beta(u_k)| = |\gamma(c_l)| = \|\gamma_{n+m-1}\|_I \|\xi - \omega_n\|_I = \|\gamma_{n+m-1}\|_I \varepsilon_n(\xi)$. Thus, if for some l (where $0 < l < n+m-1$) we have equation (11) then the point c_l is at least a double zero of the function δ_- (or δ_+ , respectively) in the case when $(-1)^k s_\xi$ equals to $(-1)^l s_\gamma$ (or $-(-1)^l s_\gamma$, respectively). In fact, the value of one of the functions δ_- , δ_+ at this point is equal to zero, and the rest follows from the property (Δ_1) given in the proof of Theorem 11:2, or from an analogous property of the function δ_+ .

We shall now establish under which conditions the end points of the interval $I = \langle -1, 1 \rangle$ are the zeros of functions δ_- or δ_+ . By the definition of the signs s_ξ and s_γ the number -1 is a zero of the function δ_- if $s_\xi = s_\gamma$, and a zero of the function δ_+ if $s_\xi = -s_\gamma$. Since it follows for $u_{n+1} = 1$ and $c_{n+m-1} = 1$ from the previously derived equations that $\text{sign } \beta(1) = (-1)^{n+1} s_\xi$, $\text{sign } \gamma(1) = (-1)^{n+m-1} s_\gamma$, then the number 1 is

a zero of the function δ_- if $(-1)^m s_\xi = s_\gamma$, and a zero of the function δ_+ if $(-1)^m s_\xi = -s_\gamma$. Thus, if $s_\xi = s_\gamma$, then

$$(12) \quad \begin{aligned} \delta_-(-1) &= \delta_+(1) = 0 & (m \text{ even}), \\ \delta_+(-1) &= \delta_-(1) = 0 & (m \text{ odd}). \end{aligned}$$

On the other hand, if $s_\xi = -s_\gamma$, then $\delta_+(-1) = \delta_-(1) = 0$ (m even), $\delta_-(-1) = \delta_+(1) = 0$ (m odd).

Theorem 11:3 will be proved only in the case when $s_\xi = s_\gamma$. The case $s_\xi = -s_\gamma$ may be investigated in a similar way, only the roles of functions δ_- and δ_+ used for estimating the number of solutions of equation (11) would then change.

Part I. $k-l$ is even. Let us suppose that there exist exactly q points $c_{l_1}, c_{l_2}, \dots, c_{l_q}$ (where $0 < l_1 < l_2 < \dots < l_q < n+m-1$) such that

$$(13) \quad u_{k_1} = c_{l_1}, \quad u_{k_2} = c_{l_2}, \quad \dots, \quad u_{k_q} = c_{l_q}$$

and that the differences $k_i - l_i$ are even for $i = 1, 2, \dots, q$. Let us introduce the notations $k_0 = l_0 = 0$ (i.e. $u_{k_0} = c_{l_0} = -1$), $k_{q+1} = n+1$, $l_{q+1} = n+m-1$ (i.e. $u_{k_{q+1}} = c_{l_{q+1}} = 1$) and

$$g_i = k_{i+1} - k_i, \quad h_i = l_{i+1} - l_i \quad (i = 0, 1, \dots, q).$$

Under these notations, from the assumptions concerning the numbers k_i, l_i it follows that

$$(14) \quad \begin{cases} g_i \equiv h_i \pmod{2} & (i = 0, 1, \dots, q-1), \\ g_q \equiv h_q \pmod{2} & (m \text{ even}), \\ g_q \not\equiv h_q \pmod{2} & (m \text{ odd}) \end{cases}$$

(the relation $g \equiv h \pmod{2}$ means that the difference $g-h$ is even, and the relation $g \not\equiv h \pmod{2}$ means that the difference $g-h$ is odd).

We may apply Theorem 10:3 to the functions β and γ and to each of the intervals $\langle a, b \rangle = \langle c_{l_i}, c_{l_{i+1}} \rangle$ ($i = 0, 1, \dots, q$). In particular assumptions (iii) and (v) of this theorem follow from the assumptions (E) and (Γ_2^{m+m-1}), and assumption (iv) follows from the fact that all the equations of the type (11) for even values of the difference $k-l$ are equations of the type (13). In Theorem 10:3 we have estimated from below the number of distinct zeros of the function $\delta_- = \beta - \gamma$ in the interval (a, b) . In the case of the interval $(c_{l_i}, c_{l_{i+1}})$, the role of the numbers k and l from this theorem are played by the numbers $g_i = k_{i+1} - k_i$, $h_i = l_{i+1} - l_i$. By relations (14) the number of zeros of the function δ_- is estimated from below by the number (5), § 10, for $i = 0, 1, \dots, q-1$ and for $i = q$ for an even m , and by the number (6), § 10 for $i = q$ and an odd m .

Let us denote by z_- the number of zeros (with their orders) of the function δ_- in the whole interval $\langle -1, 1 \rangle$. To estimate this number from below we should use—besides Theorem 10:3—the remarks given at the beginning of the proof. We have shown that all the points (13) are at least the double zeros of functions δ_- . Besides, it follows from equation (12), that the number -1 is a zero of this function, the same is true for the number 1 , provided that m is even. Considering all these facts we get the following estimate:

$$z_- \geq \sum_{i=0}^{q-1} (\max\{g_i, h_i\} - 1 - \text{sign}|g_i - h_i|) + \\ + 2q + \begin{cases} \max\{g_q, h_q\} - 1 - \text{sign}|g_q - h_q| + 2 & (m \text{ even}), \\ \max\{g_q, h_q\} - 1 + 1 & (m \text{ odd}). \end{cases}$$

If we notice that

$$\max\{g_i, h_i\} = \frac{1}{2}(g_i + h_i) + \frac{1}{2}|g_i - h_i|, \\ (15) \quad \sum_{i=0}^q g_i = k_{q+1} - k_0 = n + 1, \quad \sum_{i=0}^q h_i = l_{q+1} - l_0 = n + m - 1$$

we may transform the above inequality obtaining:

$$z_- \geq n + m/2 + \sum_{i=0}^q (\frac{1}{2}|g_i - h_i| - \text{sign}|g_i - h_i|) - \\ - q + 2q + \begin{cases} 1 & (m \text{ even}), \\ \text{sign}|g_q - h_q| & (m \text{ odd}). \end{cases}$$

In view of relation (14) we have $\text{sign}|g_q - h_q| = 1$ for an odd m . Since it has been shown in Theorem 11:1 that $z_- \leq n + m$, we get finally:

$$(16) \quad q \leq m/2 - 1 - \sum_{i=0}^q (\frac{1}{2}|g_i - h_i| - \text{sign}|g_i - h_i|).$$

Part I.1. The proof of the inequality $q \leq [(m-1)/2]$. The addends of the sum which appear in inequality (16) are non-negative for $i = 0, 1, \dots, q-1$, since $\frac{1}{2}|g_i - h_i| - \text{sign}|g_i - h_i|$ is positive if $|g_i - h_i| > 2$ and zero if $|g_i - h_i|$ equals 0 or 2. The last addend of the sum mentioned has the same property if m is even. Thus, in this case $q \leq m/2 - 1$. On the other hand, if m is odd, the difference $g_q - h_q$ is odd, and we have the inequality $\frac{1}{2}|g_q - h_q| - \text{sign}|g_q - h_q| \geq \frac{1}{2} - 1 = -\frac{1}{2}$. In this case $q \leq m/2 - \frac{1}{2}$. The inequalities $q \leq m/2 - 1$ for an even m and $q \leq m/2 - \frac{1}{2}$ for an odd m are equivalent to the inequality $q \leq [(m-1)/2]$.

Part I.2. The proof of the inequality $q \leq [(2n + m - 1)/4]$.
Subtracting equations (15) we get

$$(17) \quad \sum_{i=0}^q |g_i - h_i| \geq m - 2.$$

From (17) and (16) it follows that

$$(18) \quad q \leq \sum_{i=0}^q \text{sign} |g_i - h_i|.$$

Suppose that m is even. Then the differences $g_i - h_i$ for $i = 0, 1, \dots, q$ are also even. Thus, if $g_i \neq h_i$, we have $g_i + h_i \geq 4$ (the numbers g_i and h_i are positive by definition), and if $g_i = h_i$, then $g_i + h_i \geq 2$. According to (18), among $q + 1$ numbers $g_i - h_i$ at most one is equal to zero. Thus

$$\sum_{i=0}^q (g_i + h_i) \geq 4q + 2.$$

Considering (15) we then obtain

$$2n + m \geq 4q + 2, \quad q \leq (2n + m - 2)/4.$$

Suppose now that m is odd. The differences $g_i - h_i$ for $i = 0, 1, \dots, q - 1$ are even as before, which implies that $g_i + h_i \geq 4$ for $g_i \neq h_i$ and $g_i + h_i \geq 2$ for $g_i = h_i$. On the other hand the difference $g_q - h_q$ is odd, hence $g_q + h_q \geq 3$. This time, at most one of the numbers $g_i - h_i$ may be equal to zero, but it cannot be the number $g_q - h_q$. Thus

$$2n + m = \sum_{i=0}^q (g_i + h_i) \geq 4(q - 1) + 2 + 3 = 4q + 1, \quad q \leq (2n + m - 1)/4.$$

The inequalities $q \leq (2n + m - 2)/4$ for an even m and $q \leq (2n + m - 1)/4$ for an odd m may be written uniformly, such as $q \leq [(2n + m - 1)/4]$.

Part II. $k - l$ is odd. Suppose that there exists exactly q points $c_{l_1}, c_{l_2}, \dots, c_{l_q}$ (where $0 < l_1 < l_2 < \dots < l_q < n + m - 1$) which satisfy equations (13) and such that the differences $k_i - l_i$ are odd for $i = 1, 2, \dots, q$. The symbols $k_0, l_0, k_{q+1}, l_{q+1}, g_i, h_i$ have the same meaning as in part I of the proof. Now, instead of (14), we have the relations:

$$(19) \quad \begin{aligned} g_0 &\not\equiv h_0 \pmod{2}, \\ g_i &\equiv h_i \pmod{2} & (i = 1, 2, \dots, q - 1), \\ g_q &\not\equiv h_q \pmod{2} & (m \text{ even}), \\ g_q &\equiv h_q \pmod{2} & (m \text{ odd}). \end{aligned}$$

The points c_i are at least the double zeros of the function $\delta_+ = \beta + \gamma$. If m is odd, the number 1 is a zero of this function by (12). To estimate the number of zeros of the function δ_+ in the intervals (c_i, c_{i+1}) ($i = 0, 1, \dots, q$) we shall again use Theorem 10:3, which will now be applied to the functions $\beta^* = \beta$ and $\gamma^* = -\gamma$. By (19) for $i = q$, if m is even and for $i = 0$, the number of zeros of the function δ_+ in the interval (c_i, c_{i+1}) is estimated by expression (6), § 10, and for the remaining i 's—by expression (5), § 10. If we denote by z_+ the number of zeros (with their orders) of the function δ_+ in the whole interval $\langle -1, 1 \rangle$ we get the estimate:

$$z_+ \geq \max\{g_0, h_0\} - 1 + \sum_{i=1}^{q-1} (\max\{g_i, h_i\} - 1 - \text{sign}|g_i - h_i|) + \\ + 2q + \begin{cases} \max\{g_q, h_q\} - 1 & (m \text{ even}), \\ \max\{g_q, h_q\} - 1 - \text{sign}|g_q - h_q| + 1 & (m \text{ odd}). \end{cases}$$

Since $\text{sign}|g_q - h_q| = 1$ for an even m and $\text{sign}|g_0 - h_0| = 1$, the last inequality may be written more compactly as follows:

$$z_+ \geq n + m/2 + q + 1 + \sum_{i=0}^q (\frac{1}{2}|g_i - h_i| - \text{sign}|g_i - h_i|).$$

In Theorem 11:1, provided that $s_\varepsilon = s_\gamma$, we showed that $z \leq n + m - 1$ if m is even, and $z_+ \leq n + m$ if m is odd. Thus

$$q \leq m/2 - \sum_{i=0}^q (\frac{1}{2}|g_i - h_i| - \text{sign}|g_i - h_i|) - \begin{cases} 2 & (m \text{ even}), \\ 1 & (m \text{ odd}). \end{cases}$$

Applying the remarks from the first part of the proof we see, that if m is even, we have the inequalities

$$\frac{1}{2}|g_0 - h_0| - \text{sign}|g_0 - h_0| \geq -\frac{1}{2}, \quad \frac{1}{2}|g_q - h_q| - \text{sign}|g_q - h_q| \geq -\frac{1}{2},$$

$$\frac{1}{2}|g_i - h_i| - \text{sign}|g_i - h_i| \geq 0$$

for the remaining i 's. Thus $q \leq m/2 - 1 = [(m-1)/2]$. In a similar way for an odd m we get $q \leq m/2 - \frac{1}{2} = [(m-1)/2]$.

Using inequality (17) we may transform the previous estimation for q , obtaining:

$$q \leq \sum_{i=0}^q \text{sign}|g_i - h_i| - \begin{cases} 1 & (m \text{ even}), \\ 0 & (m \text{ odd}). \end{cases}$$

From this inequality we can deduce that for an even m all the numbers $g_i - h_i$ are different from zero, and for an odd m at most one of the numbers

$g_i - h_i$ may be equal to zero. From part I of the proof we know that $g_i + h_i \geq 4$ if $g_i \neq h_i$ and if the difference $g_i - h_i$ is even, $g_i + h_i \geq 2$ if $g_i = h_i$ and $g_i + h_i \geq 3$ if the difference $g_i - h_i$ is odd. This allows us to write for an even m the inequality

$$2n + m = \sum_{i=0}^q (g_i + h_i) \geq 3 + 4(q-1) + 3 = 4q + 2$$

and for an odd m the inequality

$$2n + m = \sum_{i=0}^q (g_i + h_i) \geq 3 + 4(q-1) + 2 = 4q + 1.$$

These are the inequalities which were obtained in part I of the proof, and which, together with the inequality $q \leq [(m-1)/2]$, give the assertion of our theorem.

11.3. The most interesting corollaries follow from Theorems 11:2 and 11:3 for $m = 1$ (Theorems 11:4 and 11:5) and for $m = 2$ (Theorem 11:7); i.e., in the case when equation (11) has no solution in numbers k and l for $0 < l < n + m - 1$. However, for greater values of m we may obtain, using these theorems, the relatively strong estimates of (n) -points u_k of the function ξ (Theorem 11:6). The first corollary occupies, in the sense of its generality, a mean position between two theorems of Bernstein ([4], p. 85 and 87).

THEOREM 11:4. *If the assumptions (Ξ_{123}^{n1}) or (Γ_{123}^{n1}) are satisfied, then we have the inequality*

$$(20) \quad -1 = u_0 = c_{n0} < u_1 < c_{n1} < \dots < u_n < c_{nn} = u_{n+1} = 1.$$

Proof. It follows from Theorem 11:3 for $m = 1$ that none of the inequalities $u_k = c_{nl}$ for $0 < l < n$ holds. If we consider the equation $u_0 = c_{n0} = -1$, $u_{n+1} = c_{nn} = 1$, we see that the assertion of Theorem 11:2 for $m = 1$ may be formulated as follows: in each of the $n+1$ disjoint intervals $\langle -1, u_1 \rangle = \langle u_0, u_1 \rangle, \langle u_1, u_2 \rangle, \dots, \langle u_{n-1}, u_n \rangle, \langle u_n, u_{n+1} \rangle = \langle u_n, 1 \rangle$ there is at least one of the $n+1$ numbers $c_{n0}, c_{n1}, \dots, c_{nn}$. Thus $c_{nl} \in (u_l, u_{l+1})$ for $l = 1, 2, \dots, n-1$, which was to be proved.

Under sufficiently strong assumptions concerning the function ξ , Theorem 11:4 gives some information about the mutual position of the (n) -points of this function (these (n) -points will now be denoted as $u_{n0}, u_{n1}, \dots, u_{n,n+1}$) and the $(n+1)$ -points $u_{n+1,0} < u_{n+1,1} < \dots < u_{n+1,n+2}$ of the same function. This information may be used directly for the estimation of the $(n+1)$ -points if we know the n -th best polynomial and the (n) -points, and indirectly for the estimation of the (n) -points (see Theorem 12:4).

THEOREM 11:5 (Paszkowski, [24], p. 53-54). *If $\xi \in \mathcal{D}_I^{n+2}$, then*

$$(21) \quad -1 = u_{n+1,0} = u_{n0} < u_{n+1,1} < u_{n1} < \dots < u_{nn} < u_{n+1,n+1} < u_{n,n+1} \\ = u_{n+1,n+2} = 1.$$

Proof. One can show that the function $\gamma_{n+1} = \xi - \omega_n$, where ω_n is the n -th best polynomial for the function ξ , satisfies the assumptions (Γ_{12}^{n+1}) and $(\Xi_{123}^{n+1,1})$. Indeed, the assumptions (Γ_{12}^{n+1}) for this function follow directly from the definition and from the assumption (Ξ) . The assumption $(\Xi_1^{n+1,1})$ follows from the fact that $\xi \in \mathcal{D}_I^{n+2}$ and $\gamma_{n+1}^{(n+2)} = \xi^{(n+2)}$. The assumption $(\Xi_2^{n+1,1})$ takes the form $\varepsilon_{n+1}(\xi) < \|\gamma_{n+1}\|_I = \|\xi - \omega_n\|_I = \varepsilon_n(\xi)$. Since only $n+2$ (n) -points of the function ξ exist, the polynomial ω_n cannot be the $(n+1)$ -st best polynomial for this function and $\varepsilon_{n+1}(\xi) < \varepsilon_n(\xi)$. Finally, the assumption $(\Xi_3^{n+1,1})$ is satisfied because $\gamma_{n+1}^{(n+2)} = \xi^{(n+2)}$.

For the function $\gamma_{n+1} = \xi - \omega_n$ the points $c_{n+1,l}$ coincide with the (n) -points $u_{n0}, u_{n1}, \dots, u_{n,n+1}$ of function ξ . Thus inequalities (21) follow from inequalities (20) after replacing n by $n+1$.

THEOREM 11:6. *If the assumptions (Ξ_{123}^{nm}) or (Γ_{123}^{nm}) are satisfied, then we have the inequalities*

$$(22) \quad u_k \begin{cases} > c_{n+m-1, (k-1)/2} & (k = 1, 3, \dots, 4r+1), \\ \geq c_{n+m-1, k/2} & (k = 2, 4, \dots, 4r), \\ > c_{n+m-1, k-2r-1} & (k = 4r+2, 4r+3, \dots, n), \end{cases}$$

$$(23) \quad u_k \begin{cases} < c_{n+m-1, k+m+2r-1} & (k = 1, 2, \dots, n-4r-1), \\ \leq c_{n+m-1, m+(k+n-3)/2} & (k = n-4r+1, n-4r+3, \dots, n-1), \\ < c_{n+m-1, m+(k+n-2)/2} & (k = n-4r, n-4r+2, \dots, n). \end{cases}$$

The number r is defined by (10). If $4r+1 \geq n$, the last group of inequalities (22) and the first group of inequalities (23) have no meaning, and the remaining groups are true for all $k = 1, 2, \dots, n$, even and odd, respectively.

Proof. Suppose that at least one of the inequalities from the second or third group in (22) is false if $4r+1 < n$:

$$c_1 \leq u_2, \quad c_2 \leq u_4, \quad \dots, \quad c_{2r-1} \leq u_{4r-2}, \quad c_{2r} \leq u_{4r}, \\ c_{2r+1} < u_{4r+2}, \quad c_{2r+2} < u_{4r+3}, \quad \dots, \quad c_{n-2r+1} < u_n,$$

or, one from the second group only, if $4r+1 \geq n$:

$$c_1 \leq u_2, \quad c_2 \leq u_4, \quad \dots, \quad c_{[n/2]} \leq u_{[2n/2]}.$$

We shall prove that if this were the case, the equation $c_l = u_k$ would have more solutions than is allowed by Theorem 11:3, or, contrary to Theorem 11:2, in one of the intervals $\langle u_1, u_2 \rangle, \langle u_2, u_3 \rangle, \dots, \langle u_{n-1}, u_n \rangle$ there would be no point c_l .

Suppose first that $c_l > u_{2l}$ for some l . According to Theorem 11:2, in each of the intervals

$$(24) \quad \langle u_1, u_2 \rangle, \langle u_2, u_3 \rangle, \dots, \langle u_{2l-1}, u_{2l} \rangle$$

there should be one of the points c_1, c_2, \dots, c_{l-1} . However, this is impossible, since there are $2l-1$ intervals, $l-1 < (2l-1)/2$ points, and one point c_j may belong to at most two neighbouring intervals (24) if it coincides with their common end-point.

In the case $4r+1 < n$ suppose that $c_l \geq u_{l+2r+1}$ for an l such that $2r+1 \leq l \leq n-2r-1$. If $c_l > u_{l+2r+1}$ then in each of $l+2r$ intervals

$$(25) \quad \langle u_1, u_2 \rangle, \langle u_2, u_3 \rangle, \dots, \langle u_{l+2r}, u_{l+2r+1} \rangle$$

there should be one of the $l-1$ points c_1, c_2, \dots, c_{l-1} . According to Theorem 11:3 at most $2r$ of these points may belong to two of the intervals (25) simultaneously, and each of the remaining points belongs to only one interval. Thus, the points c_1, c_2, \dots, c_{l-1} lie in at most $2 \cdot 2r + l - 2r - 1 = l + 2r - 1$ intervals and there exists at least one interval (25) which contains no point c_j .

If $c_l = u_{l+2r+1}$, then only the points c_1, c_2, \dots, c_l may lie in intervals (25), and the last point c_l belongs only to the last of intervals (25). Among the points c_1, c_2, \dots, c_{l-1} there may be at most $2r-1$ points which belong to two of intervals (25) simultaneously. Thus, the points c_1, c_2, \dots, c_l lie in at most $2(2r-1) + l - 2r + 1 = l + 2r - 1$ intervals (25).

In this way we proved the second and third group of inequalities (22). Since $u_0 < u_1 < \dots < u_{n+1}$, the first group of these inequalities may be easily obtained; this group gives the estimate for the points $u_1, u_3, \dots, u_{4r+1}$. Thus, for instance, $u_1 > u_0 = -1 = c_0$, $u_3 > u_2 \geq c_1$ etc.

Inequalities (23) can be proved analogously; the only difference being that Theorem 11:2, in this case, is applied to the last intervals of the system

$$\langle u_1, u_2 \rangle, \langle u_2, u_3 \rangle, \dots, \langle u_{n-1}, u_n \rangle.$$

Inequalities (22), (23) have a rather complicated form. They can be considerably simplified if $m = 2$, since in this case $r = 0$ and the system of inequalities (22) reduces to the form

$$u_k > c_{n+1, k-1} \quad (k = 1, 2, \dots, n)$$

and system (23) reduces to the form

$$u_k < c_{n+1,k+1} \quad (k = 1, 2, \dots, n).$$

As a result we obtain the following theorem:

THEOREM 11:6'. *If the assumptions (Ξ_{123}^{n2}) or (Γ_{123}^{n2}) are satisfied, then*

$$u_k \in (c_{n+1,k-1}, c_{n+1,k+1}) \quad (k = 1, 2, \dots, n).$$

11.4. Now we shall apply Theorem 10:4, in which we specify the signs of certain functional determinants in order to strengthen the corollaries which follow from Theorems 11:2 and 11:3, and also in order to obtain some new results.

THEOREM 11:7. *If the assumptions (Ξ_{12345}^{n2}) are satisfied, then we have the inequalities*

$$(26) \quad -1 = u_0 = c_{n+1,0} < u_1 < c_{n+1,1} < \dots < u_n < c_{n+1,n} \\ < u_{n+1} = c_{n+1,n+1} = 1 \quad (\xi^{(n+1)}(t) \xi^{(n+2)}(t) < 0 \text{ for } t \in I)$$

or inequalities

$$(27) \quad -1 = c_{n+1,0} = u_0 < c_{n+1,1} < u_1 < \dots < c_{n+1,n} < u_n \\ < c_{n+1,n+1} = u_{n+1} = 1 \quad (\xi^{(n+1)}(t) \xi^{(n+2)}(t) > 0 \text{ for } t \in I).$$

If the assumptions (Γ_{12345}^{n2}) are satisfied, then we have inequalities (26) for $\gamma_{n+1}^{(n+1)}(t) \gamma_{n+1}^{(n+2)}(t) > 0$ and inequalities (27) for $\gamma_{n+1}^{(n+1)}(t) \gamma_{n+1}^{(n+2)}(t) < 0$.

Proof. We remember that the assumptions (Ξ_{12345}^{n2}) and (Γ_{12345}^{n2}) have been given in § 11.1. In Theorem 11:3 we proved under the assumption (Ξ_{123}^{n2}) or (Γ_{123}^{n2}) that none of the equation $u_k = c_{n+1,l}$ holds for $0 < l < n+1$. Thus, we already know that in each of $n-1$ intervals (u_1, u_2) , (u_2, u_3) , \dots , (u_{n-1}, u_n) there lies at least one of the (n) -points c_1, c_2, \dots, c_n (where $c_l \equiv c_{n+1,l}$).

Suppose that one of these intervals contains two successive points from among c_1, c_2, \dots, c_n . Then for a certain l such that $0 < l < n$ we have

$$(28) \quad u_1 < c_1 < u_2 < c_2 < \dots < u_l < c_l < c_{l+1} < u_{l+1} < \dots < c_n < u_n.$$

Suppose also that

$$(29) \quad \text{sign}(\xi(u_k) - \omega_n(u_k)) = \text{sign} \gamma_{n+1}(c_k) \quad (k = 0, 1, \dots, n+1).$$

Under these assumptions we shall prove that the function $\delta_- = \|\gamma_{n+1}\|_I(\xi - \omega_n) - \varepsilon_n(\xi) \gamma_{n+1}$ has at least $n+3$ zeros in the interval I , which contradicts Theorem 11:1 for $m = 2$.

Using the definition of points u_k and c_k , inequalities (28) and the assumptions (Ξ) and (Γ_2^{n+1}) we see that

$$\begin{aligned} |\xi(u_k) - \omega_n(u_k)| &= \varepsilon_n(\xi), & |\gamma_{n+1}(u_k)| &< \|\gamma_{n+1}\|_I \quad (0 < k < n+1), \\ |\xi(c_k) - \omega_n(c_k)| &< \varepsilon_n(\xi), & |\gamma_{n+1}(c_k)| &= \|\gamma_{n+1}\|_I \quad (0 < k < n+1). \end{aligned}$$

Thus we obtain the inequalities

$$\begin{aligned} \text{sign } \delta_-(u_k) &= \text{sign}(\xi(u_k) - \omega_n(u_k)), & \text{sign } \delta_-(c_k) &= -\text{sign } \gamma_{n+1}(c_k) \\ & & & (0 < k < n+1) \end{aligned}$$

which, together with assumptions (29), show that the function δ_- has zeros in each of $n+1$ intervals $(u_1, c_1), (u_2, c_2), \dots, (u_l, c_l), (c_l, c_{l+1}), (c_{l+1}, u_{l+1}), \dots, (c_n, u_n)$. In addition, we have the equations $\delta_-(-1) = \delta_-(1) = 0$ which were obtained in the proof of Theorem 11:3 under the assumption $s_\xi = s_\gamma$, which is equivalent to assumption (29). Thus, we proved the existence of $n+3$ zeros of the function δ_- .

If, instead of equations (29), we have the equations $\text{sign}(\xi(u_k) - \omega_n(u_k)) = -\text{sign } \gamma_{n+1}(c_k)$, then we can prove in a similar way the existence of at least $n+3$ zeros of the function $\delta_+ = \|\gamma_{n+1}\|_I(\xi - \omega_n) + \varepsilon_n(\xi)\gamma_{n+1}$, which again contradicts Theorem 11:1. This contradiction has followed from hypothesis (28). Thus, the distribution of points c_1, c_2, \dots, c_n in the interval I is consistent with Theorems 11:2 and 11:3 only if we have inequalities (26) or (27).

We shall now assume that inequality (26) holds and we shall prove that $\xi^{(n+1)}(t)\xi^{(n+2)}(t) < 0$ if (Ξ_{12345}^{n2}) , or $\gamma_{n+1}^{(n+1)}(t)\gamma_{n+1}^{(n+2)}(t) > 0$ if (Γ_{12345}^{n2}) . The above statement will be sufficient for the proof of Theorem 11:7, since it could be shown in an analogous way that inequalities (27) imply the opposite inequalities for the derivatives of function ξ or γ_{n+1} .

Let

$$\delta = \|\gamma_{n+1}\|_I(\xi - \omega_n) - \varepsilon_n(\xi)\gamma_{n+1} \text{sign } \xi^{(n+1)}(t)\gamma_{n+1}^{(n+1)}(t) = \omega + g\gamma_{n+1} + x\xi,$$

where

$$\omega = -\|\gamma_{n+1}\|_I\omega_n, \quad g = -\varepsilon_n(\xi)\text{sign } \xi^{(n+1)}(t)\gamma_{n+1}^{(n+1)}(t), \quad x = \|\gamma_{n+1}\|_I.$$

The function δ is equal either to the function δ_- or to δ_+ , according to the sign of functions $\xi^{(n+1)}$ and $\gamma_{n+1}^{(n+1)}$. The value of function δ at each of the (n) -points u_0, u_1, \dots, u_{n+1} of the function ξ has the sign consistent with the value of the difference $\xi - \omega_n$ at the point, or it is equal to zero. Since we have assumed that $\xi^{(n+1)}$ has a constant sign (Ξ_4^{n2}) or (Γ_4^{n2}) , according to equations (12), § 10, the sign $\delta(u_k)$ is equal to $(-1)^{n+1-k}\text{sign } \xi^{(n+1)}(t)$ or zero, hence

$$\delta(u_k) = (-1)^{n+1-k}|\delta(u_k)|\text{sign } \xi^{(n+1)}(t) \quad (k = 0, 1, \dots, n).$$

Besides from assumption (1) for $p = n + 1$ and $l = n$ and from the assumption of the constancy of the sign of the function $\gamma_{n+1}^{(n+1)}$ ((Ξ_4^{n2}) or (Γ_4^{n2})), it follows that $\text{sign } \delta(c_n) = -\text{sign } \gamma_{n+1}(c_n) \text{sign } \xi^{(n+1)}(t) \gamma_{n+1}^{(n+1)}(t) = \text{sign } \xi^{(n+1)}(t)$, hence $\delta(c_n) = |\delta(c_n)| \text{sign } \xi^{(n+1)}(t)$. Finally, by definition of the function δ , we have $\delta(1) = 0$ and $\delta(1) = -|\delta(1)| \text{sign } \xi^{(n+1)}(t)$. Thus, we obtain the following system of $n + 3$ equations:

$$\begin{aligned} \omega(u_k) + g\gamma_{n+1}(u_k) + x\xi(u_k) &= (-1)^{n+1-k} |\delta(u_k)| \text{sign } \xi^{(n+1)}(t) \\ &\quad (k = 0, 1, \dots, n), \\ \omega(c_n) + g\gamma_{n+1}(c_n) + x\xi(c_n) &= |\delta(c_n)| \text{sign } \xi^{(n+1)}(t), \\ \omega(1) + g\gamma_{n+1}(1) + x\xi(1) &= -|\delta(1)| \text{sign } \xi^{(n+1)}(t). \end{aligned}$$

These equations are linear with respect to $n + 1$ coefficients of the polynomial $\omega \in \mathcal{W}_n$ of t^0, t^1, \dots, t^n , and the numbers g and x . Treating these parameters as unknowns and applying Cramer's formula we obtain

$$(30) \quad x \text{sign } \xi^{(n+1)}(t) = \frac{\begin{vmatrix} 1 & u_0 & \dots & u_0^n & \gamma_{n+1}(u_0) & (-1)^{n+1} |\delta(u_0)| \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & u_n & \dots & u_n^n & \gamma_{n+1}(u_n) & -|\delta(u_n)| \\ 1 & c_n & \dots & c_n^n & \gamma_{n+1}(c_n) & |\delta(c_n)| \\ 1 & 1 & \dots & 1 & \gamma_{n+1}(1) & -|\delta(1)| \end{vmatrix}}{\begin{vmatrix} 1 & u_0 & \dots & u_0^n & \gamma_{n+1}(u_0) & \xi(u_0) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & u_n & \dots & u_n^n & \gamma_{n+1}(u_n) & \xi(u_n) \\ 1 & c_n & \dots & c_n^n & \gamma_{n+1}(c_n) & \xi(c_n) \\ 1 & 1 & \dots & 1 & \gamma_{n+1}(1) & \xi(1) \end{vmatrix}}.$$

According to assumption (26) the points $u_0, u_1, \dots, u_n, c_n, 1$ are distinct and are arranged according to their magnitudes. Applying Theorem 10:4 for $p = n + 1$ and $\gamma = \gamma_{n+1}$ to the denominator of the right-hand side of equation (30) we see that the sign of this determinant equals $\text{sign } \gamma_{n+1}^{(n+1)}(t) \xi^{(n+2)}(t)$ if (Ξ_{1345}^{n2}) , then the assumptions (Ξ_{123}) of Theorem 10:4 are satisfied, or $-\text{sign } \gamma_{n+1}^{(n+2)}(t) \xi^{(n+1)}(t)$ if (Γ_{1345}^{n2}) , then the assumptions (Γ_{123}) of Theorem 10:4 are satisfied. The numerator of the right-hand side of (30) may be expanded with respect to the last column. Then we obtain a sum of negative numbers multiplied by the determinants of which the sign is determined by formula (10), § 10 for $p = n + 1$, $\eta = \gamma_{n+1}$. Thus, the sign of the numerator equals $-\text{sign } \gamma_{n+1}^{(n+1)}(t)$. Since $x = \|\gamma_{n+1}\|_I > 0$, from (30) follow the relations between the signs

of the values of the derivatives of functions ξ or γ_{n+1} : in the case (Ξ_{12345}^{n2})

$$\text{sign } \xi^{(n+1)}(t) = \text{sign}(-\gamma_{n+1}^{(n+1)}(t))\gamma_{n+1}^{(n+1)}(t)\xi^{(n+2)}(t) = -\text{sign } \xi^{(n+2)}(t),$$

hence $\xi^{(n+1)}(t)\xi^{(n+2)}(t) < 0$;

in the case (Γ_{12345}^{n2})

$$\begin{aligned} \text{sign } \xi^{(n+1)}(t) &= \text{sign}(-\gamma_{n+1}^{(n+1)}(t))(-\gamma_{n+1}^{(n+2)}(t)\xi^{(n+1)}(t)) \\ &= \text{sign } \gamma_{n+1}^{(n+1)}(t)\gamma_{n+1}^{(n+2)}(t)\xi^{(n+1)}(t), \quad \text{hence } \gamma_{n+1}^{(n+1)}(t)\gamma_{n+1}^{(n+2)}(t) > 0, \end{aligned}$$

which was to be proved.

THEOREM 11:8. *Let $m \geq 2$ and let the assumptions (Ξ_{1345}^{nm}) be satisfied. If $\xi^{(n+1)}(t)\xi^{(n+m)}(t) < 0$, then we have inequalities*

$$(31) \quad u_k \leq c_{n+m-1, k+m-2} \quad (k = 1, 2, \dots, n)$$

and if $(-1)^m \xi^{(n+1)}(t)\xi^{(n+m)}(t) > 0$, we have inequalities

$$(32) \quad u_k \geq c_{n+m-1, k} \quad (k = 1, 2, \dots, n).$$

Let $m \geq 2$ and let the assumptions (Γ_{1345}^{nm}) be satisfied. If $\xi^{(n+1)}(t)\xi^{(n+m-1)}(t) \times \gamma_{n+m-1}^{(n+m-1)}(t)\gamma_{n+m-1}^{(n+m)}(t) > 0$, we have inequalities (31), and if $(-1)^m \xi^{(n+1)}(t) \times \xi^{(n+m-1)}(t)\gamma_{n+m-1}^{(n+m-1)}(t)\gamma_{n+m-1}^{(n+m)}(t) < 0$, then we have inequalities (32).

Proof. Suppose that under the assumptions formulated above the part of the theorem concerning inequalities (31) is false. This means that we have $u_k > c_{k+m-2}$ (where $c_l \equiv c_{n+m-1, l}$) for some k and inequality

$$(33) \quad c_0 < c_1 < \dots < c_{k+m-2} < u_k < u_{k+1} < \dots < u_{n+1}$$

is satisfied.

Let

$$\begin{aligned} \delta &= \|\gamma_{n+m-1}\|_I (\xi - \omega_n) - \varepsilon_n(\xi) \gamma_{n+m-1} \text{sign } \xi^{(n+1)}(t) \gamma_{n+m-1}^{(n+m-1)}(t) \\ &= \omega + g\gamma_{n+m-1} + x\xi, \end{aligned}$$

where

$$\begin{aligned} \omega &= -\|\gamma_{n+m-1}\|_I \omega_n, \quad g = -\varepsilon_n(\xi) \text{sign } \xi^{(n+1)}(t) \gamma_{n+m-1}^{(n+m-1)}(t), \\ x &= \|\gamma_{n+m-1}\|_I. \end{aligned}$$

As in the proof of the preceding theorem, function δ equals δ_- or δ_+ . From the definition of this function and points c_i, u_j follow the equation

$$\begin{aligned} \delta(c_i) &= (-1)^{n+m-i} |\delta(c_i)| \text{sign } \xi^{(n+1)}(t) \quad (i = 0, 1, \dots, k+m-2), \\ \delta(u_j) &= (-1)^{n+1-j} |\delta(u_j)| \text{sign } \xi^{(n+1)}(t) \quad (j = k, k+1, \dots, n+1). \end{aligned}$$

They can be represented in the form

$$\begin{aligned} \omega(c_i) + g\gamma_{n+m-1}(c_i) + x\xi(c_i) &= (-1)^{n+m-i} |\delta(c_i)| \text{sign } \xi^{(n+1)}(t) \\ &\quad (i = 0, 1, \dots, k+m-2), \\ \omega(u_j) + g\gamma_{n+m-1}(u_j) + x\xi(u_j) &= (-1)^{n+1-j} |\delta(u_j)| \text{sign } \xi^{(n+1)}(t) \\ &\quad (j = k, k+1, \dots, n+1). \end{aligned}$$

Here we have $n + m + 1$ of these equations; they are linear with respect to $n + m - 1$ coefficients of the polynomial ω of $t^0, t^1, \dots, t^{n+m-2}$ and the numbers g and x . It follows that

$$x \operatorname{sign} \xi^{(n+1)}(t) = \frac{\begin{vmatrix} 1 & c_0 & \dots & c_0^{n+m-2} & \gamma_{n+m-1}(c_0) & (-1)^{n+m} |\delta(c_0)| \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & c_{k+m-2} & \dots & c_{k+m-2}^{n+m-2} & \gamma_{n+m-1}(c_{k+m-2}) & (-1)^{n+2-k} |\delta(c_{k+m-2})| \\ 1 & u_k & \dots & u_k^{n+m-2} & \gamma_{n+m-1}(u_k) & (-1)^{n+1-k} |\delta(u_k)| \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & u_{n+1} & \dots & u_{n+1}^{n+m-2} & \gamma_{n+m-1}(u_{n+1}) & |\delta(u_{n+1})| \end{vmatrix}}{\begin{vmatrix} 1 & c_0 & \dots & c_0^{n+m-2} & \gamma_{n+m-1}(c_0) & \xi(c_0) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & c_{k+m-2} & \dots & c_{k+m-2}^{n+m-2} & \gamma_{n+m-1}(c_{k+m-2}) & \xi(c_{k+m-2}) \\ 1 & u_k & \dots & u_k^{n+m-2} & \gamma_{n+m-1}(u_k) & \xi(u_k) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & u_{n+1} & \dots & u_{n+1}^{n+m-2} & \gamma_{n+m-1}(u_{n+1}) & \xi(u_{n+1}) \end{vmatrix}}.$$

The numerator of the right-hand side of this equation can be expanded with respect to the last column. According to inequality (33) the sign of the determinants obtained in such a way may be determined from formula (10), § 10 for $p = n + m - 1$, $\eta = \gamma_{n+m-1}$. The same inequality allows us to apply Theorem 10:4 for $p = n + m - 1$, $\gamma = \gamma_{n+m-1}$ to the denominator. Thus, in the case (Ξ_{1345}^{nm}) we obtain the equation

$$\operatorname{sign} \xi^{(n+1)}(t) = \operatorname{sign} \gamma_{n+m-1}^{(n+m-1)}(t) \gamma_{n+m-1}^{(n+m-1)}(t) \xi^{(n+m)}(t),$$

hence $\xi^{(n+1)}(t) \xi^{(n+m)}(t) > 0$,

and in the case (Γ_{1345}^{nm}) —the equality

$$\operatorname{sign} \xi^{(n+1)}(t) = \operatorname{sign} \gamma_{n+m-1}^{(n+m-1)}(t) (-\gamma_{n+m-1}^{(n+m)}(t) \xi^{(n+m-1)}(t)),$$

hence $\xi^{(n+1)}(t) \xi^{(n+m-1)}(t) \gamma_{n+m-1}^{(n+m-1)}(t) \gamma_{n+m-1}^{(n+m)}(t) < 0$.

These equations contradict the assumptions of our theorem which proves that (33) is false; i.e., inequalities (31) are true.

The proof of the part of the theorem which is connected with inequality (32) is similar; thus, we shall present it but briefly. Now we shall define the function δ by the formula $\delta = \omega + g\gamma_{n+m-1} + x\xi$, where ω and x have the same meaning as before and $g = -(-1)^m \operatorname{sign} \xi^{(n+1)}(t) \times \gamma_{n+m-1}^{(n+m-1)}(t)$. If, contrary to our theorem, we assume that for the suitable assumption we have for some k the inequality

$$u_0 < u_1 < \dots < u_k < c_k < c_{k+1} < \dots < c_{n+m-1}$$

and apply the equation

$$\omega(u_i) + g\gamma_{n+m-1}(u_i) + x\xi(u_i) = (-1)^{n+1-i} |\delta(u_i)| \text{sign } \xi^{(n+1)}(t) \quad (i = 0, 1, \dots, k),$$

$$\omega(c_j) + g\gamma_{n+m-1}(c_j) + x\xi(c_j) = (-1)^{n-j} |\delta(c_j)| \text{sign } \xi^{(n+1)}(t) \quad (j = k, k+1, \dots, n+m-1)$$

we obtain

$$x \text{sign } \xi^{(n+1)}(t) = \begin{vmatrix} 1 & u_0 & \dots & u_0^{n+m-2} & \gamma_{n+m-1}(u_0) & (-1)^{n+1} |\delta(u_0)| \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & u_k & \dots & u_k^{n+m-2} & \gamma_{n+m-1}(u_k) & (-1)^{n+1-k} |\delta(u_k)| \\ 1 & c_k & \dots & c_k^{n+m-2} & \gamma_{n+m-1}(c_k) & (-1)^{n-k} |\delta(c_k)| \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & c_{n+m-1} & \dots & c_{n+m-1}^{n+m-2} & \gamma_{n+m-1}(c_{n+m-1}) & (-1)^{m+1} |\delta(c_{n+m-1})| \\ 1 & u_0 & \dots & u_0^{n+m-2} & \gamma_{n+m-1}(u_0) & \xi(u_0) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & u_k & \dots & u_k^{n+m-2} & \gamma_{n+m-1}(u_k) & \xi(u_k) \\ 1 & c_k & \dots & c_k^{n+m-2} & \gamma_{n+m-1}(c_k) & \xi(c_k) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & c_{n+m-1} & \dots & c_{n+m-1}^{n+m-2} & \gamma_{n+m-1}(c_{n+m-1}) & \xi(c_{n+m-1}) \end{vmatrix}.$$

In view of the theorems from § 10.2 which were applied in the first part of the proof, it follows that in the case (Ξ_{1345}^{nm}) we have

$$\text{sign } \xi^{(n+1)}(t) = (-1)^{m+1} \text{sign } \gamma_{n+m-1}^{(n+m-1)}(t) \gamma_{n+m-1}^{(n+m-1)}(t) \xi^{(n+m)}(t)$$

and in the case (Γ_{1345}^{nm}) we have

$$\text{sign } \xi^{(n+1)}(t) = (-1)^{m+1} \text{sign } \gamma_{n+m-1}^{(n+m-1)}(t) (-\gamma_{n+m-1}^{(n+m)}(t) \xi^{(n+m-1)}(t)),$$

which contradicts the assumptions of our theorem. This completes the proof of Theorem 11:8.

If m is even, Theorem 11:8 gives the estimate of the (n) -points u_1, u_2, \dots, u_n of function ξ either from below or from above, no matter what the signs of the derivative are values for the functions ξ and γ_{n+m-1} . If, however, m is odd (equals 3, 5, ...), then two-sided estimates follow from this theorem. We shall formulate them in the next theorem.

THEOREM 11:8'. *Let m be an odd integer greater than 1. If either assumptions (Ξ_{1345}^{nm}) and the inequality $\xi^{(n+1)}(t) \xi^{(n+m)}(t) < 0$ or the assumptions (Γ_{1345}^{nm}) and the inequality $\xi^{(n+1)}(t) \xi^{(n+m-1)}(t) \gamma_{n+m-1}^{(n+m-1)}(t) \gamma_{n+m-1}^{(n+m)}(t) > 0$ are satisfied, then*

$$u_k \in \langle c_{n+m-1, k}, c_{n+m-1, k+m-2} \rangle \quad (k = 1, 2, \dots, n).$$

Finally, let us notice, that the estimates of (n) -points contained in Theorem 11:8 seem to be better (at least for odd m) than the analogous estimates from Theorem 11:6. Besides, their proof is simpler, for it does not require the applications of the auxiliary Theorems 11:1, 11:2 and 11:3. However, the assumptions of theorems under comparison are not identical. The common part of them consists of (Ξ_{13}^{nm}) or (Γ_{13}^{nm}) .

12. The estimates dependent upon the properties of the derivatives of the function ξ . The choice of function γ_p .

12.1. In the previous section the function γ_p was not defined uniquely. We have only assumed that, besides having other properties it is the difference between some function β_p and its $(p-1)$ -st best polynomial. The numbers c_{pl} have been defined as the $(p-1)$ -points of the function β_p . These numbers serve as the estimates of the (n) -points u_k of the function ξ . In general, finding the $(p-1)$ -points c_{pl} is as difficult as finding the (n) -points u_k . We can, however, define the function β_p in such a way that the points c_{pl} can be easily found, or we may get some simplifications in comparison with the general theory.

The simplest and most important case occurs when $\gamma_p = a\tau_p$, where $a \neq 0$ and τ_p is the p -th Chebyshev polynomial (§ 2.2). According to Theorem 2:2 this polynomial is the difference between the function $2^{p-1}t^p$ and its $(p-1)$ -st best polynomial. The p -th derivative of the function $2^{p-1}t^p$ is constant, hence it has a constant sign, and the function $a\tau_p$ satisfies the assumptions (Γ_{12}^p) from § 11.1. Moreover, for an arbitrary a we have

$$c_{pl} = t_{pl} = -\cos l\pi/p \quad (l = 0, 1, \dots, p).$$

If $\gamma_{n+m-1} = a\tau_{n+m-1}$, then the assumption (Γ_3^{nm}) takes the form $|\xi^{(n+m)}(t)| \leq |\gamma_{n+m-1}^{(n+m)}(t)| = 0$ for $t \in I$. This case is not interesting, and we shall formulate only the assumptions (Ξ_{12345}^{nm}) :

- $(\Xi_1^{nm}) \quad \xi \in \mathcal{D}_I^{n+m},$
- $(\Xi_2^{nm}) \quad \varepsilon_n(\xi) < \|\gamma_{n+m-1}\|_I = |a| \cdot \|\tau_{n+m-1}\|_I = |a|,$
- $(\Xi_3^{nm}) \quad |\xi^{(n+m)}(t)| \geq |\gamma_{n+m-1}^{(n+m)}(t)| = 0 \quad \text{for } t \in I,$
- $(\Xi_4^{nm}) \quad \text{the function } \xi^{(n+1)} \text{ has a constant sign in the interval } I,$
- $(\Xi_5^{nm}) \quad |\xi^{(n+m-1)}(t)| < |\gamma_{n+m-1}^{(n+m-1)}(t)| = 2^{n+m-2}(n+m-1)!|a| \quad \text{for } t \in I.$

The third assumption is satisfied for every function from the class \mathcal{E}_I^{n+m} . The second and the fifth assumptions are satisfied for a sufficiently large a . Thus only the first and the fourth assumptions remain. By Theorem 10:5 the fourth assumption guarantees the fulfillment of the assumption (Ξ) (see § 11.1) for the (n) -points of the function ξ ; thus we shall add it to the system (Ξ_{123}^{nm}) , obtaining the identical assumptions $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+m}$ (for a suitable m) in all theorems from § 11. Then Theorem 11:6

will be weaker than 11:8, and Theorem 11:6' will be weaker than 11:7. Thus we shall present only the corollaries from Theorems 11:4, 11:8, 11:8' and 11:7.

THEOREM 12:1 (Bernstein, [4], p. 85). *If $\xi \in \mathcal{D}_I^{n+1}$, then we have the inequality*

$$-1 = u_0 = t_{n0} < u_1 < t_{n1} < \dots < u_n < t_{nn} = u_{n+1} = 1.$$

THEOREM 12:2. *Let $m \geq 2$ and $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+m}$. If $\xi^{(n+1)}(t) \xi^{(n+m)}(t) < 0$, then we have the inequalities*

$$u_k \leq t_{n+m-1, k+m-2} \quad (k = 1, 2, \dots, n)$$

and if $(-1)^m \xi^{(n+1)}(t) \xi^{(n+m)}(t) > 0$, then

$$u_k \geq t_{n+m-1, k} \quad (k = 1, 2, \dots, n).$$

THEOREM 12:2'. *If m is an odd number greater than 1, $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+m}$ and $\xi^{(n+1)}(t) \xi^{(n+m)}(t) < 0$, then*

$$u_k \in \langle t_{n+m-1, k}, t_{n+m-1, k+m-2} \rangle \quad (k = 1, 2, \dots, n).$$

THEOREM 12:3 (Paszkowski, [24], p. 49). *If $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+2}$, then for $k = 1, 2, \dots, n$*

$$(1) \quad u_k \in (t_{n, k-1}, t_{n+1, k}) \quad (\xi^{(n+1)}(t) \xi^{(n+2)}(t) < 0),$$

$$(2) \quad u_k \in (t_{n+1, k}, t_{nk}) \quad (\xi^{(n+1)}(t) \xi^{(n+2)}(t) > 0).$$

Proof. It follows from Theorem 12:1 that $u_k \in (t_{n, k-1}, t_{nk})$. On the other hand it follows from Theorem 11:7 that $u_k \in (t_{n+1, k-1}, t_{n+1, k})$ if $\xi^{(n+1)}(t) \xi^{(n+2)}(t) < 0$ and $u_k \in (t_{n+1, k}, t_{n+1, k+1})$ if $\xi^{(n+1)}(t) \xi^{(n+2)}(t) > 0$. Thus it suffices to verify that

$$(t_{n, k-1}, t_{nk}) \cap (t_{n+1, k-1}, t_{n+1, k}) = (t_{n, k-1}, t_{n+1, k}),$$

$$(t_{n, k-1}, t_{nk}) \cap (t_{n+1, k}, t_{n+1, k+1}) = (t_{n+1, k}, t_{nk}).$$

The necessary inequalities $t_{n, k-1} \geq t_{n+1, k-1}$, $t_{n+1, k} \leq t_{nk}$, $t_{n+1, k} \geq t_{n, k-1}$ and $t_{nk} \leq t_{n+1, k+1}$ follow from the general remark, that the inequality $t_{nk} = -\cos k\pi/n \leq -\cos l\pi/p = t_{nl}$ ($0 \leq k \leq n$, $0 \leq l \leq p$) is equivalent to the inequality $k/n \leq l/p$.

12.2. Theorem 11:5 determines the mutual position of the (n) -points and the $(n+1)$ -points of the function ξ whose $(n+1)$ -st and $(n+2)$ -nd derivatives have a constant sign. Now we shall show that in some cases this theorem may allow us to sharpen the estimates (1) or (2). We shall

denote by $u_{p0}, u_{p1}, \dots, u_{p, p+1}$ the (p) -points of the function ξ arranged according to their magnitude.

THEOREM 12:4. *If $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+2} \cap \dots \cap \mathcal{D}_I^{n+l+1}$ for $l \geq 2$ and the signs of the products $\xi^{(n+1)} \xi^{(n+2)}, \dots, \xi^{(n+l-1)} \xi^{(n+l)}$ coincide, and the sign of the product $\xi^{(n+l)} \xi^{(n+l+1)}$ differs from them, then in the case $\xi^{(n+1)}(t) \xi^{(n+2)}(t) < 0$ we have*

$$u_k \in (t_{n+l, k}, t_{n+1, k}) \quad (k = 1, 2, \dots, [n/l] + 1),$$

$$u_k \in (t_{n, k-1}, t_{n+1, k}) \quad (k = [n/l] + 2, \dots, n)$$

and in the case $\xi^{(n+1)}(t) \xi^{(n+2)}(t) > 0$ we have

$$u_k \in (t_{n+1, k}, t_{nk}) \quad (k = 1, 2, \dots, n - [n/l] - 1),$$

$$u_k \in (t_{n+1, k}, t_{n+l, k+l-1}) \quad (k = n - [n/l], \dots, n).$$

Proof. Under the assumption $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+2} \cap \dots \cap \mathcal{D}_I^{n+l+1}$ Theorem 11:5 holds after changing n into $n+1, \dots, n+l-1$, also. Thus

$$u_{n+j-1, k} \in (u_{n+j, k}, u_{n+j, k+1}) \quad (k = 1, 2, \dots, n+j-1; j = 1, 2, \dots, l).$$

Hence it follows that

$$(3) \quad u_{nk} \in (u_{n+j, k}, u_{n+j, k+j}) \quad (k = 1, 2, \dots, n; j = 1, 2, \dots, l).$$

To estimate the $(n+j)$ -points ($j = 1, 2, \dots, l-1$) we apply Theorem 12:3 after changing n into $n+j$. We may do this for $j = 1, 2, \dots, l-1$. Then we obtain the estimates:

$$u_{n+j, k} \in (t_{n+j, k-1}, t_{n+j+1, k}) \quad (\xi^{(n+j+1)}(t) \xi^{(n+j+2)}(t) < 0),$$

$$u_{n+j, k} \in (t_{n+j+1, k}, t_{n+j, k}) \quad (\xi^{(n+j+1)}(t) \xi^{(n+j+2)}(t) > 0),$$

which allow us to write (3) in the form

$$(4) \quad \begin{aligned} u_{nk} &\in (t_{n+j, k-1}, t_{n+j+1, k+j}) && (\xi^{(n+j+1)}(t) \xi^{(n+j+2)}(t) < 0), \\ u_{nk} &\in (t_{n+j+1, k}, t_{n+j, k+j}) && (\xi^{(n+j+1)}(t) \xi^{(n+j+2)}(t) > 0). \end{aligned}$$

We shall compare the estimates so obtained with estimates (1) and (2). First we shall verify that estimate (4) does not provide us with anything new in comparison with (1) and (2) if the signs of the products $\xi^{(n+1)} \xi^{(n+2)}$ and $\xi^{(n+j+1)} \xi^{(n+j+2)}$ coincide (which holds for $j = 1, 2, \dots, l-2$); i.e., that

$$(t_{n, k-1}, t_{n+1, k}) \subset (t_{n+j, k-1}, t_{n+j+1, k+j}),$$

$$(t_{n+1, k}, t_{nk}) \subset (t_{n+j+1, k}, t_{n+j, k+j}).$$

This follows from the inequalities

$$\frac{k-1}{n} \geq \frac{k-1}{n+j}, \quad \frac{k}{n+1} < \frac{k+j}{n+j+1}, \quad \frac{k}{n+1} > \frac{k}{n+j+1},$$

$$\frac{k}{n} \leq \frac{k+j}{n+j} \quad (k = 1, 2, \dots, n; j \geq 1).$$

Let us now assume that

$$(5) \quad \xi^{(n+1)}(t) \xi^{(n+2)}(t) < 0, \quad \xi^{(n+l)}(t) \xi^{(n+l+1)}(t) > 0$$

(the second inequality follows from the first by the assumption of the theorem). Since $t_{n+1,k} < t_{nk} \leq t_{n+l-1,k+l-1}$, the right end point of the interval

$$(t_{n,k-1}, t_{n+1,k}) \cap (t_{n+l,k}, t_{n+l-1,k+l-1}),$$

to which u_{nk} belongs, is the point $t_{n+1,k}$. The inequality $t_{n,k-1} \leq t_{n+1,k}$ holds if and only if $(k-1)/n \leq k/(n+l)$, i.e., if $k \leq [n/l] + 1$. Thus, under assumptions (5) we have

$$u_{nk} \in (t_{n+l,k}, t_{n+1,k}) \quad \text{for } k \leq [n/l] + 1 \quad \text{and}$$

$$u_{nk} \in (t_{n,k-1}, t_{n+1,k}) \quad \text{for } k > [n/l] + 1.$$

If $\xi^{(n+1)}(t) \xi^{(n+2)}(t) > 0$ and $\xi^{(n+l)}(t) \xi^{(n+l+1)}(t) < 0$, then

$$u_{nk} \in (t_{n+1,k}, t_{nk}) \cap (t_{n+l-1,k-1}, t_{n+l,k+l-1}).$$

Since $t_{n+1,k} > t_{n+l-1,k-1}$ and since $t_{nk} \geq t_{n+l,k+l-1}$ if and only if $k \geq n - [n/l]$, we have $u_{nk} \in (t_{n+1,k}, t_{nk})$ for $k < n - [n/l]$ and $u_{nk} \in (t_{n+1,k}, t_{n+l,k+l-1})$ for $k \geq n - [n/l]$ which was to be proved.

Theorem 12:4 is strongest for $l = 2$; i.e., if the signs of the values of the functions $\xi^{(n+1)}$, $\xi^{(n+2)}$ and $\xi^{(n+3)}$ are constant and form one of the combinations $++-$, $--+$, $+--$, $-++$. Then this theorem improves estimates (1) or (2), approximately, for half of the (n) -points of the function ξ , thus eliminating the intervals, whose joint length is $1 - 2/\pi + O(n^{-2})$.

12.3. The particular case $\gamma_p = a\tau_p$, which was already considered, is important not only for its own interest, but also because it enables us to choose other functions γ_p , for which the estimates from Theorems 11:4-11:8' are essentially better than those obtained from Theorems 12:1-12:4. To achieve this, one must use functions which are connected more closely than before with the approximated function ξ . In the case when the function γ_p was the polynomial $a\tau_p$, the function ξ had an influence only upon the coefficient a , and the numbers c_{pi} were independent of this coefficient. We shall look for the function γ_p with the required

properties in the family of functions which depend upon two parameters. Now we shall define this family. We introduce the notation

$$\beta_{pgh} = 2^{p-1}gt^p - h\eta_p \quad (p \text{ positive integer}).$$

The parameters g and $h > 0$ and the function η_p satisfy the system of assumptions $(-H^p)$, consisting of the conditions (H_1^p) and $(-H_2^p)$, or the system of assumptions $(+H^p)$, consisting of the conditions (H_1^p) and $(+H_2^p)$:

$$\begin{aligned} (H_1^p) \quad & \eta_p \in \mathcal{D}_I^{p+1}, \quad \eta_p^{(p+1)}(t) < 0 \quad \text{for} \quad t \in I, \\ (-H_2^p) \quad & 2^{p-1}p!g/h < y_{p, \min}, \quad \text{where} \quad y_{p, \min} = \min_{t \in I} \eta_p^{(p)}(t), \\ (+H_2^p) \quad & 2^{p-1}p!g/h > y_{p, \max}, \quad \text{where} \quad y_{p, \max} = \max_{t \in I} \eta_p^{(p)}(t). \end{aligned}$$

Since $\beta_{pgh}^{(p)} = 2^{p-1}p!g - h\eta_p^{(p)}$, $\beta_{pgh}^{(p+1)} = -h\eta_p^{(p+1)} > 0$, then under the assumptions $(-H^p)$ or $(+H^p)$ the values of the p -th and the $(p+1)$ -st derivative of function β_{pgh} have constant signs, and

$$(6) \quad \beta_{pgh}^{(p)}(t) \beta_{pgh}^{(p+1)}(t) \begin{cases} < 0 & \text{in the case} \quad (-H^p), \\ > 0 & \text{in the case} \quad (+H^p). \end{cases}$$

According to the general definitions of § 11.1 the function γ_{pgh} will be defined as the difference between the function β_{pgh} and its $(p-1)$ -st best polynomial. The assumptions (Γ_2^p) from § 11.1 follow from the constancy of the sign of the p -th derivative of the function β_{pgh} . Since the multiplication of the function β_{pgh} by any number different from zero, e. g. by $1/h$, does not change its $(p-1)$ -points, these points depend only on the value of the ratio g/h . We shall denote them by the symbols $c_{pl}(g/h)$ ($l = 0, 1, \dots, p$).

By Theorem 12:3 for $n = p-1$ and $\xi = \beta_{pgh}$, and by inequality (6) we have the estimates

$$(7) \quad c_{pl}(g/h) \in \begin{cases} (t_{p-1, l-1}, t_{pl}) & \text{in the case} \quad (-H^p), \\ (t_{pl}, t_{p-1, l}) & \text{in the case} \quad (+H^p). \end{cases}$$

Theorems 12:1-12:4 are the corollaries from theorems of § 11 for function $\gamma_p = a\tau_p$. In this particular case the points c_{pl} were identical with the points t_{pl} . Now we would like to obtain the estimates which are more exact than the ones given by the mentioned corollaries. To do that, in estimating from above the (n) -point u_k of function ξ by the point c_{pl} (for suitable p and l) we shall choose the parameters g and h in such a way that the assumptions $(-H^p)$ will be satisfied. On the other hand, while estimating the point u_k from below by the point c_{pl} we shall choose g and h in such a way that the assumption $(+H^p)$ will be satisfied. Estimates (7) show the aim of such a procedure.

Let us investigate further how to choose the parameters g and h , in order that the sharpening of Theorems 12:1-12:4 is most essential. If we fix $g \neq 0$, and h tends to zero, the function β_{pgh} tends to $2^{p-1}gt^p$ and its $(p-1)$ -points $c_{pl}(g/h)$ tend to the $(p-1)$ -points t_{pl} of the limit function. One may expect that the converse is also true—the closer ratio g/h is to zero, the more the points $c_{pl}(g/h)$ differ from t_{pl} , hence they are better for our purposes. We shall follow this general indication; of course we shall not consider it as a theorem. We shall choose the parameters g and h in such a way that ratio $|g|/h$ will be small under the assumptions $(-H^p)$ or $(+H^p)$ and the assumptions of the corresponding theorems of § 11.

12.4. Before we formulate the corollaries from theorems of § 11, for the functions γ_{pgh} considered now, we shall prove an auxiliary inequality

$$(8) \quad ||g| - h\varepsilon_{p-1}(\eta_p)| \leq \|\gamma_{pgh}\|_I \leq |g| + h\varepsilon_{p-1}(\eta_p).$$

From the definition of the function γ_{pgh} it follows that $\|\gamma_{pgh}\|_I = \varepsilon_{p-1}(\beta_{pgh}) = \varepsilon_{p-1}(2^{p-1}gt^p - h\eta_p)$. We shall estimate this $(p-1)$ -st error of the best approximation from below and from above by inequalities (4) from § 4:

$$|\varepsilon_{p-1}(2^{p-1}gt^p) - \varepsilon_{p-1}(-h\eta_p)| \leq \|\gamma_{pgh}\|_I \leq \varepsilon_{p-1}(2^{p-1}gt^p) + \varepsilon_{p-1}(-h\eta_p).$$

Since $\varepsilon_{p-1}(-h\eta_p) = h\varepsilon_{p-1}(\eta_p)$ by Theorem 4:2, and since $\varepsilon_{p-1}(2^{p-1}gt^p) = 2^{p-1}|g|\varepsilon_{p-1}(t^p) = |g|$ by Theorem 2:2 (we remember that the approximation interval is $\langle -1, 1 \rangle$), inequality (8) follows.

THEOREM 12:5. *If $\xi \in \mathcal{D}_I^{n+1}$ and if the numbers s_+ and s_- satisfy the inequalities*

$$s_+ > \max \left\{ \frac{y_{n,\max}}{2^{n-1}n!}, \varepsilon_{n-1}(\eta_n) + \frac{\varepsilon_n(\xi)}{x_{n+1}} \right\},$$

$$s_- < \min \left\{ \frac{y_{n,\min}}{2^{n-1}n!}, - \left(\varepsilon_{n-1}(\eta_n) + \frac{\varepsilon_n(\xi)}{x_{n+1}} \right) \right\},$$

where

$$x_{n+1} = \min_{t \in I} \left| \frac{\xi^{(n+1)}(t)}{\eta_n^{(n+1)}(t)} \right|,$$

then for $k = 1, 2, \dots, n$ we have the inequalities

$$u_k \in (c_{n,k-1}(s_+), c_{nk}(s_-)).$$

Proof. Let us write for $\gamma_\nu = \gamma_{\rho\sigma h}$ the assumptions (Ξ_{123}^{n1}) and (Γ_{123}^{n1}) from § 11.1 sharpened according to inequality (8):

- $(\Xi_1^{n1}) \quad \xi \in \mathcal{D}_I^{n+1},$
- $(\Xi_2^{n1}) \quad \varepsilon_n(\xi) < |g| - h\varepsilon_{n-1}(\eta_n),$
- $(\Xi_3^{n1}) \quad |\xi^{(n+1)}(t)| \geq h |\eta_n^{(n+1)}(t)| \quad \text{for } t \in I,$
- $(\Gamma_1^{n1}) \quad \xi \in \mathcal{C}_I^{n+1},$
- $(\Gamma_2^{n1}) \quad \varepsilon_n(\xi) > |g| + h\varepsilon_{n-1}(\eta_n),$
- $(\Gamma_3^{n1}) \quad |\xi^{(n+1)}(t)| \leq h |\eta_n^{(n+1)}(t)| \quad \text{for } t \in I.$

The assumption (Γ_3^{n1}) may be written in the form $h \geq \|\xi^{(n+1)}/\eta_n^{(n+1)}\|_I$. From the assumptions (Γ_2^{n1}) and (Γ_3^{n1}) it follows that $\varepsilon_n(\xi)/\varepsilon_{n-1}(\eta_n) > \|\xi^{(n+1)}/\eta_n^{(n+1)}\|_I$. For some functions ξ and η_n this inequality may be false. Thus we shall not consider the case (Γ_{123}^{n1}) , and we shall assume that the assumptions (Ξ_{123}^{n1}) are satisfied. The assumption (Ξ_2^{n1}) will be sharpened more, by requiring that $\varepsilon_n(\xi) < |g| - h\varepsilon_{n-1}(\eta_n)$. Assuming that $h = x_{n+1}$ (then the assumption (Ξ_3^{n1}) is satisfied) we get the inequality

$$(9) \quad |g|/h > \varepsilon_{n-1}(\eta_n) + \varepsilon_n(\xi)/x_{n+1}.$$

We may apply Theorem 11:4 to the function $\gamma_{\rho\sigma h}$ with the parameters g and h which satisfy these conditions, obtaining in this way

$$(10) \quad u_k > c_{n,k-1}(g/h) \quad (k = 1, 2, \dots, n),$$

$$(11) \quad u_k < c_{nk}(g/h) \quad (k = 1, 2, \dots, n).$$

In § 12.3 we have noticed that estimate (10) should be applied when the assumptions $(^+H^n)$ are satisfied; i.e., when

$$g/h > y_{n,\max}/(2^{n-1}n!).$$

On the other hand, estimate (11) should be applied when the assumptions $(^-H^n)$ are satisfied; i.e., when

$$g/h < y_{n,\min}/(2^{n-1}n!).$$

These new conditions, together with inequality (9), applied to (10) and (11) respectively, give the conditions

$$u_k > c_{n,k-1}(s_+), \quad u_k < c_{nk}(s_-) \quad (k = 1, 2, \dots, n),$$

which were to be proved. One may expect (§ 12.3) that these inequalities would be better if the number s_+ would be smaller, and the number s_- would be greater (closer to zero).

THEOREM 12:6. *Let $m \geq 2$ and $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+m}$. If $\xi^{(n+1)}(t)\xi^{(n+m)}(t) < 0$ then we have the inequalities*

$$u_k \leq c_{n+m-1,k+m-2}(s_-) \quad (k = 1, 2, \dots, n)$$

and if $(-1)^m \xi^{(n+1)}(t) \xi^{(n+m)}(t) > 0$ then we have the inequalities

$$u_k \geq c_{n+m-1,k}(s_+) \quad (k = 1, 2, \dots, n),$$

where

$$(12) \quad s_+ > \frac{1}{2^{n+m-2}(n+m-1)!} \left(y_{n+m-1, \max} + \frac{\|\xi^{(n+m-1)}\|_I}{x_{n+m}} \right),$$

$$(13) \quad s_- < \frac{1}{2^{n+m-2}(n+m-1)!} \left(y_{n+m-1, \min} - \frac{\|\xi^{(n+m-1)}\|_I}{x_{n+m}} \right),$$

$$x_{n+m} = \min_{t \in I} \left| \frac{\xi^{(n+m)}(t)}{\eta_{n+m-1}^{(n+m)}(t)} \right|.$$

Proof is based upon Theorem 11:8 with the assumptions (Ξ_{1345}^{nm}). These assumptions, besides the requirement $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+m}$, contain the inequalities

$$|\xi^{(n+m)}(t)| \geq |\gamma_{n+m-1, a, h}^{(n+m)}(t)|,$$

$$\text{i.e.} \quad |\xi^{(n+m)}(t)| \geq h |\eta_{n+m-1}^{(n+m)}(t)| \quad \text{for } t \in I,$$

$$|\xi^{(n+m-1)}(t)| < |\gamma_{n+m-1, a, h}^{(n+m-1)}(t)|,$$

$$\text{i.e.} \quad |\xi^{(n+m-1)}(t)| < |2^{n+m-2}(n+m-1)!g - h\eta_{n+m-1}^{(n+m-1)}(t)| \quad \text{for } t \in I.$$

In order to have the first inequality satisfied, we assume that $h = x_{n+m}$. Then the second inequality will be satisfied in two cases: if

$$2^{n+m-2}(n+m-1)!g - h\eta_{n+m-1}^{(n+m-1)}(t) > \|\xi^{(n+m-1)}\|_I,$$

which is the same as

$$(14) \quad \frac{g}{h} > \frac{1}{2^{n+m-2}(n+m-1)!} \left(y_{n+m-1, \max} + \frac{\|\xi^{(n+m-1)}\|_I}{x_{n+m}} \right)$$

or if

$$2^{n+m-2}(n+m-1)!g - h\eta_{n+m-1}^{(n+m-1)}(t) < -\|\xi^{(n+m-1)}\|_I,$$

which is the same as

$$(15) \quad \frac{g}{h} < \frac{1}{2^{n+m-2}(n+m-1)!} \left(y_{n+m-1, \min} - \frac{\|\xi^{(n+m-1)}\|_I}{x_{n+m}} \right).$$

It is obvious that the assumption (H_2^{n+m-1}) follows from inequality (14) and the assumption (H_2^{n+m-1}) follows from inequality (15). The rest of the proof, similar to the preceding theorem, makes use of the general remarks concerning the choice of parameters g and h .

From Theorem 12:6 we obtain the following immediate corollary:

THEOREM 12:6'. *If m is an odd integer greater than 1, $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+m}$ and $\xi^{(n+1)}(t) \xi^{(n+m)}(t) < 0$, then*

$$u_k \in \langle c_{n+m-1,k}(s_+), c_{n+m-1,k+m-2}(s_-) \rangle \quad (k = 1, 2, \dots, n),$$

where the numbers s_+ and s_- satisfy inequalities (12) and (13).

THEOREM 12:7. *If $\xi \in \mathcal{D}_I^{n+1} \cap \mathcal{D}_I^{n+2}$, then for $k = 1, 2, \dots, n$*

$$u_k \in \langle c_{n+1,k-1}(s_+), c_{n+1,k}(s_-) \rangle \quad (\xi^{(n+1)}(t) \xi^{(n+2)}(t) < 0),$$

$$u_k \in \langle c_{n+1,k}(s_+), c_{n+1,k+1}(s_-) \rangle \quad (\xi^{(n+1)}(t) \xi^{(n+2)}(t) > 0),$$

where

$$(16) \quad s_+ > \max \left\{ \varepsilon_n(\eta_{n+1}) + \frac{\varepsilon_n(\xi)}{x_{n+2}}, \frac{1}{2^n(n+1)!} \left(y_{n+1,\max} + \frac{\|\xi^{(n+1)}\|_I}{x_{n+2}} \right) \right\},$$

$$s_- < \min \left\{ - \left(\varepsilon_n(\eta_{n+1}) + \frac{\varepsilon_n(\xi)}{x_{n+2}} \right), \frac{1}{2^n(n+1)!} \left(y_{n+1,\min} - \frac{\|\xi^{(n+1)}\|_I}{x_{n+2}} \right) \right\},$$

$$x_{n+2} = \min_{t \in I} \left| \frac{\xi^{(n+2)}(t)}{\eta_{n+1}^{(n+2)}(t)} \right|.$$

The proof is analogous to that of the two preceding theorems. Thus we shall only state that it is based on Theorem 11:7 with the assumptions (\mathcal{E}_{12345}^{n2}) which contain the inequalities

$$\varepsilon_n(\xi) < \|\gamma_{n+1,\theta,h}\|_I, \quad |\xi^{(n+2)}(t)| \geq |\gamma_{n+1,\theta,h}^{(n+2)}(t)| \quad (t \in I),$$

$$|\xi^{(n+1)}(t)| < |\gamma_{n+1,\theta,h}^{(n+1)}(t)| \quad (t \in I).$$

Together with assumptions $(+H^{n+2})$ or $(-H^{n+2})$ they determine the form of the inequalities which restrict the values s_+ and s_- of parameter g/h .

Theorems 12:5, 12:6 and 12:6' give better estimates for the (n) -points of function ξ than Theorems 12:1, 12:2 and 12:2', respectively. Similarly, Theorem 12:7 together with 12:5 is stronger than Theorem 12:3.

The points $c_{pl}(g/h)$ are determined by the values of parameters g and h and the function η_p . This function has only to satisfy the assumption (H_1^p) ; i.e., its $(p+1)$ -st derivative should be continuous, and have the negative values in the interval I . The simplest case is $\eta_p = -t^{p+1}$. Then we have $\eta_p^{(p+1)} = -(p+1)! < 0$. For such a function the determination of numbers s_+ and s_- will be considerably simplified, and in-

equality (16), for example, will be reduced to the form

$$s_{n+1} > \max \left\{ \frac{1}{2^{n+1}} + \frac{(n+2)! \varepsilon_n(\xi)}{x_{n+2}^*}, \frac{n+2}{2^n} \left(1 + \frac{\|\xi^{(n+1)}\|_I}{x_{n+2}^*} \right) \right\},$$

where $x_{n+2}^* = \min_{t \in I} |\xi^{(n+2)}(t)|$.

The application of Theorems 12:5-12:7 requires the knowledge of the points $c_{pi}(g/h)$. In the present case they are simply the $(p-1)$ -points of the function $\beta_{pgh} = ht^{p+1} + 2^{p-1}gt^p$, connected with so-called Zolotarev polynomials which will be discussed in Chapter IV of this paper.

CHAPTER IV

METHODS OF COMPUTING WELL-APPROXIMATING AND BEST POLYNOMIALS

13. Well-approximating polynomials. While reading the next paragraphs we shall see that the problem of finding the best polynomial with a given accuracy, though theoretically solvable, may be very cumbersome. Thus, it is worth-while to find methods which are more primitive in the sense that none of them gives the best polynomial but all of them provide us with the polynomials which approximate the given function ξ sufficiently well. We shall not state precisely what will be understood under the term "well-approximating polynomial τ ". In any case, it should approximate the function ξ with an error $\|\xi - \tau\|$ which does not exceed to much the error $\varepsilon_n(\xi)$ of the best approximation, and the estimate of $\|\xi - \tau\|$ should be known.

In the subsequent part of the section we shall give the following three methods of computing the well-approximating polynomials: 1. by excluding the basic part from the remainder of the power series using either the Chebyshev or the Zolotarev polynomials (§ 13.1), 2. by expanding into an orthogonal series with respect to the Chebyshev polynomials (§ 13.2) and 3. by interpolation with the use of certain nodes connected with the Chebyshev polynomials (§ 13.3 and 13.4). It can be seen from this enumeration how important is the part played by the Chebyshev polynomials also in the numerical methods of the theory of approximation.

The whole § 13 deals with the approximation of a continuous function ξ in the interval $I = \langle -1, 1 \rangle$.

13.1. For $p > 0$ the Chebyshev polynomial τ_p is expressed the formula

$$(1) \quad \tau_p = p \sum_{j=0}^{\lfloor p/2 \rfloor} \frac{(-1)^j}{p-j} \binom{p-j}{j} 2^{p-2j-1} t^{p-2j},$$

mentioned already in § 2.2. Thus, it is a polynomial of degree p with the coefficient 2^{p-1} of t^p . Hence the polynomial $\tilde{\tau}_p = 2^{-p+1} \tau_p$ is the p -th Chebyshev polynomial normed in such a way that the coefficient of t^p equals 1 (and the coefficient of t^{p-1} equals 0).

In the next two theorems we assume that

$$(2) \quad \xi = \sum_{k=0}^{n+2} x_k t^k + \varrho_{n+3} \quad (n \geq 0, \varrho_{n+3} \in \mathcal{C}_I).$$

These theorems are worth applying in practice, if the norm $\|\varrho_{n+3}\|_I$ is small in comparison with the coefficients x_k . In particular, the function ϱ_{n+3} may be equal to the remainder of the power series convergent to the function ξ in the interval I , if such a series exists.

THEOREM 13:1. *If the function ξ is expressed by formula (2) and*

$$\varphi_n = \sum_{k=0}^n x_k t^k + x_{n+1}(t^{n+1} - \tilde{\tau}_{n+1}) + x_{n+2}(t^{n+2} - \tilde{\tau}_{n+2}),$$

then $\varphi_n \in \mathcal{W}_n$ and

$$(3) \quad \|\xi - \varphi_n\|_I \leq 2^{-n} |x_{n+1}| + 2^{-n-1} |x_{n+2}| + \|\varrho_{n+3}\|_I.$$

Proof. The fact that $\varphi_n \in \mathcal{W}_n$ follows directly from the introductory remarks concerning the polynomials $\tilde{\tau}_p$. Since

$$\xi - \varphi_n = x_{n+1} \tilde{\tau}_{n+1} + x_{n+2} \tilde{\tau}_{n+2} + \varrho_{n+3}$$

and $\|\tilde{\tau}_{n+1}\|_I = 2^{-n} \|\tau_{n+1}\|_I = 2^{-n}$, $\|\tilde{\tau}_{n+2}\|_I = 2^{-n-1}$, inequality (3) also holds.

The method applied in Theorem 13:1 is based upon the fact that each of the monomials $x_{n+1}t^{n+1}$ and $x_{n+2}t^{n+2}$ is represented as a sum of a polynomial with the degree not exceeding n and a polynomial of the least possible norm in the interval I :

$$(4) \quad \begin{aligned} x_{n+1}t^{n+1} &= x_{n+1}(t^{n+1} - \tilde{\tau}_{n+1}) + x_{n+1}\tilde{\tau}_{n+1}, \\ x_{n+2}t^{n+2} &= x_{n+2}(t^{n+2} - \tilde{\tau}_{n+2}) + x_{n+2}\tilde{\tau}_{n+2}. \end{aligned}$$

The first terms of these sums are added to the polynomial $x_0 + x_1t + \dots + x_n t^n$, thus giving a polynomial φ_n which approximates the function ξ . The second terms of these sums together with ϱ_{n+3} are included in the remainder.

It is often convenient to represent the function ξ in the form

$$\xi = \sum_{k=0}^{n+1} x_k t^k + \varrho_{n+2} \quad (\varrho_{n+2} \in \mathcal{C}_I).$$

Then we apply only inequality (4), i.e., we put

$$\varphi_n^* = \sum_{k=0}^n x_k t^k + x_{n+1}(t^{n+1} - \tilde{\tau}_{n+1}).$$

The following estimate holds:

$$\|\xi - \varphi_n^*\|_I \leq 2^{-n} |x_{n+1}| + \|\varrho_{n+2}\|_I.$$

Theorem 13:1 often gives a polynomial φ_n which is sufficiently close to the n -th best polynomial ω_{nI} for the function ξ , though, of course, not identical with it. One of the sources of deviation between φ_n and ω_{nI} is the fact that each of the monomials $x_{n+1}t^{n+1}$ and $x_{n+2}t^{n+2}$ has been approximated separately. This may be avoided if we apply the so-called Zolotarev polynomials instead of the Chebyshev polynomials. The exact description of these polynomials will be given in § 14; here we shall restrict ourselves to the presentation of their definition.

Let s be an arbitrary real number, let p be an integer greater than 1 and let ω_{p-2}^s denote the $(p-2)$ -nd best polynomial of the function $t^p + st^{p-1}$ in the interval $I = \langle -1, 1 \rangle$. We shall denote the difference $t^p + st^{p-1} - \omega_{p-2}^s$ by ζ_p^s and we shall call it the p -th Zolotarev polynomial with parameter s (after a Russian mathematician who defined them first; see [38]).

The definition of the polynomial ζ_p^s may also be formulated as follows: Among the polynomials of the form $t^p + st^{p-1} + a_2t^{p-2} + \dots + a_p$ (p and s being fixed and a_2, \dots, a_p being arbitrary real numbers) the polynomial ζ_p^s equals that one which has the smallest norm in the interval I . It follows immediately that the Zolotarev polynomials are the generalization of the normed Chebyshev polynomials: $\zeta_p^0 = \tilde{\tau}_p$ (here we use the remark, made at the beginning of § 13.1, that $\tilde{\tau}_p$ does not contain the term with t^{p-1}).

Unlike the case of the Chebyshev polynomials, there are no explicit general analytical expressions which may be used in practice for the coefficients of the Zolotarev polynomials ζ_p^s . One has to use tables which contain the above-mentioned coefficients, only for a finite set of values of parameters s and p of course. Such tables are presented in the second part of this paper.

THEOREM 13:2. *If the function ξ is expressed by formula (2), $x_{n+2} \neq 0$ and, for a fixed s ,*

$$\chi_n^s = \sum_{k=0}^{n+2} x_k t^k - x_{n+2} (\zeta_{n+2}^s + (x_{n+1}/x_{n+2} - s) \tilde{\tau}_{n+1}),$$

then $\chi_n^s \in \mathcal{W}_n$ and

$$\|\xi - \chi_n^s\|_I \leq |x_{n+2}| (\|\zeta_{n+2}^s\|_I + 2^{-n} |x_{n+1}/x_{n+2} - s|) + \|e_{n+3}\|_I.$$

Proof. Under the assumption $x_{n+2} \neq 0$ we split the sum $x_{n+1}t^{n+1} + x_{n+2}t^{n+2}$ into terms in another way:

$$x_{n+1}t^{n+1} + x_{n+2}t^{n+2} = x_{n+2}(t^{n+2} + st^{n+1} + (x_{n+1}/x_{n+2} - s)t^{n+1}).$$

Using the formulas

$$\begin{aligned} t^{n+2} + st^{n+1} &= (t^{n+2} + st^{n+1} - \zeta_{n+2}^s) + \zeta_{n+2}^s, \\ t^{n+1} &= (t^{n+1} - \tilde{\tau}_{n+1}) + \tilde{\tau}_{n+1} \end{aligned}$$

we extract the polynomial of degree n from the sum $x_{n+1}t^{n+1} + x_{n+2}t^{n+2}$. In the remainder we leave a polynomial whose norm in the interval I is small, provided that $s \approx x_{n+1}/x_{n+2}$:

$$\begin{aligned} &x_{n+1}t^{n+1} + x_{n+2}t^{n+2} \\ &= x_{n+2}(t^{n+2} + st^{n+1} - \zeta_{n+2}^s + (x_{n+1}/x_{n+2} - s)(t^{n+1} - \tilde{\tau}_{n+1})) + \\ &\quad + x_{n+2}(\zeta_{n+2}^s + (x_{n+1}/x_{n+2} - s)\tilde{\tau}_{n+1}) \\ &= x_{n+1}t^{n+1} + x_{n+2}t^{n+2} - x_{n+2}(\zeta_{n+2}^s + (x_{n+1}/x_{n+2} - s)\tilde{\tau}_{n+1}) + \\ &\quad + x_{n+2}(\zeta_{n+2}^s + (x_{n+1}/x_{n+2} - s)\tilde{\tau}_{n+1}). \end{aligned}$$

This equation determines the form of the polynomial χ_n^s which is supposed to approximate the function ξ . The estimate for the norm $\|\xi - \chi_n^s\|_I$ follows immediately from the fact that

$$\xi - \chi_n^s = x_{n+2}(\zeta_{n+2}^s + (x_{n+1}/x_{n+2} - s)\tilde{\tau}_{n+1}) + \varrho_{n+3}.$$

In the proof it turned out that the relatively best approximation of the function ξ by polynomial χ_n^s may be expected when $s \approx x_{n+1}/x_{n+2}$. Of course, at the same time s should be such that the coefficients of the polynomial ζ_{n+2}^s should be present in the tables. In Theorem 13:2 we assumed that $x_{n+2} \neq 0$. If $x_{n+2} = 0$, the application of the Zolotarev polynomials is useless in comparison with Theorem 13:1.

13.2. With every function $\xi \in \mathcal{C}_I$ there is connected an orthogonal series

$$\sum_{l=0}^{\infty} a_l[\xi] \tau_l,$$

whose coefficients are expressed by the formulas

$$(5) \quad a_l[\xi] = \begin{cases} \frac{1}{\pi} \int_{-1}^1 \xi(t) (1-t^2)^{-1/2} dt & (l=0), \\ \frac{2}{\pi} \int_{-1}^1 \xi(t) \tau_l(t) (1-t^2)^{-1/2} dt & (l>0). \end{cases}$$

If the function ξ has a bounded first derivative in the interval I , then this series is uniformly convergent to the function ξ in the interval I .

THEOREM 13:3 (Bernstein, [4], p. 73). *If*

$$\xi = \sum_{l=0}^{\infty} a_l \tau_l \quad (t \in I),$$

then

$$(6) \quad \left(\frac{1}{2} \sum_{l=n+1}^{\infty} a_l^2 \right)^{1/2} \leq \|\xi - v_n\|_I \leq \sum_{l=n+1}^{\infty} |a_l|$$

where

$$(7) \quad v_n = \sum_{k=0}^n a_k \tau_k.$$

Proof. The right part of inequality (6) follows immediately from the fact that

$$\xi - v_n = \sum_{l=n+1}^{\infty} a_l \tau_l$$

since $\|\tau_l\|_{(-1,1)} = 1$ for every l . From the general theory of orthogonal series it follows that

$$\int_{-1}^1 (\xi(t) - v_n(t))^2 (1-t^2)^{-1/2} dt = \sum_{l=n+1}^{\infty} a_l^2 \int_{-1}^1 \tau_l^2(t) (1-t^2)^{-1/2} dt = \frac{\pi}{2} \sum_{l=n+1}^{\infty} a_l^2.$$

Since

$$\int_{-1}^1 (\xi(t) - v_n(t))^2 (1-t^2)^{-1/2} dt \leq \|\xi - v_n\|_I^2 \int_{-1}^1 (1-t^2)^{-1/2} dt = \pi \|\xi - v_n\|_I^2,$$

the left part of inequality (6) is also satisfied.

The difficulty of using Theorem 13:3 may lie in the computation of the coefficients $a_l \equiv a_l[\xi]$. These coefficients, however, are known for some elementary functions. Besides, for a function ξ , which can be expanded into a power series in the interval I , we may use Theorem 5:7, which gives the relation between the coefficients a_l and the coefficients of the power series. To simplify the application of Theorem 13:3 further, we shall express the coefficients of the polynomial v_n directly by the coefficients of the expansion of ξ into a power series.

THEOREM 13:4. *If*

$$\xi = \sum_{l=0}^{\infty} x_l t^l \quad (t \in I),$$

then polynomial (7) is expressed by the formula

$$v_n = \sum_{k=0}^n \left\{ x_k + \sum_{l=[(n-k)/2]+1}^{\infty} \frac{x_{k+2l}}{2^{2l}} \sum_{m=0}^{[(n-k)/2]} \frac{(-1)^m \max\{k+2m, 1\}}{\max\{k+m, 1\}} \binom{k+2l}{l-m} \binom{k+m}{m} \right\} t^k.$$

Proof. Using the notation

$$(8) \quad q_l = \begin{cases} \frac{1}{2} & (l = 0), \\ 1 & (l > 0), \end{cases} \quad a_m^j = \frac{1}{2^{m-1}} \binom{m}{j} \quad (m \geq j \geq 0),$$

$$b_j^m = \frac{(-1)^m 2^{j-1} \max\{j+2m, 1\}}{\max\{j+m, 1\}} \binom{j+m}{m} \quad (j \geq 0, m \geq 0)$$

we have the following equations

$$(9) \quad a_l[\xi] = q_l \sum_{j=0}^{\infty} a_{2j+1}^j x_{2j+1}, \quad \tau_l = \frac{1}{q_l} \sum_{m=0}^{[l/2]} b_{l-2m}^m t^{l-2m}.$$

In fact,

$$a_{2j+1}^j = \frac{1}{2^{2j+1-1}} \binom{2j+1}{j}$$

and the first of equations (9) follows from Theorem 5:7 and from the definition of number q_l . Since

$$b_{l-2m}^m = \frac{(-1)^m 2^{l-2m-1} \max\{l, 1\}}{\max\{l-m, 1\}} \binom{l-m}{m}, \quad (l \geq 2m \geq 0),$$

we have

$$b_{l-2m}^m = \begin{cases} l \frac{(-1)^m}{l-m} \binom{l-m}{m} 2^{l-2m-1} & (l > 0), \\ \frac{1}{2} & (l = 0). \end{cases}$$

Thus, for $l > 0$ the second one of equations (9) follows (after changing the notation) from formula (1), and for $l = 0$ it is consistent with the fact that $\tau_0 = 1$.

It follows from (9) that

$$v_n = \sum_{l=0}^n a_l[\xi] \tau_l = \sum_{l=0}^n \sum_{j=0}^{\infty} a_{2j+1}^j x_{2j+1} \sum_{m=0}^{[l/2]} b_{l-2m}^m t^{l-2m}.$$

Let $k = l - 2m$. Since the inequalities $0 \leq l \leq n$ and $0 \leq m \leq [l/2]$ are equivalent to the inequalities $0 \leq k \leq n$ and $0 \leq m \leq [(n-k)/2]$ (in particular $m = (l-k)/2$, whence $m \leq (n-k)/2$), we have

$$v_n = \sum_{k=0}^n \sum_{j=0}^{\infty} \sum_{m=0}^{[(n-k)/2]} a_{2j+k+2m}^j x_{2j+k+2m} b_k^m t^k.$$

It is known that if the coefficients x_{n+1}, x_{n+2}, \dots of function ξ equal zero, i.e. if the function ξ is a polynomial $x_0 + x_1 t + \dots + x_n t^n$, then the polynomial v_n coincides with it. This means that that part of the triple sum under consideration for which $2j + k + 2m \leq n$ equals $x_0 + x_1 t + \dots + x_n t^n$. Let us now consider the remaining part of this sum, i.e. that part for which $2j + k + 2m > n$, or $j + m \geq [(n-k)/2] + 1$. We shall change the summation with respect to j from 0 to ∞ into the summation with respect to $l = j + m$ from $[(n-k)/2] + 1$ to ∞ . The index m will not be changed:

$$v_n = \sum_{k=0}^n \left\{ x_k + \sum_{l=[(n-k)/2]+1}^{\infty} x_{k+2l} \sum_{m=0}^{[(n-k)/2]} a_{k+2l}^{l-m} b_k^m \right\} t^k.$$

Using (8) the last sum may be transformed as follows:

$$\begin{aligned} \sum_{m=0}^{[(n-k)/2]} a_{k+2l}^{l-m} b_k^m &= \sum_{m=0}^{[(n-k)/2]} \frac{1}{2^{k+2l-1}} \binom{k+2l}{l-m} \frac{(-1)^m 2^{k-1} \max\{k+2m, 1\}}{\max\{k+m, 1\}} \binom{k+m}{m} \\ &= \frac{1}{2^{2l}} \sum_{m=0}^{[(n-k)/2]} \frac{(-1)^m \max\{k+2m, 1\}}{\max\{k+m, 1\}} \binom{k+2l}{l-m} \binom{k+m}{m}, \end{aligned}$$

which was to be proved.

The tables which are placed (together with some others) in the second part of this paper are based upon Theorem 13:4. They reduce the amount of work which is necessary for finding the coefficients of the approximating polynomial v_n to the computation of linear combinations of certain known coefficients of a power series.

13.3. Let ψ be the interpolation polynomial constructed for the function ξ and nodes t_0, t_1, \dots, t_n (where $-1 \leq t_0 < t_1 < \dots < t_n \leq 1$):

$$\psi(t_k) = \xi(t_k) \quad (k = 0, 1, \dots, n).$$

It is known ([18], p. 6) that if $\xi \in \mathcal{C}_I^{n+1}$, then

$$(10) \quad |\xi(t) - \psi(t)| \leq \frac{|(t-t_0)(t-t_1)\dots(t-t_n)|}{(n+1)!} \|\xi^{(n+1)}\|_I \quad (t \in I).$$

The polynomial $\tilde{\tau}_p = 2^{-p+1} \cos p \arccos t$ is equal to that one among the polynomials of the form $t^p + a_1 t^{p-1} + \dots + a_p$, which has the smallest norm in the interval I (§ 2.2). All the zeros of this polynomial are:

$$z_{pk} = -\cos \frac{(2k+1)\pi}{2p} \quad (k = 0, 1, \dots, p-1)$$

and these numbers lie in the interval I . Thus, the maximum of the right-hand side of inequality (10) will be smallest if

$$t_k = z_{n+1,k} \quad (k = 0, 1, \dots, n),$$

i.e. if

$$(t-t_0)(t-t_1)\dots(t-t_n) = \tilde{\tau}_{n+1}.$$

In this case inequality (10) takes the form

$$\|\xi - \psi\|_I \leq \frac{\|\xi^{(n+1)}\|_I}{2^n (n+1)!}.$$

If we compare this inequality with Theorem 6:2, we see that the choice of points $z_{n+1,k}$ for the interpolation nodes was good—the error of the approximation of the function ξ by polynomial ψ is of the same order as the n -th error of the best approximation $\varepsilon_n(\xi; I)$. To use this fact in practice we should establish how the form of the interpolation polynomial ψ depends upon the function ξ and degree n .

We shall introduce one more symbol:

$$y_{pk} = \sin \frac{(2k+1)\pi}{2p} \quad (k = 0, 1, \dots, p-1).$$

THEOREM 13:5. If $\xi \in \mathcal{C}_I^{n+1}$ and

$$(11) \quad \psi_n = \sum_{k=0}^n \xi(z_{n+1,k}) \frac{(-1)^{n-k} y_{n+1,k} \tau_{n+1}}{(n+1)(t-z_{n+1,k})},$$

then $\psi_n \in \mathcal{H}_n$ and

$$(12) \quad \|\xi - \psi_n\|_I \leq \frac{\|\xi^{(n+1)}\|_I}{2^n (n+1)!}.$$

Proof. The polynomial ψ_n is defined by the equations

$$(13) \quad \psi_n(z_{n+1,k}) = \xi(z_{n+1,k}) \quad (k = 0, 1, \dots, n).$$

From the Lagrange interpolation formula it follows immediately that

$$\psi_n = \sum_{k=0}^n \xi(z_{n+1,k}) \lambda_k \quad \text{where} \quad \lambda_k = \frac{\tau_{n+1}}{\tau'_{n+1}(z_{n+1,k})(t-z_{n+1,k})} \quad \text{and} \quad \tau_{n+1} =$$

$= (t - z_{n+1,0})(t - z_{n+1,1}) \dots (t - z_{n+1,n})$. We already know that $\tau_{n+1} = \tilde{\tau}_{n+1} = 2^{-n} \tau_{n+1}$, and thus

$$\lambda_k = \frac{\tau_{n+1}}{\tau'_{n+1}(z_{n+1,k})(t - z_{n+1,k})}.$$

Since $\tau_{n+1} = \cos(n+1) \arccos t$, we have

$$(14) \quad \begin{aligned} \tau'_{n+1} &= (n+1)(1-t^2)^{-1/2} \sin(n+1) \arccos t, \\ (1-t^2) \tau'^2_{n+1} &= (n+1)^2 (1-\tau_{n+1}^2). \end{aligned}$$

Putting in the last identity $t = z_{n+1,k}$, and using the fact that $\tau_{n+1}(z_{n+1,k}) = 0$, we get

$$|\tau'_{n+1}(z_{n+1,k})| = (n+1)(1-z_{n+1,k}^2)^{-1/2} = (n+1)/y_{n+1,k}.$$

The polynomial $\tau_{n+1} = 2^n t^{n+1} + \dots$ is positive to the right of its greatest zero, i.e. to the right of $z_{n+1,n}$; hence its derivative at this point is positive. This at once determines the signs of the values τ'_{n+1} at the remaining zeros:

$$\text{sign } \tau'_{n+1}(z_{n+1,k}) = (-1)^{n-k} \quad (k = 0, 1, \dots, n).$$

Thus

$$\lambda_k = \frac{(-1)^{n-k} y_{n+1,k} \tau_{n+1}}{(n+1)(t - z_{n+1,k})}$$

and this proves Theorem 13:5, since inequality (12) has already been obtained earlier.

Formula (11), which determines the polynomial ψ_n defined by conditions (13), is, of course, valid not only for functions of class \mathcal{C}_I^{n+1} , but also for any other function defined on the set of points $z_{n+1,k}$. In particular, formula (11) may be applied to every function ξ such that the series

$$(15) \quad \sum_{l=0}^{\infty} a_l[\xi] \tau_l$$

with coefficients (5) converges to that function in the interval I . However, in this case it is more reasonable to represent the polynomial ψ_n in a form analogous to the form in which the function ξ is given. In fact, if we find the coefficients $a_j[\psi_n]$ of the expansion (finite, of course) of the polynomial ψ_n into an orthogonal series with respect to the Chebyshev polynomials, we shall be able to obtain the estimate for the norm $\|\xi - \psi_n\|_I$, in place of estimate (12), which is now useless.

THEOREM 13:6 (Hörnecker, [14]). *If series (15) converges to the function ξ in the interval I , then the polynomial ψ_n defined by (13) can be expressed as*

$$\psi_n = \sum_{j=0}^n a_j[\psi_n] \tau_j,$$

where

$$(16) \quad a_j[\psi_n] = \begin{cases} \sum_{i=0}^{\infty} (-1)^i a_{2i(n+1)}[\xi] & (j = 0), \\ \sum_{i=0}^{\infty} (-1)^i (a_{2i(n+1)+j}[\xi] - a_{2(i+1)(n+1)-j}[\xi]) & (j = 1, 2, \dots, n). \end{cases}$$

The norm of the difference $\xi - \psi_n$ satisfies the inequality

$$\|\xi - \psi_n\|_I \leq \sum_{i=0}^{\infty} (|a_{2i(n+1)}[\xi]| + 2 \sum_{j=1}^{2n+1} |a_{2i(n+1)+j}[\xi]|).$$

Proof. Let us first compute the value of an arbitrary Chebyshev polynomial τ_l at the point

$$z_{pk} = -\cos \frac{(2k+1)\pi}{2p} = \cos \frac{(2p-2k-1)\pi}{2p} \quad (p \text{ natural}).$$

If we represent the number l in the form $2ip + j$ where i is a non-negative integer, and $0 \leq j \leq 2p-1$, we get

$$\begin{aligned} \tau_l(z_{pk}) &= \cos(2ip + j) \operatorname{arccos} z_{pk} = \cos \frac{(2ip + j)(2p-2k-1)\pi}{2p} \\ &= \cos \left(i(2p-2k-1)\pi + \frac{j(2p-2k-1)\pi}{2p} \right) \\ &= \cos i(2p-2k-1)\pi \cdot \cos \frac{j(2p-2k-1)\pi}{2p} \\ &= (-1)^{i(2p-2k-1)} \tau_j(z_{pk}) = (-1)^i \tau_j(z_{pk}). \end{aligned}$$

Thus

$$\xi(z_{pk}) = \sum_{l=0}^{\infty} a_l \tau_l(z_{pk}) = \sum_{i=0}^{\infty} \sum_{j=0}^{2p-1} (-1)^i a_{2ip+j} \tau_j(z_{pk})$$

where $a_l \equiv a_l[\xi]$.

Applying the equation

$$\begin{aligned}\tau_{2p-j}(z_{pk}) &= \cos \frac{(2p-j)(2p-2k-1)\pi}{2p} \\ &= \cos \left((2p-2k-1)\pi - \frac{j(2p-2k-1)\pi}{2p} \right) = -\tau_j(z_{pk})\end{aligned}$$

and using the fact that $\tau_p(z_{pk}) = 0$, by definition of the numbers z_{pk} , we can write that

$$\begin{aligned}\xi(z_{pk}) &= \sum_{i=0}^{\infty} \left(\sum_{j=0}^{p-1} (-1)^i a_{2ip+j} \tau_j(z_{pk}) - \sum_{j=p+1}^{2p-1} (-1)^i a_{2ip+j} \tau_{2p-j}(z_{pk}) \right) \\ &= \sum_{i=0}^{\infty} (-1)^i \left(\sum_{j=0}^{p-1} a_{2ip+j} \tau_j(z_{pk}) - \sum_{j=1}^{p-1} a_{2(i+1)p-j} \tau_j(z_{pk}) \right) \\ &= \tau_0(z_{pk}) \sum_{i=0}^{\infty} (-1)^i a_{2ip} + \sum_{j=1}^{p-1} \tau_j(z_{pk}) \sum_{i=0}^{\infty} (-1)^i (a_{2ip+j} - a_{2(i+1)p-j}).\end{aligned}$$

Now it is already obvious that polynomial

$$\begin{aligned}\psi_n &= \tau_0 \sum_{i=0}^{\infty} (-1)^i a_{2i(n+1)}[\xi] + \\ &\quad + \sum_{j=1}^n \tau_j \sum_{i=0}^{\infty} (-1)^i (a_{2i(n+1)+j}[\xi] - a_{2(i+1)(n+1)-j}[\xi])\end{aligned}$$

of class \mathscr{W}_n satisfies condition (13). Hence the formulas for the coefficients $a_j[\psi_n]$ follow.

We easily notice that the polynomial ψ_n does not depend upon the coefficients $a_{n+1}[\xi]$, $a_{3(n+1)}[\xi]$, \dots , and that every other coefficient $a_l[\xi]$ appears in the above formula exactly once. In particular, the polynomial ψ_n contains the sum $a_0[\xi]\tau_0 + a_1[\xi]\tau_1 + \dots + a_n[\xi]\tau_n$. Thus the difference $\xi - \psi_n$ does not depend upon the coefficients $a_l[\xi]$ for $l \leq n$. The coefficients $a_{n+1}[\xi]$, $a_{3(n+1)}[\xi]$, \dots are multiplied by polynomials τ_{n+1} , $\tau_{3(n+1)}$, \dots in the difference $\xi - \psi_n$, and every other coefficient $a_l[\xi]$ for $l > n$ is multiplied by the sum of the polynomial τ_l and either τ_j or $-\tau_j$ for some $j = 0, 1, \dots, n$. The estimate of the norm $\|\xi - \psi_n\|_I$ given in the theorem follows now from the inequalities $\|\tau_j\|_I = 1$, $\|\tau_l \pm \tau_j\|_I \leq 2$.

Using Theorem 13:6 we shall now express the coefficients of the polynomial ψ_n by the coefficients of a power series convergent to the function ξ . This theorem will be analogous to Theorem 13:4, which concerns the polynomial ν_n . In the formulation of the theorem we shall apply the usual convention that $\binom{m}{j} = 0$ for any negative integer j and positive integer m .

THEOREM 13:7. *If*

$$\xi = \sum_{l=0}^{\infty} x_l t^l \quad (t \in I),$$

then the polynomial ψ_n defined by formula (13) can be expressed by the formula

$$\begin{aligned} \psi_n = & \sum_{k=0}^n \left\{ x_k + \sum_{l=(n-k)/2+1}^{\infty} \frac{x_{k+2l}}{2^{2l}} \times \right. \\ & \times \sum_{m=0}^{(n-k)/2} \frac{(-1)^m \max\{k+2m, 1\}}{\max\{k+m, 1\}} \binom{k+m}{m} \times \\ & \left. \times \sum_{i=0}^{(l-m)/(n+1)} (-1)^i \left(\binom{k+2l}{l-m-i(n+1)} - \binom{k+2l}{l+m-(i+1)(n+1)+k} \right) \right\} t^k \end{aligned}$$

and the norm of the difference $\xi - \psi_n$ satisfies the inequality

$$\|\xi - \psi_n\|_I \leq \frac{|x_{n+1}| + |x_{n+2}|}{2^n} + \sum_{k=2}^{\infty} \frac{|x_{n+k+1}|}{2^{n+k-1}} \sum_{l=\max\{[k/2]-n, 0\}}^{[k/2]} \sum_{i=0}^{[l/(n+1)]} \binom{n+k+1}{l-i(n+1)}.$$

Proof. Introducing notation (8) we see that

$$\psi_n = \sum_{l=0}^n a_l[\psi_n] \tau_l = \sum_{l=0}^n \frac{a_l[\psi_n]}{q_l} \sum_{m=0}^{[l/2]} b_{l-2m}^m t^{l-2m}.$$

The index l is replaced (as in the proof of Theorem 13:4) by the index $k = l - 2m$. This leads to the equation

$$(17) \quad \psi_n = \sum_{k=0}^n t^k \sum_{m=0}^{[(n-k)/2]} \frac{a_{k+2m}[\psi_n]}{q_{k+2m}} b_k^m.$$

Next, let us notice that formulas (16) may be written in a uniform way

$$(18) \quad \frac{a_j[\psi_n]}{q_j} = \sum_{i=0}^{\infty} (-1)^i \left(\frac{a_{2i(n+1)+j}[\xi]}{q_{2i(n+1)+j}} - \frac{a_{2(i+1)(n+1)-j}[\xi]}{q_{2(i+1)(n+1)-j}} \right).$$

Indeed, if $j > 0$, then formula (18), which contains only such coefficients $a_l[\xi]$ that $l > 0$, coincides with the second formula of (16). On the other

hand, for $j = 0$ the right-hand side of (18) equals

$$\begin{aligned} 2a_0[\xi] - a_{2(n+1)}[\xi] + \sum_{i=1}^{\infty} (-1)^i (a_{2i(n+1)}[\xi] - a_{2(i+1)(n+1)}[\xi]) \\ = 2a_0[\xi] + 2 \sum_{i=1}^{\infty} (-1)^i a_{2i(n+1)}[\xi] \end{aligned}$$

and, according to the first formula of (16), is equal to $a_0[\psi_n]/q_0$.

Changing in the first formula of (9) the summation index into k we find from (18) that

$$\begin{aligned} \frac{a_j[\psi_n]}{q_j} = \sum_{i=0}^{\infty} (-1)^i \sum_{k=0}^{\infty} a_{2i(n+1)+j+2k}^k x_{2i(n+1)+j+2k} - \\ - \sum_{i=0}^{\infty} (-1)^i \sum_{k=0}^{\infty} a_{2(i+1)(n+1)-j+2k}^k x_{2(i+1)(n+1)-j+2k}. \end{aligned}$$

Instead of k we introduce the new index of summation p , equal to $i(n+1)+k$ in the first double sum, and equal to $(i+1)(n+1)-j+k$ in the second sum. At the same time we change the order of summation, which changes the domain of index i :

$$\begin{aligned} \frac{a_j[\psi_n]}{q_j} = \sum_{p=0}^{\infty} \sum_{i=0}^{\lfloor p/(n+1) \rfloor} (-1)^i a_{j+2p}^{p-i(n+1)} x_{j+2p} - \\ - \sum_{p=0}^{\infty} \sum_{i=0}^{\lfloor (p+j)/(n+1) \rfloor - 1} (-1)^i a_{j+2p}^{p-(i+1)(n+1)+j} x_{j+2p}. \end{aligned}$$

To unify both sum with respect to the index i we shall add to the definition (8) of numbers a_m^j the condition that $a_m^j = 0$ for $j < 0$. Since $j \leq n$, we have $\lfloor (p+j)/(n+1) \rfloor - 1 \leq \lfloor p/(n+1) \rfloor$ and

$$\frac{a_j[\psi_n]}{q_j} = \sum_{p=0}^{\infty} x_{j+2p} \sum_{i=0}^{\lfloor p/(n+1) \rfloor} (-1)^i (a_{j+2p}^{p-i(n+1)} - a_{j+2p}^{p-(i+1)(n+1)+j}).$$

For $j = k + 2m$ we put this equation into (17):

$$\begin{aligned} \psi_n = \sum_{k=0}^n t^k \sum_{p=0}^{\infty} \sum_{m=0}^{\lfloor (n-k)/2 \rfloor} x_{k+2m+2p} \times \\ \times \sum_{i=0}^{\lfloor p/(n+1) \rfloor} (-1)^i (a_{k+2m+2p}^{p-i(n+1)} - a_{k+2m+2p}^{p-(i+1)(n+1)+k+2m}) b_k^m. \end{aligned}$$

Now we replace the index p by $1 = p + m$. Since $0 \leq m \leq [(n-k)/2]$ and $m \leq l$, we have

$$\psi_n = \sum_{k=0}^n t^k \sum_{l=0}^{\infty} x_{k+2l} \sum_{m=0}^{\min\{[(n-k)/2], l\}} \sum_{i=0}^{[(l-m)/(n+1)]} (-1)^i (a_{k+2l}^{l-m-i(n+1)} - a_{k+2l}^{l+m-(i+1)(n+1)+k}) b_k^m.$$

It follows from the definition of polynomial ψ_n that that part of the above quadruple sum which contains the terms x_0, x_1, \dots, x_n coincides with the sum $x_0 + x_1 t + \dots + x_n t^n$. This sum is excluded, and in the remaining part $k + 2l > n$, i.e. $l \geq [(n-k)/2] + 1$. Then the upper limit of summation with respect to m equals $[(n-k)/2]$ and

$$\psi_n = \sum_{k=0}^n \left\{ x_k + \sum_{l=[(n-k)/2]+1}^{\infty} x_{k+2l} \sum_{m=0}^{[(n-k)/2]} \sum_{i=0}^{[(l-m)/(n+1)]} (-1)^i (a_{k+2l}^{l-m-i(n+1)} - a_{k+2l}^{l+m-(i+1)(n+1)+k}) b_k^m \right\} t^k.$$

Now we apply formulas (8). The additional condition that $a_m^j = 0$ for $j < 0$ implies that $\binom{m}{j} = 0$ for $j < 0$; which was done in the proper place. Thus we finally get the formula

$$\begin{aligned} \psi_n = & \sum_{k=0}^n \left\{ x_k + \right. \\ & + \sum_{l=[(n-k)/2]+1}^{\infty} x_{k+2l} \sum_{m=0}^{[(n-k)/2]} \sum_{i=0}^{[(l-m)/(n+1)]} (-1)^i \left(\frac{1}{2^{k+2l-1}} \binom{k+2l}{l-m-i(n+1)} - \right. \\ & \left. \left. - \frac{1}{2^{k+2l-1}} \binom{k+2l}{l+m-(i+1)(n+1)+k} \right) \frac{(-1)^m 2^{k-1} \max\{k+2m, 1\}}{\max\{k+m, 1\}} \binom{k+m}{m} \right\} t^k, \end{aligned}$$

which was to be proved.

Now we shall estimate the norm of the difference $\xi - \psi_n$. We shall use the inequality proved in the preceding theorem, which can be written in the form

$$\|\xi - \psi_n\|_I \leq 2 \sum_{i=0}^{\infty} \sum_{j=0}^{2n+1} q_j |a_{(2i+1)(n+1)+j}[\xi]|.$$

The coefficient $a_{(2i+1)(n+1)+j}[\xi]$ will be replaced by the expression obtained from (9) for $l = (2i+1)(n+1)+j$. For such an l the number q_l may be omitted as being equal to 1. Thus

$$\|\xi - \psi_n\|_I \leq 2 \sum_{i=0}^{\infty} \sum_{j=0}^{2n+1} q_j \sum_{k=0}^{\infty} a_{(2i+1)(n+1)+j+2k}^k |x_{(2i+1)(n+1)+j+2k}|.$$

We now introduce the new index equal to $l = i(n+1)+k$ in place of k . At the same time we change the order of summation:

$$\|\xi - \psi_n\|_I \leq 2 \sum_{l=0}^{\infty} \sum_{j=0}^{2n+1} q_j \sum_{i=0}^{[l/(n+1)]} a_{n+j+2l+1}^{l-i(n+1)} |x_{n+j+2l+1}|.$$

Once more we perform the change of indices, putting $k = j + 2l$. Since $l = (k-j)/2$, we have $(k-2n-1)/2 \leq l \leq k/2$. At the same time $l \geq 0$; thus finally $\max\{[k/2]-n, 0\} \leq l \leq [k/2]$ and

$$\begin{aligned} \|\xi - \psi_n\|_I &\leq 2 \sum_{k=0}^{\infty} |x_{n+k+1}| \sum_{l=\max\{[k/2]-n, 0\}}^{[k/2]} q_{k-2l} \sum_{i=0}^{[l/(n+1)]} a_{n+k+1}^{l-i(n+1)} \\ &\leq |x_{n+1}| a_{n+1}^0 + 2|x_{n+2}| a_{n+2}^0 + 2 \sum_{k=2}^{\infty} |x_{n+k+1}| \sum_{l=\max\{[k/2]-n, 0\}}^{[k/2]} \sum_{i=0}^{[l/(n+1)]} a_{n+k+1}^{l-i(n+1)}, \end{aligned}$$

which, together with the definition of numbers a_m^j , gives the required estimations.

13.4. In § 13.3 we have found the polynomial ψ_n , which interpolates the function ξ at the nodes which may be considered as the best. Now, using similar methods, we shall find an other polynomial, denoted by the symbol ω_{n1} . The main application of the polynomial ω_{n1} is based upon the fact that, in general, it gives a good start for the construction of the sequence $\{\omega_{nm}\}$ by the method of Remez (§ 15), and this sequence converges to the n -th best polynomial for the function ξ in the interval I . However, the polynomial ω_{n1} may also in itself form the ultimate task of our computation if we have the estimate of the norm of the difference $\xi - \omega_{n1}$, whence if the function ξ is sufficiently regular (Theorems 13:9 and 13:10).

We start by recalling some facts which have been established in the previous chapters. If $F = \{t_0, t_1, \dots, t_{n+1}\}$ where $t_0 < t_1 < \dots < t_{n+1}$, then the coefficients of the n -th best polynomial ω_{nF} for the function ξ on the set F , and a number e_{nF} such that $|e_{nF}| = \varepsilon_n(\xi; F)$ satisfy the system of $n+2$ equations

$$\xi(t_k) - \omega_{nF}(t_k) = (-1)^k e_{nF} \quad (k = 0, 1, \dots, n+1)$$

(Theorem 2:1). It will be seen in § 15 that in approximation in any in-

terval a large part is played by the properties and methods of approximation on the $(n+2)$ -points subsets of that interval. If the function ξ is sufficiently regular in the interval $I = \langle -1, 1 \rangle$, then the subset

$$(19) \quad F = \{t_{n+1,0}, t_{n+1,1}, \dots, t_{n+1,n+1}\}$$

where $t_{n+1,k} = -\cos k\pi/(n+1)$ ($k = 0, 1, \dots, n+1$) plays a particular part in approximation. Indeed, Theorems 9:6, 12:1 and others allow us to believe that in this case the polynomial ω_{nF} is close to the polynomial ω_{nI} , and the error of the best approximation $\varepsilon_n(\xi; F)$ is close to $\varepsilon_n(\xi; I)$.

For the set (19) we introduce special notations, ω_{n1} and e_{n1} , instead of ω_{nF} and e_{nF} . Thus the polynomial ω_{n1} and the number e_{n1} satisfy the system of equations

$$(20) \quad \xi(t_{n+1,k}) - \omega_{n1}(t_{n+1,k}) = (-1)^k e_{n1} \quad (k = 0, 1, \dots, n+1).$$

The number e_{n1} has already been expressed by the values of function ξ in § 5.3 (formula (9)):

$$(21) \quad e_{n1} = \frac{1}{n+1} \sum_{k=0}^{n+1} (-1)^k q_k \xi(t_{n+1,k})$$

where

$$q_k = \begin{cases} \frac{1}{2} & (k = 0, n+1), \\ 1 & (k = 1, 2, \dots, n). \end{cases}$$

Thus, it remains to find the polynomial ω_{n1} .

THEOREM 13 : 8 (Spitzbart, Shell, [33]). *If $\xi \in \mathcal{C}_I$, then the polynomial ω_{n1} defined by conditions (20) can be expressed by formula*

$$\omega_{n1} = \frac{(-1)^{n+1}}{n+1} \sum_{k=0}^{n+1} (-1)^k q_k \xi(t_{n+1,k}) \left(\frac{(t^2-1)\tau'_{n+1}}{(n+1)(t-t_{n+1,k})} - \tau_{n+1} \right).$$

Proof. Let us write system (20) in the form

$$\omega_{n1}(t_{n+1,k}) = \xi(t_{n+1,k}) - (-1)^k e_{n1} \quad (k = 0, 1, \dots, n+1).$$

Since

$$(22) \quad \tau_{n+1}(t_{n+1,k}) = (-1)^{n+1-k} \quad (k = 0, 1, \dots, n+1)$$

the polynomial ω_{n1} may be represented as the sum

$$(23) \quad \omega_{n1} = \delta + (-1)^n e_{n1} \tau_{n+1},$$

in which the polynomial $\delta \in \mathcal{W}_{n+1}$ is determined by the conditions

$$\delta(t_{n+1,k}) = \xi(t_{n+1,k}) \quad (k = 0, 1, \dots, n+1).$$

It follows from the Lagrange interpolation formula that

$$(24) \quad \delta = \sum_{k=0}^{n+1} \xi(t_{n+1,k}) \lambda_k,$$

where

$$\lambda_k = \frac{\pi_{n+2}}{\pi'_{n+2}(t_{n+1,k})(t-t_{n+1,k})} \quad (k = 0, 1, \dots, n+1),$$

$$\pi_{n+2} = \prod_{l=0}^{n+1} (t-t_{n+1,l}).$$

It is easy to see that $t_{n+1,0} = -1$, $t_{n+1,n+1} = 1$, and the points $t_{n+1,1}, \dots, t_{n+1,n}$ are the zeros (of course all the zeros) of the derivative of polynomial τ_{n+1} . The coefficient of t^n in τ'_{n+1} equals $2^n(n+1)$, whence

$$\pi_{n+2} = \frac{(t^2-1)\tau'_{n+1}}{2^n(n+1)},$$

$$\lambda_k = \frac{(t^2-1)\tau'_{n+1}}{(2t\tau'_{n+1} + (t^2-1)\tau''_{n+1})_{t=t_{n+1,k}}(t-t_{n+1,k})}.$$

Let us differentiate both sides of identity (14) and then divide both sides of the equation obtained by $-2\tau'_{n+1}$:

$$t\tau'_{n+1} + (t^2-1)\tau''_{n+1} = (n+1)^2\tau_{n+1}.$$

According to the last equation, for $k = 0$ and $k = n+1$ we have:

$$(2t\tau'_{n+1} + (t^2-1)\tau''_{n+1})_{t=t_{n+1,k}} = 2t_{n+1,k}\tau'_{n+1}(t_{n+1,k}) = 2(n+1)^2\tau_{n+1}(t_{n+1,k})$$

and for $k = 1, 2, \dots, n$ we have

$$(2t\tau'_{n+1} + (t^2-1)\tau''_{n+1})_{t=t_{n+1,k}} = (t_{n+1,k}^2-1)\tau''_{n+1}(t_{n+1,k}) = (n+1)^2\tau_{n+1}(t_{n+1,k}).$$

By (22), for $k = 0, 1, \dots, n+1$ we have

$$(2t\tau'_{n+1} + (t^2-1)\tau''_{n+1})_{t=t_{n+1,k}} = \frac{(-1)^{n+1-k}(n+1)^2}{q_k},$$

$$\lambda_k = \frac{(-1)^{n+1-k}q_k(t^2-1)\tau'_{n+1}}{(n+1)^2(t-t_{n+1,k})}.$$

Using equations (23), (24) and (21) we see that

$$\begin{aligned}\omega_{n1} &= \delta + (-1)^n e_{n1} \tau_{n+1} \\ &= \sum_{k=0}^{n+1} \xi(t_{n+1,k}) \lambda_k + \frac{(-1)^n \tau_{n+1}}{n+1} \sum_{k=0}^{n+1} (-1)^k q_k \xi(t_{n+1,k}) \\ &= \frac{(-1)^{n+1}}{n+1} \sum_{k=0}^{n+1} (-1)^k q_k \xi(t_{n+1,k}) \left(\frac{(t^2-1) \tau'_{n+1}}{(n+1)(t-t_{n+1,k})} - \tau_{n+1} \right),\end{aligned}$$

which was to be proved.

Let us notice in connection with Theorem 13:8, that the polynomial ω_{n1} was determined from $n+2$ interpolation conditions, and that the polynomials λ_k were of degree $n+1$. However—according to the definition of ω_{n1} —the degree of any term of the sum in the formula obtained does not exceed n .

In the next theorem which is analogous to Theorem 13:6, we shall express the polynomial ω_{n1} by the coefficients of series (15) under the assumption that this series converges to the function ξ . Under the same assumption we have proved in Theorem 5:6 that the number e_{n1} which appears in equations (20) is given by the formula

$$e_{n1} = (-1)^{n+1} \sum_{i=0}^{\infty} a_{(2i+1)(n+1)}[\xi].$$

At the same place we have obtained the formulas

$$\xi(t_{n+1,k}) = \sum_{i=0}^{\infty} \sum_{j=0}^n (-1)^{i(n+1-k)} a_{i(n+1)+j}[\xi] \tau_j(t_{n+1,k}),$$

$$(25) \quad \tau_{n+1-j}(t_{n+1,k}) = (-1)^{n+1-k} \tau_j(t_{n+1,k}) \quad (j = 0, 1, \dots, n+1),$$

which will be used now.

THEOREM 13:9 (Hornecker, [14]). *If the series (15) converges to the function ξ in the interval I , then the polynomial ω_{n1} defined by (20) can be expressed as*

$$\omega_{n1} = \sum_{j=0}^n a_j[\omega_{n1}] \tau_j$$

where

$$(26) \quad a_j[\omega_{n1}] = \begin{cases} \sum_{i=0}^{\infty} a_{2i(n+1)}[\xi] & (j = 0), \\ \sum_{i=0}^{\infty} (a_{2i(n+1)+j}[\xi] + a_{2(i+1)(n+1)-j}[\xi]) & (j = 1, 2, \dots, n). \end{cases}$$

The norm of the difference $\xi - \omega_{n1}$ satisfies the inequality

$$\|\xi - \omega_{n1}\|_I \leq \sum_{i=0}^{\infty} \left(|a_{(2i+1)(n+1)}[\xi]| + 2 \sum_{j=1}^{2n+1} |a_{(2i+1)(n+1)+j}[\xi]| \right).$$

Proof. Since $\tau_0 = 1$, conditions (20), which define the polynomial ω_{n1} , may be written in the form

$$\omega_{n1}(t_{n+1,k}) = \xi(t_{n+1,k}) - (-1)^k e_{n1} \tau_0(t_{n+1,k}) \quad (k = 0, 1, \dots, n+1)$$

and also, after using the formulas for $\xi(t_{n+1,k})$ and e_{n1} —in the form

$$\begin{aligned} \omega_{n1}(t_{n+1,k}) &= \sum_{i=0}^{\infty} \sum_{j=0}^n (-1)^{i(n+1-k)} a_{i(n+1)+j}[\xi] \tau_j(t_{n+1,k}) - \\ &\quad - (-1)^{n+1-k} \sum_{i=0}^{\infty} a_{(2i+1)(n+1)}[\xi] \tau_0(t_{n+1,k}). \end{aligned}$$

The double sum, after splitting into two parts corresponding to odd and even indices i , takes the form

$$\sum_{i=0}^{\infty} \sum_{j=0}^n a_{2i(n+1)+j}[\xi] \tau_j(t_{n+1,k}) + \sum_{i=0}^{\infty} \sum_{j=0}^n (-1)^{n+1-k} a_{(2i+1)(n+1)+j}[\xi] \tau_j(t_{n+1,k}).$$

Thus

$$\begin{aligned} \omega_{n1}(t_{n+1,k}) &= \sum_{i=0}^{\infty} \sum_{j=0}^n a_{2i(n+1)+j}[\xi] \tau_j(t_{n+1,k}) + \\ &\quad + \sum_{i=0}^{\infty} \sum_{j=1}^n (-1)^{n+1-k} a_{(2i+1)(n+1)+j}[\xi] \tau_j(t_{n+1,k}). \end{aligned}$$

The second part of the right-hand side of this equation will be transformed according to formula (25), with the simultaneous replacement of the index j by the index $n+1-j$, which assumes the values $1, 2, \dots, n$:

$$\omega_{n1}(t_{n+1,k}) = \sum_{i=0}^{\infty} \sum_{j=0}^n a_{2i(n+1)+j}[\xi] \tau_j(t_{n+1,k}) + \sum_{i=0}^{\infty} \sum_{j=1}^n a_{2(i+1)(n+1)-j}[\xi] \tau_j(t_{n+1,k}).$$

Now it becomes obvious that the polynomial ω_{n1} has the form given in the theorem. Formula (26) differ from formulas (16), which correspond to polynomial ψ_n , only in the signs of coefficients of series (15). Thus the estimate of the norm of the difference $\xi - \omega_{n1}$ coincides with the estimate of the norm of the difference $\xi - \psi_n$, which was proved in Theorem 13:6. The same analogy between the polynomials ω_{n1} and ψ_n allows us to omit the proof of the following theorem:

THEOREM 13:10. *If*

$$\xi = \sum_{l=0}^{\infty} x_l t^l \quad (t \in I),$$

then the polynomial $\omega_{n,1}$, defined by conditions (20), can be expressed by formula

$$\begin{aligned} \omega_{n,1} = & \sum_{k=0}^n \left\{ x_k + \sum_{l=\lfloor (n-k)/2 \rfloor + 1}^{\infty} \frac{x_{k+2l}}{2^{2l}} \sum_{m=0}^{\lfloor (n-k)/2 \rfloor} \frac{(-1)^m \max\{k+2m, 1\}}{\max\{k+m, 1\}} \times \right. \\ & \left. \times \binom{k+m}{m} \sum_{i=0}^{\lfloor (l-m)/(n+1) \rfloor} \left(\binom{k+2l}{l-m-i(n+1)} + \binom{k+2l}{l+m-(i+1)(n+1)+k} \right) \right\} t^k \end{aligned}$$

and the norm of the difference $\xi - \omega_{n,1}$ satisfies the inequality

$$\|\xi - \omega_{n,1}\|_I \leq \frac{|x_{n+1}| + |x_{n+2}|}{2^n} + \sum_{k=2}^{\infty} \frac{|x_{n+k+1}|}{2^{n+k-1}} \sum_{l=\max\{\lfloor k/2 \rfloor - n, 0\}}^{\lfloor k/2 \rfloor} \sum_{i=0}^{\lfloor l/(n+1) \rfloor} \binom{n+k+1}{l-i(n+1)}.$$

14. Approximation of families of functions. The Zolotarev polynomials.

14.1. It sometimes happens that we face the problem of approximating by polynomials all functions of a certain family of functions. Then, it is not worth-while to approximate each function separately by the methods described in the other sections. On the other hand, it may be useful to apply a method consisting of finding and solving the differential equations where the unknowns are the quantities which determine the best polynomials (for instance their coefficients). We shall describe this method in the simplest case.

Let us suppose that we are given the family of continuous functions $\xi(t; s)$ of the variable $t \in I = \langle -1, 1 \rangle$ with the parameter s , where $s \in S$, and S is an open or closed interval. Let

$$\omega_n(t; s) = \sum_{j=0}^n a_j(s) t^j$$

be the required n -th best polynomial (of the variable t) for the function $\xi(t; s)$ in the interval I . The symbol $e_n(s)$ will denote the n -th error of the best approximation of this function, i.e. the number $\varepsilon_n(\xi(t; s); I)$.

THEOREM 14:1 (Bernstein, [4], p. 40). *If for every $s \in S$*

(i) *the derivatives $\partial^2 \xi(t; s)/\partial t^2$ and $\partial^2 \xi(t; s)/\partial t \partial s$ exist and are continuous,*

(ii) there exist exactly $n + 2$ (n, I) -points $u_0(s) < u_1(s) < \dots < u_{n+1}(s)$ of the function $\xi(t; s)$ and $u_0(s) = -1, u_{n+1}(s) = 1,$

$$(iii) \quad \left. \frac{\partial^2(\xi(t; s) - \omega_n(t; s))}{\partial t^2} \right|_{t=u_k(s)} \neq 0 \quad (k = 1, 2, \dots, n),$$

then the functions $u_1(s), u_2(s), \dots, u_n(s), e_n(s), a_0(s), a_1(s), \dots, a_n(s)$ of parameter s have continuous first derivatives and satisfy the system of differential equations

$$(1) \quad \left. \frac{\partial(\xi(t; s) - \omega_n(t; s))}{\partial s} \right|_{t=u_k(s)} = (-1)^k c e'_n(s) \quad (k = 0, 1, \dots, n+1)$$

where $|c| = 1,$

$$(2) \quad u'_k(s) = - \left. \frac{\frac{\partial^2(\xi(t; s) - \omega_n(t; s))}{\partial t \partial s}}{\frac{\partial^2(\xi(t; s) - \omega_n(t; s))}{\partial t^2}} \right|_{t=u_k(s)} \quad (k = 1, 2, \dots, n).$$

Proof. According to assumption (ii) and the definition of the best polynomial, for every $s \in S$ there exists a number $c(s)$, equal either to 1 or to -1 , such that

$$(3) \quad \xi(u_k; s) - \omega_n(u_k; s) = (-1)^k c e_n \quad (k = 0, 1, \dots, n+1)$$

(we write shortly u_k, c and e_n instead of $u_k(s), c(s)$ and $e_n(s)$). The points u_1, u_2, \dots, u_n lie inside the interval I , in which the function $\xi(t; s)$ is differentiable with respect to t . At these points the difference $\xi(t; s) - \omega_n(t; s)$ has its extremum; thus

$$(4) \quad \left. \frac{\partial(\xi(t; s) - \omega_n(t; s))}{\partial t} \right|_{t=u_k} = 0 \quad (k = 1, 2, \dots, n).$$

We have obtained $2n + 2$ equations (3) and (4) for determining the same number of unknown values: the points u_1, u_2, \dots, u_n ($u_0 = -1, u_{n+1} = 1$, by assumption (ii)), the error e_n and the coefficients a_0, a_1, \dots, a_n of the polynomial ω_n . Using the notation $f_n = c e_n,$

$$\begin{aligned} & \varrho_k(s; u_1, u_2, \dots, u_n, f_n, a_0, a_1, \dots, a_n) \\ &= \frac{\partial}{\partial t} \left(\xi(t; s) - \sum_{j=0}^n a_j t^j \right) \Big|_{t=u_k} \quad (k = 1, 2, \dots, n), \end{aligned}$$

$$\begin{aligned} & \varrho_{n+1+k}(s; u_1, u_2, \dots, u_n, f_n, a_0, a_1, \dots, a_n) \\ &= \xi(u_k; s) - \sum_{j=0}^n a_j u_k^j - (-1)^k f_n \quad (k = 0, 1, \dots, n+1) \end{aligned}$$

we can write these equations in a uniform way:

$$\varrho_l(s; u_1, u_2, \dots, u_n, f_n, a_0, a_1, \dots, a_n) = 0 \quad (l = 1, 2, \dots, 2n+2).$$

By assumption (i) the functions $\varrho_1, \varrho_2, \dots, \varrho_{2n+2}$ have continuous derivatives with respect to all arguments. Thus, to prove the existence and the continuity of the first derivatives of $u_1(s), u_2(s), \dots, u_n(s), f_n(s) = c(s)e_n(s), a_0(s), a_1(s), \dots, a_n(s)$ it suffices to verify that the Jacobian

$$\frac{\partial(\varrho_1, \varrho_2, \dots, \varrho_{2n+2})}{\partial(u_1, \dots, u_n, f_n, a_0, \dots, a_n)}$$

is different from zero for $u_k = u_k(s), f_n = f_n(s), a_j = a_j(s)$ for every $s \in \mathcal{S}$. On the other hand, from the continuity of the derivative of the auxiliary function $f_n(s)$ easily follows the continuity of the derivative of the function $e_n(s)$. In fact, we have $e_n(s) > 0$, since for $e_n(s) = 0$ the polynomial $\omega_n(t; s)$ would be identical with the function $\xi(t; s)$, which would contradict assumption (iii). Since $e_n(s) = |f_n(s)|$, we have $f_n(s) \neq 0$, and $c(s)$ cannot depend upon s and $e_n(s) = f_n(s)$ for $s \in \mathcal{S}$, or $e_n(s) = -f_n(s)$ for $s \in \mathcal{S}$.

We find without difficulty that for $k = 1, 2, \dots, n$

$$\frac{\partial \varrho_k}{\partial u_i} = \delta_{ik} \frac{\partial^2}{\partial t^2} \left(\xi(t; s) - \sum_{j=0}^n a_j t^j \right)_{t=u_k} \quad (i = 1, 2, \dots, n)$$

(δ_{ik} denotes the Kronecker symbol),

$$\frac{\partial \varrho_k}{\partial f_n} = 0,$$

$$\frac{\partial \varrho_k}{\partial a_j} = -j u_k^{j-1} \quad (j = 0, 1, \dots, n)$$

and for $k = 0, 1, \dots, n+1$

$$\frac{\partial \varrho_{n+1+k}}{\partial u_i} = \delta_{ik} \varrho_k \quad (i = 1, 2, \dots, n),$$

$$\frac{\partial \varrho_{n+1+k}}{\partial f_n} = (-1)^{k+1},$$

$$\frac{\partial \varrho_{n+1+k}}{\partial a_j} = -u_k^j \quad (j = 0, 1, \dots, n).$$

In the derivatives obtained, and later in the Jacobian, we introduce

$u_k = u_k(s)$ ($k = 0, 1, \dots, n+1$), $f_n = f_n(s)$, $a_j = a_j(s)$ ($j = 0, 1, \dots, n$). For these values of the arguments and for $k = 1, 2, \dots, n$ we have

$$\frac{\partial \varrho_k}{\partial u_i} = \delta_{ik} r_k \quad (i = 1, 2, \dots, n)$$

where

$$r_k = \frac{\partial^2 (\xi(t; s) - \omega_n(t; s))}{\partial t^2} \Big|_{t=u_k(s)}$$

and $r_k \neq 0$ by assumption (iii). Besides, according to (4) and the definition of the functions $\varrho_1, \varrho_2, \dots, \varrho_n$, we obtain for $k = 0, 1, \dots, n+1$:

$$\frac{\partial \varrho_{n+1+k}}{\partial u_i} = \delta_{ik} \varrho_k = 0 \quad (i = 1, 2, \dots, n).$$

Thus, for the values of the arguments previously mentioned, we have

$$\begin{aligned} & \frac{\partial(\varrho_1, \varrho_2, \dots, \varrho_{2n+2})}{\partial(u_1, \dots, u_n, f_n, a_0, \dots, a_n)} \\ &= \begin{vmatrix} r_1 & 0 & \dots & 0 & 0 & 0 & -1 & \dots & -nu_1^{n-1}(s) \\ 0 & r_2 & \dots & 0 & 0 & 0 & -1 & \dots & -nu_2^{n-1}(s) \\ \dots & \dots \\ 0 & 0 & \dots & r_n & 0 & 0 & -1 & \dots & -nu_n^{n-1}(s) \\ 0 & 0 & \dots & 0 & -1 & -1 & -u_0(s) & \dots & -u_0^n(s) \\ 0 & 0 & \dots & 0 & 1 & -1 & -u_1(s) & \dots & -u_1^n(s) \\ \dots & \dots \\ 0 & 0 & \dots & 0 & (-1)^n & -1 & -u_{n+1}(s) & \dots & -u_{n+1}^n(s) \end{vmatrix} \\ &= (-1)^n r_1 r_2 \dots r_n \begin{vmatrix} 1 & 1 & u_0(s) & \dots & u_0^n(s) \\ -1 & 1 & u_1(s) & \dots & u_1^n(s) \\ \dots & \dots & \dots & \dots & \dots \\ (-1)^{n+1} & 1 & u_{n+1}(s) & \dots & u_{n+1}^n(s) \end{vmatrix} \\ &= (-1)^n r_1 r_2 \dots r_n \sum_{k=0}^{n+1} \text{van}(u_0(s), \dots, u_{k-1}(s), u_{k+1}(s), \dots, u_{n+1}(s)), \end{aligned}$$

where $\text{van}(u_0(s), \dots, u_{k-1}(s), u_{k+1}(s), \dots, u_{n+1}(s))$ is the Vandermonde determinant. This determinant is positive, since $u_0(s) < u_1(s) < \dots < u_{n+1}(s)$. Besides $r_1 r_2 \dots r_n \neq 0$, whence the Jacobian is also different from zero, which was to be proved.

We complete the proof of Theorem 14:1 with the derivation of differential equations (1) and (2) from identities (3) and (4) in which $u_k = u_k(s)$, $e_n = e_n(s)$. Differentiating both sides of (3) with respect to s we get

$$\begin{aligned} u'_k(s) \frac{\partial(\xi(t; s) - \omega_n(t; s))}{\partial t} \Big|_{t=u_k(s)} + \frac{\partial(\xi(t; s) - \omega_n(t; s))}{\partial s} \Big|_{t=u_k(s)} \\ = (-1)^k c e'_n(s) \quad (k = 0, 1, \dots, n+1). \end{aligned}$$

The first-term on the left-hand side of this equation is equal to zero for $k = 1, 2, \dots, n$ because of (4), and for $k = 0$ and $k = n+1$ because $u_0(s) = -1$, $u_{n+1}(s) = 1$. This proves that the differential equations of the first order (1) which contain the derivatives of functions $a_0(s)$, $a_1(s)$, \dots , $a_n(s)$, $e_n(s)$ are valid.

Let us now differentiate identity (4) with respect to s :

$$\begin{aligned} u'_k(s) \frac{\partial^2(\xi(t; s) - \omega_n(t; s))}{\partial t^2} \Big|_{t=u_k(s)} + \frac{\partial^2(\xi(t; s) - \omega_n(t; s))}{\partial t \partial s} \Big|_{t=u_k(s)} = 0 \\ (k = 1, 2, \dots, n). \end{aligned}$$

By assumption (iii) the identities obtained may be written in the form (2). Besides the derivatives of functions $u_1(s)$, $u_2(s)$, \dots , $u_n(s)$ equations (2) contain also (on the right-hand side in the numerator) the derivatives of the coefficients $a_j(s)$.

The initial conditions for systems (1) and (2) are, for some $s_0 \in \mathcal{S}$, the (n) -points $u_k(s_0)$ of the function $\xi(t; s_0)$, the best error of this function $e_n(s_0)$ and the coefficients $a_j(s_0)$ of the best polynomial $\omega_n(t; s_0)$. Thus the value s_0 should be chosen in such a way that the function $\xi(t; s_0)$ should be easy to approximate.

One step of a numerical solution of systems (1) and (2), for instance by the Runge-Kutta method, consists of two stages. At first we compute $a'_j(s)$ and $e'_n(s)$ from system (1). It follows from the form of this system that $a'_j(s)$ are the coefficients of the n -th best polynomial for the function $\partial \xi(t; s) / \partial s$ of the variable t on $(n+2)$ -point set $\{u_0(s), u_1(s), \dots, u_{n+1}(s)\}$ and $e'_n(s)$ is the n -th error of the best approximation of this function on this set. Thus system (1) may be solved by the method described in the second part of this paper. In the second stage we compute $u'_k(s)$ for system (2), which does not require any additional explanations.

In connection with Theorem 14:1 we shall mention that the form of assumptions (ii) and (iii) corresponds to the application of this theorem in a rather particular case (see § 14.2). In general, it is more convenient to use the assumption, easy to verify in practice, that in the interval I the $(n+1)$ -st derivative of the function $\xi(t; s)$ with respect to t exists

and has a constant sign. Then assumptions (ii) and (iii) are satisfied. Indeed, assumption (ii) follows directly from Theorem 10:5. We notice earlier that the points $u_1(s), u_2(s), \dots, u_n(s)$ are the zeros of the function

$$\frac{\partial(\xi(t; s) - \omega_n(t; s))}{\partial t}.$$

If assumption (iii) were not true for some k , the point $u_k(s)$ would be the double zero of this function. Thus, this function would have at least $n+1$ zeros (together with their orders), and the n -th derivative of this function, equal to

$$\frac{\partial^{n+1}(\xi(t; s) - \omega_n(t; s))}{\partial t^{n+1}} = \frac{\partial^{n+1} \xi(t; s)}{\partial t^{n+1}},$$

would have at least one zero. This, in turn, would contradict the assumption of constancy of sign of the function $\partial^{n+1} \xi(t; s) / \partial t^{n+1}$.

14.2. Now we shall investigate more closely the Zolotarev polynomials, defined in § 13.1; while doing that we shall use the results of § 14.1. We recall that the n -th Zolotarev polynomial ζ_n^s with parameter s equals, by definition, the difference between the binomial $t^n + st^{n-1}$ and $(n-2)$ -nd best polynomial for this binomial in the interval $I = \langle -1, 1 \rangle$ (of course it is assumed that $n \geq 2$). Thus, according to the fundamental Theorem 1:2, the polynomial ζ_n^s is completely determined by the following properties: (Z_{n1}^s) it is a polynomial of degree n of the form $t^n + st^{n-1} + \dots$; (Z_{n2}^s) there exists at least n points $u_0 < u_1 < \dots < u_{n-1}$ in the interval I at which this polynomial has the extremal values equal alternately to $\|\zeta_n^s\|_I$ and $-\|\zeta_n^s\|_I$.

The following identity, valid for all n and s , follows directly from these properties:

$$(5) \quad \zeta_n^{-s}(t) = (-1)^n \zeta_n^s(-t).$$

In fact

$$(-1)^n \zeta_n^s(-t) = (-1)^n ((-t)^n + s(-t)^{n-1} + \dots) = t^n - st^{n-1} + \dots,$$

i.e. the polynomial $(-1)^n \zeta_n^s(-t)$ has the property (Z_{n1}^{-s}) . It also has the property (Z_{n2}^{-s}) , since at the points $-u_{n-1} < \dots < -u_1 < -u_0$ of the interval I , it has alternately the minimal and the maximal values.

Those of the points u_0, u_1, \dots, u_{n-1} mentioned in property (Z_{n2}^s) which lie inside the interval I are the zeros of the derivative of the polynomial ζ_n^s . Since this derivative is a polynomial of the degree $n-1$, it

has at most $n-1$ distinct zeros. Thus the inequality $-1 < u_0 < u_1 < \dots < u_{n-1} < 1$ is excluded, and we have only two possible cases:

- I. $-1 = u_0 < u_1 < \dots < u_{n-1} < 1$ or $-1 < u_0 < u_1 < \dots < u_{n-1} = 1$,
 II. $-1 = u_0 < u_1 < \dots < u_{n-1} = 1$.

Let us first consider case I, which, as will be seen, corresponds to those values of parameter s which are close to zero. The theorem below concerns also case II, provided that number -1 or 1 equals one of the zeros of the derivative of the polynomial ζ_n^s .

THEOREM 14:2 (Achiezer, [1], p. 265). *If $|s| \leq n \tan^2 \pi/2n$, then*

$$(6) \quad \zeta_n^s(t) = 2 \left(\frac{n+|s|}{2n} \right)^n \tau_n \left(\frac{nt+s}{n+|s|} \right).$$

Proof. At first we shall prove formula (6) for the polynomial ζ_n^s under the assumption that the points u_0, u_1, \dots, u_{n-1} which correspond to this polynomial satisfy the inequality

$$(7) \quad -1 = u_0 < u_1 < \dots < u_{n-1} < 1.$$

In case I also the inequality $-1 < u_0 < u_1 < \dots < u_{n-1} = 1$ is possible. However, in this case, according to (5), the polynomial ζ_n^s has its extremal values at the points $-u_{n-1}, \dots, -u_1, -u_0$. They satisfy the inequality $-1 = -u_{n-1} < \dots < -u_1 < -u_0 < 1$, whence the formula for the polynomial ζ_n^{-s} will be derived in the latter part of the proof. If we know the form of the polynomial ζ_n^{-s} , we can obtain ζ_n^s using identity (5). In order to justify our restriction to the case of (7) we should verify moreover, that formula (6) coincides with identity (5). We can do it using the identity $\tau_n(-t) = (-1)^n \tau_n(t)$, which follows from the definition of the Chebyshev polynomial τ_n :

$$\begin{aligned} (-1)^n \zeta_n^s(-t) &= 2(-1)^n \left(\frac{n+|s|}{2n} \right)^n \tau_n \left(\frac{-nt+s}{n+|s|} \right) \\ &= 2 \left(\frac{n+|-s|}{2n} \right)^n \tau_n \left(\frac{nt+(-s)}{n+|-s|} \right) = \zeta_n^{-s}(t). \end{aligned}$$

In view of (7) all the zeros of the derivative of the polynomial ζ_n^s are the points u_1, u_2, \dots, u_{n-1} . This polynomial is of the form $t^n + \dots$, whence it increases to $+\infty$ to the right of the largest zero u_{n-1} of its derivative. The value $\zeta_n^s(u_{n-1})$ must be negative, for in the opposite case the value $\zeta_n^s(1)$ would be greater than $\zeta_n^s(u_{n-1}) = \|\zeta_n^s\|_I$, which would contradict the definition of this norm. Thus there exists $u_n > u_{n-1}$ such that

$$(8) \quad \zeta_n^s(u_n) = -\zeta_n^s(u_{n-1}) > 0.$$

At the same time $u_n \geq 1$, because $\zeta_n^s(1)$ cannot exceed $\|\zeta_n^s\|_I$.

From (8) and from the property (Z_{n2}^s) of points u_0, u_1, \dots, u_{n-1} it follows that

$$\zeta_n^s(u_k) = (-1)^{n-k} \|\zeta_n^s\|_{\langle -1, u_n \rangle} \quad (k = 0, 1, \dots, n)$$

(the norm of the polynomial ζ_n^s in the interval $\langle -1, u_n \rangle$ coincides with the norm in I).

Let us now consider the polynomial

$$\tau(t) = \zeta_n^s \left(\frac{(u_n - u_0)t + u_0 + u_n}{2} \right) = \zeta_n^s \left(\frac{(u_n + 1)t + u_n - 1}{2} \right).$$

The transformation

$$(9) \quad u = \frac{(u_n + 1)t + u_n - 1}{2}$$

transforms linearly the interval $I = \langle -1, 1 \rangle$ into the interval $\langle -1, u_n \rangle$. At $n + 1$ points $-1 = t_0 < t_1 < \dots < t_n = 1$, which, under this transformation, are transformed into the points u_0, u_1, \dots, u_n , the polynomial τ has its extremal values, equal alternately to $\|\zeta_n^s\|_{\langle -1, u_n \rangle}$ and $-\|\zeta_n^s\|_{\langle -1, u_n \rangle}$, i.e. $\|\tau\|_I$ and $-\|\tau\|_I$.

The n -th Chebyshev polynomial τ_n has a similar property. It has extremal values equal to either 1 or -1 at $n + 1$ points of the interval I . This property determines uniquely τ_n ; thus the polynomial τ differs from τ_n only by a constant factor, and the points t_k are identical with the points $t_{nk} = -\cos k\pi/n$ (see § 2.2), for $k = 0, 1, \dots, n$.

Thus we have proved the identity

$$\zeta_n^s \left(\frac{(u_n + 1)t + u_n - 1}{2} \right) = p\tau_n(t)$$

which is the same as

$$(10) \quad \zeta_n^s(t) = p\tau_n \left(\frac{2t + 1 - u_n}{u_n + 1} \right),$$

where p is a certain number, dependent, like parameter s , on u_n .

Before we derive formula (6) from this identity, we shall give the conditions which should be satisfied by the point u_n . We already know that $u_n \geq 1$. The estimate from above follows from the fact that transformation (9) transforms the point $t_{n,n-1} = \cos \pi/n$ into a point u_{n-1} such that $u_{n-1} < 1$:

$$\frac{(u_n + 1)\cos \pi/n + u_n - 1}{2} < 1.$$

At the same time $u_n + 1 > 0$, and the number -1 is transformed into

itself, so that this inequality ensures the fulfilment of condition (7). Finally we get

$$1 \leq u_n < \frac{3 - \cos \pi/n}{1 + \cos \pi/n} = 1 + 2 \frac{1 - \cos \pi/n}{1 + \cos \pi/n} = 1 + 2 \tan^2 \pi/2n.$$

Let us now return to identity (10). It follows from the equation given at the beginning of § 13.1 that $\tau_n(t) = 2^{n-1}t^n + \dots$ (here and in the next equation the symbol "... represents the terms which do not contain t^n and t^{n-1}). Thus

$$\begin{aligned} \zeta_n^s(t) &= 2^{n-1}p \left(\frac{2t+1-u_n}{u_n+1} \right)^n + \dots = 2^{n-1}p \frac{2^n t^n + 2^{n-1}n(1-u_n)t^{n-1} + \dots}{(u_n+1)^n} \\ &= \frac{2^{2n-1}p}{(u_n+1)^n} (t^n + \frac{1}{2}n(1-u_n)t^{n-1} + \dots). \end{aligned}$$

Since also $\zeta_n^s(t) = t^n + st^{n-1} + \dots$, we have

$$s = \frac{1}{2}n(1-u_n) \leq 0, \quad u_n = 1 - 2s/n,$$

$$p = \frac{(u_n+1)^n}{2^{2n-1}} = \frac{(2-2s/n)^n}{2^{2n-1}} = 2 \left(\frac{n-s}{2n} \right)^n,$$

$$\zeta_n^s(t) = 2 \left(\frac{n-s}{2n} \right)^n \tau_n \left(\frac{2t+2s/n}{2-2s/n} \right) = 2 \left(\frac{n+|s|}{2n} \right)^n \tau_n \left(\frac{nt+s}{n+|s|} \right).$$

The proof of Theorem 14:2 will be completed if we verify that formula (6) is valid also for $s = -n \tan^2 \pi/2n$, i.e. for $u_{n-1} = 1$. This equation was excluded in assumption (7). However, the reasoning that led us to (10) will be valid also in the case when (as in case II) $-1 = u_0 < u_1 < \dots < u_{n-1} = 1$, provided u_{n-1} will remain the zero of the derivative of the polynomial ζ_n^s . It will be seen in the sequel that formula (6) for $|s| = n \tan^2 \pi/2n$ helps in computing the coefficients of polynomial ζ_n^s for greater values of $|s|$.

In case I the polynomial ζ_n^s had, for an arbitrary n , the exact analytical expression (6). Let us now consider case II where $-1 = u_0 < u_1 < \dots < u_{n-1} = 1$. In this case we do not have any similar expressions, and the polynomial ζ_n^s can be found only approximately, for instance, by numerical solution of a certain system of differential equations.

At the points u_1, u_2, \dots, u_{n-2} , which lie inside the interval I , the polynomial ζ_n^s assumes its maximal and minimal values alternately. Thus, these points are the zeros of an odd order of the derivative of the polynomial ζ_n^s , which has only $n-1$ zeros altogether. Thus all its zeros are single. Let us denote the zero different from u_1, u_2, \dots, u_{n-2} by the

symbol u . The zero u does not belong to any of the intervals (u_i, u_{i+1}) for $i = 0, 1, \dots, n-2$, since in this case the polynomial ζ_n^s would have extrema of the same kind at the points u_i and u_{i+1} . Thus we must have $|u| \geq 1$. If necessary, we can always assume that $|u| > 1$, since—as has been mentioned earlier — in case II for $u = -1$ and $u = 1$, i.e. for $s = n \tan^2 \pi / 2n$ and $s = -n \tan^2 \pi / 2n$ Theorem 14:2 holds.

THEOREM 14:3. *If $|s| \geq n \tan^2 \pi / 2n$, then the zeros $u_1, u_2, \dots, u_{n-2}, u$ of the derivative of the polynomial ζ_n^s , regarded as functions of the parameter s , have continuous derivatives. The zero u is a function of s which decreases from $+\infty$ to 1 for $s \leq -n \tan^2 \pi / 2n$ and from -1 to $-\infty$ for $s \geq n \tan^2 \pi / 2n$.*

It will be seen later that it is convenient to consider the polynomial ζ_n^s not as a function of s but as a function of u . Theorem 14:3 allows us to do so; it follows also from this theorem that $s, u_1, u_2, \dots, u_{n-2}$, regarded as functions of u , have continuous derivatives.

Proof. At first we shall verify that we may apply Theorem 14:1 for $\xi(t; s) = t^n + st^{n-1}$ and

$$s = (-\infty, -n \tan^2 \pi / 2n) \quad \text{or} \quad s = (n \tan^2 \pi / 2n, +\infty)$$

if we replace n by $n-2$.

Assumption (i) of Theorem 14:1 is obviously satisfied. It follows from the definition of the polynomial ζ_n^s that $\zeta_n^s(t) = \xi(t; s) - \omega_{n-2}(t; s)$, where ω_{n-2} is the $(n-2)$ -nd best polynomial for the function ξ in the interval I . The points u_0, u_1, \dots, u_{n-1} , which appear in property (Z_{n-2}^s) , are, at the same time, the $(n-2)$ -points of the function $\xi(t; s)$, and assumption (ii) follows immediately from this property and from the case II. Assumption (iii) is also satisfied, because it states that the points u_1, u_2, \dots, u_{n-2} are the single zeros of the derivative of the polynomial ζ_n^s .

From Theorem 14:1 it follows directly that the points u_1, u_2, \dots, u_{n-2} regarded as functions of s have continuous derivatives. Since

$$\begin{aligned} (\zeta_n^s)' &= n(t - u_1)(t - u_2) \dots (t - u_{n-2})(t - u) \\ &= nt^{n-1} - n(u_1 + u_2 + \dots + u_{n-2} + u)t^{n-2} + \dots, \end{aligned}$$

we have

$$\begin{aligned} \zeta_n^s &= t^n - \frac{n}{n-1}(u_1 + u_2 + \dots + u_{n-2} + u)t^{n-1} + \dots, \\ (11) \quad u &= -\frac{n-1}{n}s - u_1 - u_2 - \dots - u_{n-2} \end{aligned}$$

and also the zero u has a continuous derivative.

It remains to prove that u is a decreasing function of s . Suppose, on the contrary, that there exist two different numbers s_1 and s_2 such that the derivatives of the polynomials $\zeta_n^{s_1}$ and $\zeta_n^{s_2}$ have the same zero $u \notin (-1, 1)$. Let

$$\begin{aligned}\delta &= \zeta_n^{s_1}(u) \zeta_n^{s_2} - \zeta_n^{s_2}(u) \zeta_n^{s_1} \\ &= (\zeta_n^{s_1}(u) - \zeta_n^{s_2}(u)) t^n + (s_2 \zeta_n^{s_1}(u) - s_1 \zeta_n^{s_2}(u)) t^{n-1} + \dots\end{aligned}$$

This polynomial cannot be identically zero, as $\zeta_n^{s_1}(u) \neq \zeta_n^{s_2}(u)$ or $\zeta_n^{s_1}(u) = \zeta_n^{s_2}(u) \neq 0$ and $s_1 \neq s_2$. The number u is a double zero of the polynomial δ , whence it has at most $n-2$ zeros in the interval $(-1, 1)$

Suppose first that

$$(12) \quad |\zeta_n^{s_1}(u)| \cdot \|\zeta_n^{s_2}\|_I > |\zeta_n^{s_2}(u)| \cdot \|\zeta_n^{s_1}\|_I.$$

Retaining the notation u_0, u_1, \dots, u_{n-1} for $s = s_2$ we see that at these n points the polynomial δ has the same signs as the polynomial $\zeta_n^{s_1}(u) \zeta_n^{s_2}$, i.e. positive and negative, alternately. Thus the polynomial δ has $n-1$ distinct zeros in the interval $(-1, 1)$, which contradicts the facts established previously. Thus inequality (12) does not hold. The same is true for the opposite strict inequality. Thus

$$|\zeta_n^{s_1}(u)| \cdot \|\zeta_n^{s_2}\|_I = |\zeta_n^{s_2}(u)| \cdot \|\zeta_n^{s_1}\|_I.$$

Then the functions $\beta = \zeta_n^{s_1}(u) \zeta_n^{s_2}$, $\gamma = \zeta_n^{s_2}(u) \zeta_n^{s_1}$ whose difference is equal to δ have the same norm in the interval I . For $p = n-1$ and $c_k = u_k$ ($k = 0, 1, \dots, n-1$) we can apply to them Theorem 10:2, from which it follows that the polynomial δ has in the interval $\langle -1, 1 \rangle$ at least $n-1$ zeros, including their orders. This, however, is impossible if $|u| > 1$, and the equation $|u| = 1$ has already been investigated in the preceding theorem.

The assumption concerning the numbers s_1 and s_2 have turned out to be false. Thus we have proved that u is a continuous and univalued function of parameter s . Since $|u_1 + u_2 + \dots + u_{n-2}| < n-2$, it follows from (11) that $u \rightarrow +\infty$ for $s \rightarrow -\infty$ and $u \rightarrow -\infty$ for $s \rightarrow +\infty$. At the same time $u = 1$ and $u = -1$ for $s = -n \tan^2 \pi/2n$ and $s = n \tan^2 \pi/2n$, respectively; hence u is a decreasing function of s , which completes the proof of Theorem 14:3.

Now we shall proceed according to our earlier statement and we shall consider the parameter s and the zeros u_1, u_2, u_{n-2} of the derivative of the polynomial ζ_n^s as functions of the zero u of that derivative. Thus the symbol ζ_n^s will now be replaced by a new symbol $\zeta_n(t; u)$. We retain the notations $u_0 = -1$ and $u_{n-1} = 1$ for arbitrary u .

THEOREM 14:4 (Zolotarev, [38], Voronovskaya, [36]). *If*

$$\pi_n(t; u) = \prod_{k=0}^{n-1} (t - u_k),$$

then

$$(13) \quad (t^2 - 1) \frac{\partial \zeta_n(t; u)}{\partial t} = n(t - u) \pi_n(t; u),$$

$$(14) \quad \frac{\partial}{\partial u} \left(\frac{\zeta_n(t; u)}{\zeta_n(1; u)} \right) = \pi_n(t; u) \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right).$$

Proof. Both sides of equation (13) are polynomials of the variable t , of the degree $n + 1$, and with the coefficients n of t^{n+1} . Since

$$\frac{\partial}{\partial u} \left(\frac{\zeta_n(t; u)}{\zeta_n(1; u)} \right) = \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right) t^n + \dots$$

both sides of equation (14) are polynomials of the same, n -th, degree and with the same coefficients of t^n . Thus, it suffices to verify that both sides of these equations have the same zeros.

The zeros of the left-hand side of (13) are -1 and 1 and the zeros of the derivative with respect to t of $\zeta_n(t; u)$, that is the numbers $u_1, u_2, \dots, u_{n-2}, u$. By definition of function $\pi_n(t; u)$ the same numbers are the zeros of the right-hand side of (13).

By property (Z_{n2}^s) the values of the polynomial $\zeta_n(t; u)$ at the points $u_0, u_1, \dots, u_{n-1} = 1$ ($u_{n-1} = 1$, since we consider case II) have the same absolute value. Thus, for all u :

$$\frac{\zeta_n(u_k; u)}{\zeta_n(1; u)} = \pm 1 \quad (k = 0, 1, \dots, n-1).$$

Let us differentiate this identity with respect to u , remembering that u_k 's are functions of u :

$$\frac{\partial}{\partial u} \left(\frac{\zeta_n(u_k; u)}{\zeta_n(1; u)} \right) = \frac{du_k}{du} \cdot \frac{\partial}{\partial t} \left(\frac{\zeta_n(t; u)}{\zeta_n(1; u)} \right) \Big|_{t=u_k} + \frac{\partial}{\partial u} \left(\frac{\zeta_n(t; u)}{\zeta_n(1; u)} \right) \Big|_{t=u_k} = 0$$

$$(k = 0, 1, \dots, n-1).$$

In the sum obtained the first term is equal to zero for all k , since u_0 and u_{n-1} are constants, and for $k = 1, 2, \dots, n-2$ we have

$$\frac{\partial}{\partial t} \left(\frac{\zeta_n(t; u)}{\zeta_n(1; u)} \right) \Big|_{t=u_k} = \frac{1}{\zeta_n(1; u)} \cdot \frac{\partial \zeta_n(t; u)}{\partial t} \Big|_{t=u_k} = 0$$

according to the definition of points u_k . Hence

$$\frac{\partial}{\partial u} \left(\frac{\zeta_n(t; u)}{\zeta_n(1; u)} \right) \Big|_{t=u_k} = 0 \quad (k = 0, 1, \dots, n-1).$$

It means that the zeros u_0, u_1, \dots, u_{n-1} of the polynomial $\pi_n(t; u)$ are also the zeros of the left-hand side of (14), which was to be proved.

THEOREM 14:5. *If $x = u^{-1}$, then for $x \in (-1, 0) \cup (0, 1)$ the zeros u_1, u_2, \dots, u_{n-2} of the derivative with respect to t of the polynomial $\zeta_n(t; u)$ satisfy the system of differential equations*

$$(15) \quad \frac{du_k}{dx} = - \frac{1 - u_k^2}{(1 - xu_k) \left[2 + (1 - x^2) \sum_{j=1}^{n-2} (1 - xu_j)^{-1} \right]} \quad (k = 1, 2, \dots, n-2)$$

with the initial conditions

$$u_k(-1) = \frac{t_{n,k+1} - \sin^2 \pi/2n}{\cos^2 \pi/2n}, \quad u_k(1) = \frac{t_{nk} + \sin^2 \pi/2n}{\cos^2 \pi/2n} \quad (k = 1, 2, \dots, n-2).$$

Proof. We divide both sides of identity (13) by number $\zeta_n(1; u)$ (different from zero) and differentiate it with respect to u :

$$(t^2 - 1) \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \cdot \frac{\partial \zeta_n(t; u)}{\partial t} \right) = n \left(- \frac{\pi_n(t; u)}{\zeta_n(1; u)} + \right. \\ \left. + (t - u) \pi_n(t; u) \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right) + (t - u) \frac{\partial \pi_n(t; u)}{\partial u} \cdot \frac{1}{\zeta_n(1; u)} \right).$$

Next, we differentiate identity (14) with respect to t and multiply it by $t^2 - 1$:

$$(t^2 - 1) \frac{\partial^2}{\partial t \partial u} \left(\frac{\zeta_n(t; u)}{\zeta_n(1; u)} \right) = (t^2 - 1) \frac{\partial \pi_n(t; u)}{\partial t} \cdot \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right).$$

The left-hand sides of the identities obtained coincide; thus we compare the right-hand sides:

$$(16) \quad (t^2 - 1) \frac{\partial \pi_n(t; u)}{\partial t} \cdot \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right) = n \left(- \frac{\pi_n(t; u)}{\zeta_n(1; u)} + \right. \\ \left. + (t - u) \pi_n(t; u) \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right) + (t - u) \frac{\partial \pi_n(t; u)}{\partial u} \cdot \frac{1}{\zeta_n(1; u)} \right).$$

Let us now put $t = u$ in identity (16):

$$(u^2 - 1) \frac{\partial \pi_n(t; u)}{\partial t} \Big|_{t=u} \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right) = -n \frac{\pi_n(u; u)}{\zeta_n(1; u)}.$$

Since

$$\frac{1}{\pi_n(t; u)} \cdot \frac{\partial \pi_n(t; u)}{\partial t} = \sum_{j=0}^{n-1} (t - u_j)^{-1},$$

we have

$$(17) \quad \zeta_n(1; u) \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right) = - \frac{n}{(u^2 - 1) \sum_{j=0}^{n-1} (u - u_j)^{-1}}.$$

On the other hand, if we put $t = u_k$ ($k = 1, 2, \dots, n - 2$) into identity (16), then using the equation $\pi_n(u_k; u) = 0$, we get

$$\begin{aligned} (u_k^2 - 1) \frac{\partial \pi_n(t; u)}{\partial t} \Big|_{t=u_k} \cdot \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right) \\ = n(u_k - u) \frac{\partial \pi_n(t; u)}{\partial u} \Big|_{t=u_k} \frac{1}{\zeta_n(1; u)}. \end{aligned}$$

The derivatives of the functions $\pi_n(t; u)$ are given by the formulas

$$\begin{aligned} \frac{\partial \pi_n(t; u)}{\partial t} &= \sum_{k=0}^{n-1} \prod_{j=0, j \neq k}^{n-1} (t - u_j), \\ \frac{\partial \pi_n(t; u)}{\partial u} &= - \sum_{k=0}^{n-1} \frac{du_k}{du} \prod_{j=0, j \neq k}^{n-1} (t - u_j). \end{aligned}$$

It follows from these formulas that

$$\begin{aligned} \frac{\partial \pi_n(t; u)}{\partial t} \Big|_{t=u_k} &= \prod_{j=0, j \neq k}^{n-1} (u_k - u_j), \\ \frac{\partial \pi_n(t; u)}{\partial u} \Big|_{t=u_k} &= - \frac{du_k}{du} \prod_{j=0, j \neq k}^{n-1} (u_k - u_j). \end{aligned}$$

Hence

$$(u_k^2 - 1) \zeta_n(1; u) \frac{\partial}{\partial u} \left(\frac{1}{\zeta_n(1; u)} \right) = -n(u_k - u) \frac{du_k}{du} \quad (k = 1, 2, \dots, n - 2).$$

Comparing this equation with (17) we see that

$$\begin{aligned} \frac{n(u_k^2 - 1)}{(u^2 - 1) \sum_{j=0}^{n-1} (u - u_j)^{-1}} &= -n(u_k - u) \frac{du_k}{du}, \\ \frac{du_k}{du} &= \frac{1 - u_k^2}{(u - u_k)(u^2 - 1) \sum_{j=0}^{n-1} (u - u_j)^{-1}} \quad (k = 1, 2, \dots, n - 2). \end{aligned}$$

The independent variable u , which runs through the intervals $(-\infty, -1)$ and $(1, +\infty)$ will now be replaced by the new variable $x = u^{-1}$, which runs through the intervals $(-1, 0)$ and $(0, 1)$:

$$\begin{aligned} \frac{du_k}{dx} &= -\frac{1}{x^2} \cdot \frac{du_k}{du} = -\frac{1}{x^2} \cdot \frac{1-u_k^2}{(x^{-1}-u_k)(x^{-2}-1) \prod_{j=0}^{n-1} (x^{-1}-u_j)^{-1}} \\ &= -\frac{1-u_k^2}{(1-xu_k)(1-x^2) \prod_{j=0}^{n-1} (1-xu_j)^{-1}} \quad (k = 1, 2, \dots, n-2). \end{aligned}$$

In the denominator of the right-hand side of each of these differential equations we have $u_0 = -1$, whence $1-xu_0 = 0$ for $x = -1$. Similarly $u_{n-1} = 1$ and $1-xu_{n-1} = 0$ for $x = 1$. To remove this apparent singularity (apparent only, since at the same time $1-x^2 = 0$) we perform the transformation

$$\begin{aligned} (1-x^2) \prod_{j=0}^{n-1} (1-xu_j)^{-1} &= (1-x^2) \left((1+x)^{-1} + (1-x)^{-1} + \sum_{j=1}^{n-2} (1-xu_j)^{-1} \right) \\ &= 2 + (1-x^2) \sum_{j=1}^{n-2} (1-xu_j)^{-1} \end{aligned}$$

and we obtain the system of differential equations in its final form (15).

To determine the initial conditions for the solutions of system (15) we use formula (6). This formula is valid in particular for $s = n \tan^2 \pi/2n$, i.e. for $u = x = -1$

$$\begin{aligned} \zeta_n^{n \tan^2 \pi/2n}(t) &= 2 \left(\frac{n + n \tan^2 \pi/2n}{2n} \right)^n \tau_n \left(\frac{nt + n \tan^2 \pi/2n}{n + n \tan^2 \pi/2n} \right) \\ &= 2(2 \cos^2 \pi/2n)^{-n} \tau_n(t \cos^2 \pi/2n + \sin^2 \pi/2n). \end{aligned}$$

The zeros of the derivative of the polynomial τ_n are the points $t_{nk} = -\cos k\pi/n$ for $k = 1, 2, \dots, n-1$. Thus the zeros of the derivative of the polynomial $\zeta_n^{n \tan^2 \pi/2n}$ are the points

$$\frac{t_{nk} - \sin^2 \pi/2n}{\cos^2 \pi/2n} \quad (k = 1, 2, \dots, n-1);$$

the first among them equals $-1 = u$, and the remaining ones lie in the interval $(-1, 1)$. Thus they are equal to $u_1(-1), u_2(-1), \dots, u_{n-2}(-1)$, respectively. The formulas for $u_k(1)$ can be obtained in an analogous way, by applying formula (6) for $s = -n \tan^2 \pi/2n$, i.e. for $u = x = 1$.

We conclude this section with a few remarks concerning Theorem 14:5, which has just been proved.

The functions $u_k(x)$ could be defined also at the point $x = 0$, i.e. $u = \pm\infty$, assuming that $u_k(0) = t_{n-1,k}$ ($k = 1, 2, \dots, n-2$). This is a consequence of the fact that

$$\lim_{|u| \rightarrow \infty} s^{-1} \zeta_n(t; u) = \tau_{n-1}(t).$$

However, the proof of this equation would be rather cumbersome.

Equations (15) (given for the first time in my paper [28]) have arisen as a result of the elimination of function $\zeta_n(t; u)$ from identities (13) and (14). We can eliminate function $\pi_n(t; u)$ from these identities, thus obtaining the equations directly for the required coefficients of the polynomial $\zeta_n(t; u)$. Voronovskaya in [36] proceeds thus. However, in this way one gets much more complicated equations, which have to be solved by iterational methods. On the other hand it does not present any difficulties to obtain the coefficients of the polynomial $\zeta_n(t; u)$ after the numerical solution of system (15). It suffices to use the formula

$$\frac{d\zeta_n(t; u)}{dt} = u(t - x^{-1}) \prod_{k=1}^{n-2} (t - u_k),$$

compute the coefficients of the right-hand side of this equation and then integrate this polynomial with respect to t . This method will provide us with all the coefficients of the polynomial $\zeta_n(t; u)$ except the absolute term a_n . This term may be computed by using, for instance, the equation $\zeta_n(u_{n-2}; u) = -\zeta_n(1; u)$. In fact, all the coefficients of the polynomial $\zeta_n(t; u) - a_n$ are known and

$$a_n = -\frac{1}{2} \left((\zeta_n(u_{n-2}; u) - a_n) + (\zeta_n(1; u) - a_n) \right).$$

Let us now consider system (15) in the simplest case. For $n = 3$ it reduces to the single equation

$$\frac{du_1}{dx} = \frac{1 - u_1^2}{(1 - xu_1)[2 + (1 - x^2)(1 - xu_1)^{-1}]} = -\frac{1 - u_1^2}{3 - 2xu_1 - x^2}.$$

The solution of this equation, which satisfies the initial condition

$$u_1(-1) = \frac{t_{32} - \sin^2 \pi/6}{\cos^2 \pi/6} = \frac{1}{3}$$

(and the equivalent condition

$$u_1(1) = \frac{t_{31} + \sin^2 \pi/6}{\cos^2 \pi/6} = -\frac{1}{3}),$$

is the function $u_1 = -x/3 = -1/3u$. Hence

$$\begin{aligned}\frac{\partial \zeta_3(t; u)}{\partial t} &= 3(t-u)(t+1/3u) = 3t^2 + (1/u - 3u)t - 1, \\ \zeta_3(t; u) &= t^3 + (1/2u - 3u/2)t^2 - t + a_3, \\ \zeta_3(u_1; u) - a_3 &= 1/54u^3 + 1/6u, \quad \zeta_3(1; u) - a_3 = 1/2u - 3u/2, \\ a_3 &= 3u/4 - 1/3u - 1/108u^3\end{aligned}$$

and finally

$$\zeta_3(t; u) = t^3 + (1/2u - 3u/2)t^2 - t + 3u/4 - 1/3u - 1/108u^3.$$

It is easy to notice that the return to the original parameter $s = 1/2u - 3u/2$ would considerably complicate the form of the absolute term in the polynomial $\zeta_3(t; u)$.

15. The theory of the method of Remez.

15.1. We do not know any analytical formula which would express the best polynomial for a given continuous function by some quantities connected with this function. The only formulas that are known are some approximate formulas which allow us to compute the approximation to the best polynomial with, in general, an arbitrarily small error, given in advance. In practice, of course, this error is restricted by rounding-off errors and the accuracy of the computer used.

Remez ([29], [30], see also [31]) has given two methods of approximate computation of the best polynomial. In § 15 we shall discuss in detail the method which he calls in [31] the "second algorithm", and which he himself considers as the better one. This method is widely applied, known sometimes as the method of equalizing the maxima. It has been generalized by other authors (see for instance [21]) and modified. These modifications will be described in § 16.

We shall see that the method of Remez is based upon the properties of the best approximation of the $(n+2)$ -point sets. Some of these properties were already given earlier (Theorems 2:1 and 5:1). Here we shall present one more auxiliary theorem, connected with this subject. The symbol F will denote a closed set consisting of at least $n+2$ points.

THEOREM 15:1 ([34], p. 84). *If $\xi \in \mathcal{C}_F$ and if for the sequence of sets*

$$U_m = \{u_{m0}, u_{m1}, \dots, u_{m, n+1}\} \subset F$$

where $u_{m0} < u_{m1} < \dots < u_{m, n+1}$ ($m = 1, 2, \dots$) we have the inequality

$$(1) \quad \inf_m \varepsilon_n(\xi; U_m) > 0,$$

then

$$\inf_m \min_{0 \leq k \leq n} (u_{m,k+1} - u_{mk}) > 0.$$

Proof (see [31], p. 40). Suppose that the theorem is false. Then we can select a subsequence $\{U_{m_l}\}$ from the sequence $\{U_m\}$ satisfying (1), such that

$$\lim_{l \rightarrow \infty} \min_{0 \leq k \leq n} (u_{m_l,k+1} - u_{m_l,k}) = 0.$$

It follows—in view of the inequality for the points $u_{m_0}, u_{m_1}, \dots, u_{m,n+1}$ —that there exists a subsequence of the sequence of sets $\{U_{m_l}\}$ convergent to the set $U \subset F$, consisting of $p < n + 2$ points. Thus there exists a polynomial $\omega \in \mathcal{W}_n$ identical with the function ξ on the whole set U . The difference $\xi - \omega$ is a continuous function, whence in some neighbourhoods of the points of set U we have

$$|\xi(t) - \omega(t)| < \inf_m \varepsilon_n(\xi; U_m).$$

On the other hand, sufficiently large terms of the above-mentioned subsequence of the sequence $\{U_{m_l}\}$ belong to this neighbourhood, i.e. for some m and $t \in U_m$ we have

$$|\xi(t) - \omega(t)| < \varepsilon_n(\xi; U_m).$$

However, this contradicts the definition of the error of the best approximation $\varepsilon_n(\xi; U_m)$.

15.2. We shall describe the method of Remez for functions continuous on a closed set F . Let ω_{nF} be the n -th best polynomial for the function $\xi \in \mathcal{C}_F$ on this set:

$$\varepsilon_n(\xi; F) \stackrel{\text{def}}{=} \min_{\omega \in \mathcal{W}_n} \|\xi - \omega\|_F = \|\xi - \omega_{nF}\|_F.$$

We shall assume that $\varepsilon_n(\xi; F) > 0$, i.e. that the function ξ is not a polynomial of class \mathcal{W}_n on the set F . Otherwise we could find the coefficients of the polynomial ω_{nF} (if at all necessary) by any interpolation formula.

The method of Remez consists of a recursive computation of a sequence of polynomials $\omega_{n1}, \omega_{n2}, \dots$ of class \mathcal{W}_n which converges to ω_{nF} . The polynomial ω_{nm} ($m = 1, 2, \dots$) is, by definition, the n -th best polynomial on the $(n + 2)$ -point subset $U_m = \{u_{m0}, u_{m1}, \dots, u_{m,n+1}\}$ of the set F , where

$$(2) \quad u_{m0} < u_{m1} < \dots < u_{m,n+1}.$$

Thus, by Theorem 2:1, the coefficients of the polynomial ω_{nm} and a number e_{nm} such that $|e_{nm}| = \varepsilon_n(\xi; U_m)$ satisfy the system of linear equations

$$(3) \quad \xi(u_{mk}) - \omega_{nm}(u_{mk}) = (-1)^k e_{nm} \quad (k = 0, 1, \dots, n+1).$$

The second element of the method of Remez is the construction of the sets U_1, U_2, \dots , which will now be described. Later, Theorem 15:2 will show the feasibility of this construction. In Theorem 15:3 we shall prove the convergence of the numerical sequence $\{\|\xi - \omega_{mn}\|_F\}$ to $\varepsilon_n(\xi; F)$; hence the convergence of the sequence $\{\omega_{nm}\}$ to ω_{nF} on the set F will follow from Theorem 15:4. Theorem 15:3 will also allow us to estimate the speed of this convergence.

The subsets U_1, U_2, \dots of the set F which satisfy inequalities (2) are defined as follows:

I. The set U_1 is an arbitrary set such that $e_{n1} \neq 0$. Thus, to verify whether this set has been defined properly, one has to solve the system of equations (3) for $m = 1$. Under the assumption that $\varepsilon_n(\xi; F) > 0$, the set U_1 exists. It can be formed, for instance, from (n, F) -points of function ξ ; then $|e_{n1}| = \varepsilon_n(\xi; F)$. In practice, almost every subset of F consisting of $n+2$ points may serve as the set U_1 . If, however, it turns out that $e_{n1} = 0$, it suffices to find a point $t \in F$ at which $\xi(t) - \omega_{nm}(t) \neq 0$, and replace by it a suitable point of U_1 . It is easy to verify, using Theorems 4:3 and 5:1, that for the set U_1 thus modified, the inequality $e_{n1} \neq 0$ will hold.

II. For a certain number $r \in (0, 1)$ and for every $m > 0$ the set U_{m+1} is an arbitrary set with the following properties:

(Ω_1) the numbers $\xi(u_{m+1,k}) - \omega_{nm}(u_{m+1,k})$ ($k = 0, 1, \dots, n+1$) are positive and negative alternately,

(Ω_2) $|\xi(u_{m+1,k}) - \omega_{nm}(u_{m+1,k})| \geq |e_{nm}|$ for $k = 0, 1, \dots, n+1$,

(Ω_3) there exists a j (where $0 \leq j \leq n+1$) such that

$$(4) \quad |\xi(u_{m+1,j}) - \omega_{nm}(u_{m+1,j})| \geq \|\xi - \omega_{nm}\|_F - r(\|\xi - \omega_{nm}\|_F - |e_{nm}|).$$

THEOREM 15:2. For any function $\xi \in \mathcal{C}_F$, any real number $r \in (0, 1)$ and any set U_1 the sets U_2, U_3, \dots exist.

Proof. We shall prove that the existence of the set U_m (m natural) implies the existence of the set U_{m+1} .

Let us first notice that properties (Ω_{12}) would apply to the set U_{m+1} identical with U_m ; this follows from system (3). If $\|\xi - \omega_{nm}\|_F = |e_{nm}|$, then for $U_{m+1} = U_m$ inequality (4) is satisfied for all $j = 0, 1, \dots, n+1$. It is worth-while to add that in this case the polynomial ω_{nm} (together with polynomials $\omega_{n,m-1}, \dots$) is identical with the required polynomial, and the points of the set U_m are the (n, F) -points of the function ξ .

If, however, $\|\xi - \omega_{nm}\|_F > |e_{nm}|$ (the opposite strict inequality is impossible, as $|e_{nm}| = |\xi(u_{mk}) - \omega_{nm}(u_{mk})| \leq \|\xi - \omega_{nm}\|_F$), then a point $t_{\max} \in F$ such that

$$|\xi(t_{\max}) - \omega_{nm}(t_{\max})| = \|\xi - \omega_{nm}\|_F$$

does not belong to the set U_m . We shall prove that a set U_{m+1} defined by one of the six formulas given below has the properties (Ω_{12a}) .

If $t_{\max} < u_{m0}$ and the values of the difference $\xi - \omega_{nm}$ at the points t_{\max} and u_{m0} have the same signs, then

$$U_{m+1} = \{t_{\max}, u_{m1}, \dots, u_{m,n+1}\}$$

and if these signs are different, then

$$U_{m+1} = \{t_{\max}, u_{m0}, \dots, u_{mn}\}.$$

Similarly, if $t_{\max} > u_{m,n+1}$ and the values of the difference $\xi - \omega_{nm}$ have the same signs at the points t_{\max} and $u_{m,n+1}$, then

$$U_{m+1} = \{u_{m0}, \dots, u_{mn}, t_{\max}\}$$

and if these signs are different, then

$$U_{m+1} = \{u_{m1}, \dots, u_{m,n+1}, t_{\max}\}.$$

Finally, if $u_{mk} < t_{\max} < u_{m,k+1}$ for some k ($0 \leq k \leq n$), then the value of the difference $\xi - \omega_{nm}$ at the point t_{\max} has a sign identical with the value of this difference at the point u_{mk} , or with its value at the point $u_{m,k+1}$. In the first case

$$U_{m+1} = \{u_{m0}, \dots, u_{m,k-1}, t_{\max}, u_{m,k+1}, \dots, u_{m,n+1}\}$$

and in the second case

$$U_{m+1} = \{u_{m0}, \dots, u_{mk}, t_{\max}, u_{m,k+2}, \dots, u_{m,n+1}\}.$$

In each of these six cases the inequality $u_{m+1,0} < u_{m+1,1} < \dots < u_{m+1,n+1}$ holds. The properties (Ω_{12}) follow from system (3), the definition of the point t_{\max} and the method of joining it to the set U_m . The set U_{m+1} also has property (Ω_3) . In fact, for a j such that $u_{m+1,j} = t_{\max}$ we have

$$|\xi(u_{m+1,j}) - \omega_{nm}(u_{m+1,j})| = \|\xi - \omega_{nm}\|_F$$

and this inequality is stronger than inequality (4).

This method of construction of the set U_{m+1} , in which it differs from the set U_m only by one point is not applied in computational practice for the following two reasons. First, usually one cannot determine the

point t_{\max} exactly, but only with a certain error, whose admissible magnitude is given by inequality (4). Secondly, while constructing the set U_{m+1} one usually changes all the points of the set U_m in such a way that the value $|\xi - \omega_{nm}|$ exceeds the number $|e_{nm}|$ at as many points $u_{m+1,k}$ as possible. Then, as follows from the proof of Theorem 15:3, we may expect faster convergence of the sequence $\{\omega_{nm}\}$ to ω_{nF} .

15.3. Now we shall investigate the properties of polynomials ω_{nm} which correspond to the sets U_m .

THEOREM 15:3 (Remez, [30]). *For any function $\xi \in \mathcal{C}_F$, any real number $r \in (0, 1)$ and any sequence of sets $\{U_m\}$ there exist numbers $g > 0$ and $h \in (0, 1)$ such that*

$$(5) \quad \|\xi - \omega_{nm}\|_F - \varepsilon_n(\xi; F) \leq gh^m \quad (m = 1, 2, \dots).$$

(From the definition of the error $\varepsilon_n(\xi; F)$ it follows that $\varepsilon_n(\xi; F) \leq \|\xi - \omega_{nm}\|_F$; thus inequality (5) means the convergence of the sequence $\{\|\xi - \omega_{nm}\|_F\}$ to $\varepsilon_n(\xi; F)$.)

Proof. According to Theorem 4:3

$$|e_{n,m+1}| = \varepsilon_n(\xi; U_{m+1}) = \varepsilon_n(\xi - \omega_{nm}; U_{m+1}).$$

We may apply Theorem 5:1 to the function $\xi - \omega_{nm}$ and the set $F^* = U_{m+1}$. We shall use the fact that, according to property (Ω_1) , the values of the function $\xi - \omega_{nm}$ on the set F^* are alternately positive and negative. Hence

$$|e_{n,m+1}| = \frac{\sum_{k=0}^{n+1} |\xi(u_{m+1,k}) - \omega_{nm}(u_{m+1,k})|/w_{mk}}{\sum_{k=0}^{n+1} 1/w_{mk}}$$

where

$$w_{mk} = \prod_{i=0, i \neq k}^{n+1} |u_{m+1,k} - u_{m+1,i}|.$$

The number $|e_{n,m+1}|$ is the weighted mean of the numbers $|\xi(u_{m+1,k}) - \omega_{nm}(u_{m+1,k})|$ with the positive weights $1/w_{mk}$. Thus, under the notation

$$p_m = \min_{0 \leq k \leq n+1} |\xi(u_{m+1,k}) - \omega_{nm}(u_{m+1,k})|,$$

$$q_m = \max_{0 \leq k \leq n+1} |\xi(u_{m+1,k}) - \omega_{nm}(u_{m+1,k})|$$

we get the inequality

$$p_m \leq |e_{n,m+1}| \leq q_m \quad (m = 1, 2, \dots).$$

At the same time, property (Ω_2) implies that

$$(6) \quad p_m \geq |e_{nm}|.$$

Thus, $0 < |e_{n1}| \leq \dots \leq |e_{nm}| \leq |e_{n,m+1}| \leq \dots$ and the assumptions of Theorem 15:1 are satisfied. It follows from this theorem that the differences $u_{m,k+1} - u_{mk}$ may be estimated from below by a positive number, common for all $k = 0, 1, \dots, n$ and all natural m . On the other hand, $|u_{mk} - u_{ml}| \leq b - a$. Thus there exist positive numbers w_{\min} and w_{\max} such that

$$w_{\min} \leq w_{mk} \leq w_{\max} \quad (k = 0, 1, \dots, n+1; m = 1, 2, \dots).$$

Now we may estimate from below the difference

$$|e_{n,m+1}| - p_m = \frac{\sum_{k=0}^{n+1} (|\xi(u_{m+1,k}) - \omega_{nm}(u_{m+1,k})| - p_m) / w_{mk}}{\sum_{k=0}^{n+1} 1 / w_{mk}}.$$

According to the definition of number q_m there exists an i such that $q_m = |\xi(u_{m+1,i}) - \omega_{nm}(u_{m+1,i})|$. At the same time $|\xi(u_{m+1,k}) - \omega_{nm}(u_{m+1,k})| \geq p_m$ for $k = 0, 1, \dots, n+1$. Thus in the numerator of the expression equal to $|e_{n,m+1}| - p_m$, all the terms of the sum for $k \neq i$ may be estimated from below by 0, and for $k = i$ by $(q_m - p_m) / w_{mi}$:

$$(7) \quad |e_{n,m+1}| - p_m \geq \frac{(q_m - p_m) / w_{mi}}{\sum_{k=0}^{n+1} 1 / w_{mk}} \geq \frac{w_{\min}(q_m - p_m)}{(n+2)w_{\max}}.$$

It follows that

$$q_m \leq \frac{(n+2)w_{\max}}{w_{\min}} |e_{n,m+1}| - \left(\frac{(n+2)w_{\max}}{w_{\min}} - 1 \right) p_m.$$

In view of (6) and inequality $|e_{n,m+1}| = \varepsilon_n(\xi; U_{m+1}) \leq \varepsilon_n(\xi; F)$ we have also

$$q_m \leq \frac{(n+2)w_{\max}}{w_{\min}} \varepsilon_n(\xi; F) - \left(\frac{(n+2)w_{\max}}{w_{\min}} - 1 \right) |e_{nm}|.$$

From property (Ω_3) and the definition of the number q_m it follows that

$$(8) \quad q_m \geq (1-r) \|\xi - \omega_{nm}\|_F + r |e_{nm}|,$$

which can be also written as

$$\|\xi - \omega_{nm}\|_F - \varepsilon_n(\xi; F) \leq \frac{q_m}{1-r} - \frac{r}{1-r} |e_{nm}| - \varepsilon_n(\xi; F).$$

We apply the estimates for q_m previously obtained:

$$\begin{aligned}
 (9) \quad & \|\xi - \omega_{nm}\|_F - \varepsilon_n(\xi; F) \\
 & \leq \frac{(n+2)w_{\max}}{(1-r)w_{\min}} \varepsilon_n(\xi; F) - \left(\frac{(n+2)w_{\max}}{(1-r)w_{\min}} - \frac{1}{1-r} \right) |e_{nm}| - \frac{r}{1-r} |e_{nm}| - \varepsilon_n(\xi; F) \\
 & = \left(\frac{(n+2)w_{\max}}{(1-r)w_{\min}} - 1 \right) (\varepsilon_n(\xi; F) - |e_{nm}|).
 \end{aligned}$$

Now we shall estimate from above the difference $\varepsilon_n(\xi; F) - |e_{nm}|$. If we replace in (7) the number q_m by the right-hand side of inequality (8), we get

$$|e_{n,m+1}| \geq \frac{w_{\min}}{(n+2)w_{\max}} \cdot ((1-r)\varepsilon_n(\xi; F) + r|e_{nm}|) + \left(1 - \frac{w_{\min}}{(n+2)w_{\max}} \right) p_m.$$

According to (6) we can replace p_m by $|e_{nm}|$:

$$\begin{aligned}
 |e_{n,m+1}| & \geq \frac{(1-r)w_{\min}}{(n+2)w_{\max}} \varepsilon_n(\xi; F) + \left(1 - \frac{(1-r)w_{\min}}{(n+2)w_{\max}} \right) |e_{nm}|, \\
 \varepsilon_n(\xi; F) - |e_{n,m+1}| & \leq h (\varepsilon_n(\xi; F) - |e_{nm}|)
 \end{aligned}$$

where

$$h = 1 - \frac{(1-r)w_{\min}}{(n+2)w_{\max}}.$$

Of course $h \in (0, 1)$. Applying analogous inequalities, we obtain

$$\varepsilon_n(\xi; F) - |e_{nm}| \leq h (\varepsilon_n(\xi; F) - |e_{n,m-1}|) \leq \dots \leq h^{m-1} (\varepsilon_n(\xi; F) - |e_{n1}|).$$

From the last inequality and from (9) it follows that

$$\|\xi - \omega_{nm}\|_F - \varepsilon_n(\xi; F) \leq \left(\frac{1}{1-h} - 1 \right) h^{m-1} (\varepsilon_n(\xi; F) - |e_{n1}|) = gh^m$$

where

$$g = \frac{\varepsilon_n(\xi; F) - |e_{n1}|}{1-h}.$$

Thus we have proved inequality (5). This inequality shows that the error of the approximation of the function ξ by polynomial ω_{nm} exceeds the error of the best approximation by at most gh^m , which tends to zero as $m \rightarrow \infty$. In the computational practice this is sufficient. We can, however, prove more, namely the convergence of the sequence $\{\omega_{nm}\}$ to ω_{nF} . This convergence will follow from the theorem of Vallée Poussin ([34], p. 87), which will be presented here in a slightly stronger form.

THEOREM 15:4. *For any function $\xi \in \mathcal{C}_F$ there exists a positive number x such that for any polynomial $\omega \in \mathcal{W}_n$ we have*

$$\|\omega - \omega_{nF}\|_F \leq x(\|\xi - \omega\|_F - \varepsilon_n(\xi; F)).$$

Proof. Let the points u_0, u_1, \dots, u_{n+1} , where $u_0 < u_1 < \dots < u_{n+1}$, be alternately the $(n, +, F)$ -points and $(n, -, F)$ -points of the function ξ . The best polynomial ω_{nF} and the number e_n satisfy the system of equations

$$(10) \quad \xi(u_k) - \omega_{nF}(u_k) = (-1)^k e_n \quad (k = 0, 1, \dots, n+1).$$

Besides

$$e_n = \frac{\sum_{l=0}^{n+1} (-1)^l \xi(u_l)/w_l}{\sum_{l=0}^{n+1} 1/w_l} \quad (w_l > 0)$$

(§ 5, (2)) and $\varepsilon_n(\xi; F) = s e_n$, where $s = 1$ or $s = -1$.

Let ω be an arbitrary polynomial of class \mathcal{W}_n . From the definition of the norm of a function it follows that there exist numbers r_k ($k = 0, 1, \dots, n+1$) from the interval $\langle -1, 1 \rangle$ such that

$$(11) \quad \xi(u_k) - \omega(u_k) = (-1)^k r_k s \|\xi - \omega\|_F \quad (k = 0, 1, \dots, n+1).$$

Regarding the coefficients of the polynomial ω and the number $\|\xi - \omega\|_F$ as unknowns in this system of equations we may derive the formula

$$\|\xi - \omega\|_F = \frac{s \sum_{l=0}^{n+1} (-1)^l \xi(u_l)/w_l}{\sum_{l=0}^{n+1} r_l/w_l}$$

in the same way as in § 5 we obtained the formula for e_n from system (10).

There exists a number $f \geq 0$ such that $\|\xi - \omega\|_F = (1+f)\varepsilon_n(\xi; F) = (1+f)se_n$, and comparing formulas for $\|\xi - \omega\|_F$ and e_n we get

$$\frac{1}{\sum_{l=0}^{n+1} r_l/w_l} = \frac{1+f}{\sum_{l=0}^{n+1} 1/w_l}, \quad \sum_{l=0}^{n+1} (1-r_l)/w_l = f \sum_{l=0}^{n+1} r_l/w_l.$$

Since $w_k > 0$ and $r_k \leq 1$ for $k = 0, 1, \dots, n+1$, we have

$$1 - r_k \leq f w_k \sum_{l=0}^{n+1} r_l/w_l \leq f \max_j \sum_{l=0}^{n+1} w_j/w_l.$$

It follows from systems (10) and (11) that

$$|\omega(u_k) - \omega_{nF}(u_k)| = |\varepsilon_n(\xi; F) - r_k \|\xi - \omega\|_F|.$$

Using the equation $\|\xi - \omega\|_F = (1 + f)\varepsilon_n(\xi; F)$ we easily see that

$$\begin{aligned} \varepsilon_n(\xi; F) - r_k \|\xi - \omega\|_F &\geq \varepsilon_n(\xi; F) - \|\xi - \omega\|_F, \\ \varepsilon_n(\xi; F) - r_k \|\xi - \omega\|_F &= (1 - r_k - fr_k) \varepsilon_n(\xi; F) \\ &\leq f \left(1 + \max_j \sum_{l=0}^{n+1} w_j/w_l\right) \varepsilon_n(\xi; F) \\ &= \left(1 + \max_j \sum_{l=0}^{n+1} w_j/w_l\right) (\|\xi - \omega\|_F - \varepsilon_n(\xi; F)) \\ &\quad (k = 0, 1, \dots, n+1), \end{aligned}$$

$$\max_{0 \leq k \leq n+1} |\omega(u_k) - \omega_{nF}(u_k)| \leq \left(1 + \max_j \sum_{l=0}^{n+1} w_j/w_l\right) (\|\xi - \omega\|_F - \varepsilon_n(\xi; F)).$$

If the values of two polynomials of an degree not exceeding n are close to one another at $n+2$ points, then they are close also at any point of a fixed closed set. More precisely, from the above inequalities follows the existence of a constant $x > 0$ (which may depend upon n and the points u_0, u_1, \dots, u_{n+1} , and indirectly upon the function ξ), such that

$$\|\omega - \omega_{nF}\|_F \leq x (\|\xi - \omega\|_F - \varepsilon_n(\xi; F)),$$

which was to be proved.

The following theorem is a direct consequence of the last two theorems:

THEOREM 15:5. *For any function $\xi \in \mathcal{C}_F$ and any sequence of sets $\{U_m\}$ there exist numbers $f > 0$ and $h \in (0, 1)$ such that*

$$\|\omega_{nm} - \omega_{nF}\|_F \leq fh^m \quad (m = 1, 2, \dots).$$

There are known even stronger results which concern the convergence of the sequence of norms $\|\omega_{nm} - \omega_{nF}\|_F$ to zero. If the set F coincides with the interval $\langle a, b \rangle$, the function ξ is sufficiently regular (it has, among other properties, a continuous second derivative in the interval (a, b)), and if we replace property (Ω_3) by the requirement that

$$(12) \quad \xi'(u_{m+1,k}) - \omega'_{nm}(u_{m+1,k}) = 0$$

for all k such that $u_{m+1,k} \in (a, b)$, then there exist positive numbers $h_0 < 1$ and f_0 such that

$$\|\omega_{nm} - \omega_{nF}\|_F \leq f_0 h_0^{2^m}$$

(Veidinger, [35]). This shows the unquestionable similarity between the method of Remez and the method of Newton for the solution of non-linear equations. It seems, however, that the operational importance of

the Veidinger inequality for computational practice is small, as we cannot, in general, achieve the exact fulfilment of equation (12).

15.4. Let $F = A$, where A is a finite set, and let the table of values of function ξ on this set be given. Comparing the finite number of values $|\xi - \omega_{nm}|$ we easily find the point at which this function assumes its maximal value on the set A . Thus we may require that for every $m = 1, 2, \dots$ the set U_{m+1} should have property (Ω_3) in its strongest form, i.e. we may require the existence of a point $u_{m+1,j}$ such that

$$|\xi(u_{m+1,j}) - \omega_{nm}(u_{m+1,j})| = \|\xi - \omega_{nm}\|_A.$$

The case where F is a finite set A has also a second, even more important feature: there exists an m such that $\omega_{nA} = \omega_{nm}$. This follows from the fact that the set A has a finite number of subsets consisting of $n+2$ points each, whence the sequence $\omega_{n1}, \omega_{n2}, \dots$ of the best polynomials on U_1, U_2, \dots , which converges to ω_{nA} , contains only a finite number of distinct polynomials.

For a finite set A , in view of the simplicity of this case, we may make a completely universal routine, which would serve for computing the polynomial ω_{nA} for any function ξ on an digital computer. If the set F is an interval $I = \langle a, b \rangle$ we may suggest a procedure which reduces this case, in a certain sense, to the previous case. We choose a finite set A which covers the interval I sufficiently densely. One may expect (and the next theorem will justify this expectation under certain assumptions concerning the function ξ) that in this case the best n -th polynomial ω_{nI} is approximately equal to the n -th best polynomial ω_{nA} . It is more convenient to use the approximate equation $\omega_{nI} \approx \omega_{nA}$, than the approximate equation $\omega_{nI} \approx \omega_{nm}$, where $\omega_{n1}, \omega_{n2}, \dots, \omega_{nm}$ are the polynomials found by the method of Remez for $F = I$. Of course, the polynomial ω_{nA} will also be found by the method of Remez, but this—as we have already established—is relatively simple.

THEOREM 15:6 (Burov, [7]). *If $A = \{a_0, a_1, \dots, a_p\}$ where $a = a_0 < a_1 < \dots < a_p = b$ and if $\xi \in \mathcal{C}_I^2$, then*

$$\|\omega_{nA} - \omega_{nI}\|_I = O(r^2)$$

where $r = \max_{0 \leq l \leq p-1} (a_{l+1} - a_l)$.

Proof. Let $\delta = \xi - \omega_{nA}$ and let φ_l be a polynomial of class \mathcal{W}_1 satisfying the conditions $\varphi_l(a_l) = \delta(a_l)$, $\varphi_l(a_{l+1}) = \delta(a_{l+1})$. Thus we have $|\varphi_l(a_l)| = |\xi(a_l) - \omega_{nA}(a_l)| \leq \varepsilon_n(\xi; A)$, $|\varphi_l(a_{l+1})| \leq \varepsilon_n(\xi; A)$ and $\|\varphi_l\|_{\langle a_l, a_{l+1} \rangle} \leq \varepsilon_n(\xi; A)$. From the well-known estimate of the remainder of the inter-

polution formula it follows that for $t \in \langle a_l, a_{l+1} \rangle$:

$$\begin{aligned} |\delta(t) - \varphi_l(t)| &\leq \frac{1}{2} |(t - a_l)(t - a_{l+1})| \cdot \|\delta''\|_{\langle a_l, a_{l+1} \rangle} \\ &\leq \frac{1}{8} (a_{l+1} - a_l)^2 \|\delta''\|_{\langle a_l, a_{l+1} \rangle}, \\ \|\delta - \varphi_l\|_{\langle a_l, a_{l+1} \rangle} &\leq \frac{1}{8} r^2 \|\delta''\|_I \leq \frac{1}{8} r^2 (\|\xi''\|_I + \|\omega''_{nA}\|_I), \\ \|\delta\|_{\langle a_l, a_{l+1} \rangle} &\leq \varepsilon_n(\xi; A) + \frac{1}{8} r^2 (\|\xi''\|_I + \|\omega''_{nA}\|_I). \end{aligned}$$

Let us now estimate $\|\omega''_{nA}\|_I$. For every $t \in I$ we denote by $a_{l(t)}$ the point of the set A which is at a distance of at most $r/2$ from t (such a point exists by definition of r). Then

$$\begin{aligned} |\omega_{nA}(t) - \omega_{nA}(a_{l(t)})| &\leq |t - a_{l(t)}| \cdot \|\omega'_{nA}\|_I, \\ \|\omega_{nA}\|_I &\leq \|\omega_{nA}\|_A + \frac{1}{2} r \|\omega'_{nA}\|_I. \end{aligned}$$

Markov has proved (see for instance [20], p. 178) the inequality

$$\|\omega'\|_I \leq \frac{2n^2 \|\omega\|_I}{b - a}$$

valid for $\omega \in \mathcal{W}_n$. Hence we have

$$\|\omega'_{nA}\|_I \leq \frac{2n^2 \|\omega_{nA}\|_A + n^2 r \|\omega'_{nA}\|_I}{b - a}$$

and if $r \leq (b - a)/2n^2$, then

$$\begin{aligned} \|\omega'_{nA}\|_I &\leq \frac{2n^2}{b - a} \|\omega_{nA}\|_A + \frac{1}{2} \|\omega'_{nA}\|_I, \\ \|\omega'_{nA}\|_I &\leq \frac{4n^2}{b - a} \|\omega_{nA}\|_A \leq \frac{4n^2}{b - a} (\|\xi - \omega_{nA}\|_A + \|\xi\|_A) \\ &\leq \frac{4n^2}{b - a} (\varepsilon_n(\xi; I) + \|\xi\|_I). \end{aligned}$$

Applying the Markov inequality once more, this time to the polynomial $\omega = \omega'_{nA}$, we get

$$\|\omega''_{nA}\|_I \leq \frac{8n^4}{(b - a)^2} (\varepsilon_n(\xi; I) + \|\xi\|_I),$$

$$\begin{aligned} \|\xi - \omega_{nA}\|_I &= \|\delta\|_I = \max_l \|\delta\|_{\langle a_l, a_{l+1} \rangle} \\ &\leq \varepsilon_n(\xi; A) + \frac{1}{8} r^2 \left(\|\xi''\|_I + \frac{8n^4}{(b - a)^2} (\varepsilon_n(\xi; I) + \|\xi\|_I) \right). \end{aligned}$$

Since

$$(13) \quad \varepsilon_n(\xi; A) \leq \varepsilon_n(\xi; I) \leq \|\xi - \omega_{nA}\|_I,$$

we have proved that $\|\xi - \omega_{nA}\|_I - \varepsilon_n(\xi; I) = O(r^2)$. The assertion of the theorem which we are now proving follows from the last inequality, and from Theorem 15:4 for $F = I$, and $\omega = \omega_{nA}$.

Theorem 15:6 gives only a general view of the relation between the norm $\|\omega_{nA} - \omega_{nI}\|_I$ and the distribution of points of the set A in the interval I . In practice we measure the deviation of the polynomial ω_{nA} from the polynomial ω_{nI} as follows: after computing the coefficients of the polynomial ω_{nA} we estimate from above the norm $\|\xi - \omega_{nA}\|_I$, for instance by using Theorem 15:7. The two-sided estimate (13) for the n -th error of the best approximation of function ξ in the interval I allows us to accept as a certain (relative) measure of the above-mentioned deviation the ratio $(\|\xi - \omega_{nA}\|_I - \varepsilon_n(\xi; A)) / \varepsilon_n(\xi; A)$.

THEOREM 15:7. *If $b_1 - b_0 = b_2 - b_1 = \dots = b_m - b_{m-1} = r > 0$ and $\delta \in \mathcal{C}_{\langle b_0, b_m \rangle}^{m+1}$, then for $j = 0, 1, \dots, m-1$ and $t \in \langle b_j, b_{j+1} \rangle$ we have the inequality*

$$(14) \quad \delta(t) \leq \max_{0 \leq l \leq m} \delta(b_l) + \sum_{k \in K_j} p_{mjk} (\max_{0 \leq l \leq m} \delta(b_l) - \delta(b_k)) + p_{mj} r^{m+1} d_j^{m+1}$$

where

$$K_j = \{k: 0 \leq k < j \text{ and } k-j \text{ is odd}$$

$$\text{or } j < k \leq m \text{ and } k-j \text{ is even}\},$$

$$p_{mik} = \frac{1}{k!(m-k)!} \left\| \prod_{i=0, i \neq k}^m (j+h-i) \right\|_{\langle 0,1 \rangle} \quad (k \in K_j, j = 0, 1, \dots, m-1),$$

$$p_{mj} = \frac{1}{(m+1)!} \left\| \prod_{i=0}^m (j+h-i) \right\|_{\langle 0,1 \rangle} \quad (j = 0, 1, \dots, m-1),$$

$$d_j^{m+1} = \max \{0, \max_{t \in \langle b_0, b_m \rangle} (-1)^{m-j} \delta^{(m+1)}(t)\}.$$

In the formulas for p_{mjk} and p_{mj} the variable with respect to which we compute the norm is h .

We apply Theorem 15:7 to the function $\delta = \xi - \omega_{nA}$ or to the function $\delta = \omega_{nA} - \xi$ in order to estimate its values between those successive points b_0, b_1, \dots, b_m of the set A at which this function is close to $\varepsilon_n(\xi; A)$. It is most convenient to assume that $m = n$, since this is the smallest value of m for which the $(m+1)$ -st derivative of the function δ does not depend upon the coefficients of ω_{nA} .

Proof. Let $\bar{d} = \max_{0 \leq l \leq m} \delta(b_l)$. We use the Lagrange interpolation formula for the function $\delta - \bar{d}$ and for the points b_0, b_1, \dots, b_m , obtaining

$$\delta(t) - \bar{d} = \sum_{k=0}^m \frac{\prod_{i=0, i \neq k}^m (b_i - t)}{\prod_{i=0, i \neq k}^m (b_i - b_k)} (\delta(b_k) - \bar{d}) + \frac{\delta^{(m+1)}(s)}{(m+1)!} \prod_{i=0}^m (t - b_i)$$

where $s, t \in \langle b_0, b_m \rangle$. Using the fact that the distances between the successive points b_0, b_1, \dots, b_m are equal, we introduce the new variable h , by formula $t = b_j + h(b_{j+1} - b_j) = b_0 + (j+h)r$. Then we obtain

$$\delta(t) = \bar{d} + \sum_{k=0}^m \left(- \frac{\prod_{i=0, i \neq k}^m (i - j - h)}{\prod_{i=0, i \neq k}^m (i - k)} \right) (\bar{d} - \delta(b_k)) + \frac{r^{m+1} \delta^{(m+1)}(s)}{(m+1)!} \prod_{i=0}^m (j + h - i).$$

The first term on the right-hand side of the above equation is equal to the first term on the right-hand side of inequality (14). We shall prove that the second and third terms of the above equation may be estimated from above for $t \in (b_j, b_{j+1})$, i.e. for $h \in (0, 1)$ by the second and third term of the right-hand side of inequality (14), respectively.

We have

$$\prod_{i=0, i \neq k}^m (i - j - h) \begin{cases} > 0 & (k \leq j \text{ and } j \text{ even or } k > j \text{ and } j \text{ odd}), \\ < 0 & (k \leq j \text{ and } j \text{ odd or } k > j \text{ and } j \text{ even}), \end{cases}$$

$$\prod_{i=0, i \neq k}^m (i - k) \begin{cases} > 0 & (k \text{ even}), \\ < 0 & (k \text{ odd}). \end{cases}$$

The ratio of the left-hand sides of these inequalities is negative if and only if $k \leq j$ and $k - j$ is odd or $k > j$ and $k - j$ is even, i.e. if $k \in K_j$. Besides it is known that

$$\left| \prod_{i=0, i \neq k}^m (i - k) \right| = k!(m - k)!$$

and $\bar{d} - \delta(b_k) \geq 0$; hence for $k \in K_j$,

$$- \frac{\prod_{i=0, i \neq k}^m (i - j - h)}{\prod_{i=0, i \neq k}^m (i - k)} (\bar{d} - \delta(b_k))$$

$$\leq \frac{\left\| \prod_{i=0, i \neq k}^m (i - j - h) \right\|_{\langle 0, 1 \rangle}}{k!(m - k)!} (\bar{d} - \delta(b_k)) = p_{mj k} (\bar{d} - \delta(b_k)).$$

For the remaining values of k analogous expressions may be estimated from above by 0, which is the best estimate.

It remains to estimate from above the expression

$$(15) \quad \delta^{(m+1)}(s) \prod_{i=0}^m (j+h-i) = (-1)^{m-j} \delta^{(m+1)}(s) \prod_{i=0}^m |j+h-i|.$$

If $\max_{s \in \langle b_0, b_m \rangle} (-1)^{m-j} \delta^{(m+1)}(s) < 0$, then we estimate expression (15) from

above by 0, since the product $\prod_{i=0}^m |j+h-i|$ tends to zero for $h \rightarrow 0$. If this maximum is non-negative, then the expression (15) can be estimated from above by the number

$$\left\| \prod_{i=0}^m (j+h-i) \right\|_{\langle 0,1 \rangle} \max_{s \in \langle b_0, b_m \rangle} (-1)^{m-j} \delta^{(m+1)}(s).$$

Introducing the notation \bar{d}_j^{m+1} we get the third term of the right-hand side of inequality (14).

16. Other methods of computing the best polynomials. In this section we shall describe some modifications of the methods of Remez and certain complements to it; besides we shall give a brief description of one other method of computing the best polynomials. We shall, in general, assume that $\xi \in \mathcal{E}_I$, where $I = \langle -1, 1 \rangle$, and that we look for the best polynomial on the interval I .

16.1. We have already mentioned several times, recently in § 13.4, that for the approximation of the function in the interval I an important part is played by the subset

$$(1) \quad \{t_{n+1,0}, t_{n+1,1}, \dots, t_{n+1,n+1}\}$$

of this interval, where $t_{n+1,k} = -\cos k\pi/(n+1)$ ($k = 0, 1, \dots, n+1$). In § 5.3 we have given the formula for the n -th error of the best approximation of function ξ on the set (1); in § 13.4 we have given the formulas for the n -th best polynomial for function ξ on this set. Theorems of §§ 9 and 12 allow us to believe that for a sufficiently regular function ξ this polynomial is close to the polynomial ω_{nI} . Thus, it is recommended to choose as the set U_1 in the method of Remez the set (1):

$$(2) \quad u_{1k} = t_{n+1,k} \quad (k = 0, 1, \dots, n+1).$$

One should check, however, whether this definition is correct, i.e. whether the number e_{n1} which satisfies together with the coefficients of polynomial ω_{n1} the system of equations

$$(3) \quad \xi(u_{1k}) - \omega_{n1}(u_{1k}) = (-1)^k e_{n1} \quad (k = 0, 1, \dots, n+1)$$

is different from zero. We may do it applying the formula from § 5.3:

$$e_{n1} = \frac{1}{n+1} \left(\frac{1}{2} \xi(t_{n+1,0}) + \sum_{k=1}^n (-1)^k \xi(t_{n+1,k}) + \frac{(-1)^{n+1}}{2} \xi(t_{n+1,n+1}) \right)$$

or other formulas from §§ 5.3 and 5.4 which were obtained as its transformations.

If we approximate the function ξ not in the interval I but on a finite subset A of that interval (the importance of this case has been pointed out in § 15.4), then the set U_1 may be formed from the points of the set A which lie close to the points $t_{n+1,k}$.

Let us now go back once more to approximation in the interval I . For the case where the set U_1 is defined by formulas (2), Hornecker in [14] and [15] has given a method of approximate determination of the set U_2 which does not require the investigation of the difference $\xi - \omega_{n1}$. This method is correct under rather strong assumptions concerning the function ξ . If these assumptions are satisfied, the construction of the sequence of polynomials $\{\omega_{nm}\}$ starts from the polynomial ω_{n2} (and it may end with it, since this polynomial is often sufficiently close to the required polynomial ω_{nI}).

We assume about the function ξ that 1. the series $\sum_{l=0}^{\infty} a_l[\xi] \tau_l$ (see § 13.2) converges to it in the whole interval I , 2. $a_{n+1}[\xi] \neq 0$, 3. a positive number m such that $a_{n+2}[\xi] = \dots = a_{n+m-1}[\xi] = 0$, $a_{n+m}[\xi] \neq 0$ is smaller than $n+2$, 4. the coefficient $|a_{n+m}[\xi]|$ is considerably smaller than $|a_{n+1}[\xi]|$ and the coefficients $|a_l[\xi]|$ for $l > n+m$ are considerably smaller than $|a_{n+m}[\xi]|$. This last assumption will not be stated more precisely.

Under assumption 3. it follows from Theorem 13:9 that the polynomial ω_{n1} defined by (3) and (2) equals

$$\omega_{n1} = \sum_{k=0}^n a_k[\xi] \tau_k + a_{n+m}[\xi] \tau_{n-m+2} + \varrho$$

where ϱ is the sum of the terms which contain only the coefficients $a_{n+m+1}[\xi], a_{n+m+2}[\xi], \dots$, whence which are considerably smaller as compared with other terms. Thus

$$\xi - \omega_{n1} = a_{n+1}[\xi] \tau_{n+1} + a_{n+m}[\xi] (\tau_{n+1} - \tau_{n-m+2}) + \varrho^*$$

where ϱ^* has the same property as the function ϱ .

The term $a_{n+1}[\xi] \tau_{n+1}$ of the difference $\xi - \omega_{n1}$ assumes its extremal values equal alternately to $a_{n+1}[\xi]$ and to $-a_{n+1}[\xi]$ at the points $-1 = t_{n+1,0}, t_{n+1,1}, \dots, t_{n+1,n+1} = 1$. According to assumption 4. the re-

maining two terms of the difference $\xi - \omega_{n+1}$ have little influence upon it. Thus we may state that this difference has its extremal values at the points $u_{20}, u_{21}, \dots, u_{2, n+1}$ such that $u_{20} = t_{n+1,0} = -1$, $u_{2, n+1} = t_{n+1, n+1} = 1$ (at these two points the derivative of the polynomial τ_{n+1} is different from zero), and the points u_{21}, \dots, u_{2n} lie close to the points $t_{n+1,1}, \dots, t_{n+1,n}$. Thus the differences $d_k = u_{2k} - t_{n+1,k}$ are close to zero. We shall modify slightly the definition of the points $u_{2k} = t_{n+1,k} + d_k$. At first, we replace the difference $\xi' - \omega'_{n+1}$ by the sum $a_{n+1}[\xi] \tau'_{n+1} + a_{n+m}[\xi] (\tau'_{n+m} - \tau'_{n-m+2})$, which is close to it. Secondly, we expand this sum into a Taylor series in the neighbourhood of the point $t_{n+1,k}$, retaining two terms of the expansion in the polynomial τ'_{n+1} , and only one term in the polynomials τ'_{n+m} and τ'_{n-m+2} . Thus we shall get the equation

$$a_{n+1}[\xi] (\tau'_{n+1}(t_{n+1,k}) + d_k \tau''_{n+1}(t_{n+1,k})) + a_{n+m}[\xi] (\tau'_{n+m}(t_{n+1,k}) - \tau'_{n-m+2}(t_{n+1,k})) = 0.$$

Considering the fact that

$$(4) \quad \tau'_{n+1}(t_{n+1,k}) = 0 \quad (k = 1, 2, \dots, n)$$

we transform this equation into the form

$$u_{2k} = t_{n+1,k} + \frac{a_{n+m}[\xi]}{a_{n+1}[\xi]} \cdot \frac{\tau'_{n-m+2}(t_{n+1,k}) - \tau'_{n+m}(t_{n+1,k})}{\tau''_{n+1}(t_{n+1,k})} \quad (k = 1, 2, \dots, n).$$

Since

$$\cos(n-m+2)u - \cos(n+m)u = 2\sin(m-1)u \sin(n+1)u$$

we have

$$\tau_{n-m+2} - \tau_{n+m} = 2(\sin(m-1) \operatorname{arc} \cos t)(\sin(n+1) \operatorname{arc} \cos t).$$

At the same time we have for every positive integer p :

$$\tau'_p = \frac{p}{\sqrt{1-t^2}} \sin p \operatorname{arc} \cos t,$$

whence

$$\tau_{n-m+2} - \tau_{n+m} = \frac{2(1-t^2)}{(m-1)(n+1)} \tau'_{m-1} \tau'_{n+1}.$$

Differentiating both sides of this equation and considering (4) we obtain

$$\tau'_{n-m+2}(t_{n+1,k}) - \tau'_{n+m}(t_{n+1,k}) = \frac{2(1-t_{n+1,k}^2) \tau'_{m-1}(t_{n+1,k}) \tau'_{n+1}(t_{n+1,k})}{(m-1)(n+1)},$$

$$u_{2k} = t_{n+1,k} + \frac{a_{n+m}[\xi]}{a_{n+1}[\xi]} \cdot \frac{2(1-t_{n+1,k}^2) \tau'_{m-1}(t_{n+1,k})}{(m-1)(n+1)} \quad (k = 1, 2, \dots, n).$$

Most frequently of course $m = 2$ or $m = 3$. The last case holds, for instance, if the number $n+1$ and the function ξ are both odd or both even. Since $\tau_1(t) = t$ and $\tau_2(t) = 2t^2 - 1$ we have for $m = 2$:

$$u_{2k} = t_{n+1,k} + \frac{a_{n+2}[\xi]}{a_{n+1}[\xi]} \cdot \frac{2(1-t_{n+1,k}^2)}{n+1} \quad (k = 1, 2, \dots, n)$$

and for $m = 3$:

$$u_{2k} = t_{n+1,k} + \frac{a_{n+3}[\xi]}{a_{n+1}[\xi]} \cdot \frac{4(1-t_{n+1,k}^2)t_{n+1,k}}{n+1} \quad (k = 1, 2, \dots, n).$$

We check the usefulness of these formulas for every given function ξ by finding the polynomial ω_{n2} and the number e_{n2} and investigating the difference $\xi - \omega_{n2}$.

A further extension of this method is given in [15]. There are given approximate formulas which express the polynomial ω_{n2} (the best for the set $\{u_{20}, u_{21}, \dots, u_{2,n+1}\}$) by coefficients $a_1[\xi]$.

16.2. Let U_m be the $(n+2)$ -points subset of the interval I obtained for the function ξ by the method of Remez, or by its possible modification. The symbol ω_{nm} will denote, as before, the n -th best polynomial for the function ξ on the set U_m . At successive points of this set the difference $\xi - \omega_{nm}$ is positive and negative, alternately.

Let us assume that the function ξ has a continuous second derivative in the interval I , and that for every polynomial $\omega \in \mathcal{W}_n$ the first derivative of the difference $\xi - \omega$ has at most n zeros in this interval. Thus, there exist exactly $n+2$ points $u_{m+1,0} < u_{m+1,1} < \dots < u_{m+1,n+1}$ at which the difference $\xi - \omega_{nm}$ assumes alternately the local maxima and minima. The points $u_{m+1,1}, \dots, u_{m+1,n}$ lie inside the interval I and are equal to the roots of the equation

$$(5) \quad \xi'(t) - \omega'_{nm}(t) = 0.$$

The extremely points coincide with the end points of the interval I : $u_{m+1,0} = -1$, $u_{m+1,n+1} = 1$. It is easy to see that the above-mentioned $n+2$ points form a set U_{m+1} with the properties (Ω_{123}) from § 15.2, and we may put $r = 0$ in (Ω_3) .

Equation (5) may, in general, be solved only by approximate methods. In particular, if we take the number u_{mk} as the first approximation of the root $u_{m+1,k}$ of equation (5), then one iteration by Newton's method will give the result

$$u_{m+1,k} \approx u_{mk} - \frac{\xi'(u_{mk}) - \omega'_{nm}(u_{mk})}{\xi''(u_{mk}) - \omega''_{nm}(u_{mk})} \quad (k = 1, 2, \dots, n).$$

One can often accept this approximate definition of the set U_{m+1} . Then, provided the choice of the set U_1 was successful, the sequence $\{\omega_{nm}\}$ should converge to the n -th best polynomial ω_{nI} .

16.3. Some methods of iterational computing of best polynomials are known which are essentially different from the method of Remez. Most of them are connected with a more general problem — namely of solving (in a suitable sense) inconsistent systems of algebraic linear equations. I shall not present these methods. I shall briefly sketch the method of Bratton [6], whose generalizations and some new modifications have been given in [12] and [11]. Curtiss and Frank in [12] admit that the method of Remez leads to shorter computations than the method of Bratton nevertheless the latter method being very natural, also deserves mentioning.

Bratton gives a method of constructing a sequence of $(n + 2)$ -point subsets U_m of the interval I such that

$$\lim_{n \rightarrow \infty} \varepsilon_n(\xi; U_m) = \varepsilon_n(\xi; I).$$

By Theorem 15:4 (slightly modified) this equation means that the sequence of n -th best polynomials on the sets U_m converges to the polynomial ω_{nI} , which is the best polynomial for the function ξ on the interval I . We shall see, however that the computation of all the polynomials of the above-mentioned sequence is not necessary.

Now we shall give the definition of the sets U_2, U_3, \dots, U_{n+3} for a fixed set U_1 such that $\varepsilon_n(\xi; U_1) > 0$. In a similar way, for the set U_{n+3} already found we define the next $n + 2$ sets $U_{n+4}, U_{n+5}, \dots, U_{2n+5}$ etc. We introduce the auxiliary function for the set U_1 :

$$\delta_1(u) = \begin{cases} \varepsilon_n(\xi; \{u, u_{11}, \dots, u_{1,n+1}\}) & (-1 \leq u < u_{11}) \\ \varepsilon_n(\xi; \{u_{11}, \dots, u_{1,n+1}, u\}) & (u_{1,n+1} < u \leq 1). \end{cases}$$

According to Theorem 15:1 we have $\delta_1(u) \rightarrow 0$ for $u \rightarrow u_{11}$ and $u \rightarrow u_{1,n+1}$. Thus in the sum S of intervals $\langle -1, u_{11} \rangle$ and $\langle u_{1,n+1}, 1 \rangle$ there exists a point u^* such that $\delta_1(u^*) = \max_{u \in S} \delta_1(u)$. The set U_2 will be defined as follows:

$$U_2 = \begin{cases} \{u^*, u_{11}, \dots, u_{1,n+1}\} & (-1 \leq u^* < u_{11}), \\ \{u_{11}, \dots, u_{1,n+1}, u^*\} & (u_{1,n+1} < u^* \leq 1). \end{cases}$$

After finding the set U_j for $2 \leq j \leq n + 1$ we introduce the function

$$\delta_j(u) = \varepsilon_n(\xi; \{u_{j0}, \dots, u_{j,j-2}, u, u_{jj}, \dots, u_{j,n+1}\}) \quad (u_{j,j-2} < u < u_{jj})$$

and for the point $u^* \in (u_{j,j-2}, u_{jj})$, at which the function δ_j has its maximum we define the set U_{j+1} :

$$U_{j+1} = \{u_{j0}, \dots, u_{j,j-2}, u^*, u_{jj}, \dots, u_{j,n+1}\}.$$

After finding the set U_{n+2} we introduce the function

$$\delta_{n+2}(u) = \begin{cases} \varepsilon_n(\xi; \{u, u_{n+2,0}, \dots, u_{n+2,n}\}) & (-1 \leq u < u_{n+2,0}), \\ \varepsilon_n(\xi; \{u_{n+2,0}, \dots, u_{n+2,n}, u\}) & (u_{n+2,n} < u \leq 1). \end{cases}$$

If the function δ_{n+2} has its maximum on the set $\langle -1, u_{n+2,0} \rangle \cup \langle u_{n+2,n}, 1 \rangle$ at the point u^* , we put

$$U_{n+3} = \begin{cases} \{u^*, u_{n+2,0}, \dots, u_{n+2,n}\} & (-1 \leq u^* < u_{n+2,0}), \\ \{u_{n+2,0}, \dots, u_{n+2,n}, u^*\} & (u_{n+2,n} < u^* \leq 1). \end{cases}$$

It follows that the essence of the method of Bratton is such a change of the m -th point of the sets U_m, U_{m+n+2}, \dots as would give the maximum of the n -th error of the best approximation on the next set.

After finding several successive sets of the sequence $\{U_m\}$ one has to compute the coefficients of the n -th best polynomial on the last set and verify how well this polynomial approximates the required polynomial ω_{nI} . This verification is not necessary if we seek the best polynomial not on the interval I but on a finite set A . Then the signal of terminating the computations is the equation

$$\varepsilon_n(\xi; U_{(n+2)l+2}) = \varepsilon_n(\xi; U_{(n+2)l+3}) = \dots = \varepsilon_n(\xi; U_{(n+2)(l+1)+1})$$

valid for some l . The polynomial ω_{nA} is equal to the n -th best polynomial on the set $U_{(n+2)(l+1)+1}$ and we need to compute only that polynomial.

REFERENCES

- [1] N. I. Achiezer, *Teoria aproksymacji*, Warszawa 1957.
- [2] S. Bernstein, *Sur le problème inverse de la théorie de la meilleure approximation des fonctions continues*, *Compt. Rend. Hébd. Acad. Sci.* 206 (1938), p. 1520-1523.
- [3] — (С. Н. Бернштейн), *Доказательство теоремы Вейерштрасса, основанное на теории вероятностей*, *Собрание сочинений*, т. I, Москва 1952, p. 105-106.
- [4] — (С. Н. Бернштейн), *Экстремальные свойства полиномов и наилучшее приближение непрерывных функций одной вещественной переменной*, ч. I, Ленинград 1937.
- [5] E. Borel, *Leçons sur les fonctions de variables réelles*, Paris 1905.
- [6] D. Bratton, *New results in the theory and techniques of Chebyshev fitting*, *Abstract No. 564-34, Notices Amer. Math. Soc.* 5 (1958), p. 210.
- [7] V. N. Вуров (В. Н. Буров), *Некоторые эффективные способы решения задачи П. Л. Чебышева о наилучшем приближении*, *Изв. Высших Учебных Заведений, Математика* 1 (1957), p. 67-79.
- [8] P. L. Chebyshev (П. Л. Чебышев), *Вопросы о наименьших величинах, связанные с приближенным представлением функций*, *Полное собрание сочинений*, т. II, Москва 1947, p. 151-235.
- [9] — (П. Л. Чебышев), *О функциях мало удаляющихся от нуля при некоторых величинах переменной*, *Полное собрание сочинений*, т. III, Москва 1948, p. 108-127.
- [10] S. N. Chernikov (С. Н. Черников), *Системы линейных неравенств*, *Успехи Мат. Наук* 8.2 (1953), p. 7-73.
- [11] P. C. Curtis Jr., *n-parameter families and best approximation*, *Pacific Journ. Math.* 9 (1959), p. 1013-1027.
- [12] — and W. L. Frank, *An algorithm for the determination of the polynomial of best minimum approximation to a function defined on a finite point set*, *Journ. Assoc. Comput. Machinery* 6 (1959), p. 395-404.
- [13] P. Erdős, *Note on some elementary properties of polynomials*, *Bull. Amer. Math. Soc.* 46 (1940), p. 954-958.
- [14] G. Hornecker, *Détermination approchée, à précision numérique élevée, du polynôme de meilleure approximation d'ordre n, au sens de Tchébicheff, d'une fonction bornée continue, sur un segment fini*, *Compt. Rend. Hébd. Acad. Sci.* 246 (1958), p. 43-46.
- [15] — *Evaluation approchée de la meilleure approximation polynomiale d'ordre n de f(x) sur un segment fini [a, b]*, *Chiffres* 1 (1958), p. 157-169.
- [16] H. Lebesgue, *Sur l'approximation des fonctions*, *Bull. Sci. Math.* 22.1 (1898), p. 278-287.

- [17] V. S. Linskiy (В. С. Линский), *Вычисление элементарных функций на автоматических цифровых машинах*, Вычислительная Математика 2 (1957), p. 90-119.
- [18] L. M. Milne-Thomson, *The calculus of finite differences*, London 1951.
- [19] Jan Mycielski and S. Paszkowski, *A generalization of Chebyshev polynomials*, Bull. Acad. Polon. Sci. 8 (1960), p. 433-438.
- [20] I. P. Natanson (И. П. Натансон), *Конструктивная теория функций*, Москва 1949.
- [21] E. P. Novodvorskiiy and I. Sh. Pinsker (Е. П. Новодворский и И. Ш. Пинскер), *Процесс уравнивания максимумов*, Успехи Мат. Наук 6.6 (1951), p. 174-181.
- [22] S. Paszkowski, *On the approximation with nodes*, Warszawa 1957.
- [23] — (С. Пашковский), *О расположении (e)-точек полиномов наилучшего приближения*, Доклады АН СССР 117 (1957), p. 576-577.
- [24] —, *Zagadnienia numeryczne aproksymacji jednostajnej, I (O alternansie wielomianów optymalnych)*, Zastosow. Mat. 4 (1958), p. 42-58, 279-280 (errata).
- [25] —, *Zagadnienia numeryczne aproksymacji jednostajnej, II (O szacowaniu błędu aproksymacji optymalnej)*, Zastosow. Mat. 4 (1958), p. 59-74.
- [26] — (С. Пашковский), *О полиномах, все корни которых вещественны*, Ann. Polon. Math. 5 (1958), p. 165-194.
- [27] — (С. Пашковский), *О некотором свойстве наилучшей равномерной аппроксимации*, Ann. Polon. Math. 5 (1958), p. 195-208.
- [28] —, *Automatyczna elektronowa maszyna licząca „Striela-4”*, Zastosow. Mat. 5 (1960), p. 67-96.
- [29] E. Remez, *Sur un procédé convergent d'approximations successives pour déterminer les polynomes d'approximation*, Compt. Rend. Hébd. Acad. Sci. 198 (1934), p. 2063-2065.
- [30] —, *Sur le calcul effectif des polynomes d'approximation de Tchebichef*, Compt. Rend. Hébd. Acad. Sci. 199 (1934), p. 337-340.
- [31] — (Е. Я. Ремез), *Общие вычислительные методы чебышевского приближения*, Киев 1957.
- [32] I. M. Ryzhik and I. S. Gradstein (И. М. Рыжик и И. С. Градштейн), *Таблицы интегралов, сумм, рядов и произведений*, Москва 1951.
- [33] A. Spitzbart and D. L. Shell, *A Chebyshev fitting criterion*, Journ. Assoc. Comput. Machinery 5 (1958), p. 22-31.
- [34] Ch. de la Vallée Poussin, *Leçons sur l'approximation des fonctions d'une variable réelle*, Paris 1919.
- [35] L. Veidinger, *On the numerical determination of the best approximations in the Chebyshev sense*, Numer. Math. 2 (1960), p. 99-105.
- [36] E. V. Voronovskaya (Е. В. Вороновская), *О системе дифференциальных уравнений некоторых экстремальных полиномов*, Успехи Мат. Наук 11.1 (1956), p. 251-254.
- [37] K. Weierstrass, *Über die analytische Darstellbarkeit sogenannter willkürlicher Funktionen einer reellen Veränderlichen*, Sitzungsber. Deutsch. Akad. Wiss. Berlin (1885), p. 633-639, 789-805.
- [38] E. I. Zolotarev (Е. И. Золотарев), *Приложение эллиптических функций к вопросам о функциях, наименее и наиболее уклоняющихся от нуля*, Полное собрание сочинений, т. 2, 1932.

CONTENTS

Introduction	3
CHAPTER I. Basic properties of the best polynomials	
1. Existence, uniqueness and the characteristic properties of the best polynomials	6
2. The direct application of the theorem concerning the (n, F) -points to computing the best polynomials. The Chebyshev polynomials	15
3. The characteristic properties of pairs of successive best polynomials	26
CHAPTER II. Estimation of error of the best approximation	
4. Basic theorems	33
5. Vallée Poussin type estimations	38
6. Estimation of the error of the best approximation for functions which are differentiable several times	53
7. Estimation of the $(n+1)$ -st error of the best approximation based upon the knowledge of the n -th best polynomial	57
8. The relation between the error of the best approximation and the interval of approximation	65
CHAPTER III. The distribution of (n)-points in the interval of approximation	
9. The estimates dependent upon the value of the ratio $\varepsilon_{n+1}(\xi)/\varepsilon_n(\xi)$. .	70
10. Auxiliary theorems	84
11. The estimates dependent upon the properties of the derivatives of the function ξ . General theorems	92
12. The estimates dependent upon the properties of the derivatives of the function ξ . The choice of function γ_p	111
CHAPTER IV. Methods of computing well-approximating and best polynomials	
13. Well-approximating polynomials	121
14. Approximation of families of functions. The Zolotarev polynomials	140
15. The theory of the method of Remez	156
16. Other methods of computing the best polynomials	169
References	175