1

Approximation of Univariate Functions

1.1 Introduction

The primary problem in approximation theory is the choice of a successful method of approximation. In this chapter and in Chapter 2 we test various approaches, based on the concept of width, to the evaluation of the quality of a method of approximation. We take as an example the approximation of periodic functions of a single variable. The two main parameters of a method of approximation are its accuracy and complexity. These concepts may be treated in various ways depending on the particular problems involved. Here we start from classical ideas about the approximation of functions by polynomials. After Fourier's 1807 article the representation of a 2π-periodic function by its Fourier series became natural. In other words, the function $f(x)$ is approximately represented by a partial sum $S_n(f,x)$ of its Fourier series:

$$S_n(f,x) := a_0/2 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

$$a_k := \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx, \qquad b_k := \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\,dx.$$

We are interested in the approximation of a function f by a polynomial $S_n(f)$ in some L_p-norm, $1 \le p \le \infty$. In the case $p = \infty$ we assume that we are dealing with the uniform norm. As a measure of the accuracy of the method of approximating a periodic function by means of its Fourier partial sum we consider the quantity $\|f - S(f)\|_p$. The complexity of this method of approximation contains the following two characteristics. The order of the trigonometric polynomial $S_n(f)$ is the quantitative characteristic. The following observation gives us the qualitative characteristic. The coefficients of this polynomial are found by the Fourier formulas, which means that the operator S_n is the orthogonal projector onto the subspace of trigonometric polynomials of order n.

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

In 1854 Chebyshev suggested representing continuous function f by its polynomial of best approximation, namely, by the polynomial $t_n(f)$ such that

$$\|f - t_n(f)\|_\infty = E_n(f)_\infty := \inf_{\alpha_k,\beta_k} \left\| f(x) - \sum_{k=0}^{n} (\alpha_k \cos kx + \beta_k \sin kx) \right\|_\infty .$$

He proved the existence and uniqueness of such a polynomial. We consider this method of approximation not only in the uniform norm but in all L_p-norms, $1 \le p < \infty$. The accuracy of the Chebyshev method can be easily compared with the accuracy of the Fourier method:

$$E_n(f)_p \le \left\| f - S_n(f) \right\|_p .$$

However, it is difficult to compare the complexities of these two methods. The quantitative characteristics coincide but the qualitative characteristics are different (for example, it is not difficult to understand that for $p = \infty$ the mapping $f \to t_n(f)$ is not a linear operator). The Du Bois–Reymond 1873 example of a continuous function f such that $\left\| f - S_n(f) \right\|_\infty \to \infty$ when $n \to \infty$, and the Weierstrass theorem which says that for each continuous function f we have $E_n(f)_\infty \to 0$ as $n \to \infty$, showed the advantage of the Chebyshev method over the Fourier method from the point of view of accuracy.

The desire to construct methods of approximation which have the advantages of both the Fourier and Chebyshev methods has led to the study of various methods of summation of Fourier series. The most important among them from the point of view of approximation are the de la Vallée Poussin, Fejér, and Jackson methods, which were constructed early in the twentieth century. All these methods are linear. For example, in the de la Vallée Poussin method a function f is approximated by the polynomial

$$V_n(f) := \frac{1}{n} \sum_{l=n}^{2n-1} S_l(f)$$

of order $2n - 1$.

From the point of view of accuracy this method is close to the Chebyshev method; de la Vallée Poussin proved that

$$\left\| f - V_n(f) \right\|_p \le 4E_n(f)_p, \qquad 1 \le p \le \infty.$$

From the point of view of complexity it is close to the Fourier method, and the property of linearity essentially distinguishes it from the Chebyshev method.

We see that common to all these methods is approximation by trigonometric polynomials. However, the methods of constructing these polynomials differ: some

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

methods use orthogonal projections on to the subspace of trigonometric polynomials of fixed order, some use best-approximation operators, and some use linear operators.

Thus, the approximation of periodic functions by trigonometric polynomials is natural and this problem has been thoroughly studied. The approximation of functions by algebraic polynomials has been studied in parallel with approximation by trigonometric polynomials. We now point out some results, which determined the style of investigation of a number of problems in approximation theory. These problems are of interest even today.

It was proved by de la Vallée Poussin (1908) that, for best approximation of the function $|x|$ in the uniform norm on $[-1,1]$ by algebraic polynomials of degree n, the following upper *estimate* or *bound* holds:

$$e_n\big(|x|\big) \le C/n.$$

He raised the question of the possibility of an improvement of this estimate in the sense of order. In other words, could the function C/n be replaced by a function that decays faster to zero? Bernstein (1912) proved that this order estimate is sharp. Moreover, he then established the asymptotic behavior of the sequence $\big\{e_n\big(|x|\big)\big\}$ (see Bernstein, 1914):

$$e_n\big(|x|\big) = \mu/n + o(1/n), \qquad \mu = 0.282 \pm 0.004.$$

These results initiated a series of investigations into best approximations of individual functions having special singularities.

At this stage of investigation the natural conjecture arose that the smoother a function, the more rapidly its sequence of best approximations decreases.

In 1911 Jackson proved the inequality

$$E_n(f)_\infty \le Cn^{-r}\omega(f^{(r)}, 1/n)_\infty.$$

The relations which give upper estimates for the best approximations of a function in terms of its smoothness are now called the *Jackson inequalities*, and in a wider sense such relations are called *direct theorems* of approximation theory.

As a result of Bernstein's (1912) and de la Vallée Poussin's (1908, 1919) investigations we can formulate the following assertion, which is now called the *inverse theorem* of approximation theory. If

$$E_n(f)_\infty \le Cn^{-r-\alpha}, \qquad 0 \le r \text{ integer}, \qquad 0 < \alpha < 1,$$

then f has a continuous derivative of order r which belongs to the class Lip α; that is, $f \in W^r H^\alpha$ (in the notation of this book it is the class $H_\infty^{r+\alpha}$). Thus, the results of Jackson, Bernstein, and de la Vallée Poussin show that functions from the class $W^r H^\alpha$, $0 < \alpha < 1$, can be characterized by the order of decrease of its sequences of best approximations.

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

We remark that at that time, early in the twentieth century, classes similar to $W^r H^\alpha$ were used in other areas of mathematics for obtaining the orders of decrease of various quantities. As an example we formulate a result of Fredholm (1903). Let $f(x,y)$ be continuous on $[a,b]\times[a,b]$ and

$$\max_{x,y}\big|f(x,y+t)-f(x,y)\big| \le C|t|^\alpha, \qquad 0<\alpha\le 1.$$

Then for eigenvalues $\lambda(J_f)$ of the integral operator

$$(J_f\psi)(x) = \int_a^b f(x,y)\psi(y)dy$$

the following relation is valid for any $\rho > 2/(2\alpha+1)$:

$$\sum_{n=1}^{\infty} \big|\lambda_n(J_f)\big|^\rho < \infty.$$

The investigation of the upper bounds or estimates of errors of approximation of functions from a fixed class by some method of approximation began with an article by Lebesgue (1910). In particular, Lebesgue proved that

$$S_n(\mathrm{Lip}\,\alpha)_\infty := \sup_{f\in \mathrm{Lip}\,\alpha} \big\|f-S_n(f)\big\|_\infty \asymp n^{-\alpha}\ln n.$$

Here and later we write $a_n \asymp b_n$ for two sequences a_n and b_n if there are two positive constants C_1 and C_2 such that $C_1 b_n \le a_n \le C_2 b_n$ for all n.

The problem of approximation of functions in the classes $W^r H^\alpha$ by trigonometric polynomials was so natural that a tendency to find either asymptotic or exact values of the following quantities appeared:

$$S_n(W^r H^\alpha)_\infty := \sup_{f\in W^r H^\alpha} \|f-S_n(f)\|_\infty, \qquad E_n(W^r H^\alpha)_\infty := \sup_{f\in W^r H^\alpha} E_n(f)_\infty.$$

We now formulate the first results in this direction. Kolmogorov (1936) proved the relation (in our notation $W^r = W^r_{\infty,r}$, see §1.4)

$$S_n(W^r)_\infty = \frac{4}{\pi^2}\frac{\ln n}{n^r} + O(n^{-r}), \qquad n\to\infty.$$

Independently, Favard (1937) and Akhiezer and Krein (1937) proved the equality

$$E_n(W^r)_\infty = K_r(n+1)^{-r},$$

where K_r is a number depending on the natural number r.

In 1936 Kolmogorov introduced the concept of the width d_n of a class F in a space X:

$$d_n(F,X) := \inf_{\{\phi_j\}_{j=1}^n} \sup_{f\in F} \inf_{\{c_j\}_{j=1}^n} \left\| f - \sum_{j=1}^n c_j\phi_j \right\|_X.$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

This concept allows us to find, for a fixed n and for a class F, a subspace of dimension n that is optimal with respect to the construction of a best approximating element. In other words, the concept of width allows us to choose from among various Chebyshev methods having the same quantitative characteristic of complexity (the dimension of the approximating subspace) the one which has the greatest accuracy.

The first result about widths (Kolmogorov, 1936), namely

$$d_{2n+1}(W_2^r, L_2) = (n+1)^{-r},$$

showed that the best subspace of dimension $2n+1$ for the approximation of classes of periodic functions is the subspace of trigonometric polynomials of order n. This result confirmed that the approximation of functions in the class W_2^r by trigonometric polynomials is natural. Further estimates of the widths $d_{2n+1}(W^r_{q,\alpha}, L_p)$, $1 \leq q$, $p \leq \infty$, some of which are discussed in §2.1 below, showed that, for some values of the parameters $q,\ p$, the subspace of trigonometric polynomials of order n is optimal (in the sense of the order of decay) but for other values of q, p this subspace is not optimal.

The Ismagilov (1974) estimate for the quantity $d_n(W_1^r, L_\infty)$ gave the first example, where the subspace of trigonometric polynomials of order n is not optimal. This phenomenon was thoroughly studied by Kashin (1977).

In analogy to the problem of the *Kolmogorov width*, that is, to the problem concerning the best Chebyshev method, problems concerning the best linear method and the best Fourier method were considered.

Tikhomirov (1960b) introduced the *linear width*:

$$\lambda_n(F, L_p) := \inf_{A:\operatorname{rank} A \leq n} \sup_{f \in F} \|f - Af\|_p,$$

and Temlyakov (1982a) introduced the *orthowidth* (Fourier width):

$$\varphi_n(F, L_p) := \inf_{\text{orthonormal system } \{u_i\}_{i=1}^n} \sup_{f \in F} \left\| f - \sum_{i=1}^n \langle f, u_i \rangle u_i \right\|_p .$$

A discussion and comparison of results concerning $d_n(W_q^r, L_p)$, $\lambda_n(W_q^r, L_p)$ and $\varphi_n(W_q^r, L_p)$ can be found in §2.1. Here we remark that, from the point of view of the orthowidth, the Fourier operator S_n is optimal (in the sense of order) for all $1 \leq q$, $p \leq \infty$ with the exception of the two cases $(1,1)$ and (∞,∞).

Keeping in mind the primary question about the selection of an optimal subspace of approximating functions, we now draw some conclusions from this brief historical survey.

(1) The trigonometric polynomials have been considered as a natural means of approximation of periodic functions during the whole period of development of approximation theory.

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

(2) In approximation theory (as well as in other fields of mathematics) it has turned out that it is natural to unite functions with the same smoothness into a class.
(3) The subspace of trigonometric polynomials has been obtained in many cases as the solution of problems regarding the most precise method for the classes of smooth functions: the Chebyshev method (which uses the Kolmogorov width), the linear method (which uses the linear width), or the Fourier method (which uses the orthowidth).

On the basis of these remarks we may formulate the following general strategy for investigating approximation problems; we remark that this strategy turns out to be most fruitful in those cases where we do not know *a priori* a natural method of approximation. First, we solve the width problem for a class of interest in the simplest case, that of approximation in Hilbert space, L_2. Second, we study the system of functions obtained and apply it to approximation in other spaces L_p. This strategy will be used in Chapters 3, 4, and 5.

1.2 Trigonometric Polynomials

Functions of the form

$$t(x) = \sum_{|k|\le n} c_k e^{ikx} = a_0/2 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

(c_k, a_k, b_k are complex numbers) will be called trigonometric polynomials of order n. We denote the set of such polynomials by $\mathscr{T}(n)$ and the subset of $\mathscr{T}(n)$ of real polynomials by $\mathscr{R}\mathscr{T}(n)$.

We first consider a number of concrete polynomials that play an important role in approximation theory.

1.2.1 The Dirichlet Kernel of Order n

The classical univariate Dirichlet kernel of order n is defined as follows:

$$\begin{aligned}\mathscr{D}_n(x) &:= \sum_{|k|\le n} e^{ikx} = e^{-inx}(e^{i(2n+1)x} - 1)(e^{ix} - 1)^{-1} \\ &= \frac{\sin(n+1/2)x}{\sin(x/2)}. \end{aligned} \tag{1.2.1}$$

The Dirichlet kernel is an even trigonometric polynomial with the majorant

$$\left|\mathscr{D}_n(x)\right| \le \min\bigl(2n+1, \pi/|x|\bigr), \qquad |x| \le \pi. \tag{1.2.2}$$

The estimate

$$\|\mathscr{D}_n\|_1 \le C \ln n, \qquad n = 2, 3, \ldots, \tag{1.2.3}$$

follows from (1.2.2).

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

We mention the well-known relation (see Dzyadyk, 1977, p. 112)

$$\|\mathscr{D}_n\|_1 = \frac{4}{\pi^2}\ln n + R_n, \qquad |R_n| \le 3, \qquad n = 1,2,3,\dots$$

For any trigonometric polynomial $t \in \mathscr{T}(n)$ we have

$$\mathscr{D}_n * t := (2\pi)^{-1} \int_{\mathbb{T}} \mathscr{D}_n(x-y)t(y)dy = t.$$

Denote

$$x^l := 2\pi l/(2n+1), \qquad l = 0,1,\dots,2n.$$

Clearly, the points x^l, $l = 1,\dots,2n$, are zeros of the Dirichlet kernel $\mathscr{D}_n$ on $[0,2\pi]$. For any $|k| \le n$ we have

$$\sum_{l=0}^{2n} e^{ikx^l}\mathscr{D}_n(x-x^l) = \sum_{|m|\le n} e^{imx} \sum_{l=0}^{2n} e^{i(k-m)x^l} = e^{ikx}(2n+1).$$

Consequently, for any $t \in \mathscr{T}(n)$,

$$t(x) = (2n+1)^{-1} \sum_{l=0}^{2n} t(x^l)\mathscr{D}_n(x-x^l). \tag{1.2.4}$$

Further, it is easy to see that for any $u,\ v \in \mathscr{T}(n)$ we have

$$\langle u,v\rangle := (2\pi)^{-1} \int_{-\pi}^{\pi} u(x)\overline{v(x)}dx = (2n+1)^{-1} \sum_{l=0}^{2n} u(x^l)\overline{v(x^l)} \tag{1.2.5}$$

and, for any $t \in \mathscr{T}(n)$,

$$\|t\|_2^2 = (2n+1)^{-1} \sum_{l=0}^{2n} \left|t(x^l)\right|^2. \tag{1.2.6}$$

For $1 < q \le \infty$ the estimate

$$\|\mathscr{D}_n\|_q \le C(q)n^{1-1/q} \tag{1.2.7}$$

follows from (1.2.2). Applying the Hölder inequality (see (A.1.1) in the Appendix) for estimating $\|\mathscr{D}_n\|_2^2$ we get

$$2n+1 = \|\mathscr{D}_n\|_2^2 \le \|\mathscr{D}_n\|_q \|\mathscr{D}_n\|_{q'}. \tag{1.2.8}$$

Relations (1.2.7) and (1.2.8) imply for $1 < q < \infty$ the relation

$$\|\mathscr{D}_n\|_q \asymp n^{1-1/q}. \tag{1.2.9}$$

Relation (1.2.9) for $q = \infty$ is obvious.

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

We denote by S_n the operator taking a partial sum of order n. Then for $f \in L_1$ we have

$$S_n(f) := \mathscr{D}_n * f = (2\pi)^{-1} \int_{-\pi}^{\pi} \mathscr{D}_n(x-y) f(y) dy.$$

Theorem 1.2.1 *The operator S_n does not change polynomials from $\mathscr{T}(n)$ and for $p = 1$ or ∞ we have*

$$\|S_n\|_{p\to p} \le C \ln n, \qquad n = 2, 3, \dots,$$

and for $1 < p < \infty$ for all n we have

$$\|S_n\|_{p\to p} \le C(p).$$

This theorem follows from (1.2.3) and the Marcinkiewicz multiplier theorem (see Theorem A.3.6).

For $t \in \mathscr{T}(n)$,

$$t(x) = a_0/2 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx),$$

we call the polynomial $\tilde{t} \in \mathscr{T}(n)$, where

$$\tilde{t}(x) := \sum_{k=1}^{n} (a_k \sin kx - b_k \cos kx)$$

the polynomial conjugate to t.

Corollary 1.2.2 *For $1 < p < \infty$ and all n we have*

$$\|\tilde{t}\|_p \le C(p) \|t\|_p.$$

Proof Let $t \in \mathscr{T}(n)$. It is not difficult to see that $\tilde{t} = t * \widetilde{\mathscr{D}}_n$, where

$$\widetilde{\mathscr{D}}_n(x) := 2 \sum_{k=1}^{n} \sin kx.$$

Clearly, it suffices to consider the case of odd n. Let this be the case and set $m := (n+1)/2$, $l := (n-1)/2$. Representing $\tilde{\mathscr{D}}_n(x)$ in the form

$$\tilde{\mathscr{D}}_n(x) = \frac{1}{i}\left(\sum_{k=1}^{n} e^{ikx} - \sum_{k=-n}^{-1} e^{ikx}\right) = \frac{1}{i}\left(e^{imx}\mathscr{D}_l(x) - e^{-imx}\mathscr{D}_l(x)\right),$$

we obtain the corollary. □

A trigonometric conjugate operator maps a function $f(x)$ to a function

$$\sum_k (\operatorname{sign} k) \hat{f}(k) e^{ikx}.$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

The Marcinkiewicz multiplier theorem A.3.6 implies that this operator is bounded as an operator from L_p to L_p for $1 < p < \infty$. We denote by $\tilde{f}$ the conjugate function.

1.2.2 The Fejér Kernel of Order $n-1$

The classical univariate Fejér kernel of order $n-1$ is defined as follows:

$$\begin{aligned}\mathscr{K}_{n-1}(x) &:= n^{-1}\sum_{m=0}^{n-1}\mathscr{D}_m(x) = \sum_{|m|\le n}\bigl(1-|m|/n\bigr)e^{imx}\\ &= \frac{\bigl(\sin(nx/2)\bigr)^2}{n\bigl(\sin(x/2)\bigr)^2}.\end{aligned}$$

The Fejér kernel is an even nonnegative trigonometric polynomial in $\mathscr{T}(n-1)$ with majorant

$$\bigl|\mathscr{K}_{n-1}(x)\bigr| = \mathscr{K}_{n-1}(x) \le \min\bigl(n, \pi^2/(nx^2)\bigr), \qquad |x| \le \pi. \tag{1.2.10}$$

From the obvious relations

$$\|\mathscr{K}_{n-1}\|_1 = 1, \qquad \|\mathscr{K}_{n-1}\|_\infty = n$$

and the inequality, see (A.1.6),

$$\|f\|_q \le \|f\|_1^{1/q}\|f\|_\infty^{1-1/q}$$

we get in the same way as we obtained (1.2.9),

$$Cn^{1-1/q} \le \|\mathscr{K}_{n-1}\|_q \le n^{1-1/q}, \qquad 1 \le q \le \infty. \tag{1.2.11}$$

1.2.3 The de la Vallée Poussin Kernels

The classical univariate de la Vallée Poussin kernel with parameters m, n is defined as follows:

$$\mathscr{V}_{m,n}(x) := (n-m)^{-1}\sum_{l=m}^{n-1}\mathscr{D}_l(x), \qquad n > m.$$

It is convenient to represent this kernel in terms of Fejér kernels:

$$\begin{aligned}\mathscr{V}_{m,n}(x) &= (n-m)^{-1}\bigl(n\mathscr{K}_{n-1}(x) - m\mathscr{K}_{m-1}(x)\bigr)\\ &= (\cos mx - \cos nx)\bigl(2(n-m)\bigl(\sin(x/2)\bigr)^2\bigr)^{-1}.\end{aligned}$$

The de la Vallée Poussin kernels $\mathscr{V}_{m,n}$ are even trigonometric polynomials of order $n-1$ with majorant

$$\bigl|\mathscr{V}_{m,n}(x)\bigr| \le C\min\bigl(n,\ 1/|x|, 1/\bigl((n-m)x^2\bigr)\bigr), \qquad |x| \le \pi. \tag{1.2.12}$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

Relation (1.2.12) implies the estimate

$$\|\mathcal{V}_{m,n}\|_1 \le C \ln\big(1 + n/(n-m)\big).$$

We often use the de la Vallée Poussin kernel with $n = 2m$ and denote it by

$$\mathcal{V}_m(x) := \mathcal{V}_{m,2m}(x), \qquad m \ge 1, \qquad \mathcal{V}_0(x) := 1.$$

Then for $m \ge 1$ we have

$$\mathcal{V}_m = 2\mathcal{K}_{2m-1} - \mathcal{K}_{m-1},$$

which, with the properties of $\mathcal{K}_n$, implies

$$\|\mathcal{V}_m\|_1 \le 3. \tag{1.2.13}$$

In addition,

$$\|\mathcal{V}_m\|_\infty \le 3m.$$

Consequently, in the same way as above, see (1.2.9) and (1.2.11), we get

$$\|\mathcal{V}_m\|_q \asymp m^{1-1/q}, \qquad 1 \le q \le \infty. \tag{1.2.14}$$

Denote

$$x(l) := \pi l/2m, \qquad l = 1, \dots, 4m.$$

Then, analogously to (1.2.4), for each $t \in \mathcal{T}(m)$ we have

$$t(x) = (4m)^{-1} \sum_{l=1}^{4m} t\big(x(l)\big) \mathcal{V}_m\big(x - x(l)\big). \tag{1.2.15}$$

The operator $\mathcal{V}_m$ defined on L_1 by the formula

$$V_m(f) := f * \mathcal{V}_m$$

is called the de la Vallée Poussin operator.

The following theorem is a corollary of the definition of the kernels $\mathcal{V}_m$ and the relation (1.2.13).

Theorem 1.2.3 *The operator V_m does not change polynomials from $\mathcal{T}(m)$, and for all $1 \le p \le \infty$ we have*

$$\|V_m\|_{p \to p} \le 3, \qquad m = 1, 2, \dots$$

In addition, we formulate two properties of the de la Vallée Poussin kernels.

(1) Relation (1.2.12) with $n = 2m$ implies the inequality

$$\big|\mathcal{V}_m(x)\big| \le C \min\big(m, 1/(mx^2)\big), \qquad |x| \le \pi.$$

It is easy to derive from this inequality the following property.

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

(2) For h satisfying the condition $C_1 \le mh \le C_2$ we have

$$\sum_{0\le l\le 2\pi/h} \left|\mathscr{V}_m(x-lh)\right| \le Cm.$$

We remark that property (2) is valid for the Fejér kernel $\mathscr{K}_m$.

1.2.4 The Jackson Kernel

The classical univariate Jackson kernel with parameters n, a is defined as follows:

$$J_n^a(x) := \gamma_{a,n}^{-1} \left(\frac{\sin(nx/2)}{\sin(x/2)}\right)^{2a}, \qquad a \in \mathbb{N},$$

where $\gamma_{a,n}$ is selected in such a way that

$$\|J_n^a\|_1 = 1. \tag{1.2.16}$$

Let us estimate $\gamma_{a,n}$ from below. We have

$$\begin{aligned}\gamma_{a,n} &= (2\pi)^{-1}\int_{-\pi}^{\pi}\left(\frac{\sin(nx/2)}{\sin(x/2)}\right)^{2a} dx\\ &\ge \pi^{-1}\int_0^{\pi/n}\left(\frac{nx/\pi}{x/2}\right)^{2a} dx \ge Cn^{2a-1}.\end{aligned} \tag{1.2.17}$$

The Jackson kernel is an even nonnegative trigonometric polynomial of order $a(n-1)$. It follows from (1.2.17) that

$$J_n^a(x) \le C\min(n, n^{1-2a}x^{-2a}), \qquad |x| \le \pi. \tag{1.2.18}$$

Relation (1.2.18) implies that for $0 \le r < 2a-1$,

$$\int_0^{\pi} J_n^a(x)x^r dx \le C(r)n^{-r}. \tag{1.2.19}$$

1.2.5 The Rudin–Shapiro Polynomials

We define recursively pairs of trigonometric polynomials $P_j(x)$ and $Q_j(x)$ of order 2^j-1:

$$P_0 := Q_0 := 1,$$

$$P_{j+1}(x) := P_j(x) + e^{i2^jx}Q_j(x), \qquad Q_{j+1}(x) := P_j(x) - e^{i2^jx}Q_j(x).$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

Then at each point x we have

$$\begin{aligned}|P_{j+1}|^2+|Q_{j+1}|^2 &= (P_j+e^{i2^jx}Q_j)(\overline{P}_j+e^{-i2^jx}\overline{Q}_j)\\ &\quad+(P_j-e^{i2^jx}Q_j)(\overline{P}_j-e^{-i2^jx}\overline{Q}_j)\\ &= 2\big(|P_j|^2+|Q_j|^2\big).\end{aligned}$$

Therefore, for all x

$$\big|P_j(x)\big|^2+\big|Q_j(x)\big|^2=2^{j+1}.$$

Thus, for example,

$$\|P_n\|_\infty \le 2^{(n+1)/2}. \tag{1.2.20}$$

It is clear from the definition of the polynomials P_n that

$$P_n(x)=\sum_{k=0}^{2^n-1}\varepsilon_k e^{ikx}, \qquad \varepsilon_k=\pm 1, \quad \varepsilon_0=1.$$

Let N be a natural number and

$$N=\sum_{j=1}^{m}2^{n_j}, \qquad n_1>n_2>\cdots>n_m\ge 0,$$

its binary representation. We set

$$\begin{aligned}R'_N(x)&:=P_{n_1}(x)+\sum_{j=2}^{m}P_{n_j}(x)e^{i(2^{n_1}+\cdots+2^{n_{j-1}})x},\\ R_N(x)&:=R'_N(x)+R'_N(-x)-1.\end{aligned}$$

Then $R_N(x)$ has the form

$$R_N(x)=\sum_{|k|<N}\varepsilon_k e^{ikx}, \qquad \varepsilon_k=\pm 1,$$

and for this polynomial the estimate

$$\|R_N\|_\infty \le CN^{1/2} \tag{1.2.21}$$

holds.

1.2.6 A Modification of the Fejér Kernel

We consider the polynomials

$$G_n(x):=\sum_{|m|<n}\big(1-|m|/n\big)^{1/2}e^{imx}.$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

These are even trigonometric polynomials of order $n-1$ with the properties

$$(2\pi)^{-1}\int_0^{2\pi} G_n(x-a)G_n(x-b)dx = \mathscr{K}_{n-1}(a-b), \tag{1.2.22}$$

$$\|G_n\|_1 \le C. \tag{1.2.23}$$

Relation (1.2.22) is obvious. It implies that the system of polynomials $G_n(x-2\pi l/n)$, $l=1,\ldots,n$ is an orthogonal system in $\mathscr{T}(n-1)$.

Let us prove the relation (1.2.23). Denote

$$\phi(u) := \big(1-|u|\big)^{1/2}, \qquad |u|\le 1.$$

Then we have on $[-1,1]$

$$\phi(u) = \sum_l a_l e^{i\pi l x}$$

and it is not hard to prove that

$$|a_l| \le C\big(|l|+1\big)^{-3/2}. \tag{1.2.24}$$

Further,

$$\begin{aligned} G_n(x) &= \sum_{|m|\le n} \phi(m/n)e^{imx} \\ &= \sum_l a_l \sum_{|m|\le n} e^{im(x+\pi l/n)} = \sum_l a_l \mathscr{D}_n(x+\pi l/n). \end{aligned} \tag{1.2.25}$$

Let us consider the function

$$g_{n,l}(x) := \mathscr{D}_n(x+\pi l/n) - (-1)^l \mathscr{D}_n(x).$$

Using the representation (1.2.1) one can obtain the estimate

$$\|g_{n,l}\|_1 \le C\ln\big(|l|+2\big). \tag{1.2.26}$$

Further, owing to the equality

$$\phi(1) = \sum_l a_l(-1)^l = 0,$$

relation (1.2.25) can be rewritten in the form

$$G_n(x) = \sum_l a_l g_{n,l}(x).$$

Using relations (1.2.24) and (1.2.26) we then obtain relation (1.2.23).

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

1.2.7 A Generalization of the Rudin–Shapiro Polynomials

The trigonometric polynomials considered above (see §§1.2.1–1.2.6) were constructively obtained: either they are given by a formula (§§1.2.1–1.2.4, 1.2.6) or a method of construction is supplied (§1.2.5). In this subsection we formulate a theorem that proves the existence of polynomials with given properties.

Theorem 1.2.4 *Let $\varepsilon > 0$, and let a subspace $\Psi \subset \mathscr{T}(n)$ be such that* $\dim \Psi \geq \varepsilon(2n+1)$. *Then there exists a $t \in \Psi$ such that*

$$\|t\|_\infty = 1,$$

and

$$\|t\|_2 \geq C(\varepsilon) > 0.$$

An analogous statement is valid for the multivariable trigonometric polynomials, will be proved in Chapter 3 (see Theorem 3.2.1).

We remark that the polynomial t from Theorem 1.2.4, by virtue of the inequality

$$\|t\|_2^2 \leq \|t\|_1 \|t\|_\infty,$$

satisfies the condition

$$\|t\|_1 \geq C(\varepsilon)^2 > 0. \tag{1.2.27}$$

1.2.8 An Application of the Gaussian Sums

In this subsection we construct polynomials that we will use in studying linear widths. This construction is based on properties of the Gaussian sums:

$$S(q,l) := \sum_{j=1}^{q} e^{i2\pi l j^2/q},$$

where q is a natural number and l, q are coprime; that is, $(l,q) = 1$. We confine ourselves to the case where q is an odd prime.

Theorem 1.2.5 *Let $q > 2$ be a prime, $l \neq 0$ an integer, and k an integer. Then, for*

$$S(q,l,k) := \sum_{j=1}^{q} e^{i2\pi(l j^2 + kj)/q},$$

the following equality is true:

$$\left|S(q,l,k)\right| = q^{1/2}.$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

Proof We first consider the case $k = 0$. Note that the quantity $S(q,l)$ does not change if we sum over the complete system of remainders modulo q instead of the segment $[1,q]$. Consequently, for any integer h,

$$S(q,l) = \sum_{j=1}^{q} e^{i2\pi l(j+h)^2/q}. \tag{1.2.28}$$

Further,

$$\left|S(q,l)\right|^2 = \left(\sum_{h=1}^{q} e^{-i2\pi l h^2/q}\right)\left(\sum_{j=1}^{q} e^{i2\pi l j^2/q}\right).$$

Using (1.2.28), we see that this is equal to

$$\sum_{h=1}^{q} e^{-i2\pi l h^2/q} \sum_{j=1}^{q} e^{i2\pi l(j+h)^2/q} = \sum_{h=1}^{q}\sum_{j=1}^{q} e^{i2\pi l(j^2+2jh)/q}. \tag{1.2.29}$$

Taking into account that

$$\sum_{h=1}^{q} e^{i2\pi l 2jh/q} = \begin{cases} q & \text{for } \; j = q, \\ 0 & \text{for } \; j \in [1,q), \end{cases}$$

we get from (1.2.29),

$$\left|S(q,l)\right|^2 = q. \tag{1.2.30}$$

Now let k be nonzero. Since q is a prime different from 2, the numbers $2lb$, $b = 1,\dots,q$, run through a complete system of remainders modulo q. Consequently, there is a b such that

$$2lb \equiv k \pmod q.$$

Then

$$lj^2 + kj \equiv l(j+b)^2 - lb^2 \pmod q$$

and, consequently,

$$\left|S(q,l,k)\right| = \left|S(q,l)\right| = q^{1/2}.$$

The theorem is proved. □

Theorem 1.2.6 *Let q be a prime and $q = 2a+1$. For any $n \in [1,a]$ there is a trigonometric polynomial $t_n \in \mathcal{T}(a)$ such that only n Fourier coefficients of t_n are nonzero and for all k we have $\left|\hat{t}(k)\right| \le 1$ and in addition*

$$t_n(0) \ge (n+1)/2, \qquad \left|t(2\pi l/q)\right| \le Cq^{1/2}, \qquad l = 1,\dots,2a.$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

Proof The proof of this theorem can easily be derived from a deep number theoretical result due to Hardy and Littlewood about estimating incomplete Gaussian sums: for any $n \in [1,q]$

$$\left|\sum_{j=1}^{n} e^{i2\pi l j^2/q}\right| \le Cq^{1/2}, \qquad (l,q)=1. \tag{1.2.31}$$

Indeed, let k_j denote the smallest nonnegative remainder of the number j^2 modulo q, $j=1,\dots,n$, and let

$$G := \{k_j - a,\ j=1,\dots,n\}.$$

We set

$$t_n(x) := \sum_{k\in G} e^{ikx}.$$

Then

$$\left|t_n(2\pi l/q)\right| = \left|\sum_{k\in G} e^{i2\pi lk/q}\right| = \left|\sum_{j=1}^{n} e^{i2\pi l j^2/q}\right|,$$

which by (1.2.31) implies the required estimates for $t_n(2\pi l/q)$. The bound $t_n(0) = n \ge (n+1)/2$ is obvious. $\square$

For the sake of completeness we will prove Theorem 1.2.6 using Theorem 1.2.5. Instead of (1.2.31) we prove the inequality

$$\left|\sum_{j} \left|(1-|j-a|/n)_{+} e^{i2\pi l j^2/q}\right|\right| \le q^{1/2}, \qquad (l,q)=1. \tag{1.2.32}$$

Let $l \in [1,q-1]$. Consider the trigonometric polynomial

$$t(x) := \sum_{j=0}^{q-1} e^{i2\pi l j^2/q} e^{i(j-a)x}.$$

Then at the points $x^k = 2\pi k/(2a+1) = 2\pi k/q$ we have

$$\left|t(x^k)\right| = \left|S(q,l,k)\right| = q^{1/2}, \qquad k=0,\dots,2a. \tag{1.2.33}$$

We set

$$u_n(x) := t(x) * \mathscr{K}_{n-1}(x).$$

Then by (1.2.5),

$$u_n(x) = q^{-1} \sum_{k=0}^{2a} t(x^k)\mathscr{K}_{n-1}(x-x^k),$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

and, using (1.2.33) we find that

$$\left|u_n(x)\right| \le q^{-1/2} \sum_{k=0}^{2a} \mathscr{K}_{n-1}(x - x^k) = q^{1/2}. \tag{1.2.34}$$

Further,

$$u_n(0) = \sum_j \left(1 - |j-a|/n\right)_+ e^{i2\pi l j^2/q},$$

where $(a)_+ := \max(a,0)$. By (1.2.34) this implies (1.2.32).

Setting

$$t_n(x) := \sum_j \left(1 - |j-a|/n\right)_+ e^{i(k_j - a)x},$$

where the k_j are the same as in the beginning of the proof of this theorem, we get

$$\left|t_n(2\pi l/q)\right| = \left|\sum_j \left(1 - |j-a|/n\right)_+ e^{i2\pi l k_j/q}\right| = \left|\sum_j \left(1 - |j-a|/n\right)_+ e^{i2\pi l j^2/q}\right|,$$

which by (1.2.32) implies the conclusion of Theorem 1.2.6, with $2n-1$ nonzero Fourier coefficients instead of n.

1.3 The Bernstein–Nikol'skii Inequalities. The Marcienkiewicz Theorem

The Bernstein–Nikol'skii inequalities connect the L_p-norms of a derivative of some polynomial with the L_q-norm, $1 \le q \le p \le \infty$, of this polynomial. We obtain here inequalities for a derivative that is slightly more general than the Weyl fractional derivative. We first make some auxiliary considerations.

For a sequence $\{a_\nu\}_{\nu=0}^{\infty}$ we write

$$\Delta a_\nu := a_\nu - a_{\nu+1}; \qquad \Delta^2 a_\nu := \Delta(\Delta a_\nu) = a_\nu - 2a_{\nu+1} + a_{\nu+2}.$$

Theorem 1.3.1 *We have*

$$(\pi)^{-1} \int_{-\pi}^{\pi} \left| a_0/2 + \sum_{\nu=1}^{n} a_\nu \cos \nu x \right| dx \le \sum_{\nu=0}^{n} (\nu+1)|\Delta^2 a_\nu|.$$

Proof Applying twice the Abel transformation (see (A.1.18) in the Appendix) with $a_\nu = 0$ for $\nu > n$, we obtain

$$\begin{aligned} t(x) &:= a_0 + \sum_{\nu=1}^{n} a_\nu 2\cos \nu x = \sum_{\nu=0}^{n} \mathscr{D}_\nu(x) \Delta a_\nu \\ &= \sum_{\nu=0}^{n} \left(\sum_{\mu=0}^{\nu} \mathscr{D}_\mu(x) \right) \Delta^2 a_\nu = \sum_{\nu=0}^{n} (\nu+1) \mathscr{K}_\nu(x) \Delta^2 a_\nu. \end{aligned} \tag{1.3.1}$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

From (1.3.1), using $\|\mathscr{K}_\nu\|_1 = 1$ we find

$$\|t\|_1 \le \sum_{\nu=0}^{n} (\nu+1)|\Delta^2 a_\nu|,$$

as required. □

1.3.1 The Bernstein inequality

We first prove the Bernstein inequality. Let us consider the following special trigonometric polynomials. Let s be a nonnegative integer. We define

$$\mathscr{A}_0(x) := 1, \quad \mathscr{A}_1(x) := \mathscr{V}_1(x) - 1, \quad \mathscr{A}_s(x) := \mathscr{V}_{2^{s-1}}(x) - \mathscr{V}_{2^{s-2}}(x), \quad s \ge 2,$$

where the $\mathscr{V}_m$ are the de la Vallée Poussin kernels (see §1.2.3). Then $\mathscr{A}_s \in \mathscr{T}(2^s)$ and by (1.2.13),

$$\|\mathscr{A}_s\|_1 \le 6. \tag{1.3.2}$$

Let $r \ge 0$ and α be real numbers. We consider the polynomials

$$\mathscr{V}_n^r(x,\alpha) := 1 + 2\sum_{k=1}^{n} k^r \cos(kx + \alpha\pi/2)$$
$$+ 2\sum_{k=n+1}^{2n-1} k^r \big(1 - (k-n)/n\big) \cos(kx + \alpha\pi/2).$$

Let us prove that, for all $r > 0$ and α,

$$\big\|\mathscr{V}_n^r(x,\alpha)\big\|_1 \le C(r)n^r, \qquad n = 1, 2, \dots \tag{1.3.3}$$

Since for an arbitrary α

$$\mathscr{V}_n^r(x,\alpha) - 1 = \big(\mathscr{V}_n^r(x,0) - 1\big)\cos(\alpha\pi/2) + \big(\mathscr{V}_n^r(x,1) - 1\big)\sin(\alpha\pi/2),$$

it suffices to prove (1.3.3) for $\alpha = 0$ and for $\alpha = 1$. We first consider the case $\alpha = 0$. Let v_k be the Fourier cosine coefficients of the function $\mathscr{V}_n^r(x,0)$. Then, by Theorem 1.3.1,

$$\big\|\mathscr{V}_n^r(x,0)\big\|_1 \le \sum_{k=0}^{2n-1} (k+1)|\Delta^2 v_k|. \tag{1.3.4}$$

It is easy to see that, for $1 \le k \le n-2$,

$$|\Delta^2 v_k| \le C(r)k^{r-2}. \tag{1.3.5}$$

By the identity

$$\Delta^2(a_k b_k) = (\Delta^2 a_k)b_k + 2(\Delta a_{k+1})(\Delta b_k) + a_{k+2}(\Delta^2 b_k)$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

with $a_k = k^r$ and $b_k = 1-(k-n)/n$, we see that the inequality (1.3.5) will be valid for $n \le k \le 2n-3$ too. For the remaining values of $k \neq 0$ we have

$$|\Delta^2 v_k| \le |\Delta v_k| + |\Delta v_{k+1}| \le C(r)n^{r-1}. \tag{1.3.6}$$

From the inequality $|\Delta^2 v_0| \le C(r)$ and relations (1.3.4)–(1.3.6) we get the relation (1.3.3) for $r > 0$ and $\alpha = 0$.

Let $\alpha = 1$ and let $\tilde{\mathscr{A}}_s(x)$ denote the polynomial which is the trigonometric conjugate to $\mathscr{A}_s(x)$, which means that in the expression for $\mathscr{A}_s(x)$ the functions $\cos kx$ are substituted by $\sin kx$. We prove that

$$\|\tilde{\mathscr{A}}_s\|_1 \le C. \tag{1.3.7}$$

Clearly, it suffices to consider $s \ge 3$. It is not difficult to see that the equality

$$\tilde{\mathscr{A}}_s(x) = 2\,\mathrm{Im}\Big(\mathscr{A}_s(x) * \Big(\big(4\mathscr{K}_{2^{s-1}-1}(x) - 3\mathscr{K}_{2^{s-1}-2^{s-3}-1}(x)\big)e^{i(2^{s-1}+2^{s-3})x}\Big)\Big),$$

holds. From this equality, by virtue of the Young inequality with $p = q = a = 1$ (see A.1.16)) and the properties of the functions $\mathscr{K}_n$ and $\mathscr{A}_s$, we obtain (1.3.7).

Further, for $n = 2^m$, we have

$$\begin{aligned}\mathscr{V}_n^r(x,1) - 1 &= \big(\mathscr{V}_{2n}^r(x,0) - 1\big) * \mathscr{V}_n^0(x,1)\\ &= -\sum_{s=1}^{m+1} \mathscr{V}_{2n}^r(x,0) * \tilde{\mathscr{A}}_s(x) = -\sum_{s=1}^{m+1} \mathscr{V}_{2^s}^r(x,0) * \tilde{\mathscr{A}}_s(x).\end{aligned} \tag{1.3.8}$$

From (1.3.8) by means of the Young inequality and using (1.3.7) and relation (1.3.3), which has been proved for $\alpha = 0$, we get

$$\big\|\mathscr{V}_n^r(x,\alpha)\|_1 \le C(r)\sum_{s=0}^{m+1} 2^{rs} \le C(r)n^r. \tag{1.3.9}$$

Now let $2^{m-1} \le n < 2^m$; then

$$\mathscr{V}_n^r(x,1) = \mathscr{V}_{2^{m+1}}^r(x,1) * \mathscr{V}_n(x),$$

which by (1.3.9) and the Young inequality gives the required estimate for all n. Relation (1.3.3) is proved.

We define the operator D_α^r, $r \ge 0$, $\alpha \in \mathbb{R}$, on the set of trigonometric polynomials as follows. Let $f \in \mathscr{T}(n)$; then

$$D_\alpha^r f := f^{(r)}(x,\alpha) := f(x) * \mathscr{V}_n^r(x,\alpha), \tag{1.3.10}$$

and $f^{(r)}(x,\alpha)$ is called the (r,α) derivative. It is clear that for $f(x)$ such that $\hat{f}(0) = 0$ we have for natural numbers r,

$$D_r^r f = \frac{d^r}{dx^r} f.$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

The operator D^r_α is defined in such a way that it has an inverse for each $\mathscr{T}(n)$. This property distinguishes D^r_α from the differential operator and it will be convenient for us. On the other hand it is clear that

$$\frac{d^r f}{dx^r} = D^r_r f - \hat{f}(0).$$

Theorem 1.3.2 *For any $t \in \mathscr{T}(n)$ we have, for $r > 0$, $\alpha \in \mathbb{R}$, $1 \le p \le \infty$,*

$$\left\| t^{(r)}(x,\alpha) \right\|_p \le C(r) n^r \|t\|_p, \qquad n = 1, 2, \ldots$$

Proof By the definition (1.3.10),

$$t^{(r)}(x,\alpha) = t(x) * \mathscr{V}^r_n(x,\alpha).$$

Therefore, by the Young inequality (A.1.16) with $p = q$, $a = 1$ for all $1 \le p \le \infty$ and r we have

$$\left\| t^{(r)}(x,\alpha) \right\|_p \le \|t\|_p \left\| \mathscr{V}^r_n(x,\alpha) \right\|_1.$$

To conclude the proof we just use inequality (1.3.3). □

Let us discuss the case $r = 0$, which is excluded from Theorem 1.3.2. In the case where $r = 0$ and α is an even integer we have

$$\left| t^{(0)}(x,\alpha) \right| = \left| t(x) \right|$$

and, consequently,

$$\left\| t^{(0)}(x,\alpha) \right\|_p = \|t\|_p, \qquad 1 \le p \le \infty. \tag{1.3.11}$$

To investigate the general case it suffices to study the trigonometric conjugate operator. Theorem 1.2.1 and its corollary show that for all α and $1 < p < \infty$ the inequality

$$\left\| t^{(0)}(x,\alpha) \right\| \le C(p) \|t\|_p$$

holds.

It remains to consider the cases $p = 1, \infty$. It is sufficient to consider $\alpha = 1$. We have for $t \in \mathscr{T}(n)$,

$$t^{(0)}(x,1) = \hat{t}(0) - \tilde{t}(x) = \hat{t}(0) - t(x) * \widetilde{\mathscr{D}}_{2n+1}(x).$$

Further,

$$\widetilde{\mathscr{D}}_{2n+1}(x) = 2 \sum_{k=1}^{2n+1} \sin kx = 2 \operatorname{Im} \mathscr{D}_n(x) e^{i(n+1)x};$$

consequently,

$$\|\widetilde{\mathscr{D}}_{2n+1}\|_1 \le C \ln(n+2).$$

Thus, for $t \in \mathscr{T}(n)$,

$$\left\|t^{(0)}(x,1)\right\|_p \le C\ln(n+2)\|t\|_p, \qquad p=1,\infty. \tag{1.3.12}$$

The relation (1.3.11) with $\alpha = 0$ and (1.3.12) imply for all α the inequality

$$\left\|t^{(0)}(x,\alpha)\right\|_p \le C\ln(n+2)\|t\|_p, \qquad p=1,\infty. \tag{1.3.13}$$

Remark 1.3.3 We have the relation

$$\sup_{t\in\mathscr{T}(n)} \|t^{(0)}(x,1)\|_p \,/\, \|t\|_p \asymp \ln(n+2), \qquad p=1,\infty.$$

The upper estimate follows from (1.3.12). Let us prove the lower estimate. We first consider the case $p=\infty$. Let $f(x) = (\pi - x)/2$, $0 < x < 2\pi$, be a 2π-periodic function; then

$$f(x) = \sum_{k=1}^{\infty} (\sin kx)/k.$$

Let $m = [n/2]$. Then

$$t(x) := f(x) * \mathscr{V}_m(x)$$

has the following properties: $t \in \mathscr{T}(n)$,

$$\|t\|_\infty \le 3\pi/2, \qquad t^{(0)}(0,1) \ge \sum_{k=1}^{m} 1/k \ge C\ln(m+2), \tag{1.3.14}$$

which imply the required lower estimate in the case $p=\infty$.

Let $p=1$ and $m=[n/2]$. Then the function $\mathscr{V}_m \in \mathscr{T}(n)$ has the following properties:

$$\|\mathscr{V}_m\|_1 \le 3, \tag{1.3.15}$$

$$\left\|\mathscr{V}_m^{(0)}(x,1)\right\|_1 \ge C\ln(m+2). \tag{1.3.16}$$

Let us prove (1.3.16). For t we have from the above consideration for $p=\infty$,

$$\sigma = |\langle \mathscr{V}_m^{(0)}(x,1), t\rangle| \le \left\|V_m^{(0)}(x,1)\right\|_1 \|t\|_\infty \tag{1.3.17}$$

and

$$\sigma \ge \sum_{k=1}^{m} 1/k \ge C\ln(m+2). \tag{1.3.18}$$

From relations (1.3.14), (1.3.17) and (1.3.18) we obtain (1.3.16). Then (1.3.15) and (1.3.16) give the required lower estimate for $p=1$.

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

1.3.2 The Nikol'skii Inequality

Let us now prove the Nikol'skii inequality.

Theorem 1.3.4 *For any $t \in \mathscr{T}(n)$, $n > 0$, we have the inequality*

$$\|t\|_p \le Cn^{1/q-1/p}\|t\|_q, \qquad 1 \le q < p \le \infty.$$

Proof First let $p = \infty$; then

$$t = t * \mathscr{V}_n$$

and by the Hölder inequality (A.1.1) we have

$$\|t\|_\infty \le \|t\|_q \|\mathscr{V}_n\|_{q'},$$

which, by (1.2.14), implies that

$$\|t\|_\infty \le C\|t\|_q n^{1/q}. \tag{1.3.19}$$

Further, let $q < p < \infty$. Then by (A.1.6),

$$\|t\|_p \le \|t\|_q^{q/p}\|t\|_\infty^{1-q/p}. \tag{1.3.20}$$

The theorem follows from relations (1.3.19) and (1.3.20). $\square$

We now formulate a corollary of Theorems 1.3.2 and 1.3.4.

Corollary 1.3.5 (The Bernstein–Nikol'skii inequality) *For $t \in \mathscr{T}(n)$ and arbitrary $r > 0$, α, $1 \le q \le p \le \infty$, we have the inequality*

$$\left\|t^{(r)}(x,\alpha)\right\|_p \le C(r)n^{r+1/q-1/p}\|t\|_q, \qquad n = 1,2,\dots$$

1.3.3 The Marcinkiewicz Theorem

The set $\mathscr{T}(n)$ of trigonometric polynomials is a space of dimension $2n+1$. Each polynomial $t \in \mathscr{T}(n)$ is uniquely defined by its Fourier coefficients $\left\{\hat{t}(k)\right\}_{|k|\le n}$, and by the Parseval identity we have

$$\|t\|_2^2 = \sum_{|k|\le n} \left|\hat{t}(k)\right|^2, \tag{1.3.21}$$

which means that the set $\mathscr{T}(n)$ as a subspace of L_2 is isomorphic to ℓ_2^{2n+1}. Relation (1.2.6) shows that a similar isomorphism can be set up in another way: by mapping a polynomial $t \in \mathscr{T}(n)$ to the vector $m(t) := \left\{t(x^l)\right\}_{l=0}^{2n}$ of its values at the points

$$x^l := 2\pi l/(2n+1), \qquad l = 0,\dots,2n.$$

Relation (1.2.6) gives

$$\|t\|_2 = (2n+1)^{-1/2}\left\|m(t)\right\|_2.$$

The following statement is the Marcinkiewicz theorem.

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

Theorem 1.3.6 *Let $1 < p < \infty$; then for $t \in \mathscr{T}(n)$, $n > 0$, we have the relation*

$$C_1(p)\|t\|_p \le n^{-1/p}\big\|m(t)\big\|_p \le C_2(p)\|t\|_p.$$

Proof We first prove a lemma.

Lemma 1.3.7 *Let $1 \le p \le \infty$; then, for $n > 0$,*

$$\left\|\sum_{l=0}^{2n} a_l \mathscr{V}_n(x-x^l)\right\|_p \le Cn^{1-1/p}\|\mathbf{a}\|_{\ell_p^{2n+1}}, \qquad \mathbf{a} := (a_0,\dots,a_{2n}).$$

Proof Let V be an operator on ℓ_p^{2n+1} defined as follows:

$$V(\mathbf{a}) := \sum_{l=0}^{2n} a_l \mathscr{V}_n(x-x^l).$$

It is obvious that (see (1.2.13))

$$\|V\|_{\ell_1^{2n+1}\to L_1} \le 3. \tag{1.3.22}$$

Using the estimate (see (1.2.12))

$$\big|\mathscr{V}_n(x)\big| \le C\min\big(n,(nx^2)^{-1}\big)$$

it is not hard to prove that

$$\|V\|_{\ell_\infty^{2n+1}\to L_\infty} \le Cn. \tag{1.3.23}$$

From relations (1.3.22) and (1.3.23), using the Riesz–Torin theorem (see Theorem A.3.2) we find that

$$\|V\|_{\ell_p^{2n+1}\to L_p} \le Cn^{1-1/p},$$

which implies the lemma. □

We now continue the proof of Theorem 1.3.6. Let S_n be the operator that takes the partial Fourier sum of order n. Using Theorem 1.2.1 we derive from Lemma 1.3.7 the upper estimate (the first inequality in Theorem 1.3.6):

$$\begin{aligned} t(x) &= (2n+1)^{-1}\sum_{l=0}^{2n} t(x^l)\mathscr{D}_n(x-x^l) \\ &= S_n\left((2n+1)^{-1}\sum_{l=0}^{2n} t(x^l)\mathscr{V}_n(x-x^l)\right). \end{aligned}$$

Consequently,

$$\|t\|_p \le C(p)n^{-1/p}\big\|m(t)\big\|_p.$$

We now prove the lower estimate (the second inequality in Theorem 1.3.6) for $1 \le p < \infty$. We have

$$\|m(t)\|_p^p = \sum_{l=0}^{2n} \left|t(x^l)\right|^p = \sum_{l=0}^{2n} t(x^l)\varepsilon_l \left|t(x^l)\right|^{p-1}$$
$$= (2\pi)^{-1} \int_0^{2\pi} t(x) \sum_{l=0}^{2n} \varepsilon_l \left|t(x^l)\right|^{p-1} \mathscr{V}_n(x - x^l)dx$$
$$\le \|t\|_p \left\| \sum_{l=0}^{2n} \varepsilon_l \left|t(x^l)\right|^{p-1} \mathscr{V}_n(x - x^l) \right\|_{p'},$$

using Lemma 1.3.7 we see that the last expression is

$$\le C\|t\|_p n^{1/p} \|m(t)\|_p^{p-1},$$

which implies the required lower estimate and the theorem is proved. □

Remark 1.3.8 In the proof of Theorem 1.3.6 we also proved the inequality

$$\|m(t)\|_1 \le Cn\|t\|_1.$$

We now prove a statement that is analogous to Theorem 1.3.6 but, in contrast to it, includes the cases $p = 1$ and $p = \infty$. Instead of the vector $m(t)$ we now consider the vector

$$M(t) := \big(t\big(x(1)\big), \dots, t\big(x(4n)\big)\big), \qquad x(l) := \pi l/(2n), \qquad l = 1, \dots, 4n.$$

Theorem 1.3.9 *For an arbitrary $t \in \mathscr{T}(n)$, $n > 0$, $1 \le p \le \infty$, we have*

$$C_1\|t\|_p \le n^{-1/p}\|M(t)\|_p \le C_2\|t\|_p.$$

Proof In the same way as for Lemma 1.3.7 one can prove:

Lemma 1.3.10 *Let $1 \le p \le \infty$, then, for $n > 0$,*

$$\left\| \sum_{l=1}^{4n} a_l \mathscr{V}_n\big(x - x(l)\big) \right\|_p \le Cn^{1-1/p}\|\mathbf{a}\|_{\ell_p^{4n}}.$$

Lemma 1.3.10 with $\mathbf{a} = M(t)$ and relation (1.2.15) implies the upper estimate

$$\|t\|_p \le Cn^{-1/p}\|M(t)\|_p.$$

The corresponding lower estimate for $1 \le p < \infty$ can be proved in the same way as above for $m(t)$, substituting x^l by $x(l)$.

The lower estimate for $p = \infty$ is obvious. □

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

1.4 Approximation of Functions in the Classes $W^r_{q,\alpha}$ and H^r_q

1.4.1 Some Properties of the Bernoulli Kernels

For $r > 0$ and $\alpha \in \mathbb{R}$ the functions

$$F_r(x,\alpha) = 1 + 2\sum_{k=1}^{\infty} k^{-r}\cos(kx - \alpha\pi/2)$$

are called *Bernoulli kernels*.

We define the following operator in the space L_1,

$$(I^r_\alpha \phi)(x) := (2\pi)^{-1}\int_0^{2\pi} F_r(x-y,\alpha)\phi(y)dy. \tag{1.4.1}$$

Let us prove that the definition of this operator is reasonable. To establish this it suffices to prove that $F_r \in L_1$.

Theorem 1.4.1 *For $r > 0$, $\alpha \in \mathbb{R}$ we have*

$$F_r \in L_1, \qquad E_n(F_r)_1 \le C(r)(n+1)^{-r}, \qquad n = 0,1,\dots$$

Proof Let us consider the functions

$$f^r_s(x,\alpha) := \mathscr{A}_s(x) * \left(1 + 2\sum_{k=1}^{2^s} k^{-r}\cos(kx - \alpha\pi/2)\right),$$

where the $\mathscr{A}_s$ are defined in §1.3.

We first consider the case $\alpha = 0$. Using Theorem 1.3.1 in the same way as in the proof of inequality (1.3.3) we get

$$\left\|f^r_s(x,0)\right\|_1 \le C(r)2^{-rs}. \tag{1.4.2}$$

Further,

$$f^r_s(x,\alpha) = D^r_{-\alpha} f^{2r}_s(x,0),$$

and, consequently, from (1.4.2) and Theorem 1.3.2 we find that

$$\left\|f^r_s(x,\alpha)\right\|_1 \le C(r)2^{-rs}. \tag{1.4.3}$$

Thus the series

$$\sum_{s=0}^{\infty} f^r_s(x,\alpha)$$

converges in L_1 to some function $f(x)$ and

$$\left\|\sum_{s=m}^{\infty} f^r_s(x,\alpha)\right\|_1 \le C(r)2^{-rm}. \tag{1.4.4}$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

From the definition of the function $f_s^r(x,\alpha)$ we get

$$S_n(f) = 1 + 2\sum_{k=1}^{n} k^{-r}\cos(kx - \alpha\pi/2)$$

and

$$\begin{aligned}\|f - S_n(f)\|_1 &\le \sum_{s=0}^{\infty} \|f_s^r(x,\alpha) - S_n(f_s^r(x,\alpha))\|_1 \\ &\le \sum_{s:2^s>n} \|f_s^r - S_n(f_s^r)\|_1 \le C\ln(n+2)\sum_{2^s>n}\|f_s^r\|_1 \\ &\le C(r)n^{-r}\ln(n+2). \qquad (1.4.5)\end{aligned}$$

Here we have used Theorem 1.2.1 and relation (1.4.3). Relation (1.4.5) shows that the series defining the function $F_r(x,\alpha)$ converges in L_1 to $f(x)$. The first part of the theorem is proved. The second part of the theorem follows from relation (1.4.4). □

We now proceed to formulate some properties of the operators D_α^r and I_α^r. From the equality ($\phi \in L_1$)

$$\int_0^{2\pi}\left(\pi^{-1}\int_0^{2\pi}\phi(u)\cos\big(k(y-u)+\alpha\pi/2\big)\cos\big(k(x-y)+\beta\pi/2\big)dy\right)du$$
$$= \int_0^{2\pi}\phi(u)\cos\big(k(x-u)+(\alpha+\beta)\pi/2\big)du,$$

which is valid for any nonzero k, the equalities

$$D_{\alpha_1}^{r_1}D_{\alpha_2}^{r_2} = D_{\alpha_1+\alpha_2}^{r_1+r_2}, \qquad (1.4.6)$$
$$I_{\alpha_1}^{r_1}I_{\alpha_2}^{r_2} = I_{\alpha_1+\alpha_2}^{r_1+r_2}, \qquad (1.4.7)$$
$$D_\alpha^r I_\alpha^r = I_\alpha^r D_\alpha^r = I \qquad (1.4.8)$$

follow (we assume that the operators act on a set of trigonometric polynomials).

Denote by $W_{q,\alpha}^r B$, $r > 0$, $\alpha \in \mathbb{R}$, $1 \le q \le \infty$, the class of functions $f(x)$ representable in the form

$$f = I_\alpha^r\phi, \qquad \|\phi\|_q \le B. \qquad (1.4.9)$$

For such functions, with some q and B. we define (see (1.4.8))

$$D_\alpha^r f = \phi.$$

Let $1 < q < p < \infty$, $\beta := 1/q - 1/p$. From Corollary A.3.8 of the Hardy–Littlewood inequality (see the Appendix) and the boundedness of the trigonometric conjugate operator as an operator from L_p to L_p for $1 < p < \infty$ (see Corollary 1.2.2), it follows that

$$\|I_\alpha^\beta\|_{q\to p} \le C(q,p). \qquad (1.4.10)$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

Relations (1.4.7) and (1.4.10) imply the following embedding theorem.

Theorem 1.4.2 *Let $1 < q < p < \infty$, $\beta = 1/q - 1/p$, $r > \beta$; then*

$$W^r_{q,\alpha_1} \subset W^{r-\beta}_{p,\alpha_2} B, \qquad \alpha_1,\ \alpha_2 \in \mathbb{R}.$$

1.4.2 Approximation for Smoothness Classes

Let us define the classes $H^r_q B$, $r > 0$, $1 \le q \le \infty$ as follows:

$$H^r_q B := \left\{ f \in L_q : \|f\|_q \le B,\ \left\|\Delta^a_t f(x)\right\|_q \le B|t|^r,\ a = [r] + 1 \right\},$$
$$\Delta_t f(x) := f(x) - f(x+t), \qquad \Delta^a_t := (\Delta_t)^a.$$

For the case $B = 1$ we simply write $H^r_q := H^r_q 1$, i.e., we drop the constant B.

Let us study these classes from the point of view of their approximation by trigonometric polynomials.

Theorem 1.4.3 *Let $r > 0$, $1 \le q \le \infty$, then*

$$E_n(H^r_q)_q \asymp (n+1)^{-r}, \qquad n = 0, 1, \ldots$$

Proof Let us prove the upper estimate. Clearly, it suffices to consider the case $n > 0$. Let $f \in H^r_q$. We consider (see §1.2.4)

$$t(x) := (2\pi)^{-1} \int_{-\pi}^{\pi} \left(f(x) - \Delta^a_y f(x) \right) J^a_n(y) dy.$$

Then $t \in \mathscr{T}(an)$ and

$$f(x) - t(x) = (2\pi)^{-1} \int_{-\pi}^{\pi} \Delta^a_y f(x) J^a_n(y) dy.$$

By a generalization of the Minkowskii inequality, (A.1.9), we have

$$\|f - t\|_q \le (2\pi)^{-1} \int_{-\pi}^{\pi} \left\|\Delta^a_y f(x)\right\|_q J^a_n(y) dy,$$

which by the definition of the class H^r_q and relation (1.2.19) implies that

$$\|f - t\|_q \le C(r) n^{-r}.$$

The upper estimate is proved.

We now prove the lower estimate. We construct functions which will be used in the proof of the more general Theorem 1.4.9. Let $n > 0$ be given and s be such that

$$4n \le 2^s \le 8n.$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

We consider

$$f(x) := 2^{-(r+1-1/q)s}\mathscr{A}_s(x) \tag{1.4.11}$$

and remark that to prove the theorem it suffices to consider the simpler function $f(x) = (n+1)^{-r}e^{i(n+1)x}$. Then, for any $t \in \mathscr{T}(n)$, we have on the one hand

$$\langle f-t, \mathscr{A}_s\rangle = \langle f, \mathscr{A}_s\rangle = 2^{-(r+1-1/q)s}\|\mathscr{A}_s\|_2^2 \geq C2^{-(r-1/q)s}. \tag{1.4.12}$$

On the other hand using the definition of $\mathscr{A}_s$ and (1.2.14) we get

$$\langle f-t, \mathscr{A}_s\rangle \leq \|f-t\|_q\|\mathscr{A}_s\|_{q'} \leq C2^{s/q}\|f-t\|_q. \tag{1.4.13}$$

From relations (1.4.12) and (1.4.13) we obtain

$$E_n(f)_q \geq C2^{-rs} \geq Cn^{-r}.$$

To show that $f \in H_q^r B$, we prove the following auxiliary statement.

Lemma 1.4.4 *Let $g(x)$ be an a-times continuously differentiable 2π-periodic function. Then for all $1 \leq q \leq \infty$ we have*

$$\left\|\Delta_y^a g(x)\right\|_q \leq |y|^a \left\|g^{(a)}(x)\right\|_q.$$

Proof Clearly it suffices to consider the case $a = 1$. We have

$$\left\|\Delta_y g(x)\right\|_q = \left\|\int_x^{x+y} g'(u)du\right\|_q = \left\|\int_0^y g'(x+u)du\right\|_q \leq |y|\|g'\|_q,$$

as required. □

From (1.4.11), (1.2.14), and the Bernstein inequality (Theorem 1.3.2) we get

$$\|f^{(a)}\|_q \leq C(a)2^{(a-r)s}. \tag{1.4.14}$$

Using Lemma 1.4.4 and the simple inequality

$$\left\|\Delta_y^a f(x)\right\|_q \leq 2^a\|f\|_q,$$

we obtain

$$\left\|\Delta_y^a f(x)\right\|_q \leq C(a)\min\left(|y|^a n^{a-r}, n^{-r}\right), \tag{1.4.15}$$

which implies that $f \in H_q^r B$ with some B that is independent of n, and this proves the lower estimate. □

Let us now prove a representation theorem for the class $H_q^r B$. Let

$$A_s(f) := \mathscr{A}_s * f$$

and denote the value of $A_s(f)$ at a point x by $A_s(f,x)$.

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

Theorem 1.4.5 *Let $f \in L_q$, $1 \le q \le \infty$, $\|f\|_q \le 1$. For $\|\Delta_t^a f\|_q \le |t|^r$, $a = [r]+1$ it is necessary and sufficient that the following conditions be satisfied:*

$$\left\|A_s(f)\right\|_q \le C(r,q)2^{-rs}, \qquad s = 0,1,\dots .$$

(*The constants $C(r,q)$ may be different for the cases of necessity and sufficiency.*)

Proof

Necessity. Let $f \in H^r_q$; then for any $t_s \in \mathscr{T}(2^{s-2})$, $s \ge 2$ we have

$$A_s(f) = A_s(f - t_s)$$

and

$$\|A_s(f)\|_q \le \|\mathscr{A}_s\|_1 \|f - t_s\|_q.$$

Applying Theorem 1.4.3 and using relation (1.3.2) we get

$$\left\|A_s(f)\right\|_q \le C(r,q)2^{-rs}.$$

Sufficiency. Let

$$\left\|A_s(f)\right\|_q \le \gamma 2^{-rs}, \tag{1.4.16}$$

then using Corollary 2.2.7 we get

$$f = \sum_{s=0}^{\infty} A_s(f),$$

in the sense of convergence in L_q, and

$$\|\Delta_t^a f\|_q \le \sum_{s=1}^{\infty} \left\|\Delta_t^a A_s(f)\right\|_q. \tag{1.4.17}$$

From Lemma 1.4.4 we find, in the same way as in (1.4.15),

$$\left\|\Delta_t^a A_s(f)\right\|_q \le C(a)2^{-rs}\min\left(1, \left(|t|2^s\right)^a\right). \tag{1.4.18}$$

From (1.4.17) and (1.4.18) we obtain

$$\|\Delta_t^a f\|_q \le C(r)\gamma |t|^r,$$

which concludes the proof of the theorem if we take $\gamma < 1/C(r)$. □

Denote

$$\delta_0(f) := S_0(f), \qquad \delta_s(f) := S_{2^s-1}(f) - S_{2^{s-1}-1}(f), \qquad s = 1,2,\dots$$

Corollary 1.4.6 *In the case $1 < q < \infty$ the functions $A_s(f)$ in Theorem 1.4.5 can be replaced by $\delta_s(f)$.*

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

Proof For $1 < q < \infty$ the conditions

(1) $\|A_s(f)\|_q \le C(q)2^{-rs}$,
(2) $\|\delta_s(f)\|_q \le C(q)2^{-rs}$

are equivalent for all s. Indeed,

$$A_s(f) = \mathscr{A}_s * \big(\delta_{s-1}(f) + \delta_s(f)\big),$$
$$\delta_s(f) = \delta_s\big(A_s(f) + A_{s+1}(f)\big),$$

which by (1.3.2) and the boundedness of the operator δ_s as an operator from L_q to L_q, $1 < q < \infty$ (see Corollary A.3.4) implies the equivalence of conditions (1) and (2). □

Corollary 1.4.7 *Let* $1 \le q \le \infty$, $\|f\|_q \le 1$ *and*

$$E_n(f)_q \ll (n+1)^{-r}, \qquad n = 0, 1, \dots;$$

then $f \in H^r_q B$ *for some* B.

Indeed, in the same way as in the proof of the necessity in Theorem 1.4.5 we get

$$\|A_s(f)\|_q \ll 2^{-rs},$$

which by Theorem 1.4.5 (regarding the sufficiency) implies that $f \in H^r_q B$.

Statements of the type of Theorem 1.4.3 are called *direct* theorems of approximation theory, and statements of the type of Corollary 1.4.7 are called *inverse* theorems of approximation theory.

Theorem 1.4.1 and Corollary 1.4.7 imply that

$$F_r(x,\alpha) \in H^r_1 B. \tag{1.4.19}$$

Consequently, for $f \in W^r_{q,\alpha}$ we have

$$\|\Delta^a_t f(x)\|_q \le \|\Delta^a_t F_r(x,\alpha)\|_1 \|D^r_\alpha f\|_q \le B|t|^r;$$

that is, $f \in H^r_q B$.

Thus, we have proved that

$$W^r_{q,\alpha} \subset H^r_q B. \tag{1.4.20}$$

Let us prove an embedding theorem for the H classes.

Theorem 1.4.8 *Let* $1 \le q \le p \le \infty$, $\beta := 1/q - 1/p$, $r > \beta$. *We have the inclusion*

$$H^r_q \subset H^{r-\beta}_p B$$

(in the case $p = \infty$ *this means that for any* $f \in H^r_q$ *there is an equivalent* $g \in H^{r-\beta}_\infty B$*).*

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

Proof Let $f \in H^r_q$. By Theorem 1.4.5

$$\|A_s(f)\|_q \le C(r,q)2^{-rs}.$$

Therefore, by the Nikol'skii inequality (Theorem 1.3.4) we have

$$\|A_s(f)\|_p \le C(r,q)2^{-(r-\beta)s}. \tag{1.4.21}$$

Let $g(x)$ denote the sum of the series $\sum_{s=0}^{\infty} A_s(f,x)$ in the sense of convergence in L_p. From Corollary 2.2.7 below it follows that f and g are equivalent. From (1.4.21) and the equality $A_s(f) = A_s(g)$, by Theorem 1.4.5 we obtain $g \in H_p^{r-\beta}B$.

The theorem is proved. □

With the aid of Theorem 1.4.8 we can prove the following statement.

Theorem 1.4.9 *Let* $1 \le q,\ p \le \infty$, $r > (1/q - 1/p)_+$. *Then*

$$E_n(W^r_{q,\alpha})_p \asymp E_n(H^r_q)_p \asymp n^{-r+(1/q-1/p)_+}.$$

Proof By relation (1.4.20) it suffices to prove the upper estimate for the H classes and the lower estimate for the W classes. We first prove the upper estimate. Let $1 \le q \le p \le \infty$. Then Theorems 1.4.8 and 1.4.3 give

$$E_n(H^r_q)_p \ll n^{-r+1/q-1/p}. \tag{1.4.22}$$

For $1 \le p < q \le \infty$ we have, by the monotonicity of the L_p-norms and Theorem 1.4.3,

$$E_n(H^r_q)_p \le E_n(H^r_q)_q \ll n^{-r}.$$

From this and relation (1.4.22) the required upper estimates follow.

Let us prove the lower estimate. Let n and s be the same as in the proof of the lower estimate in Theorem 1.4.3 and let f be defined by (1.4.11). Then by the Bernstein inequality,

$$\|D^r_\alpha f\|_q \le C(r),$$

and $f \in W^r_{q,\alpha}C(r)$.

Let $1 \le q \le p \le \infty$. From relation (1.4.12) and relation (1.4.13) with p instead of q we get

$$E_n(f)_p \ge Cn^{-r+1/q-1/p}. \tag{1.4.23}$$

For $1 \le p \le q \le \infty$ it suffices to consider as an example

$$f(x) = 2(n+1)^{-r}\cos(n+1)x.$$

Then $f \in W^r_{\infty,\alpha}$ and, for any $t \in \mathscr{T}(n)$,

$$\sigma = \langle f(x) - t(x),\ \cos(n+1)x\rangle = (n+1)^{-r}, \qquad \sigma \le \|f-t\|_1,$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

which implies the estimate

$$E_n(W^r_{\infty,\alpha})_1 \geq (n+1)^{-r}. \tag{1.4.24}$$

The required lower estimates follow from (1.4.23) and (1.4.24) and the theorem is proved. □

Remark 1.4.10 Theorem 1.2.3 implies that for any $f \in L_p$ the de la Vallée Poussin inequality holds:

$$\|f - V_n(f)\|_p \leq 4E_n(f)_p, \quad 1 \leq p \leq \infty. \tag{1.4.25}$$

This inequality and Theorem 1.4.9 show that, for all $1 \leq q, p \leq \infty$,

$$V_n(H^r_q)_p := \sup_{f \in H^r_q} \|f - V_n(f)\|_p \asymp E_{2n}(H^r_q)_p, \tag{1.4.26}$$

and an analogous relation is valid for the W classes.

Thus, for the classes $W^r_{q,\alpha}$ and H^r_q there exist linear methods giving an approximation of the same order as the best approximation.

Remark 1.4.11 From Theorem 1.2.1 it follows that for all $1 < p < \infty$ and $f \in L_p$,

$$\|f - S_n(f)\|_p \leq C(p)E_n(f)_p. \tag{1.4.27}$$

Consequently, if we are interested only in the dependence of the approximation of a function $f \in L_p$ on n then it suffices, in the case $1 < p < \infty$, to consider the simplest method of approximation, namely, the Fourier method.

This remains true for the classes $W^r_{q,\alpha}$ and H^r_q for all $1 \leq q, p \leq \infty$, excepting the cases $q = p = 1$ and $q = p = \infty$. For the function class F let us denote

$$S_n(F)_p := \sup_{f \in F} \|f - S_n(f)\|_p.$$

Theorem 1.4.12 *Let $1 \leq q,\ p \leq \infty$, $(q,p) \neq (1,1)$ or (∞,∞), and $r > (1/q - 1/p)_+$. Then*

$$S_n(W^r_{q,\alpha})_p \asymp S_n(H^r_q)_p \asymp n^{-r+(1/q-1/p)_+}.$$

Proof In the case $1 < p < \infty$ the theorem follows from Theorem 1.4.9 and relation (1.4.27). It remains to consider the cases $p = 1$, $q > 1$ and $1 \leq q < p = \infty$. In the case $p = 1$, $q > 1$ we have

$$S_n(H^r_q)_1 \leq S_n(H^r_{q^*})_{q^*} \ll n^{-r},$$

where $q^* = \min(q,2)$.

Now let $1 \leq q < p = \infty$. In the case $1 \leq q < 2$, by Theorem 1.4.8 we have

$$H^r_q \subset H_2^{r-(1/q-1/2)}B,$$

which indicates that it suffices to consider the case $2 \le q < \infty$. In this case by Theorem 1.2.1 and Corollary 1.4.6 we have for $s > s_n$, where s_n is such that $2^{s_n-1} \le n < 2^{s_n}$,

$$\|\delta_s(f)\|_q \le C(r,q)2^{-rs},$$
$$\|\delta_{s_n}(f) - S_n(\delta_{s_n}(f))\|_q \le C(r,q)2^{-rs_n}.$$

From these inequalities, using the Nikol'skii inequality, we get

$$\begin{aligned}\|f - S_n(f)\|_\infty &\le \|\delta_{s_n}(f) - S_n(\delta_{s_n}(f))\|_\infty + \sum_{s>s_n} \|\delta_s(f)\|_\infty \\ &\le C(r,q) \sum_{s\ge s_n} 2^{-(r-1/q)s} \le C(r,q)n^{-r+1/q},\end{aligned}$$

which concludes the proof of the theorem. □

We proceed to the cases $q = p = 1$ or ∞, which were excluded in Theorem 1.4.12. For these cases we obtain from Theorem 1.2.1 the following *Lebesgue inequality*: for $f \in L_p$, $p = 1$, or ∞,

$$\|f - S_n(f)\|_p \le C(\ln n)E_n(f)_p, \qquad n = 2,3,\ldots \tag{1.4.28}$$

Theorem 1.4.13 *Let $p = 1$, or ∞ and $r > 0$; then*

$$S_n(W^r_{p,\alpha})_p \asymp S_n(H^r_p)_p \asymp n^{-r}\ln n, \qquad n = 2,3,\ldots$$

Proof The upper estimates follow from Theorem 1.4.9 and the inequality (1.4.28). Owing to (1.4.20) it suffices to prove the lower estimates for the W classes. We first remark that

$$S_n(W^r_{1,\alpha})_1 = S_n(W^r_{\infty,-\alpha})_\infty. \tag{1.4.29}$$

Indeed (see Theorem A.2.1),

$$\begin{aligned}S_n(W^r_{1,\alpha})_1 &= \sup_{\|\phi\|_1\le 1} \|F_r(x,\alpha) * (\phi - S_n(\phi))\|_1 \\ &= \sup_{\|\phi\|_1\le 1}\ \sup_{\|\psi\|_\infty\le 1} |\langle F_r(x,\alpha) * (\phi - S_n(\phi)), \psi\rangle| \\ &= \sup_{\|\phi\|_1\le 1}\ \sup_{\|\psi\|_\infty\le 1} |\langle \phi, F_r(x,-\alpha) * (\psi - S_n(\psi))\rangle| \\ &= S_n(W^r_{\infty,-\alpha})_\infty.\end{aligned}$$

Therefore, to obtain the lower estimate it suffices to consider the case $p = 1$. Let n be given. We consider

$$f(x) := e^{inx}\mathcal{K}_{n-1}(x);$$

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

then, by the Bernstein inequality,

$$\|D^r_\alpha f\|_1 \le C(r)n^r\|\mathscr{K}_{n-1}\|_1 = C(r)n^r. \tag{1.4.30}$$

Further (see the analogous reasoning in the proof of (1.3.16)),

$$\begin{aligned}\|f - S_n(f)\|_1 &= \left\|\sum_{k=1}^{n}(1-k/n)e^{ikx}\right\|_1 \ge \left\|\sum_{k=1}^{n}(1-k/n)\sin kx\right\|_1 \\ &\ge \left(\sum_{k=1}^{n}(1-k/n)k^{-1}\right)\|\pi - x\|_\infty^{-1} \ge C\ln n. \end{aligned}\tag{1.4.31}$$

Relations (1.4.29)–(1.4.31) imply the theorem. □

1.5 Historical Remarks

In §1.1, along with classical results of Fourier, Du Bois-Reymond, and Weierstrass, which are usually included in a standard course of mathematical analysis, the following papers are cited: Chebyshev (1854), de la Vallée Poussin (1908, 1919), Bernstein (1912, 1914), Jackson (1911), Fredholm (1903), Lebesgue (1910), Kolmogorov (1936, 1985), Favard (1937), Akhiezer and Krein (1937), Ismagilov (1974), Kashin (1977), Tikhomirov (1960b), and Temlyakov (1982a).

Theorem 1.2.1 and its corollary were obtained by Riesz (see Zygmund, 1959, vol. 1). A more detailed treatment of properties of the kernels of Dirichlet, Fejér, de la Vallée Poussin, and Jackson can be found in Dzyadyk (1977). The Rudin–Shapiro polynomials were constructed in Shapiro (1951) and Rudin (1952). The polynomials $G_n(x)$ were considered in Temlyakov (1989b). The proof of relation (1.2.23) is analogous to reasoning from Trigub (1971). Theorem 1.2.5 is a classical result of Gauss. Relation (1.2.31) was obtained by Hardy and Littlewood (1966).

Theorem 1.3.1 was obtained by Kolmogorov (1985), vol. 1, pp. 12–14. Theorem 1.3.2 in the case $p = \infty$, $r = 1$, $\alpha = r$ was proved by Bernstein (1952), vol. 1, pp. 11–104. After this paper appeared, inequalities of this type began to be known as Bernstein inequalities. Today in a number of cases the Bernstein inequalities are known with explicit constants $C(r)$. Theorem 1.3.4 in the case $p = \infty$ was obtained by Jackson (1933) and in the general case by Nikol'skii (1951). Such inequalities are known as Jackson–Nikol'skii or simply Nikol'skii inequalities. Theorem 1.3.6 was obtained by Marciekiewicz (see Zygmund, 1959, vol. 2).

In a number of cases of Theorem 1.4.1 the exact values are known (see the survey Telyakovskii, 1988). Theorem 1.4.2 was proved by Hardy and Littlewood (1928). The classes H^r_q coincide with the Lipschitz classes for $0 < r < 1$ and with the Zygmund classes for $r = 1$. For r non-natural, the classes H^r_q are analogous to the

https://doi.org/10.1017/9781108689687.002 Published online by Cambridge University Press

classes $W^{[r]}H_q^{r-[r]}$. This statement follows from both direct and inverse theorems for these classes because these theorems have the same form (see Theorem 1.4.3 and Corollary 1.4.6 as well as the survey Telyakovskii, 1988). Theorem 1.4.3 for $q = \infty$ is a simple consequence of the results of Stechkin (1951). The proof in the general case $1 \leq q \leq \infty$ is carried out in the same way as in the case $q = \infty$. In fact, Theorem 1.4.5 includes both the direct and inverse theorems for the approximation of the classes $H_q^r B$. Theorem 1.4.8 was obtained by Nikol'skii (see his 1969 book). Theorem 1.4.9 is well known but it is not easy to assign priority; the situation is similar for Theorem 1.4.12. Theorem 1.4.13 is due to Lebesgue (1910) for $p = \infty$ and to Nikol'skii for $p = 1$ (see the survey Telyakovskii, 1988).