INDUSTRIAL
MATHEMATICS
INSTITUTE

2002:06

Cubature formulas and related questions

V.N. Temlyakov

IMI
Preprint Series

Department of Mathematics
University of South Carolina

CUBATURE FORMULAS AND RELATED QUESTIONS[1]

V. N. Temlyakov

Department of Mathematics, University of South Carolina, Columbia, SC 29208
E-mail: temlyak@math.sc.edu

Abstract. The main goal of this paper is to demonstrate connections between the following three big areas of research: the theory of cubature formulas (numerical intergation), the discrepancy theory, and nonlinear approximation. In Section 1 we discuss a relation between results on cubature formulas and on discrepancy. In particular, we show how standard in the theory of cubature formulas settings can be translated into the discrepancy problem and into a natural generalization of the discrepancy problem. This leads to a concept of the r-discrepancy. In Section 2 we present results on a relation between construction of an optimal cubature formula with m knots for a given function class and best nonlinear m-term approximation of a special function determined by the function class. The nonlinear m-term approximation is taken with regard to a redundant dictionary also determined by the function class. Sections 3 and 4 contain some known results on the lower and the upper estimates of errors of optimal cubature formulas for the class of functions with bounded mixed derivative. One of the important messages of this paper is that the theory of discrepancy is closely connected with the theory of cubature formulas for the classes of functions with bounded mixed derivative. We have included in the paper both new results and known results. We included some of known results with their proofs for the following two reasons. First of all we want to make the paper selfcontained (within reasonable limits). Secondly, we selected the proofs which demonstrate different methods and are not very much technically involved. Section 5 contains historical notes on discrepancy and cubature formulas, some further comments and remarks. Historical remarks on nonlinear approximation are included in Section 2. We want to point out that this paper is not a survey in any of the above mentioned areas. We did not even try to provide a complete list of results in those areas. We rather wanted to highlight the most typical results in cubature formulas (Sections 3 and 4) and show their relation to the discrepancy theory.

1. Cubature formulas and discrepancy

Numerical integration seeks good ways of approximating an integral

$$\int_\Omega f(x)d\mu$$

[1]This research was supported by the National Science Foundation Grant DMS 9970326 and by ONR Grant N00014-96-1-1003

by an expression of the form

$$\Lambda_m(f,\xi) := \sum_{j=1}^{m} \lambda_j f(\xi^j), \quad \xi = (\xi^1,\dots,\xi^m), \quad \xi^j \in \Omega, \quad j = 1,\dots,m. \tag{1.1}$$

It is clear that we must assume that f is integrable and defined at the points $\xi^1,\dots,\xi^m$. The expression (1.1) is called a cubature formula (Λ,ξ) (if $\Omega \subset \mathbb{R}^d$, $d \ge 2$) or a quadrature formula (Λ,ξ) (if $\Omega \subset \mathbb{R}$) with knots $\xi = (\xi^1,\dots,\xi^m)$ and weights $\Lambda = (\lambda_1,\dots,\lambda_m)$. For a function class W we introduce a concept of error of the cubature formula $\Lambda_m(\cdot,\xi)$ by

$$\Lambda_m(W,\xi) := \sup_{f\in W} \left| \int_\Omega f d\mu - \Lambda_m(f,\xi) \right|. \tag{1.2}$$

In order to orient the reader we will begin with the case of univariate periodic functions. Let for $r > 0$

$$F_r(x) := 1 + 2\sum_{k=1}^{\infty} k^{-r} \cos(kx - r\pi/2) \tag{1.3}$$

and

$$W_p^r := \{f : f = \varphi * F_r, \quad \|\varphi\|_p \le 1\} \tag{1.4}$$

where $*$ means convolution and $\|\cdot\|_p$ is the standard L_p-norm. It is well known that for $r > 1/p$ the class W_p^r is embedded into the space of continuous functions $C(\mathbb{T})$. In a particular case of W_1^1 we also have embedding into $C(\mathbb{T})$. From the definitions (1.1), (1.2), and (1.4) we see that for the normalized measure $d\mu = \frac{1}{2\pi}dx$

$$\Lambda_m(W_p^r,\xi) = \sup_{\|\varphi\|_p\le 1} \left|\frac{1}{2\pi}\int_{\mathbb{T}}\left(\int_{\mathbb{T}} F_r(x-y)d\mu - \sum_{j=1}^{m}\lambda_j F_r(\xi^j - y)\right)\varphi(y)dy\right|$$

$$= \|1 - \sum_{j=1}^{m} \lambda_j F_r(\xi^j - \cdot)\|_{p'}, \quad p' := \frac{p}{p-1}. \tag{1.5}$$

Thus the quality of the quadrature formula $\Lambda_m(\cdot,\xi)$ for the function class W_p^r is controlled by the quality of $\Lambda_m(\cdot,\xi)$ for the representing kernel $F_r(x-y)$. In the particular case of W_1^1 we have

$$\Lambda_m(W_1^1,\xi) = \max_y |1 - \sum_{j=1}^{m} \lambda_j F_1(\xi^j - y)|. \tag{1.6}$$

In this case the function

$$F_1(x) = 1 + 2\sum_{k=1}^{\infty} \sin kx = 1 + S(x)$$

has a simple form: $S(x) = 0$ for $x = l\pi$ and $S(x) = \pi - x$ for $x \in (0,2\pi)$. This allows us to associate the quantity $\Lambda_m(W_1^1,\xi)$ with the one that has simple geometrical interpretation. Denote by χ the class of all characteristic functions $\chi_{[0,a]}(x)$, $a \in [0,2\pi)$. Then we have the following property.

Proposition 1.1. *There exist two positive absolute constants C_1 and C_2 such that for any $\Lambda_m(\cdot,\xi)$ with a property $\sum_j \lambda_j = 1$ we have*

$$C_1\Lambda_m(\chi,\xi) \le \Lambda_m(W^1_1,\xi) \le C_2\Lambda_m(\chi,\xi). \tag{1.7}$$

Proof. We have for any $a \in [0,2\pi)$

$$\chi_{[0,a]}(x) = \frac{1}{2\pi}(a + F_1(x) - F_1(x-a)). \tag{1.8}$$

Thus using (1.6) we get

$$\Lambda_m(\chi,\xi) \le \frac{1}{\pi}\Lambda_m(W^1_1,\xi),$$

which proves the left inequality in (1.7).

Let us prove the right inequality in (1.7). Denote $\epsilon := \Lambda_m(\chi,\xi)$. Then by (1.8) we get for any $a \in [0,2\pi)$

$$-2\pi\epsilon \le \int_{\mathbb{T}} (F_1(x) - F_1(x-a))d\mu - \Lambda_m(F_1(x) - F_1(x-a),\xi) \le 2\pi\epsilon. \tag{1.9}$$

Intergating these inequalities against a over $\mathbb{T}$ we get

$$|\int_{\mathbb{T}} F_1(x)d\mu - \Lambda_m(F_1,\xi)| \le 2\pi\epsilon.$$

This inequality combined with (1.9) implies

$$\Lambda_m(W^1_1,\xi) \le 4\pi\Lambda_m(\chi,\xi).$$

We proceed to the multivariate case. For $x = (x_1,\dots,x_d)$ denote

$$F_r(x) := \prod_{j=1}^{d} F_r(x_j)$$

and

$$MW^r_p := \{f : f = \varphi * F_r, \quad \|\varphi\|_p \le 1\}.$$

For $f \in MW^r_p$ we will denote $f^{(r)} := \varphi$ where φ is such that $f = \varphi * F_r$. The letter M in the notation MW^r_p stands for "mixed", because in the case of integer r the class MW^r_p is very close to the class of functions f, satisfying $\|f^{(r,\dots,r)}\|_p \le 1$, where $f^{(r,\dots,r)}$ is the mixed derivative of f of order rd. A multivariate analog of the class χ is the class

$$\chi^d := \{\chi_{[0,a]}(x) := \prod_{j=1}^{d} \chi_{[0,a_j]}(x_j), \quad a_j \in [0,2\pi), \quad j = 1,\dots,d\}.$$

Similarly to the univariate case one obtains analogs of (1.5) and (1.6)

$$\Lambda_m(MW^r_p,\xi) = \|1 - \sum_{j=1}^{m} \lambda_j F_r(\xi^j - \cdot)\|_{p'}; \tag{1.10}$$

$$\Lambda_m(MW^1_1,\xi) = \max_y |1 - \sum_{j=1}^{m} \lambda_j F_1(\xi^j - y)|. \tag{1.11}$$

We prove a multivariate analog of Proposition 1.1.

Proposition 1.2. *There exist two positive constants $C_1(d)$ and $C_2(d)$ such that for any $\Lambda_m(\cdot, \xi)$ with a property $\sum_j \lambda_j = 1$ we have*

$$C_1(d)\Lambda_m(\chi^d, \xi) \le \Lambda_m(MW_1^1, \xi) \le C_2(d)\Lambda_m(\chi^d, \xi). \tag{1.12}$$

Proof. First we prove the left inequality. Denote $\epsilon := \Lambda_m(MW_1^1, \xi)$. Then by (1.11) we have

$$-\epsilon \le 1 - \sum_{\mu=1}^{m} \lambda_\mu F_1(\xi^\mu - y) \le \epsilon, \qquad y \in \mathbb{T}^d.$$

Take any subset $e \subset [1, d]$ and integrate the above inequality against y_j, $j \notin e$, over $\mathbb{T}^{d-|e|}$. We get

$$-\epsilon \le 1 - \sum_{\mu=1}^{m} \lambda_\mu \prod_{j\in e} F_1(\xi_j^\mu - y_j) \le \epsilon.$$

Thus for any $e \subset [1, d]$, $e \ne \emptyset$, and any z_j, y_j, $j \in e$, we have

$$|\sum_{\mu=1}^{m} \lambda_\mu \big(\prod_{j\in e} F_1(\xi_j^\mu - z_j) - \prod_{j\in e} F_1(\xi_j^\mu - y_j)\big)| \le 2\epsilon.$$

This inequality and the representation (1.8) imply the left inequality in (1.12).

We prove the right inequality in (1.12) by induction. In the case $d = 1$ it follows from Proposition 1.1. Assume we have it for all dimensions $d' < d$. Denote $\epsilon := \Lambda_m(\chi^d, \xi)$. For a subset $e \subset [1, d]$ we denote

$$\xi(e) := (\xi^1(e), \dots, \xi^m(e)), \quad \xi^\mu(e) \in \mathbb{R}^{|e|}, \quad \xi^\mu(e)_j = \xi_j^\mu, \quad j \in e.$$

Consider the class $\chi^{|e|}(e)$ that is the class $\chi^{|e|}$ corresponding to the coordinated x_j with $j \in e$. It is clear that we have

$$\Lambda_m\big(\chi^{|e|}(e), \xi(e)\big) \le \epsilon.$$

Therefore, by the induction assumption and by (1.11) we have for all $e \subset [1, d]$, $|e| < d$, that

$$|1 - \sum_{\mu=1}^{m} \lambda_\mu \prod_{j\in e} F_1(\xi_j^\mu - y_j)| \le C(d)\epsilon, \quad y_j \in \mathbb{T}, \quad j \in e. \tag{1.13}$$

It is easy to see that for functions of the form

$$f_{a,y}(x) := \prod_{j=1}^{d} \big(\chi_{[0,a_j]}(x_j - y_j) - \frac{a_j}{2\pi}\big)$$

where $\chi_{[0,a_j]}(x_j - y_j)$ is 2π periodic we have

$$|\Lambda_m(f_{a,y},\xi)| \le C(d)\epsilon.$$

Using this estimate and the representation (1.8) we obtain

$$|\Lambda_m\big(\prod_{j=1}^d (F_1(x_j-y_j) - F_1(x_j-y_j-a_j)\big)| \le C(d)\epsilon$$

what means

$$-C(d)\epsilon \le \sum_{\mu=1}^m \lambda_\mu \prod_{j=1}^d F_1(\xi_j^\mu - y_j)$$

$$+\sum_{e:|e|<d}\sum_{\mu=1}^m \lambda_\mu \prod_{j\in e} F_1(\xi_j^\mu - y_j)\prod_{j\notin e}(-F_1(\xi_j^\mu - y_j - a_j)) \le C(d)\epsilon.$$

Integrating against a over $\mathbb{T}^d$ we get from here

$$|\sum_{\mu=1}^m \lambda_\mu \prod_{j=1}^d F_1(\xi_j^\mu - y_j) + \sum_{e:|e|<d}(-1)^{d-|e|}\sum_{\mu=1}^m \lambda_\mu \prod_{j\in e} F_1(\xi_j^\mu - y_j)| \le C(d)\epsilon.$$

Using (1.13) and the identity

$$\sum_{e:|e|<d}(-1)^{d-|e|} = -1$$

we obtain

$$|\sum_{\mu=1}^m \lambda_\mu \prod_{j=1}^d F_1(\xi_j^\mu - y_j) - 1| \le C(d)\epsilon.$$

We complete the proof by applying (1.11).

The classical definition of discrepancy (in the convenient for us form) of a set X of points $x^1,\dots,x^m \subset [0,1]^d$ is as follows

$$D(X,m,d)_\infty := \max_{a\in[0,1]^d} |\prod_{j=1}^d a_j - \frac{1}{m}\sum_{\mu=1}^m \chi_{[0,a]}(x^\mu)|.$$

It is clear that

$$D(X,m,d)_\infty = \Lambda_m(\chi^d, 2\pi X) \quad \text{with} \quad \lambda_1 = \dots = \lambda_m = 1/m.$$

Thus by Proposition 1.2 the classical concept of discrepancy is directly related to the efficiency of the corresponding cubature formula for a special function class MW_1^1. It is well known that the W_1^1 is very close to the class of functions of bounded variation and the MW_1^1 is very close to the class of functions with bounded variation in the sense of Hardy-Vitali. In the beginning of 20th century D. Vitali and G. Hardy generalized the definition of variation to the multivariate case. Roughly speaking, in the one-dimensional case the condition that f be of bounded variation is close to the condition $\|f'\|_1 < \infty$. In the multidimensional case the condition that a function have bounded variation in the sense of Hardy-Vitali is close to that requiring $\|f^{(1,\dots,1)}\|_1 < \infty$, where $f^{(1,\dots,1)}$ is a mixed derivative of f.

In the thirties in connection with applications in mathematical physics, S.L. Sobolev introduced the classes of functions by imposing the following restrictions

$$\|f^{(n_1,\dots,n_d)}\|_p \leq 1 \tag{1.14}$$

for all $n = (n_1, \dots, n_d)$ such that $n_1 + \dots + n_d \leq R$. These classes appeared as natural ways to measure smoothness in many multivariate problems including the numerical integration. It was established that for Sobolev classes the optimal error of numerical integration by formulas with m knots is of order $m^{-R/d}$. Assume now for the sake of simplicity (to avoid fractional defferentiation) that $R = rd$, r natural number. At the end of fifties, N.M. Korobov discovered the following phenomenon. Let us consider the class of functions which satisfy (1.14) for all n such that $n_j \leq r, \quad j = 1, \dots, d$ (compare to the above classes MW_p^r). It is clear that this new class (class of functions with dominating mixed derivative) is wider then the Sobolev class with $R = rd$. For example, all functions of the form

$$f(x) = \prod_{j=1}^{d} f_j(x_j), \quad \|f_j^{(r)}\|_p \leq 1,$$

are in this class, while not necessarily in the Sobolev class (it would require, roughly, $\|f_j^{(rd)}\|_p \leq 1$). Korobov constructed a cubature formula with m knots which guaranteed the accuracy of numerical integration for this class of order $m^{-r}(\log m)^{rd}$, i.e. almost the same accuracy that we had for the Sobolev class. Korobov's discovery pointed out the importance of the classes of functions with dominating mixed derivative in fields such as approximation theory and numerical analysis. The convenient for us definition of these classes (classes of functions with bounded mixed derivative) is given above (see the definition of MW_p^r).

In addition to the classes of 2π-periodic functions it will be convenient for us to consider the classes of nonperiodic functions defined on $\Omega_d = [0,1]^d$.

Let r be a natural number and $MW_p^r(\Omega_d)$, $1 \leq p \leq \infty$ denote the closure in the uniform metric of the set of rd-times continuously differentiable functions $f(x)$ such that

$$\|f\|_{MW_p^r(\Omega_d)} := \sum_{\substack{0 \leq n_j \leq r \\ j=1,\dots,d}} \left\| \frac{\partial^{n_1+\dots+n_d} f}{\partial x_1^{n_1} \dots \partial x_d^{n_d}} \right\|_p \leq 1,$$

where

$$\|g\|_p = \left(\int_{\Omega_d} |g(x)|^p dx \right)^{1/p}.$$

It will be convenient for us to consider the subclass $M\dot{W}^r_p(\Omega_d)$ of the class $MW^r_p(\Omega_d)$ consisting of the functions $f(x)$ representable in the form

$$f(x) = \int_{\Omega_d} B_r(t,x)\varphi(t)dt, \qquad \|\varphi\|_p \le 1,$$

where

$$B_r(t,x) = \prod_{j=1}^{d} \bigl((r-1)!\bigr)^{-1}(t_j - x_j)_+^{r-1},$$
$$t, x \in \Omega_d, \qquad (a)_+ = \max(a,0).$$

In the connection with the definition of the class $M\dot{W}^r_p(\Omega_d)$ we remark here that for the error of the cubature formula (Λ,ξ) with weights $\Lambda = (\lambda_1,\dots,\lambda_m)$ and knots $\xi = (\xi^1,\dots,\xi^m)$ the following relation holds. Let

$$\left|\Lambda_m(f,\xi) - \int_{\Omega_d} f(x)dx\right| =: R_m(\Lambda,\xi,f),$$

then similarly to (1.5) and (1.10) one obtains $(p' = p/(p-1))$

$$\Lambda_m\bigl(M\dot{W}^r_p(\Omega_d),\xi\bigr) \stackrel{\text{def}}{=} \sup_{f\in M\dot{W}^r_p(\Omega_d)} R_m(\Lambda,\xi,f) =$$
$$= \left\|\sum_{\mu=1}^{m}\lambda_\mu B_r(t,\xi^\mu) - \prod_{j=1}^{d}(t_j^r/r!)\right\|_{p'} \stackrel{\text{def}}{=} D_r(\xi,\Lambda,m,d)_{p'}. \quad (1.15)$$

The quantity $D_r(\xi,\Lambda,m,d)_q$ in the case $r=1$, $\Lambda = (1/m,\dots,1/m)$ is the classical discrepancy of the set of points $\{\xi^\mu\}$. In the case $\Lambda = (1/m,\dots,1/m)$ we denote $D_r(\xi,m,d)_q := D_r(\xi,(1/m,\dots,1/m),m,d)_q$ and call it the r-discrepancy. Thus the quantity $D_r(\xi,\Lambda,m,d)_q$ defined in (1.15) is a natural generalization of the concept of discrepancy. This generalization contains two ingredients: general weights Λ instead of the special case of equal weights $(1/m,\dots,1/m)$ and any natural number r instead of $r=1$. We note that in approximation theory we usually study the whole scale of smoothness classes rather than an individual smoothness class. The above generalization of discrepancy for arbitrary positive integer r allows us to study a question of how does smoothness r affect the rate of decay of generalized discrepancy.

Remark 1.1. *In the case of natural r the class MW^r_p turns into the subclass of the class $MW^r_p(\Omega_d)B := \{f : f/B \in MW^r_p(\Omega_d)\}$, after the linear change of variables*

$$x_j = -\pi + 2\pi t_j, \qquad j = 1,\dots,d.$$

We are interested in dependence on m of the quantities

$$\delta_m(W) = \inf_{\substack{\lambda_1,\dots,\lambda_m\\ \xi^1,\dots,\xi^m}} \Lambda_m(W,\xi)$$

for the classes W defined above.

It will be convenient for us to use the following notations. For two nonnegative sequences $\{a_m\}_{m=1}^\infty$ and $\{b_m\}_{m=1}^\infty$ we write $a_m \ll b_m$ or $b_m \gg a_m$ if there exists a positive C independent of m such that $a_m \le C b_m$, $m = 1, 2, \dots$. In the case $a_m \ll b_m$ and $a_m \gg b_m$ we write $a_m \asymp b_m$.

Remark 1.1 shows that

$$\delta_m(MW_p^r) \ll \delta_m\big(MW_p^r(\Omega_d)\big). \tag{1.16}$$

Let $\overset{0}{MW}{}_p^r(\Omega_d)$ denote the subset of functions in $MW_p^r(\Omega_d)$ which is the closure in the uniform metric of the set of functions f which have the following property: $f(x)$ is rd times continuously differentiable and $\operatorname{supp} f(x) \subset \Omega_d$.

Theorem 1.1. *Let $1 \le p \le \infty$. Then*

$$\delta_m\big(\overset{0}{MW}{}_p^r(\Omega_d)\big) \asymp \delta_m\big(M\dot{W}_p^r(\Omega_d)\big) \asymp \delta_m\big(MW_p^r(\Omega_d)\big). \tag{1.17}$$

Proof. Let Λ and ξ be given. We will prove that

$$\sup_{g \in \overset{0}{MW}{}_p^r(\Omega_d)} \left| \Lambda_m(g, \xi) - \int_{\Omega_d} g(t) dt \right| \gg \delta_m\big(MW_p^r(\Omega_d)\big). \tag{1.18}$$

Suppose an infinitely differentiable function $\psi(x)$ is such that $\psi(x) = 0$ for $x \le 0$, $\psi(x) = 1$ for $x \ge 1$ and $\psi(x)$ strictly increases on $[0, 1]$. For the cubature formula (Λ, ξ), defined on the class $\overset{0}{MW}{}_p^r(\Omega_d)$, we define the cubature formula (Λ', η), whose error will be investigated for the class $MW_p^r(\Omega_d)$ as folows:

$$\eta_j^\mu = \psi(\xi_j^\mu), \qquad j = 1, \dots, d;$$

$$\lambda'_\mu = \lambda_\mu \prod_{j=1}^d \psi'(\xi_j^\mu), \qquad \mu = 1, \dots, m.$$

Then for the functions f and g related as

$$g(t) = f(\psi(t_1), \dots, \psi(t_d)) \prod_{j=1}^d \psi'(t_j) \tag{1.19}$$

we have

$$\int_{Q_d} f(x)\, dx = \int_{Q_d} g(t)\, dt,$$

$$\sum_{\mu=1}^m \lambda_\mu g(\xi^\mu) = \sum_{\mu=1}^m \lambda'_\mu f(\eta^\mu).$$

It remains to check that there exist a number $\delta > 0$ which does not depend on m such that $\delta g \in \overset{0}{MW}{}^r_p(\Omega_d)$ provided $f \in MW^r_p(\Omega_d)$.

In fact, differentiating (1.19) we see that the expression for $g^{(s)}(t)$, $s = (s_1, \dots, s_d)$, $0 \le s_j \le r$, $j = 1, \dots, d$ will contain terms of the form

$$\omega(t,k) = f^{(k)}\big(\psi(t)\big) \prod_{j=1}^{d} \prod_{i} \big(\psi^{(l^i_j)}(t_j)\big)^{m^i_j}, \quad \psi(t) = \big(\psi(t_1), \dots, \psi(t_d)\big),$$
$$k = (k_1, \dots, k_d), \quad 0 \le k_j \le s_j, \quad \sum_i l^i_j m^i_j = s_j + 1, \quad j = 1, \dots, d \tag{1.20}$$

the number of which depends on the vector s.

It is obvious that in the case $p = \infty$ we have

$$\|\omega(t,k)\|_\infty \le C(r) \|f\|_{MW^r_\infty(\Omega_d)}.$$

To estimate $\|\omega(\cdot,k)\|_p$, $1 \le p < \infty$ we use the following simple lemma.

Lemma 1.1. *Suppose that $f \in MW^r_p(\Omega_d)$ is rd times continuously differentiable, and the vector $k \in \mathbb{Z}^d_+$ is such that $k_j = r$ for $j \in e_1$ and $k_j \le r-1$ for $j \in e_2 = [1,d] \setminus e_1$. Then*

$$\sup_{x_j, j \in e_2} \int\limits_{[0,1]^{|e_1|}} |f^{(k_1,\dots,k_d)}(x)|^p \big(\prod_{j \in e_1} dx_j\big) \le C(p,r,d) \|f\|^p_{MW^r_p(\Omega_d)}, \qquad 1 \le p < \infty.$$

Proof. We will first prove the following statement. Let f be such that f, $\frac{\partial f}{\partial x_j}$ are continuous. Then

$$\sup_{x_j} \int\limits_{[0,1]^{d-1}} |f(x)|^p \big(\prod_{i \ne j} dx_i\big) \le 2^p \big(\|f\|^p_p + \|\frac{\partial f}{\partial x_j}\|^p_p\big). \tag{1.21}$$

Indeed, there is an $a \in [0,1]$ such that

$$\int\limits_{[0,1]^{d-1}} |f(x_1, \dots, x_{j-1}, a, x_{j+1}, \dots, x_d)|^p \big(\prod_{i \ne j} dx_i\big) \le \|f\|^p_p.$$

We represent the function $f(x)$ in the form

$$f(x) = f(x_1, \dots, x_{j-1}, a, x_{j+1}, \dots, x_d) +$$
$$\int\limits_a^{x_j} \frac{\partial f}{\partial x_j}(x_1, \dots, x_{j-1}, u, x_{j+1}, \dots, x_d)\, du.$$

Then, for any $x_j \in [0,1]$, we have

$$\int\limits_{[0,1]^{d-1}} |f(x)|^p (\prod_{i \neq j} dx_i) \leq 2^p (\|f\|_p^p + \|\frac{\partial f}{\partial x_j}\|_p^p).$$

This proves (1.21).

Applying successively relation (1.21), we obtain the statement of the lemma.

We return to estimating $\|\omega(\cdot,k)\|_p$. By Lemma 1.1 and by the uniform boundedness of the functions $|\psi^{(l)}(t_j)| \leq C(r)$, $l \leq r+1$, we get for k such that $k_j < r$, $j = 1,\dots,d$ that

$$\|\omega(\cdot,k)\|_p \leq \|\omega(\cdot,k)\|_\infty \ll \|f^{(k_1,\dots,k_d)}\|_\infty \leq \|f\|_{MW_p^r(\Omega_d)}.$$

Thus it remains to consider only those k for which there is a j such that $k_j = r$. Then, with respect to this variable, $(\psi'(t_j))^{r+1}$ participates as an additional cofactor in expression (1.20). Taking into account that

$$(\psi'(t_j))^{p(r+1)} \leq C(p,r)\psi'(t_j),$$

we obtain

$$\int\limits_0^1 |\omega(t,k)|^p \, dt_j \leq$$

$$\leq C(p,r)\Big(\int\limits_0^1 |f^{(k)}(\psi(t_1),\dots,\psi(t_{j-1}),x_j,\psi(t_{j+1}),\dots,\psi(t_d))|^p \, dx_j)\times$$

$$\times \prod_{\nu \neq j} \prod_i (\psi^{(l_\nu^i)}(t_\nu)\Big)^{m_\nu^i p}.$$

Reasoning in this way for all j such that $k_j = r$ and applying Lemma 1.1, we find that

$$\|\omega(\cdot,k)\|_p \leq C(p,r,d)\|f\|_{MW_p^r(\Omega_d)}$$

for all k. This implies that $c(p,r,d)g \in M\overset{0}{W}{}_p^r(\Omega_d)$ with some positive $c(p,r,d)$.

The arguments we presented show that

$$\sup_{g \in M\overset{0}{W}{}_p^r(\Omega_d)} |\int\limits_{Q_d} g(t)\,dt - \sum_{\mu=1}^m \lambda_\mu g(\xi^\mu)| \geq$$

$$\geq c(p,r,d) \sup_{f \in MW_p^r(\Omega_d)} |\int\limits_{Q_d} f(x)\,dx - \sum_{\mu=1}^m \lambda'_\mu f(\eta^\mu)|. \tag{1.22}$$

Relation (1.22) and the embeddings $M\overset{0}{W}{}_p^r(\Omega_d) \hookrightarrow M\dot{W}_p^r(\Omega_d) \hookrightarrow MW_p^r(\Omega_d)$ yield the statement of Theorem 1.1.

Remark 1.2. *Let $1 \le p \le \infty$ and $r > 1/p$. Then*

$$\delta_m(MW^r_p) \asymp \delta_m\big(MW^r_p(\Omega_d)\big).$$

The upper estimate follows from (1.16). The lower estimate follows from Theorem 1.1.

2. Optimal cubature formulas and nonlinear approximation

The relations (1.10) and (1.11) can be interpreted as a connection between the error of the cubature formula (Λ, ξ) on the class MW^r_p and the approximation error of a special function $1 = \int_{\mathbb{T}^d} F_r(x)d\mu$ by m-term linear combination of functions $F_r(\xi^j - \cdot)$, $j = 1, \dots, m$. The latter problem is a problem of nonlinear m-term approximation with regard to a given system of functions, in the above case with regard to the system $\{F_r(x - \cdot), x \in \mathbb{T}^d\}$. The problem of nonlinear m-term approximation has attracted a lot of attention during last ten years because of its importance in numerical applications (see surveys [15] and [51]). In this section we will use some known results from m-term approximation in Banach spaces for estimating the error of optimal cubature formulas. Let $1 \le q \le \infty$. We define a set $\mathcal{K}_q$ of kernels possessing the following properties. Let Ω_x and Ω_y be measurable sets for variables x and y respectively. Let $K(x, y)$ be a measurable function on $\Omega_x \times \Omega_y$. We assume that for any $x \in \Omega_x$ $K(x, \cdot) \in L_q(\Omega_y)$, for any $y \in \Omega_y$ the $K(\cdot, y)$ is integrable over Ω_x and $\int_{\Omega_x} K(x, \cdot)dx \in L_q(\Omega_y)$. For a kernel $K \in \mathcal{K}_{p'}$ we define the class

$$W^K_p := \{f : f = \int_{\Omega_y} K(x, y)\varphi(y)dy, \quad \|\varphi\|_{L_p(\Omega_y)} \le 1\}. \tag{2.1}$$

Then each $f \in W^K_p$ is integrable on Ω_x (by Fubini's theorem) and defined at each point of Ω_x. We denote for convenience

$$J(y) := J_K(y) := \int_{\Omega_x} K(x, y)dx.$$

For a cubature formula $\Lambda_m(\cdot, \xi)$ we have

$$\Lambda_m(W^K_p, \xi) = \sup_{\|\varphi\|_{L_p(\Omega_y)} \le 1} \Big|\int_{\Omega_y} \big(J(y) - \sum_{\mu=1}^m \lambda_\mu K(\xi^\mu, y)\big)\varphi(y)dy\Big| =$$

$$= \|J(\cdot) - \sum_{\mu=1}^m \lambda_\mu K(\xi^\mu, \cdot)\|_{L_{p'}(\Omega_y)}. \tag{2.2}$$

We use the same as above definition of the error of optimal cubature formula with m knots for a class W

$$\delta_m(W) := \inf_{\substack{\lambda_1, \dots, \lambda_m \\ \xi^1, \dots, \xi^m}} \Lambda_m(W, \xi).$$

Thus by (2.2)

$$\delta_m(W_p^K) = \inf_{\substack{\lambda_1,\dots,\lambda_m \\ \xi^1,\dots,\xi^m}} \|J(\cdot) - \sum_{\mu=1}^{m} \lambda_\mu K(\xi^\mu, \cdot)\|_{L_{p'}(\Omega_y)}.$$

We will now introduce some notations and concepts from the theory of m-term approximation in Banach spaces.

Let X be a Banach space with norm $\|\cdot\|$. We say that a set of elements (functions) $\mathcal{D}$ from X is a dictionary if each $g \in \mathcal{D}$ has norm one ($\|g\| = 1$),

$$g \in \mathcal{D} \quad \text{implies} \quad -g \in \mathcal{D},$$

and $\overline{\text{span}}\mathcal{D} = X$.

We will discuss in this section two types of greedy algorithms with regard to $\mathcal{D}$. For an element $f \in X$ we denote by N_f a norming (peack) functional for f:

$$\|N_f\| = 1, \qquad N_f(f) = \|f\|.$$

The existence of such a functional is guaranteed by Hahn-Banach theorem. Let $\tau := \{t_k\}_{k=1}^{\infty}$ be a given sequence of positive numbers $t_k \le 1$, $k = 1, \dots$. We define first the Weak Chebyshev Greedy Algorithm (WCGA) that is a generalization for Banach spaces of Weak Orthogonal Greedy Algorithm defined and studied in [49] (see also [16] for Orthogonal Greedy Algorithm).

Weak Chebyshev Greedy Algorithm (WCGA). We define $f_0^c := f_0^{c,\tau} := f$. Then for each $m \ge 1$ we inductively define

1). $\varphi_m^c := \varphi_m^{c,\tau} \in \mathcal{D}$ is any satisfying

$$N_{f_{m-1}^c}(\varphi_m^c) \ge t_m \sup_{g\in\mathcal{D}} N_{f_{m-1}^c}(g).$$

2). Define

$$\Phi_m := \Phi_m^\tau := \text{span}\{\varphi_j^c\}_{j=1}^m,$$

and define $G_m^c := G_m^{c,\tau}$ to be the best approximant to f from Φ_m.

3). Denote

$$f_m^c := f_m^{c,\tau} := f - G_m^c.$$

We define now the generalization for Banach spaces of the Weak Relaxed Greedy Algorithm studied in [49] in the case of Hilbert space.

Weak Relaxed Greedy Algorithm (WRGA). We define $f_0^r := f_0^{r,\tau} := f$ and $G_0^r := G_0^{r,\tau} := 0$. Then for each $m \ge 1$ we inductively define

1). $\varphi_m^r := \varphi_m^{r,\tau} \in \mathcal{D}$ is any satisfying

$$N_{f_{m-1}^r}(\varphi_m^r - G_{m-1}^r) \ge t_m \sup_{g\in\mathcal{D}} N_{f_{m-1}^r}(g - G_{m-1}^r).$$

2). Find $0 \le \lambda_m \le 1$ such that

$$\|f - ((1-\lambda_m)G^r_{m-1} + \lambda_m \varphi^r_m)\| = \inf_{0\le\lambda\le 1} \|f - ((1-\lambda)G^r_{m-1} + \lambda\varphi^r_m)\|$$

and define

$$G^r_m := G^{r,\tau}_m := (1-\lambda_m)G^r_{m-1} + \lambda_m\varphi^r_m.$$

3). Denote

$$f^r_m := f^{r,\tau}_m := f - G^r_m.$$

The term "weak" in both definitions means that at the step 1). we do not shoot for the optimal element of the dictionary which realizes the corresponding sup but are satisfied with weaker property than being optimal. The obvious reason for this is that we don't know in general that the optimal one exists. Another practical reason is that the weaker the assumption the easier to satisfy it and therefore easier to realize in practice.

We present in this section results of convergence and the rate of convergence for the two above defined methods of approximation. It is clear that in the case of WRGA the assumption that f belongs to the closure of convex hull of $\mathcal{D}$ is natural. We denote the closure of convex hull of $\mathcal{D}$ by $\mathcal{A}_1(\mathcal{D})$. It has been proven in [49] that in the case of Hilbert space both algorithms WCGA and WRGA give the approximation error for the class $\mathcal{A}_1(\mathcal{D})$ of the order

$$(1 + \sum_{k=1}^m t_k^2)^{-1/2}.$$

We discuss here approximation in uniformly smooth Banach spaces. For a Banach space X we define the modulus of smoothness

$$\rho(u) := \sup_{\|x\|=\|y\|=1} (\frac{1}{2}(\|x+uy\| + \|x-uy\|) - 1).$$

The uniformly smooth Banach space is the one with the property

$$\lim_{u\to 0} \rho(u)/u = 0.$$

The following convergence result has been proven in [50]. In the formulation of this result we need a special sequence which is defined for a given modulus of smoothness $\rho(u)$ and a given $\tau = \{t_k\}_{k=1}^\infty$.

Definition 2.1. *Let $\rho(u)$ be an even convex function on $(-\infty,\infty)$ with the property: $\rho(2) \ge 1$ and*

$$\lim_{u\to 0} \rho(u)/u = 0.$$

For any $\tau = \{t_k\}_{k=1}^\infty$, $0 < t_k \le 1$, and $0 < \theta \le 1/2$ we define $\xi_m := \xi_m(\rho,\tau,\theta)$ as a number u satisfying the equation

$$\rho(u) = \theta t_m u. \tag{2.3}$$

Theorem 2.1. *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u)$. Assume that a sequence $\tau := \{t_k\}_{k=1}^{\infty}$ satisfies the condition: for any $\theta > 0$ we have*

$$\sum_{m=1}^{\infty} t_m \xi_m(\rho, \tau, \theta) = \infty.$$

Then for any $f \in X$ we have

$$\lim_{m\to\infty} \|f_m^{c,\tau}\| = 0.$$

Corollary 2.1. *Let a Banach space X have modulus of smoothness $\rho(u)$ of power type $1 < q \le 2$; $(\rho(u) \le \gamma u^q)$. Assume that*

$$\sum_{m=1}^{\infty} t_m^p = \infty, \quad p = \frac{q}{q-1}. \tag{2.4}$$

Then WCGA converges for any $f \in X$.

It is well known (see for instance [14], Lemma B.1) that in the case $X = L_p$, $1 \le p < \infty$ we have

$$\rho(u) \le \begin{cases} u^p/p & \text{if} \quad 1 \le p \le 2, \\ (p-1)u^2/2 & \text{if} \quad 2 \le p < \infty. \end{cases}$$

It is also known (see [28], p.63) that for any X with $\dim X = \infty$ one has

$$\rho(u) \ge (1+u^2)^{1/2} - 1$$

and for every X, $\dim X \ge 2$,

$$\rho(u) \ge Cu^2, \quad C > 0.$$

This limits power type modulus of smoothness of nontrivial Banach spaces to the case $1 \le q \le 2$.

Let us apply Corollary 2.1 for numerical integration. Consider a dictionary

$$\mathcal{D} := \mathcal{D}(K, p') := \{g : g(x,y) = K(x,y)/\|K(x,\cdot)\|_{L_{p'}(\Omega_y)}\}$$

(in case $\|K(x,\cdot)\|_{L_{p'}(\Omega_y)} = 0$ we set $g(x,\cdot) = 0$), and define a Banach space $X := X(K,p')$ as the $L_{p'}(\Omega_y)$-closure of span of $\mathcal{D}$. Assume now that $J_K \in X$. Then for $1 < p' < \infty$ the WCGA satisfying (2.4) with $q = \min(2, p')$ provides a deterministic algorithm of constructing a sequence of cubature formulas $\Lambda_m^c(\cdot, \xi)$ such that

$$\Lambda_m^c(W_p^K, \xi) \to 0 \quad \text{as} \quad m \to \infty.$$

We will discuss in more detail a question of the rate of convergence. The following theorem has been proven in [50].

Theorem 2.2. *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u) \le \gamma u^q$, $1 < q \le 2$. Then for a sequence $\tau := \{t_k\}_{k=1}^\infty$, $t_k \in [0,1]$, $k = 1, 2, \dots$, we have for any $f \in \mathcal{A}_1(\mathcal{D})$ that*

$$\|f_m^{c,\tau}\| \le C_1(q,\gamma)(1 + \sum_{k=1}^m t_k^p)^{-1/p},$$

$$\|f_m^{r,\tau}\| \le C_2(q,\gamma)(1 + \sum_{k=1}^m t_k^p)^{-1/p}, \quad p := \frac{q}{q-1},$$

with constants $C_i(q,\gamma)$, $i = 1, 2$, which may depend only on q and γ.

Corollary 2.2. *In a particular case $\tau = \{t_k\}_{k=1}^\infty$, $t_k = t$, $k = 1, 2, \dots$, with some $t \in (0,1]$, we have under assumptions of Theorem 2.2 that*

$$\|f_m^{c,\tau}\| \le C_1(q,\gamma,t)m^{-1/p},$$

$$\|f_m^{r,\tau}\| \le C_2(q,\gamma,t)m^{-1/p}, \quad p := \frac{q}{q-1}.$$

In order to apply Theorem 2.2 in numerical integration for W_p^K we need to check that $J_K \in \mathcal{A}_1(\mathcal{D}(K,p'))$ (or there exists a positive constant c such that $cJ_K \in \mathcal{A}_1(\mathcal{D}(K,p'))$. It could be a difficult problem. An inspection of the proof of Theorem 2.2 shows that it is sufficient to check that

$$\int_{\Omega_x} \|K(x,\cdot)\|_{L_{p'}(\Omega_y)} dx < \infty.$$

We formulate this as a theorem.

Theorem 2.3. *Let W_p^K be a class of functions defined by (2.1). Assume that $K \in \mathcal{K}_{p'}$ satisfies the condition*

$$\int_{\Omega_x} \|K(x,\cdot)\|_{L_{p'}(\Omega_y)} dx \le M$$

and $J_K \in X(K,p')$. Then for any m there exists (provided by WRGA with $\tau = \{t\}$) a cubature formula $\Lambda_m(\cdot,\xi)$ with

$$\sum_{\mu=1}^m |\lambda_\mu| \le M$$

and

$$\Lambda_m(W_p^K,\xi) \le MC(p,\Omega_x,\Omega_y,t)\begin{cases} m^{-1/2}, & 1 < p \le 2, \\ m^{-1/p}, & 2 \le p < \infty. \end{cases}$$

Let us consider a particular example of $K(x,y) = (2\pi)^{-d}F(x-y)$, $\Omega_x = \Omega_y = \mathbb{T}^d$. We denote the corresponding class W_p^K by W_p^F.

Proposition 2.1. *Let $1 < p < \infty$ and $\|F\|_{p'} \le M$. Then the kernel $K(x,y) = (2\pi)^{-d}F(x-y)$ satisfies the assumptions of Theorem 2.3.*

Proof. It is obvious that $K \in \mathcal{K}_{p'}$. Next,

$$(2\pi)^{-d}\int_{\mathbb{T}^d} \|F(x-\cdot)\|_{p'}dx = \|F\|_{p'} \le M.$$

It remains to check that $J_K \in X(K,p')$. We have

$$J_K(y) = (2\pi)^{-d}\int_{\mathbb{T}^d} F(x-y)dx = \hat{F}(0).$$

Denote

$$S_N(F,x) := \sum_{|k_j|\le N, j=1,\dots,d} \hat{F}(k)e^{i(k,x)}.$$

Then it is well known (by the M. Riesz theorem) that

$$\|F - S_N(F)\|_{p'} \to 0 \quad \text{as} \quad N \to \infty.$$

For a given N we consider a cubature formula

$$q_N(f) := N^{-d}\sum_{\mu_d=1}^{N}\cdots\sum_{\mu_1=1}^{N} f(2\pi\mu_1/N,\dots,2\pi\mu_d/N).$$

Then we have

$$\|\hat{F}(0) - q_N\big(F(\cdot-y)\big)\|_{p'} =$$
$$= \|\hat{F}(0) - q_N\big(S_{N-1}(F,\cdot-y)\big) + q_N\big(S_{N-1}(F,\cdot-y) - F(\cdot-y)\big)\|_{p'} =$$
$$= \|q_N\big(S_{N-1}(F,\cdot-y) - F(\cdot-y)\big)\|_{p'} \le \|S_{N-1}(F) - F\|_{p'} \to 0$$

as $N \to \infty$. This proves the proposition.

Theorem 2.3 and Proposition 2.1 yield the following result.

Theorem 2.4. *Let $1 < p < \infty$. Assume $F \in L_{p'}(\mathbb{T}^d)$, $p' = p/(p-1)$. Consider the class*

$$W^F_p := \{f : f = F * \varphi, \quad \|\varphi\|_p \le 1\}.$$

Then for any m there exists a cubature formula $\Lambda_m(\cdot,\xi)$ with $\sum_{\mu=1}^m |\lambda_\mu| \le \|F\|_{p'}$ and

$$\Lambda_m(W^F_p,\xi) \le C(p,d)\|F\|_{p'}\begin{cases} m^{-1/2}, & 1 < p \le 2, \\ m^{-1/p}, & 2 \le p < \infty. \end{cases}$$

A sequence of $\{\Lambda_m(\cdot,\xi)\}$ from Theorem 2.4 can be obtained by applying WRGA with a fixed $\tau = \{t\}$. We will describe this procedure in detail. Denote $q := p'$. We have $1 < q < \infty$. Then

$$X(K,q) = \overline{\operatorname{span}}\{F(x-\cdot), x \in \mathbb{T}^d\},$$

with the closure in $L_q(\mathbb{T}^d)$. It is well known that

$$\overline{\operatorname{span}}\{F(x-\cdot), x \in \mathbb{T}^d\} = L_q(\mathbb{T}^d) \cap \overline{\operatorname{span}}\{e^{i(k,x)}, k : \hat{F}(k) \neq 0\}.$$

As a dictionary we have

$$\mathcal{D} = \{\pm F(x-\cdot)/\|F\|_{p'}, x \in \mathbb{T}^d\}.$$

We note that in the proof of Proposition 2.1 we actually proved that

$$\hat{F}(0)/\|F\|_{p'} \in \mathcal{A}_1(\mathcal{D}).$$

Let us for simplicity give an algorithm for F satisfying $\|F\|_{p'} = 1$ (otherwise we take $F/\|F\|_{p'}$). We begin with $f = \hat{F}(0)$. First, we construct a norming functional N_f. It is known that for $1 < q < \infty$ the N_f acts as

$$N_f(g) = (2\pi)^{-d} \int_{\mathbb{T}^d} \|f\|_q^{1-q} |f|^{q-2} \bar{f} g dx. \tag{2.5}$$

Fix $t \in (0,1)$ and find ξ^1 satisfying

$$|N_f(F(\xi^1 - \cdot))| \geq t \sup_x |N_f(F(x-\cdot))|.$$

Find $\lambda_1 \in [0,1]$ and $\epsilon_1 = \pm 1$ such that

$$\|\hat{F}(0) - \lambda_1 \epsilon_1 F(\xi^1 - \cdot)\|_q = \inf_{|\lambda| \leq 1} \|\hat{F}(0) - \lambda F(\xi^1 - \cdot)\|_q.$$

Denote

$$G_1 := \lambda_1 \epsilon_1 F(\xi^1 - \cdot) \quad \text{and} \quad f_1 := \hat{F}(0) - G_1.$$

We now describe the mth step. Assume we have already obtained G_{m-1} and f_{m-1}. Define by (2.5) the norming functional $N_{f_{m-1}}$ anf find ξ^m satisfying

$$|N_{f_{m-1}}(F(\xi^m - \cdot))| \geq t \sup_x |N_{f_{m-1}}(F(x-\cdot))|.$$

Denote $\epsilon_m = \operatorname{sign} N_{f_{m-1}}(F(\xi^m - \cdot))$ and find $\lambda_m \in [0,1]$ such that

$$\|\hat{F}(0) - ((1-\lambda_m)G_{m-1} + \lambda_m \epsilon_m F(\xi^m - \cdot))\|_q =$$

$$= \inf_{\lambda \in [0,1]} \|\hat{F}(0) - ((1-\lambda)G_{m-1} + \lambda \epsilon_m F(\xi^m - \cdot))\|_q.$$

Denote

$$G_m := (1-\lambda_m)G_{m-1} + \lambda_m \epsilon_m F(\xi^m - \cdot), \qquad f_m := f - G_m.$$

After m steps we obtain a cubature formula with knots $\xi = (\xi^1, \dots, \xi^m)$ and weights $\Lambda = (\epsilon_1\lambda_1, \dots, \epsilon_m\lambda_m)$. Nice properties of this construction are: (1) $\sum_{\mu=1}^m \lambda_\mu \leq 1$; (2) proceeding from the step $m-1$ to the step m we add one new knot ξ^m and change in a simple way weights $\epsilon_k\lambda_k$ from the previous step.

3. Lower estimates for the classes MW_p^r

We will present here some methods of obtaining lower estimates of $\Lambda_m(W,\xi)$ for the classes MW_p^r.

It will be convenient for us to assume that the functions are 1-periodic and to keep the notation $\Omega_d = [0,1]^d$. We begin with the following result.

Theorem 3.1. *For any cubature formula (Λ,ξ) with m knots the following relation holds $(r>1/2)$*

$$\Lambda_m(MW_2^r,\xi) \ge C(r,d)m^{-r}(\log m)^{\frac{d-1}{2}}.$$

Proof. Let us denote

$$\Lambda(k) = \Lambda_m(e^{i2\pi(k,x)},\xi) = \sum_{\mu=1}^{m} \lambda_\mu e^{i2\pi(k,\xi^\mu)}$$

for the cubature formula (Λ,ξ) and $k\in\mathbb{Z}^d$.

Lemma 3.1. *The following inequality is valid for any $r>1/2$*

$$\sum_{k\ne 0} \left|\Lambda(k)\right|^2 \nu(\overline{k})^{-2r} \ge C(r,d)\left|\Lambda(0)\right|^2 m^{-2r}(\log m)^{d-1},$$

where $\overline{k_j} := \max(|k_j|,1)$ and $\nu(\overline{k}) = \prod_{j=1}^d \overline{k_j}$.

First we will deduce Theorem 3.1 from Lemma 3.1 and then prove this lemma.

We assume that $\left|\Lambda(0)\right| \ge 1/2$ because otherwise it is sufficient to take as an example the function $f(x)\equiv 1$. Let us consider the function

$$f(x) = \sum_{k\ne 0} \overline{\Lambda(k)}\nu(\overline{k})^{-2r} e^{i2\pi(k,x)}$$

where $\overline{\Lambda(k)}$ is the complex conjugate to the $\Lambda(k)$. Then

$$\|f^{(r)}\|_2 = \left(\sum_{k\ne 0}\left|\Lambda(k)\right|^2 \nu(\overline{k})^{-2r}\right)^{1/2} \tag{3.1}$$

and

$$\Lambda_m(f,\xi) - \hat{f}(0) = \sum_{k\ne 0}\left|\Lambda(k)\right|^2 \nu(\overline{k})^{-2r}. \tag{3.2}$$

By (3.1) and (3.2)

$$\Lambda_m(MW_2^r,\xi) \ge \left(\sum_{k\ne 0}\left|\Lambda(k)\right|^2 \nu(\overline{k})^{-2r}\right)^{1/2},$$

using Lemma 3.1 we get

$$\Lambda_m(MW_2^r,\xi) \geq C(r,d)m^{-r}(\log m)^{\frac{d-1}{2}},$$

which proves Theorem 3.1.

We turn to the proof of Lemma 3.1. Let $b(x)$ be an infinitly differentiable function such that $b(x)=0$ for $x \notin (0,1)$ and $b(x)>0$ for $x \in (0,1)$. Let m be given, choose $n \in \mathbb{N}$ such that

$$2m \leq 2^n < 4m \tag{3.3}$$

Denote for $s=(s_1,\dots,s_d)$, s_j nonnegative integers,

$$b_s(x) = \prod_{j=1}^d b(2^{s_j+2}x_j),$$

and

$$Y_s = \bigl\{y \in \Omega_d \text{ such that } \Lambda_m\bigl(b_s(x-y),\xi\bigr)=0\bigr\}.$$

It is easy to verify that for all s with $\|s\|_1=n$, the estimate

$$|Y_s| \geq C(d) > 0$$

is valid for the measure $|Y_s|$ of the set Y_s. Further,

$$\begin{aligned} \bigl|\hat{b}_s(0)\Lambda(0)\bigr|^2|Y_s| &= \int_{Y_s}\bigl|\Lambda_m\bigl(b_s(x-y),\xi\bigr)-\hat{b}_s(0)\Lambda(0)\bigr|^2dy \leq \\ &\leq \int_{\Omega_d}\bigl|\Lambda_m\bigl(b_s(x-y),\xi\bigr)-\hat{b}_s(0)\Lambda(0)\bigr|^2dy = \sum_{k\neq 0}\bigl|\Lambda(k)\bigr|^2\bigl|\hat{b}_s(k)\bigr|^2. \end{aligned} \tag{3.4}$$

Let $a:=[r]+1$. Then for s such that $\|s\|_1=n$ we have

$$\begin{aligned} \bigl|\hat{b}_s(k)\bigr| &\leq C(d,a)\prod_{j=1}^d\bigl(2^{-s_j}\min(1,2^{as_j}(\overline{k}_j)^{-a}\bigr) = \\ &= C(d,a)2^{(r-1)n}\prod_{j=1}^d 2^{-rs_j}\min\bigl(1,2^{as_j}(\overline{k}_j)^{-a}\bigr). \end{aligned} \tag{3.5}$$

By summing the relation (3.5) over all s such that $\|s\|_1=n$ and using the inequalities

$$\begin{aligned} &\sum_{\|s\|_1=n}\prod_{j=1}^d 2^{-rs_j}\min\bigl(1,2^{as_j}(\overline{k}_j)^{-a}\bigr) \leq \\ &\leq \prod_{j=1}^d\sum_{s_j=0}^{\infty}2^{-rs_j}\min\bigl(1,2^{as_j}(\overline{k}_j)^{-a}\bigr) \ll \prod_{j=1}^d(\overline{k}_j)^{-r} \end{aligned}$$

we get from (3.4)

$$n^{d-1}2^{-2n}\big|\Lambda(0)\big|^2 C(d) \le 2^{2(r-1)n} C(d,r) \sum_{k\neq 0} \big|\Lambda(k)\big|^2 \nu(\overline{k})^{-2r},$$

which proves the lemma.

We proceed to a stronger result than Theorem 3.1. We will replace the class MW_2^r in Theorem 3.1 by MW_p^r with $2 \le p < \infty$. We prove the following proposition.

Theorem 3.2. *The following lower estimate is valid for any cubature formula (Λ, ξ) with m knots $(r > 1/p)$*

$$\Lambda_m(MW_p^r, \xi) \ge C(r,d,p) m^{-r} (\log m)^{\frac{d-1}{2}}, \qquad 1 \le p < \infty.$$

Proof. The proof of this theorem is based on the following theorem on existence of generalized Rademacher type polynomials. We introduce first some notation. Let

$$\Pi(\mathbf{N}, d) = \big\{(a_1, \dots, a_d) \in R^d : |a_j| \le N_j,\ j = 1, \dots, d\big\},$$

where N_j are nonnegative integers and $\mathbf{N} := (N_1, \dots, N_d)$. We denote

$$T(\mathbf{N}, d) := \{t : t = \sum_{k \in \Pi(\mathbf{N},d)} c_k e^{i(k,x)}\}.$$

Then

$$\dim T(\mathbf{N}, d) = \prod_{j=1}^{d} (2N_j + 1) =: \vartheta(\mathbf{N}).$$

Theorem 3.3. *Let $\varepsilon > 0$ and a subspace $\Psi \subset T(\mathbf{N}, d)$ be such that $\dim \Psi \ge \varepsilon \vartheta(\mathbf{N})$. Then there is a $t \in \Psi$ such that*

$$\|t\|_\infty = 1, \qquad \|t\|_2 \ge C(\varepsilon, d) > 0.$$

The proof of this theorem is based on the lower estimates of the volumes of the sets of Fourier coefficients of bounded trigonometric polynomials from $T(\mathbf{N}, d)$ (see [45] and [47]).

First, we prove the following assertion.

Lemma 3.2. *Let the coordinates of the vector s be natural numbers and $\|s\|_1 = n$. Then for any $N \le 2^{n-1}$ and an arbitrary cubature formula (Λ, ξ) with N knots there is a $t_s \in T((2^{s_1}, \dots, 2^{s_d}), d)$ such that $\|t_s\|_\infty \le 1$ and*

$$\hat{t}_s(0) - \Lambda_N(t_s, \xi) \ge C(d) > 0. \tag{3.6}$$

Proof of Lemma 3.2. Let N and (Λ,ξ) be given. Let us consider in $T((2^{s_1-1},\dots,2^{s_d-1}),d)$ the linear subspace Ψ of polynomials t satisfying the conditions

$$t(\xi^j)=0,\qquad j=1,\dots,N. \tag{3.7}$$

Then

$$\dim\Psi\ge 2^n-N\ge 2^{n-1}.$$

Consequently, by Theorem 3.3 there is $t^1\in\Psi$ such that $\|t^1\|_\infty=1$ and

$$\|t^1\|_2\ge C(d)>0. \tag{3.8}$$

We difine

$$t_s(x)=\begin{cases}1 & \text{if}\qquad \sum_{j=1}^N\lambda_j\le 1/2,\\ \left|t^1(x)\right|^2 & \qquad\text{otherwise.}\end{cases}$$

The relations (3.7) and (3.8) prove the lemma.

We now complete the proof of Theorem 3.2. Let m be given. We choose n such that

$$m\le 2^{n-1}<2m.$$

We consider the polynomial

$$t(x)=\sum_{\|s\|_1=n}t_s(x),$$

where t_s are polynomials from Lemma 3.2 with $N=m$. Then

$$\hat t(0)-\Lambda_m(t,\xi)\ge C(d)n^{d-1}. \tag{3.9}$$

Let us estimate $\left\|t^{(r)}\right\|_p$, $2\le p<\infty$. We will use the following corollary of the Littlewood-Paley theorem. Let $f\in L_1$; denote for $u\in\mathbb{Z}_+^d$

$$\rho(u):=\{k:k=(k_1,\dots,k_d),[2^{u_j-1}]\le|k_j|<2^{u_j},j=1,\dots,d\}, \tag{3.10}$$

$$\delta_u(f):=\sum_{k\in\rho(u)}\hat f(k)e^{i(k,x)}.$$

Then for $f\in L_p$, $2\le p<\infty$ one has the inequality

$$\|f\|_p\le C(p,d)\big(\sum_{u\in\mathbb{Z}_+^d}\|\delta_u(f)\|_p^2\big)^{1/2}.$$

We have

$$\left\|t^{(r)}\right\|_p\ll\left(\sum_{\|u\|_1\le n+d}\left\|\delta_u\left(t^{(r)}\right)\right\|_p^2\right)^{1/2}.$$

Using the Bernstein inequality we continue the estimate

$$\ll \left(\sum_{\|u\|_1 \le n+d} 2^{2r\|u\|_1} \left\| \delta_u(t) \right\|_p^2 \right)^{1/2}. \tag{3.11}$$

Next we have

$$\delta_u(t) = \sum_{\|s\|_1 = n} \delta_u(t_s) = \sum_{\substack{\|s\|_1 = n \\ s+1 \ge u}} \delta_u(t_s).$$

By the inequality $\|t_s\|_\infty \le 1$ we get from here

$$\left\| \delta_u(t) \right\|_p \ll \left(n + d + 1 - \|u\|_1 \right)^{d-1}. \tag{3.12}$$

The estimates (3.11), (3.12) result in

$$\left\| t^{(r)} \right\|_p \ll \left(\sum_{\|u\|_1 \le n+d} 2^{2r\|u\|_1} \left(n + d + 1 - \|u\|_1 \right)^{2(d-1)} \right)^{1/2} \ll 2^{rn} n^{\frac{d-1}{2}}. \tag{3.13}$$

Comparing (3.9) and (3.13) we get the conclusion of Theorem 3.3 for $2 \le p < \infty$. Clearly, the lower estimate for $1 \le p < 2$ follows from the estimate which we have just proved.

Theorem 3.2 gives the same lower estimate for different parameters $1 \le p < \infty$. It is clear that the bigger the p the stronger the statement. We now discuss an improvement of Theorem 3.2 in the particular case $p = 1$. We will improve the lower estimate by replacing the exponent $(d-1)/2$ by $d-1$. However, this improvement will be proved under some (mild) assumptions on the weights of a cubature formula (Λ, ξ) and also for a slight modification of the classes MW_1^r. Denote

$$F_{r,0}(x) := 1 + 2 \sum_{k=1}^{\infty} k^{-r} \cos kx, \qquad x \in \mathbb{T};$$

$$F_{r,0}(x) := \prod_{j=1}^{d} F_{r,0}(x_j), \qquad x = (x_1, \dots, x_d) \in \mathbb{T}^d;$$

$$MW_{1,0}^r := \{ f : f = \varphi * F_{r,0}, \quad \|\varphi\|_1 \le 1 \}.$$

It is clear that in the case r is an even integer we have $MW_{1,0}^r = MW_1^r$. Let B be a positive number and $Q(B, m)$ be the set of cubature formulas $\Lambda_m(\cdot, \xi)$ satisfying the additional condition

$$\sum_{\mu=1}^{m} |\lambda_\mu| \le B.$$

We will obtain the lower estimates for the quantities

$$\delta_m^B(W) := \inf_{\Lambda_m(\cdot,\xi) \in Q(B,N)} \Lambda_m(W, \xi).$$

We will prove the following relation.

Theorem 3.4. *Let $r > 1$. Then*

$$\delta_m^B(MW_{1,0}^r) \geq C(r,B,d)m^{-r}(\log m)^{d-1}, \qquad C(r,B,d) > 0.$$

Proof. We use a similar to the above notation

$$\Lambda(k) = \Lambda_m(e^{i(k,x)},\xi) = \sum_{\mu=1}^{m} \lambda_\mu e^{i(k,\xi^\mu)}.$$

In the case $|\Lambda(0)| < 1/2$ it is sufficient to consider a function $f(x) \equiv 1$ as an example, and therefore we will assume that $|\Lambda(0)| \geq 1/2$. Considering the cubature formula $\tilde{\Lambda}_m(\cdot,\xi) = \Lambda_m(\cdot,\xi)/\Lambda(0)$, we see that

$$\Lambda_m(W,\xi) \geq \tilde{\Lambda}_m(W,\xi)/4$$

for W such that $\frac{1}{2}(f - \hat{f}(0)) \in W$ provided $f \in W$ and $\tilde{\Lambda}_m(\cdot,\xi)$ is exact on the function $f(x) \equiv 1$, i.e. $\tilde{\Lambda}(0) = 1$. Thus it is sufficient for our purpose to consider the cubature formulas $\Lambda_m(\cdot,\xi)$ satisfying the additional condition $\Lambda(0) = 1$. Let us consider the cubature formula Λ' constructed with the use of $\Lambda_m(\cdot,\xi)$ as follows:

$$\Lambda'(f) = \sum_{\nu=1}^{m} \bar{\lambda}_\nu \Lambda_m(f(x - \xi^\nu),\xi)$$

where $\bar{\lambda}_\nu$ is the complex conjugate to the λ_ν. Then

$$\Lambda'(k) := \Lambda'(e^{i(k,x)}) = \sum_{\nu=1}^{m} \bar{\lambda}_\nu \sum_{\mu=1}^{m} \lambda_\mu e^{i(k,(\xi^\mu - \xi^\nu))} = |\Lambda(k)|^2. \tag{3.15}$$

The function $F_{r,0}$ belongs to the closure in the uniform norm ($r > 1$) of the class $MW_{1,0}^r$. Consequently, by (3.15) and Lemma 3.1, we obtain

$$\Lambda'(MW_{1,0}^r) \geq \Lambda'(F_{r,0}) - \hat{F}_{r,0}(0) = \sum_{k\neq 0} \Lambda'(k)\hat{F}_{r,0}(k) =$$

$$\sum_{k\neq 0} |\Lambda(k)|^2 \hat{F}_{r,0}(k) \geq C(r,d)m^{-r}(\log m)^{d-1}. \tag{3.16}$$

On the other hand, for the cubature formula Λ' we have

$$\Lambda'(f) - \hat{f}(0) = \sum_{\nu=1}^{m} \bar{\lambda}_\nu (\Lambda_m(f(x - \xi^\nu),\xi) - \hat{f}(0))$$

which, for $\Lambda_m(\cdot,\xi) \in Q(B,N)$, implies the inequality

$$\Lambda'(MW_{1,0}^r) \leq B\Lambda_m(MW_{1,0}^r,\xi). \tag{3.17}$$

Relations (3.16) and (3.17) yield the required lower estimate for $\delta_m^B(MW_{1,0}^r)$.

Let us discuss how Theorems 3.2 and 3.4 can be used for estimating from below the generalized discrepancy $D_r(\xi,\Lambda,m,d)_q$. Theorem 3.2 combined with Theorem 1.1 and Remark 1.2 implies the following result.

Theorem 3.5. *Let $1 < q < \infty$ and r be a positive integer. Then for any points $\xi = (\xi^1, \dots, \xi^m) \subset \Omega_d$ and any weights $\Lambda = (\lambda_1, \dots, \lambda_m)$ we have*

$$D_r(\xi, \Lambda, m, d)_q \geq C(d,r) m^{-r} (\log m)^{(d-1)/2}$$

with a positive constant $C(d,r)$.

We now turn to application of Theorem 3.4. Let r be an even integer. then $MW^r_{1,0} = MW^r_1$. Assume that the given cubature formula $\Lambda_m(\cdot, \xi) \in Q(B,m)$. Then using the definition of $D_r(\xi, \Lambda, m, d)_\infty$ (see (1.15)) and the embedding $M\overset{0}{W}{}^r_p(\Omega_d) \hookrightarrow M\dot{W}^r_p(\Omega_d)$ we get

$$D_r(\xi, \Lambda, m, d)_\infty \gg \Lambda_m(M\overset{0}{W}{}^r_1(\Omega_d), \xi). \tag{3.18}$$

By (1.22) and the embedding $MW^r_1 \hookrightarrow MW^r_1(\Omega_d)$ we obtain

$$\Lambda_m(M\overset{0}{W}{}^r_1(\Omega_d), \xi) \gg \Lambda'_m(MW^r_1, \theta), \tag{3.19}$$

where $\theta = (\theta^1, \dots, \theta^m)$, $\theta^\mu = -\pi + 2\pi\eta^\mu$,

$$\eta^\mu_j = \psi(\xi^\mu_j), \qquad j = 1, \dots, d;$$

$$\lambda'_\mu = \lambda_\mu \prod_{j=1}^{d} \psi'(\xi^\mu_j), \qquad \mu = 1, \dots, m.$$

Next, it is clear that $\Lambda_m(\cdot, \xi) \in Q(B,m)$ implies that $\Lambda'_m(\cdot, \theta) \in Q(C(d)B, m)$. Therefore, by Theorem 3.4 we get

$$\Lambda'_m(MW^r_1, \theta) \gg m^{-r} (\log m)^{d-1}. \tag{3.20}$$

Combining (3.18)–(3.20) we obtain the following statement.

Theorem 3.6. *Let B be a positive number. For any points $\xi^1, \dots, \xi^m \subset \Omega_d$ and any weights $\Lambda = (\lambda_1, \dots, \lambda_m)$ satisfying the condition*

$$\sum_{\mu=1}^{m} |\lambda_\mu| \leq B$$

we have for even integers r

$$D_r(\xi, \Lambda, m, d)_\infty \geq C(d,B,r) m^{-r} (\log m)^{d-1}$$

with a positive constant $C(d,B,r)$.

Corollary 3.1. *Let r be an even integer. Then we have for the r-discrepancy*

$$D_r(\xi,m,d)_\infty := D_r(\xi,(1/m,\dots,1/m),d,m)_\infty \gg m^{-r}(\log m)^{d-1}.$$

The case $p=\infty$ is excluded in Theorem 3.2. There is no nontrivial general lower estimates in this case. We will give one conditional result in this direction.

Theorem 3.7. *Let the cubature formula (Λ,ξ) be such that the inequality*

$$\Lambda_m(MW^r_p,\xi) \ll m^{-r}(\log m)^{(d-1)/2}, \qquad r>1/p,$$

holds for some $1<p<\infty$.

Then

$$\Lambda_m(MW^r_\infty,\xi) \gg m^{-r}(\log m)^{(d-1)/2}.$$

Proof. We denote as above

$$\Lambda(k) = \sum_{j=1}^m \lambda_j e^{i(k,\xi^j)}.$$

Let us consider the function

$$g_{\Lambda,\xi,r}(x) = \sum_k \Lambda(k)\hat F_r(k)e^{i(k,x)} - 1.$$

Then for the quantity $\Lambda_m(MW^r_p,\xi)$ we have

$$\begin{aligned}\Lambda_m(MW^r_p,\xi) &= \sup_{f\in MW^r_p}\bigl|\Lambda_m(f,\xi)-\hat f(0)\bigr| = \\ &= \sup_{\|\varphi\|_p\le 1}\bigl|\Lambda_m\bigl(F_r(x)*\varphi(x),\xi\bigr)-\hat\varphi(0)\bigr| = \\ &= \sup_{\|\varphi\|_p\le 1}\bigl|\bigl(g_{\Lambda,\xi,r}(-y),\overline{\varphi(y)}\bigr)\bigr| = \|g_{\Lambda,\xi,r}\|_{p'}, \qquad p'=p/(p-1).\end{aligned} \tag{3.18}$$

Consequently, by hypothesis of Theorem 3.7, for some $1<q<\infty$, ($q=p'$) we have

$$\|g_{\Lambda,\xi,r}\|_q \ll m^{-r}(\log m)^{(d-1)/2}. \tag{3.19}$$

Further, for arbitrary $1<a<b$ and $f\in L_b$ the following inequality holds

$$\|f\|_a \le \|f\|_1^\kappa \|f\|_b^{1-\kappa}, \qquad \kappa = \left(\frac1a-\frac1b\right)\left(1-\frac1b\right)^{-1}. \tag{3.20}$$

By Theorem 3.2 we have for any $1\le z<\infty$

$$\Lambda_m(MW^r_{z,\alpha},\xi) \gg m^{-r}(\log m)^{(d-1)/2}. \tag{3.21}$$

Therefore, by (3.18)

$$\|g_{\Lambda,\xi,r}\|_{z'} \gg m^{-r}(\log m)^{(d-1)/2}. \tag{3.22}$$

Setting now $b=q$, $a=\frac12(b+1)$, $z'=a$ we get from relations (3.20), (3.19), (3.22)

$$\|g_{\Lambda,\xi,r}\|_1 \gg m^{-r}(\log m)^{(d-1)/2}.$$

It suffices to apply the relation (3.18) to complete the proof.

Theorem 3.7 is proved.

Remark 3.2. *We have actually proved the following inequality. Let $1 \le p_1 < p_2 < \infty$, then for any (Λ, ξ)*

$$\Lambda_m(MW^r_{p_2}, \xi) \le \Lambda_m(MW^r_{p_1}, \xi)^{\frac{p_1}{p_2}} \Lambda_m(MW^r_{\infty}, \xi)^{1-\frac{p_1}{p_2}}, \qquad (r > 1/p_1).$$

4. Upper estimates for the classes MW^r_p

4.1. The Fibonacci cubature formulas. For periodic functions of two variables we consider the Fibonacci cubature formulas

$$\Phi_n(f) = b_n^{-1} \sum_{\mu=1}^{b_n} f(2\pi\mu/b_n,\, 2\pi\{\mu b_{n-1}/b_n\}),$$

where $b_0 = b_1 = 1$, $b_n = b_{n-1} + b_{n-2}$ are the Fibonacci numbers and $\{x\}$ is the fractional part of the number x.

For a function class W we denote

$$\Phi_n(W) = \sup_{f \in W} |\Phi_n(f) - (2\pi)^{-2} \int_{\mathbb{T}^2} f(x)\, dx|,$$

where $\mathbb{T}^2 = [0, 2\pi]^2$ is the period square.

The following known result gives the order of $\Phi_n(W^r_p)$ for all parameters $1 \le p \le \infty$, $r > 1/p$.

Theorem 4.1. *We have*

$$\Phi_n(MW^r_p) \asymp \begin{cases} b_n^{-r}(\log b_n)^{1/2}, & 1 < p \le \infty, r > \max\left(\frac{1}{p}, \frac{1}{2}\right); \\ b_n^{-r} \log b_n, & p = 1, r > 1; \\ b_n^{-r}(\log b_n)^{1-r}, & 2 < p \le \infty, \frac{1}{p} < r < \frac{1}{2}; \\ b_n^{-r}((\log b_n)(\log\log b_n))^{1/2}, & 2 < p \le \infty, r = 1/2. \end{cases}$$

The lower estimates provided by Theorem 3.2 and the upper estimates from Theorem 4.1 show that the Fibonacci cubature formulas are optimal (in the sense of order) among all cubature formulas in the case $1 < p < \infty$, $r > \max(1/p, 1/2)$:

$$\delta_{b_n}(MW^r_p) \asymp \Phi_n(MW^r_p) \asymp b_n^{-r}(\log b_n)^{1/2}.$$

Theorem 3.4 combined with Theorem 4.1 implies that the Fibonacci cubature formulas are optimal (in the sense of order) among formulas satisfying an additional restriction $\sum_{\mu=1}^m |\lambda_\mu| \le B$ in the case r an even integer and $p = 1$

$$\delta^B_{b_n}(MW^r_1) \asymp \Phi_n(MW^r_1) \asymp b_n^{-r} \log b_n.$$

For all other values of parameters p and r, $r > 1/p$, the right order of $\delta_m(MW^r_p)$ is unknown.

4.2. The Korobov cubature formulas. It will be convenient for us to denote vectors in $\mathbb{Z}^d$ and $\mathbb{R}^d$ by bold letters. Let $m \in \mathbb{N}$, $\mathbf{a} = (a_1, \dots, a_d) \in \mathbb{Z}^d$. We consider the cubature formulas

$$P_m(f, \mathbf{a}) = m^{-1} \sum_{\mu=1}^{m} f\left(2\pi\left\{\frac{\mu a_1}{m}\right\}, \dots, 2\pi\left\{\frac{\mu a_d}{m}\right\}\right)$$

which will be called the Korobov cubature formulas.

In the case $d = 2$, $m = b_n$, $\mathbf{a} = (1, b_{n-1})$ we have

$$P_m(f, \mathbf{a}) = \Phi_n(f).$$

We note that in the case $d > 2$ the problem of finding concrete cubature formulas of the type $P_m(f, \mathbf{a})$ as good as the Fibonacci cubature formulas in the case $d = 2$ is unsolved. The results of this subsection deal with the case $d > 2$ and are not as complete as the results of Subsection 4.1.

We first prove an auxiliary assertion. Denote

$$\Gamma(N) := \{\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{Z}^d : \prod_{j=1}^{d} \max(|k_j|, 1) \le N\}.$$

For a finite set E the cardinality of E will be denoted by $|E|$.

Lemma 4.1. *Let n, κ, L be a prime, a positive real and a natural number, respectively, such that*

$$\left|\Gamma(L)\right| < (n-1)(1 - 2^{-\kappa})/d. \tag{4.1}$$

Then there is a natural number $a \in I_n := [1, n)$ such that for all $\mathbf{m} \in \Gamma(L)$, $\mathbf{m} \ne \mathbf{0}$

$$m_1 + a m_2 + \dots + a^{d-1} m_d \not\equiv 0 \qquad (\mathrm{mod}\ n), \tag{4.2}$$

and the relation (4.2) will be valid for all $\mathbf{m} \in F_l(L) := \Gamma(L2^l) \setminus \Gamma(L2^{l-1})$, $\mathbf{m} \ne n\mathbf{m}'$, with the exception of no more than

$$A_l^L := \left|F_l(L)\right| d 2^{(l+1)\kappa} (2^\kappa - 1)^{-1} (n-1)^{-1}, \qquad l = 1, 2, \dots .$$

Proof. Let $a \in I_n$ be a natural number. We consider the congruence

$$m_1 + a m_2 + \dots + a^{d-1} m_d \equiv 0 \qquad (\mathrm{mod}\ n). \tag{4.3}$$

For a fixed vector $\mathbf{m} = (m_1, \dots, m_d)$ we denote by $A_n(\mathbf{m})$ the set of natural numbers $a \in I_n$ which are solutions of the congruence (4.3). It is well-known that for $\mathbf{m} \ne \mathbf{0}$, $|m_j| < n$, $j = 1, \dots, d$ the number $\left|A_n(\mathbf{m})\right|$ of the elements of the set $A_n(\mathbf{m})$ satisfies the inequality

$$\left|A_n(\mathbf{m})\right| \le d - 1 < d. \tag{4.4}$$

We denote by G_1 the set of the numbers a which are solutions of the congruence (4.3) for at least one of $\mathbf{m} \in \Gamma(L)$, that is

$$G_1 = \cup_{\mathbf{m}\in\Gamma(L)} A_n(\mathbf{m}).$$

Let us estimate the number $|G_1|$ of elements of the set G_1. By (4.4) and (4.1) we have

$$|G_1| \le \sum_{\mathbf{m}\in\Gamma(L)} \big|A_n(\mathbf{m})\big| < d\big|\Gamma(L)\big| < (n-1)(1-2^{-\kappa}). \tag{4.5}$$

For any $a \in I_n \backslash G_1$ for all $\mathbf{m} \in \Gamma(L)$ we have

$$m_1 + am_2 + \cdots + a^{d-1}m_d \not\equiv 0 \pmod{n}.$$

Let G_{l+1}, $l = 1, 2, \dots$ denote the set of those a for which the number of elements of the set

$$M_a^l := \big\{\mathbf{m} : \mathbf{m} \in F_l(L), \quad \mathbf{m} \ne n\mathbf{m}', \quad m_1 + am_2 + \cdots + a^{d-1}m_d \equiv 0 \pmod{n}\big\}$$

satisfies the inequality

$$|M_a^l| > A_l^L. \tag{4.6}$$

Then by (4.6)

$$\sum_{a\in G_{l+1}} |M_a^l| > A_l^L |G_{l+1}|. \tag{4.7}$$

On the other hand, by (4.4) each $\mathbf{m}$ can belong to at most $d-1$ different sets M_a^l and therefore

$$\sum_{a\in G_{l+1}} |M_a^l| < d\big|F_l(L)\big|. \tag{4.8}$$

Comparing (4.7) and (4.8) we find

$$|G_{l+1}| < d\big|F_l(L)\big|/A_l^L = (n-1)(2^\kappa - 1)2^{-\kappa(l+1)}. \tag{4.9}$$

From relations (4.5) and (4.9) it follows that

$$\sum_{l=1}^{\infty} |G_l| < n-1.$$

This means that there exists a number $a \in I_n$ which does not belong to any set G_l, $l = 1, \dots$. This a is the required number by the definition of the sets G_l. The lemma is proved.

For $\mathbf{a}$ of the form $\mathbf{a} = (1, a, \dots, a^{d-1})$ denote $P_n(f, a) := P_n(f, \mathbf{a})$.

Theorem 4.2. *For any r such that $r > \max(1/p; 1/2)$ and any prime number n there is a natural number $a \in [1, n)$ for which*

$$P_n(MW_p^r, a) \le C(r,p,d) n^{-r} (\log n)^{r(d-1)}, \qquad 1 < p \le \infty.$$

Proof. Clearly, it suffices to consider the case $1 < p \le 2$. Let $\kappa := (r - 1/p)/2$, n be given and L satisfies (4.1) and in addition $|\Gamma(L)| \asymp n$. Let a be the number from Lemma 4.1, depending on n, κ, L and

$$Z_l = \{\mathbf{m} : \mathbf{m} \in F_l(L), \quad \mathbf{m} \ne n\mathbf{m}', \quad m_1 + am_2 + \dots + a^{d-1} m_d \equiv 0 \pmod{n}\}.$$

Then by Lemma 4.1

$$|Z_l| \le A_l^L, \tag{4.10}$$

and the error of the cubature formula can be estimated as follows

$$|P_n(f,a) - \hat{f}(\mathbf{0})| \le \left| \sum_{l=1}^{\infty} \sum_{\mathbf{m} \in Z_l} \hat{f}(\mathbf{m}) \right| + \left| \sum_{\mathbf{m} = n\mathbf{m}'} \hat{f}(\mathbf{m}) \right| =: \sigma_1 + \sigma_2. \tag{4.11}$$

Let us estimate σ_1, σ_2 from (4.11) for $f \in MW_p^r$.

We denote

$$\psi_l(x) = \sum_{\mathbf{m} \in Z_l} e^{i(k,x)}.$$

We have

$$\sigma_1 = \left| \sum_{l=1}^{\infty} (f, \psi_l) \right| \le \sum_{l=1}^{\infty} E_{L2^{l-1}}(f)_p \|\psi_l\|_{p'}, \tag{4.12}$$

where $E_N(f)_p$ denotes best approximation of f in L_p by trigonometric polynomials with frequencies in $\Gamma(N)$. It is known (see [47,Ch.3,Th.3.2]) that for $f \in MW_p^r$, $1 < p < \infty$ we have

$$E_N(f)_p \le C(r,p,d) N^{-r}, \qquad r > 0.$$

Further

$$\|\psi_l\|_{p'} \le \|\psi_l\|_2^{2/p'} \|\psi_l\|_\infty^{1-2/p'} \le |Z_l|^{1/p}. \tag{4.13}$$

From (4.12) by the estimate of $E_N(f)_p$ and by (4.13) and (4.10) we get

$$\sigma_1 \ll L^{-r} \ll n^{-r} (\log n)^{r(d-1)}.$$

For σ_2 we have

$$\sigma_2 \le \sum_{\mathbf{m}' \ne \mathbf{0}} |\hat{f}(n\mathbf{m}')| \ll n^{-r} \sum_{\mathbf{m}'} \nu(\overline{\mathbf{m}}')^{-r} |\hat{\varphi}(n\mathbf{m}')|, \tag{4.14}$$

where φ is such that $f = F_r * \varphi$, $\|\varphi\|_p \le 1$.

From (4.14), applying the Hölder inequality and the Hausdorff-Young theorem, we get

$$\sigma_2 \ll n^{-r} \left(\sum_{\mathbf{m}'} \nu(\overline{\mathbf{m}}')^{-rp} \right)^{1/p} \left(\sum_{\mathbf{m}'} |\hat{\varphi}(n\mathbf{m}')|^{p'} \right)^{1/p'} \ll n^{-r} \|\varphi\|_p \le n^{-r}.$$

The conclusion of the theorem follows from the estimates for σ_1 and σ_2.

4.3. The Frolov cubature formulas. In this subsection we construct the optimal (in the sense of order) cubature formulas for the classes $M\overset{0}{W}{}^r_p(\Omega_d)$, $2 \le p < \infty$.

The following lemma plays a fundamental role in the construction of such cubature formulas.

Lemma 4.2. *There exists a matrix A such that the lattice $L(\mathbf{m}) = A\mathbf{m}$, where $\mathbf{m}$ is a (column) vector with integer coordinates, has the following properties* $\left(L(\mathbf{m}) = \begin{pmatrix} L_1(\mathbf{m}) \\ \vdots \\ L_d(\mathbf{m}) \end{pmatrix} \right)$

1^0 $\left|\prod_{j=1}^d L_j(\mathbf{m})\right| \ge 1$ *for all* $\mathbf{m} \ne \mathbf{0}$*;*

2^0 *each parallelepiped P with volume $|P|$ whose edges are parallel to the coordinate axes contains no more than $|P| + 1$ lattice points.*

Let $a > 1$ and A be the matrix from Lemma 4.2. We consider the cubature formula

$$\Phi(a, A)(f) = \left(a^d |\det A|\right)^{-1} \sum_{\mathbf{m} \in \mathbb{Z}^d} f\left(\frac{(A^{-1})^T \mathbf{m}}{a}\right)$$

for $f \in M\overset{0}{W}{}^r_2(\Omega_d)$. Clearly, the number N of points of this cubature formula does not exceed $C(A)a^d |\det A|$.

Theorem 4.3. *Let a matrix A be from Lemma 4.2 and let r be a natural number. Then*

$$\Phi(a, A)\left(M\overset{0}{W}{}^r_2(\Omega_d)\right) \le C(A, d) a^{-rd} (\log a)^{\frac{d-1}{2}}.$$

Proof. We will use the Poisson formula which we formulate in a form convenient for us. We denote for $f \in L_1(\mathbb{R}^d)$

$$\hat{f}(\mathbf{y}) = \int_{\mathbb{R}^d} f(\mathbf{x}) e^{-2\pi i (\mathbf{y}, \mathbf{x})} d\mathbf{x}.$$

Lemma 4.3. *Let $f(\mathbf{x})$ be continuous and have compact support and the series $\sum_{\mathbf{k} \in \mathbb{Z}^d} \hat{f}(\mathbf{k})$ converges. Then*

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} \hat{f}(\mathbf{k}) = \sum_{\mathbf{n} \in \mathbb{Z}^d} f(\mathbf{n}).$$

By Lemma 4.3 the identity

$$\Phi(a, A)(f) = \left(a^d |\det A|\right)^{-1} \sum_{\mathbf{m}} \int_{\mathbb{R}^d} f\left(\frac{(A^{-1})^T \mathbf{x}}{a}\right) e^{-2\pi i (\mathbf{m}, \mathbf{x})} d\mathbf{x} =$$

$$= \sum_{\mathbf{m}} \int_{\mathbb{R}^d} f(\mathbf{y}) e^{-2\pi i (aA\mathbf{m}, \mathbf{y})} d\mathbf{y} = \sum_{\mathbf{m}} \hat{f}(aA\mathbf{m}), \tag{4.15}$$

holds under the assumption that the series in the right side of (4.15) converges. The convergence of this series will follow from further consideration. In the relation (4.15) we carried out the linear change of variables $\mathbf{y} = \frac{(A^{-1})^T\mathbf{x}}{a}$.

We have for the error of this cubature formula

$$\delta = \Phi(a, A)(f) - \hat{f}(\mathbf{0}) = \sum_{\mathbf{m}\neq\mathbf{0}} \hat{f}(aA\mathbf{m}). \tag{4.16}$$

We need the following simple assertion.

Lemma 4.4. *Let $\|\varphi\|_2 \le 1$ and the support of φ be contained in Ω_d. Then for any $a > 1$ and nonsingular matrix A we have*

$$\sum_{\mathbf{m}} \left|\hat{\varphi}(aA\mathbf{m})\right|^2 \le C(A).$$

Proof. Similarly to (4.15) we have

$$\hat{\varphi}(aA\mathbf{m}) = \left(a^d|\det A|\right)^{-1} \int_{\Omega_d} \varphi\left(\frac{(A^{-1})^T\mathbf{x}}{a}\right) e^{-2\pi i(\mathbf{m},\mathbf{x})} d\mathbf{x}. \tag{4.17}$$

Let

$$\Omega_d(\mathbf{n}) := \{\mathbf{x} : \mathbf{x} = \mathbf{y} + \mathbf{n}, \qquad \mathbf{y} \in \Omega_d\}$$

and

$$G := \left\{\mathbf{n} : \left(\operatorname{supp}\varphi\left(\frac{(A^{-1})^T\mathbf{x}}{a}\right)\right) \cap \Omega_d(\mathbf{n}) \neq \varnothing\right\}.$$

From the hypothesis of the lemma it follows that

$$|G| \le C_1(A)a^d. \tag{4.18}$$

We denote

$$c_{\mathbf{m}}(\mathbf{n}) = \int_{\Omega_d(\mathbf{n})} \varphi\left(\frac{(A^{-1})^T\mathbf{x}}{a}\right) e^{-2\pi i(\mathbf{m},\mathbf{x})} d\mathbf{x}.$$

By the Parseval identity

$$\sum_{\mathbf{m}} \left|c_{\mathbf{m}}(\mathbf{n})\right|^2 = \int_{\Omega_d(\mathbf{n})} \left|\varphi\left(\frac{(A^{-1})^T\mathbf{x}}{a}\right)\right|^2 d\mathbf{x}. \tag{4.19}$$

From the relation (4.17), using the Cauchy inequality and the inequality (4.18), we get

$$\left|\hat{\varphi}(aA\mathbf{m})\right|^2 = \left(a^d|\det A|\right)^{-2} \left|\sum_{\mathbf{n}\in G} c_{\mathbf{m}}(\mathbf{n})\right|^2 \le$$
$$\le \left(a^d|\det A|\right)^{-2}|G| \sum_{\mathbf{n}\in G} \left|c_{\mathbf{m}}(\mathbf{n})\right|^2 \le \left(a^d|\det A|\right)^{-1} C_2(A) \sum_{\mathbf{n}\in G} \left|c_{\mathbf{m}}(\mathbf{n})\right|^2.$$

Performing the summation over $\mathbf{m}$ and taking into account the relation (4.19) we get

$$\sum_{\mathbf{m}} \big|\hat{\varphi}(aA\mathbf{m})\big|^2 \le C_2(A)\big(a^d|\det A|\big)^{-1} \int_{\mathbb{R}^d} \left|\varphi\left(\frac{(A^{-1})^T\mathbf{x}}{a}\right)\right|^2 d\mathbf{x} =$$

$$= C_2(A)\int_{\mathbb{R}^d} \big|\varphi(\mathbf{y})\big|^2 d\mathbf{y} \le C_2(A).$$

The lemma is proved.

We continue the proof of Theorem 4.3.

Let $f \in \overset{0}{MW}{}^r_2(\Omega_d)$ be rd times continuously differentiable. We denote $\varphi(x) = \frac{\partial^{rd} f}{\partial x_1^r \dots \partial x_d^r}$. Then for $\mathbf{m} \neq \mathbf{0}$

$$\hat{f}(aA\mathbf{m}) = \hat{\varphi}(aA\mathbf{m}) \prod_{j=1}^d \big(2\pi i a L_j(\mathbf{m})\big)^{-r}.$$

Let l be such that

$$2^{l-1} < a^d \le 2^l.$$

Then by the property 1^0 of Lemma 4.2 the inequality $\|\mathbf{s}\|_1 \ge l$ holds for $\mathbf{s}$ such that $\rho(\mathbf{s})$ (see the definition of $\rho(\mathbf{s})$ in (3.10)) contains a point $aA\mathbf{m}$ with $\mathbf{m} \neq \mathbf{0}$.

Then

$$(2\pi)^{-rd}\delta \le \sum_{\|\mathbf{s}\|_1 \ge l} \sum_{aL(\mathbf{m}) \in \rho(\mathbf{s})} \big|\hat{\varphi}(aA\mathbf{m})\big| \prod_{j=1}^d \big|aL_j(\mathbf{m})\big|^{-r} \le$$

$$\le \left(\sum_{\|\mathbf{s}\|_1 \ge l} \sum_{aL(\mathbf{m}) \in \rho(\mathbf{s})} \prod_{j=1}^d \big|aL_j(\mathbf{m})\big|^{-2r}\right)^{1/2} \left(\sum_{\mathbf{m}} \big|\hat{\varphi}(aA\mathbf{m})\big|^2\right)^{1/2}. \tag{4.20}$$

Applying Lemma 4.4 and using the relation

$$\sum_{\|\mathbf{s}\|_1 \ge l} \sum_{aL(\mathbf{m}) \in \rho(\mathbf{s})} \prod_{j=1}^d \big|aL_j(\mathbf{m})\big|^{-2r} \ll \sum_{\|\mathbf{s}\|_1 \ge l} 2^{-2r\|\mathbf{s}\|_1} \left(\frac{|\rho(\mathbf{s})|}{a^d} + 1\right) \ll$$

$$\ll 2^{-2rl} l^{d-1} \ll a^{-2rd} (\log a)^{d-1}.$$

we get from here and (4.20)

$$\delta \ll a^{-rd} (\log a)^{\frac{d-1}{2}}.$$

Theorem 4.3 is proved.

Theorem 4.3 combined with Theorem 1.1 and Theorem 3.2 implies the following theorem.

Theorem 4.4. *Let $r \in \mathbb{N}$ and $2 \le p < \infty$. Then*

$$\delta_m(MW_p^r) \asymp m^{-r} (\log m)^{\frac{d-1}{2}}.$$

5. Historical notes, comments, and some open problems

First we will give a brief historical survey on the discrepancy. We refer the reader for a complete survey to the following books on discrepancy and related topics L. Kuipers and H. Niederreiter [27], J. Beck and W. Chen [5], J. Matoušek [29], and B. Chazelle [9]. We formulate all results in the notations of this paper and in the form convenient for us. We use the following notation

$$D(X,m,d)_q := \Big(\int_{\Omega_d} |\prod_{j=1}^d a_j - \frac{1}{m}\sum_{\mu=1}^m \chi_{[0,a]}(x^\mu)|^q da\Big)^{1/q}, \quad 1<q<\infty,$$

$$D(X,m,d)_\infty := \max_{a\in[0,1]^d} |\prod_{j=1}^d a_j - \frac{1}{m}\sum_{\mu=1}^m \chi_{[0,a]}(x^\mu)|,$$

where $X=(x^1,\dots,x^m)$. The first result in this area was the following conjecture of van der Corput [11,12] formulated in 1935. Let $\xi^j\in[0,1]$, $j=1,2,\dots$, then we have

$$\limsup_{m\to\infty} mD((\xi^1,\dots,\xi^m),m,1)_\infty = \infty.$$

This conjecture was proved by van Aardenne-Ehrenfest [1] in 1945:

$$\limsup_{m\to\infty} \frac{\log\log\log m}{\log\log m} mD((\xi^1,\dots,\xi^m),m,1)_\infty > 0.$$

Let us denote

$$D(m,d)_q := \inf_X D(X,m,d)_q, \qquad 1\le q\le\infty.$$

In 1954 K. Roth [32] proved that

$$D(m,d)_2 \ge C(d)m^{-1}(\log m)^{(d-1)/2}. \tag{5.1}$$

In 1972 W. Schmidt [36] proved

$$D(m,2)_\infty \ge Cm^{-1}\log m. \tag{5.2}$$

In 1977 W. Schmidt [37] proved

$$D(m,d)_q \ge C(d,q)m^{-1}(\log m)^{(d-1)/2}, \qquad 1<q\le\infty. \tag{5.3}$$

In 1981 G. Halász [21] proved

$$D(m,d)_1 \ge C(d)m^{-1}(\log m)^{1/2}. \tag{5.4}$$

The following conjecture has been formulated in [5] as an excruciatingly difficult great open problem.

Conjecture 5.1 ([5]). *We have for $d \geq 3$*

$$D(m,d)_\infty \geq C(d)m^{-1}(\log m)^{d-1}.$$

This problem is still open.

We now present the results on the lower estimates for the r-discrepancy. We denote

$$D_r(m,d)_q := \inf_\xi D_r(\xi,(1/m,\dots,1/m),m,d)_q$$

where $D_r(\xi,\Lambda,m,d)_q$ is defined in (1.15) and also denote

$$D_r^o(m,d)_q := \inf_{\xi,\Lambda} D_r(\xi,\Lambda,m,d)_q.$$

It is clear that

$$D_r^o(m,d)_q \leq D_r(m,d)_q.$$

The first result in estimating the generalized discrepancy was obtained in 1985 by V.A. Bykovskii [6]

$$D_r^o(m,d)_2 \geq C(r,d)m^{-r}(\log m)^{(d-1)/2}. \tag{5.5}$$

This result is a generalization of the Roth's result (5.1). The generalization of the Schmidt's result (5.3) was obtained by the author in 1990, [45], (see Theorem 3.5 of this paper)

$$D_r^o(m,d)_q \geq C(r,d,q)m^{-r}(\log m)^{(d-1)/2}, \qquad 1 < q \leq \infty. \tag{5.6}$$

In 1994, [48], the author proved that for r even integers we have for the r-discrepancy (see Theorem 3.6 and Corollary 3.1 of this paper)

$$D_r(m,d)_\infty \geq C(r,d)m^{-r}(\log m)^{d-1}. \tag{5.7}$$

This result encourages us to formulate the following generalization of the Conjecture 5.1.

Conjecture 5.2. *For all $d,r \in \mathbb{N}$ we have*

$$D_r^o(m,d)_\infty \geq C(r,d)m^{-r}(\log m)^{d-1}.$$

The above lower estimates for $D_1^o(m,d)_q$ are formally stronger than the corresponding estimates for $D(m,d)_q$ because in $D_1^o(m,d)_q$ we are optimizing over the weights Λ. However, the proofs for $D(m,d)_q$ could be adjusted to give the estimates for $D_1^o(m,d)_q$. The results (5.5)–(5.7) for the generalized discrepancy were obtained as a corollary of the corresponding results on cubature formulas (see Theorem 1.1 and Theorems 3.5, 3.6). We do not know if existing methods for $D(m,d)_q$ could be modified to obtain the estimates for $D_r^o(m,d)_q$, $r \geq 2$.

We proceed to the lower estimates for the cubature formulas. Theorem 3.1 and Lemma 3.1 were established in [6]. The proof is taken from [47]. Theorems 3.2, 3.3, and Lemma 3.2 were proved in [45]. Theorem 3.4 was proved in [48]. Theorem 3.7 is from [46]. There are two big open problems in this area. We formulate them as conjectures.

Conjecture 5.3. *For any $d \geq 2$ and any $r \geq 1$ we have*

$$\delta_m(MW_1^r) \geq C(r,d)m^{-r}(\log m)^{d-1}.$$

Conjecture 5.4. *For any $d \geq 2$ and any $r > 0$ we have*

$$\delta_m(MW_\infty^r) \geq C(r,d)m^{-r}(\log m)^{(d-1)/2}.$$

We note that by Proposition 1.2, Theorem 1.1, and (1.15) Conjecture 5.3 implies Conjecture 5.2 and Conjecture 5.4 implies

$$D_r^o(m,d)_1 \geq C(r,d)m^{-r}(\log m)^{(d-1)/2}. \tag{5.8}$$

We turn to the upper estimates. We begin with the cubature formulas. We have already made a historical remark on classes with bounded mixed derivative in Section 1. We will discuss only these classes here. For results on cubature formulas for the Sobolev type classes we refer the reader to the books of S.L. Sobolev [40], E. Novak [31], and the author [47,Ch.2]. The first result in this direction was obtained by N.M. Korobov [25] in 1959. He used the cubature formulas $P_m(f,\mathbf{a})$ defined in Subsection 4.2. We note that similar cubature formulas were also used by E. Hlawka [22]. The Korobov's results lead to the following estimate

$$\delta_m(MW_1^r) \leq C(r,d)m^{-r}(\log m)^{rd}, \qquad r > 1. \tag{5.9}$$

In 1959 N.S. Bakhvalov [2] improved (5.9) to

$$\delta_m(MW_1^r) \leq C(r,d)m^{-r}(\log m)^{r(d-1)}, \qquad r > 1.$$

The first best possible upper estimate for the classes MW_p^r was obtained by N.S. Bakhvalov [3] in 1963. He proved in the case $d=2$ that

$$\delta_m(MW_2^r) \leq C(r)m^{-r}(\log m)^{1/2}, \qquad r \in \mathbb{N}. \tag{5.10}$$

N.S. Bakhvalov used the Fibonacci cubature formulas defined in Subsection 4.1.

In 1976 K.K. Frolov [18] used the cubature formulas defined in Subsection 4.3 to extend (5.10) to the case $d > 2$:

$$\delta_m(MW_2^r) \leq C(r,d)m^{-r}(\log m)^{(d-1)/2}, \qquad r \in \mathbb{N}. \tag{5.11}$$

In 1985 this estimate was further generalized by V.A. Bykovskii [6] to $r \in \mathbb{R}$, $r \geq 1$. Bykovskii also used the Frolov cubature formulas. One can find these results in Section 4 of this paper. Theorem 4.1 was proved in [46] and [48]. Theorem 4.2 and Lemma 4.1 are from [44]. Theorem 4.2 addresses the case of small smoothness: $r > \max(1/p, 1/2)$ instead of $r \geq 1$. Lemma 4.2 is a well known result in algebraic number theory (see [8]). Theorem 4.3 was obtained by K.K. Frolov [18, 19]. The proof is taken from [47]. We note that there is no sharp results for $\delta_m(MW_p^r)$ in the case of small smoothness $1/p < r < 1$. It is an interesting open problem. The approach based on nonlinear m-term approximation (see Section 2) can be useful in this case.

The Frolov cubature formulas [19] give the following estimate

$$\delta_m(MW_1^r) \le C(r,d)m^{-r}(\log m)^{d-1}, \qquad r > 1. \tag{5.12}$$

Thus the lower estimate in Conjecture 5.3 is the best possible.

In 1994 M.M. Skriganov [39] proved the following estimate

$$\delta_m(MW_p^r) \le C(r,d,p)m^{-r}(\log m)^{(d-1)/2}, \quad 1 < p \le \infty, \quad r \in \mathbb{N}. \tag{5.13}$$

This estimate combined with Theorem 3.2 implies

$$\delta_m(MW_p^r) \asymp m^{-r}(\log m)^{(d-1)/2}, \quad 1 < p < \infty, \quad r \in \mathbb{N}. \tag{5.14}$$

Another proofs of (5.13) and Theorem 3.2 were given in 1995 by V.A. Bykovskii [7].

We now present the upper estimates for the discrepancy. In 1956 H. Davenport [13] proved that

$$D(m,2)_2 \le Cm^{-1}(\log m)^{1/2}.$$

Another proofs of this estimate were later given by I.V. Vilenkin [52], J.H. Halton and S.K. Zaremba [24], and K. Roth [33]. In 1979 K. Roth [34] proved

$$D(m,3)_2 \le Cm^{-1}\log m$$

and in 1980 K. Roth [35] and K.K. Frolov [20] proved

$$D(m,d)_2 \le C(d)m^{-1}(\log m)^{(d-1)/2}.$$

In 1980 W. Chen [10] proved

$$D(m,d)_q \le C(d)m^{-1}(\log m)^{(d-1)/2}, \qquad q < \infty.$$

The estimate (5.12) and Theorem 1.1 imply

$$D_r^o(m,d)_\infty \le C(r,d)m^{-r}(\log m)^{d-1}, \qquad r \ge 2.$$

We note that the upper estimates for $D(m,d)_q$ are stronger than the same upper estimates for $D_1^o(m,d)_q$.

Let us also mention a classical book of S.M. Nikol'skii [30] on quadrature formulas and books of N.M. Korobov [26], W. Schmidt [38], and Hua Loo Keng and Wang Yuan [23] on discrepancy and related topics. We discussed in this paper only the case of the class MW_p^r of functions with bounded mixed derivative. There are analogs of classes MW_p^r which are also natural in the theory of cubature formulas. One can find results on numerical integration of functions with bounded mixed difference in the papers [4], [17], [43], and in the book [47]. A different method of constructing cubature formulas for functions with bounded mixed derivative was suggested by S.A. Smolyak [41] (for further results see [42]).

Acknowledgments. The author would like to thank H. Niederreiter, K. Petras, and H. Wozniakovski the organizers of the workshop "Numerical Integration and its Complexity", Oberwolfach, November 18–24, 2001, which stimulated writing this paper. Also, the author is grateful to E. Novak for useful discussions of relations between cubature formulas and nonlinear approximation.

References

[1] T. van Aardenne-Ehrenfest, *Proof of the impossibility of a just distribution of an infinite sequence of points over an interval*, Proc. Kon. Ned. Akad. v. Wetensch **48** (1945), 266–271.

[2] N.S. Bakhvalov, *On the approximate computation of multiple integrals*, Vestnik Moskov. Univ. Ser. Mat. Mekh. Astr. Fiz. Khim. (1959), no. 4, 3–18.

[3] N.S. Bakhvalov, *Optimal convergence bounds for quadrature processes and integration methods of Monte Carlo type for classes of functions*, Zh. Vychisl. Mat. i Mat. Fiz. Suppl. **4** (1963), no. 4, 5–63.

[4] N.S. Bakhvalov, *Lower estimates of assimptotic characteristics of classes of functions with dominant mixed derivative*, Matem. Zametki **12** (1972), 655–664; English transl. in Math. Notes **12** (1972).

[5] J. Beck and W. Chen, *Irregularities of distribution*, Cambridge University Press, Cambridge, 1987.

[6] V.A. Bykovskii, *On the correct order of the error of optimal cubature formulas in spaces with dominant derivative, and on quadratic deviations of grids*, Preprint, Computing Center Far-Eastern Scientific Center, Akad. Sci. USSR, Vladivostok, 1985.

[7] V.A. Bykovskii, *Estimates of deviations of optimal lattices in the L_p-norm and the theory of quadrature formulas*, Preprint, Applied Math. Inst., Far-Eastern Scientific Center, Akad. Sci. Russia, Khabarovsk, 1995.

[8] J.W.S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, Berlin, 1971.

[9] B. Chazelle, *The Discrepancy Method*, Cambridge University Press, Cambridge, 2000.

[10] W.W.L. Chen, *On irregularities of distribution*, Mathematika **27** (1980), 153–170.

[11] J.G. van der Corput, *Verteilungsfunktionen.I*, Proc. Kon. Ned. Akad. v. Wetensch. **38** (1935), 813–821.

[12] J.G. van der Corput, *Verteilungsfunktionen.II*, Proc. Kon. Ned. Akad. v. Wetensch. **38** (1935), 1058–1066.

[13] H. Davenport, *Note on irregularities of distribution*, Mathematika **3** (1956), 131–135.

[14] M. Donahue, L. Gurvits, C. Darken, E. Sontag, *Rate of convex approximation in non-Hilbert spaces*, Constructive Approximation **13** (1997), 187–220.

[15] R.A. DeVore, *Nonlinear Approximation*, Acta Numerica (1998), 51–150.

[16] R.A. DeVore and V.N. Temlyakov, *Some remarks on Greedy Algorithms*, Advances in Computational Mathematics **5** (1996), 173–187.

[17] V.V. Dubinin, *Cubature formulas for classes of functions with bounded mixed difference*, Matem. Sbornik **183** (1992), no. 7; English transl. in Math. USSR Sbornik **76** (1993), 283–292.

[18] K.K. Frolov, *Upper bounds on the error of quadrature formulas on classes of functions*, Dokl. Akad. Nauk SSSR **231** (1976), 818–821; English transl. in Soviet Math. Dokl. **17** (1976).

[19] K.K. Frolov, *Quadrature formulas on classes of functions*, Candidate's dissertation, Vychisl. Tsentr Acad. Nauk SSSR, Moscow, 1979. (Russian)

[20] K.K. Frolov, *An upper estimate of the discrepancy in the L_p-metric,* $2 \leq p < \infty$, Dokl. Akad. Nauk SSSR **252** (1980), 805–807; English transl. in Soviet Math. Dokl. **21** (1980).

[21] G. Halász, *On Roth's method in the theory of irregularities of points distributions*, Recent Progress in Analytic Number Theory **2** (1981), 79–94.

[22] E. Hlawka, *Zur angenaherten Berechnung mehrfacher Integrale*, Manatsh. Math. **B66** (1962), 140–151.

[23] Hua Loo Keng and Wang Yuan, *Applications of number theory to numerical analysis*, Springer, Berlin, 1981.

[24] J.H. Halton and S.K. Zaremba, *The extreme and L_2 discrepancies of some plane sets*, Monats. für Math. **73** (1969), 316–328.

[25] N.M. Korobov, *On the approximate computation of multiple integrals*, Dokl. Akad. Nauk SSSR **124** (1959), 1207–1210.

[26] N.M. Korobov, *Number-theoretic methods in numerical analysis*, Fizmatgis, Moscow, 1963.

[27] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley, 1974.
[28] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces II*, Springer-Verlag, Berlin, 1979.
[29] J. Matoušek, *Geometric Discrepancy*, Springer-Verlag, Heidelberg, New York, 1999.
[30] S.M. Nikol'skii, *Quadrature formulas*, Nauka, Moscow, 1979.
[31] E. Novak, *Deterministic and Stochastic Error Bounds in Numerical Analysis*, Lecture Notes in Math., Vol. 1349, Springer-Verlag, Berlin, 1988.
[32] K.F. Roth, *On irregularities of distribution*, Mathematica **1** (1954), 73–79.
[33] K.F. Roth, *On irregularities of distribution. II*, Communications on Pure and Applied Math. **29** (1976), 749–754.
[34] K.F. Roth, *On irregularities of distribution. III*, Acta Arith. **35** (1979), 373–384.
[35] K.F. Roth, *On irregularities of distribution. IV*, Acta Arith. **37** (1980), 67–75.
[36] W.M. Schmidt, *Irregularities of distribution. VII*, Acta Arith. **21** (1972), 45–50.
[37] W.M. Schmidt, *Irregularities of distribution.X*, in Number Theory and Algebra, Acad.Press, New York, 1977, pp. 311–329.
[38] W.M. Schmidt, *Lectures on irregularities of distribution*, Tata Institute of fundamental research, Bombay, 1977.
[39] M.M. Skriganov, *Constructions of uniform distributions in terms of geometry of numbers*, Algebra and Analysis **6** (1994), 200–230.
[40] S.L. Sobolev, *Introduction to the theory of cubature formulas*, Nauka, Moscow, 1974.
[41] S.A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR **148** (1963), 1042–1045; English transl. in Soviet Math. Dokl. **4** (1963).
[42] V.N. Temlyakov, *Approximate recovery of periodic functions of several variables*, Matem. Sbornik **128(170)** (1985), 256–268; English transl. in Math. USSR Sbornik **56** (1987).
[43] V.N. Temlyakov, *Quadrature formulas and recovery on the values at the knots of number-theoretical nets for classes of functions of small smoothness*, Uspekhi Matem. Nauk **40** (1985), 203–204.
[44] V.N. Temlyakov, *On reconstruction of multivariate periodic functions based on their values at the knots of number-theoretical nets*, Analysis Mathematica **12** (1986), 287-305. (Russian)
[45] V.N. Temlyakov, *On a way of obtaining lower estimates for the errors of quadrature formulas*, Matem. Sbornik **181** (1990), 1403–1413; English transl. in Math. USSR Sbornik **71** (1992)).
[46] V.N. Temlyakov, *Error estimates for Fibonacci quadrature formulas for classes of functions with bounded mixed derivative*, Trudy Matem. Inst. Steklova **200** (1991), 327–335; English transl. in Proc. Steklov Inst. Math. (1993), no. 2.
[47] V.N. Temlyakov, *Approximation of periodic functions*, Nova Science Publishes, Inc., New York, 1993.
[48] V.N. Temlyakov, *On error estimates for cubature formulas*, Trudy Matem. Inst. Steklova **207** (1994), 326–338; English transl. in Proc. Steklov Inst. Math. (1995), no. 6.
[49] V.N. Temlyakov, *Weak Greedy Algorithms*, Advances in Computational Mathematics **12** (2000), 213–227.
[50] V.N. Temlyakov, *Greedy algorithms in Banach spaces*, Advances in Comp. Math. **14** (2001), 277–292.
[51] V.N. Temlyakov, *Nonlinear Methods of Approximation*, IMI-Preprints series **9** (2001), 1–57.
[52] I.V. Vilenkin, *Plane nets of integration*, Zhur. Vychisl. Mat. i Mat. Fis. **7** (1967), 189–196; English transl. in USSR Comp. Math. and Math. Phys. **7** (1967), 258–267.