SMAI-JCM
SMAI Journal of Computational Mathematics

Connections between numerical integration, discrepancy, dispersion, and universal discretization

Vladimir Temlyakov

Volume S5 (2019), p. 185-209.
<http://smai-jcm.centre-mersenne.org/item?id=SMAI-JCM_2019__S5__185_0>

SMAI Journal of Computational Mathematics
Vol. S5, 185-209 (2019)

Connections between numerical integration, discrepancy, dispersion, and universal discretization

Vladimir Temlyakov [1]

[1] University of South Carolina, USA; Steklov Institute of Mathematics and Lomonosov Moscow State University, Russia.
E-mail address: temlyakovv@gmail.com.

Abstract. The main goal of this paper is to provide a brief survey of recent results, which connect together results from different areas of research. It is well known that numerical integration of functions with mixed smoothness is closely related to the discrepancy theory. We discuss this connection in detail and provide a general view of this connection. It was established recently that the new concept of *fixed volume discrepancy* is very useful in proving the upper bounds for the dispersion. Also, it was understood recently that point sets with small dispersion are very good for the universal discretization of the uniform norm of trigonometric polynomials.

2010 Mathematics Subject Classification. 41A65, 42A10, 46B20.

Keywords. numerical integration, discrepancy, dispersion, discretization, greedy algorithm.

1. Introduction

The problem of discretizing the d-dimensional unit cube $[0,1]^d$ is a fundamental problem of mathematics. Certainly, we should clarify what do we mean by *discretization*. There are different ways of doing that. We can interpret $[0,1]^d$ as a compact set of $\mathbb{R}^d$ and use the idea of *covering numbers* (metric entropy). With such approach, for instance in the case of ℓ_∞ norm, we can find optimal coverings. For a given $n \in \mathbb{N}$ the regular grid with coordinates at the centers of intervals $[(k-1)/n, k/n]$, $k = 1, \dots, n$, provides an optimal ℓ_∞ covering with the number of points $N = n^d$. Very often the unit cube $[0,1]^d$ plays the role of a domain, where smooth functions of d variables are defined and we are interested in discretizing some continuous operations with these functions. A classical example of such a problem is the problem of numerical integration of functions. It turns out that the mentioned above regular grids are very far from being good economical discretizations of $[0,1]^d$ for numerical integration purposes. It is a fundamental problem of computational mathematics. Several areas of mathematical research are devoted to this problem: numerical integration, discrepancy, dispersion, sampling. Many nontrivial examples of good (in different sense) point sets are known (see, for instance, [2], [12], [18], [19], [20], [29], [31], [40]). The main goal of this paper is to provide a brief survey of recent results, which connect together results from different areas of research. It is well known that numerical integration of functions with mixed smoothness is closely related to the discrepancy theory. We discuss this connection in detail and provide a general view of this connection. It was established recently (see [42]) that the new concept of *fixed volume discrepancy* is very useful in proving the upper bounds for the dispersion. Also, it was understood recently that point sets with small dispersion are very good for the universal discretization of the uniform norm of trigonometric polynomials (see [41]).

For the reader's orientation we now present a very brief informal description of the content of the paper. The main topic of the paper is numerical integration, which is discussed in all sections except Sections 7 and 8. Numerical integration takes different forms and in some cases it is studied under other name – discrepancy. One of the goals of this paper is to provide a unified way of studying numerical integration. In Section 2 we show that the classical discrepancy problem can be seen as a

The work was supported by the Russian Federation Government Grant № 14.W03.31.0031.

numerical integration problem for a special class of functions. In Section 9 we further develop this idea and show that even in a very general setting numerical integration of a function class $\mathbf{W}_p^K$ and the (K,q)-discrepancy are tied by the duality principle (see Section 9 for details).

A very important issue in numerical integration is evaluation of quality of a given cubature formula. A standard criterion is to compare performance of a cubature formula $\Lambda_m(\,\cdot\,,\xi)$ on a given function class $\mathbf{W}$ (the quantity $\Lambda_m(\mathbf{W},\xi)$) with the optimal performance on $\mathbf{W}$ of cubature formulas with m knots (the quantity $\kappa_m(\mathbf{W})$). This criterion is applied in both numerical integration and discrepancy theory. However, the testing function classes, used in application of this criterion are somewhat different in numerical integration and discrepancy theory. Typically, in numerical integration smoothness classes are used and in discrepancy theory the class of characteristic functions of boxes (parallelepipeds with edges parallel to the coordinate axes) are used. Further, in addition to the above criterion, researchers in numerical integration would like to have a cubature formula, which is good (near optimal) not only for one given smoothness class but for a collection of similar smoothness classes with different smoothness parameter (unsaturated cubature formulas). Moreover, they would like to have a cubature formula, which is good (near optimal) for a collection of different type of smoothness classes (universal cubature formulas). For more detailed discussion of unsaturated and universal cubature formulas see [40, Ch. 6]. Implementation of a unified way of studying numerical integration in different settings immediately brings to our attention an extension of the discrepancy theory setting from characteristic functions of boxes to smooth hat functions supported on boxes. We discuss this setting – the smooth discrepancy setting – in Sections 4 and 5. We present there two different versions of smooth discrepancy. We note that there are outstanding open problems in this area (see Conjectures 5.2, 5.4, 9.5, and 9.6).

Numerical integration is a classical discretization problem. In addition to this problem we discuss two more discretization problems – dispersion and the Marcinkiewicz-type discretization. Formally, numerical integration and dispersion are not closely connected. In both problems we evaluate quality of a given finite set of points in the unit cube, but the criteria for that evaluation are different – the error of numerical integration of functions from a given function class in numerical integration and the volume of the largest box in $[0,1]^d$, which does not contain points of the set in dispersion. However, it turns out that the point sets, which are good for numerical integration, are also good for dispersion. The optimal dispersion of N points is of order $1/N$ (optimal in N, we do not discuss dependence on d). It is easy to understand that a direct use of the upper bounds on discrepancy (r-smooth discrepancy), which have an extra $(\log N)^c$, $c>0$, factor, will not give an optimal upper bound for dispersion. An important new observation in discrepancy theory, which is discussed in Section 6, claims that a new version of discrepancy – the r-smooth fixed volume discrepancy – allows us to obtain optimal rate of dispersion from numerical integration results. The r-smooth fixed volume discrepancy takes into account two characteristics of a smooth hat function h_B^r – its smoothness r and the volume of its support $\mathrm{vol}(B)$.

As we already pointed out above numerical integration provides results on decay of the error under an assumption that a function belongs to a smoothness class. For instance, classical results on discrepancy guarantee the following rate of decay $D(m,d)_2 \asymp m^{-1}(\log m)^{(d-1)/2}$ for numerical integration of d-variate characteristic functions of boxes. These functions have smoothness (see Section 4 below) and special geometric structure. In Section 10 we report a recent surprising observation that we can obtain the rate of decay of order $m^{-1/2}$ for characteristic functions of subsets of a unit cube without any assumption on smoothness or geometric structure.

We discuss several areas of research – numerical integration, discrepancy, dispersion – which are (to a certain extent) independent areas. Therefore, notations used in these areas are not always compatible. On the one hand, we try to use standard notations for each area and, on the other hand, we try to unify those notations. Sometimes, it results in the use of the same symbol in different meaning, for instance, the letter r is mostly used for smoothness, but in Sections 7 and 8, where we do not discuss smoothness at all, it is used in the notation for the number of points – 2^r. We hope that it will not cause any confusion. We use a number of notations for different concepts of discrepancy. We now make a remark on these notations, which, hopefully, will help to digest these notations. A classical L_q discrepancy of a point set ξ is denoted by $D_q(\xi)$, where the subscript q refers to the L_q norm used in the definition. Sometimes, in the case $q = \infty$ we drop it from the notation. In this definition every point of the set ξ of cardinality m is counted with the same weight $1/m$. The optimal discrepancy for sets of cardinality m is denoted by $D(m,d)_q$. Here, d refers to the dimension of the unit cube $[0,1]^d$. These notations are standard in discrepancy theory. As it is explained above, we develop this theory in different directions: consider general weights $\Lambda = \{\lambda_j\}_{j=1}^m$, consider two versions of smooth discrepancy, consider fixed volume discrepancy, and, finally, consider the periodic analogs of some of these concepts. Naturally, we try to reflect these new ingredients in the notations. In order to distinguish the periodic case from the non-periodic case we put tilde over D to get $\widetilde{D}$. Smoothness is reflected by parameter $r \in \mathbb{N}$ in the superscript. For instance, r-discrepancy in the L_q norm of the pair (ξ, Λ) is denoted by $D_q^r(\xi, \Lambda)$. We add a symbol "o" to the superscript to point out that we optimize over point sets of cardinality m and over weights Λ. For instance, the optimal r-discrepancy in the L_q norm is denoted by $D_q^{r,o}(m,d)$. In the same way, symbol "V" in the notation $D^r(\xi, V)$ refers to the fixed volume setting (with volume V) of the r-smooth discrepancy with equal weights for a set ξ. The optimized over weights version is denoted by $D^{r,o}(\xi, V)$.

2. Discrepancy as a special case of numerical integration

We formulate the numerical integration problem in a general setting. Numerical integration seeks good ways of approximating an integral

$$\int_\Omega f(\mathbf{x}) \mathrm{d}\mu$$

by an expression of the form

$$\Lambda_m(f,\xi) := \sum_{j=1}^m \lambda_j f(\xi^j), \quad \xi = (\xi^1,\dots,\xi^m), \quad \xi^j \in \Omega, \quad j = 1,\dots,m. \tag{2.1}$$

It is clear that we must assume that f is integrable and defined at the points $\xi^1,\dots,\xi^m$. Expression (2.1) is called a *cubature formula* (ξ,Λ) (if $\Omega \subset \mathbb{R}^d$, $d \geq 2$) or a *quadrature formula* (ξ,Λ) (if $\Omega \subset \mathbb{R}$) with knots $\xi = (\xi^1,\dots,\xi^m)$ and weights $\Lambda := (\lambda_1,\dots,\lambda_m)$. Usually, in this paper $\Omega = \Omega_d := [0,1)^d$ is the unit cube and μ is the Lebesgue measure on it ($\mathrm{d}\mu = \mathrm{d}\mathbf{x}$). In some cases $\Omega = \mathbb{T}^d := [0,2\pi)^d$ and μ is a normalized Lebesgue measure on $[0,2\pi)^d$ ($\mathrm{d}\mu = (2\pi)^{-d}\mathrm{d}\mathbf{x}$)).

Some classes of cubature formulas are of special interest. For instance, the Quasi-Monte Carlo cubature formulas, which have equal weights $1/m$, are important in applications. We use a special notation for these cubature formulas

$$Q_m(f,\xi) := \frac{1}{m}\sum_{j=1}^m f(\xi^j).$$

The following class is a natural subclass of all cubature formulas. Let B be a positive number and $Q(B,m)$ be the set of cubature formulas $\Lambda_m(\,\cdot\,,\xi)$ satisfying the additional condition

$$\sum_{\mu=1}^m |\lambda_\mu| \leq B. \tag{2.2}$$

Clearly, the Quasi-Monte Carlo cubature formulas $Q_m(\,\cdot\,,\xi)$ belong to the class $Q(1,m)$. Condition (2.2) can be seen as a "stability" condition imposed on a cubature formula $\Lambda_m(\,\cdot\,,\xi) \in Q(B,m)$.

For a function class $\mathbf{W}$ we introduce a concept of error of the cubature formula $\Lambda_m(\,\cdot\,,\xi)$ by

$$\Lambda_m(\mathbf{W},\xi) := \sup_{f\in\mathbf{W}} \left| \int_\Omega f \mathrm{d}\mu - \Lambda_m(f,\xi) \right|. \tag{2.3}$$

The quantity $\Lambda_m(\mathbf{W},\xi)$ is a classical characteristic of the quality of a given cubature formula $\Lambda_m(\,\cdot\,,\xi)$. This setting is called *the worst case setting* in the Information Based Complexity. Typically, in approximation theory we study the behavior of the quantity $\Lambda_m(\mathbf{W},\xi)$ for classes $\mathbf{W}$ of smooth functions, in particular, for the unit balls of different spaces of smooth functions – Sobolev, Nikol'skii, Besov spaces and spaces with mixed smoothness (see [40] and [12]). The problem of finding optimal in the sense of order cubature formulas for a given class is of special importance. This means that we are looking for a cubature formula $\Lambda_m^{opt}(\mathbf{W},\xi)$ such that

$$\Lambda_m^{opt}(\mathbf{W},\xi) \asymp \inf_{\xi,\Lambda} \Lambda_m(\mathbf{W},\xi) =: \kappa_m(\mathbf{W}). \tag{2.4}$$

We now describe a typical class χ^d, which is of interest in numerical integration and in discrepancy theory. The following way is used in discrepancy theory for evaluation of a quality of replacing the Lebesgue measure μ on $[0,1]^d$ by a discrete measure μ_m such that $\mu_m(\xi^j) = 1/m$, $j = 1,\dots,m$. We begin with a classical definition of discrepancy ("star discrepancy", L_∞-discrepancy) of a point set $\xi := \{\xi^\mu\}_{\mu=1}^m \subset [0,1)^d$. Let $d \geq 2$ and $[0,1)^d$ be the d-dimensional unit cube. For convenience we sometimes use the notation $\Omega_d := [0,1)^d$. For $\mathbf{x},\mathbf{y} \in [0,1)^d$ with $\mathbf{x} = (x_1,\dots,x_d)$ and $\mathbf{y} = (y_1,\dots,y_d)$ we write $\mathbf{x} < \mathbf{y}$ if this inequality holds coordinate-wise. For $\mathbf{x} < \mathbf{y}$ we write $[\mathbf{x},\mathbf{y})$ for the axes-parallel box $[x_1,y_1) \times \cdots \times [x_d,y_d)$ and define

$$\mathcal{B} := \{[\mathbf{x},\mathbf{y}) : \mathbf{x},\mathbf{y} \in [0,1)^d, \mathbf{x} < \mathbf{y}\}.$$

Introduce a class of special d-variate characteristic functions

$$\chi^d := \left\{ \chi_{[\mathbf{0},\mathbf{b})}(\mathbf{x}) := \prod_{j=1}^d \chi_{[0,b_j)}(x_j), \quad b_j \in [0,1), \quad j = 1,\dots,d \right\}$$

where $\chi_{[a,b)}(x)$ is a univariate characteristic function of the interval $[a,b)$. Then for $\mathbf{b} \in \Omega_d$ the quantity

$$\int_{\Omega_d} \chi_{[\mathbf{0},\mathbf{b})} \mathrm{d}\mathbf{x} - \int_{\Omega_d} \chi_{[\mathbf{0},\mathbf{b})} \mathrm{d}\mu_m = \prod_{j=1}^d b_j - \frac{1}{m} \sum_{\mu=1}^m \chi_{[\mathbf{0},\mathbf{b})}(\xi^\mu)$$

gives a discrepancy between the Lebesgue measure of the box $[\mathbf{0},\mathbf{b})$ and its discrete measure $\mu_m([\mathbf{0},\mathbf{b}))$. The classical definition of discrepancy of a set ξ of points $\{\xi^1,\dots,\xi^m\} \subset [0,1)^d$ is as follows

$$D_\infty(\xi) := \sup_{\mathbf{b}\in[0,1)^d} \left| \prod_{j=1}^d b_j - \frac{1}{m} \sum_{\mu=1}^m \chi_{[\mathbf{0},\mathbf{b})}(\xi^\mu) \right|. \tag{2.5}$$

It is clear from the definition of $\Lambda_m(\mathbf{W},\xi)$ in (2.3) that

$$D_\infty(\xi) = Q_m(\chi^d,\xi)$$

with $Q_m(\,\cdot\,,\xi)$ being the Quasi-Monte Carlo cubature formula.

The class χ^d is parametrized by the parameter $\mathbf{b} \in [0,1)^d$. Therefore, we can define the L_q-discrepancy, $1 \leq q \leq \infty$, of ξ as follows

$$D_q(\xi) := \left\| \prod_{j=1}^d b_j - \frac{1}{m} \sum_{\mu=1}^m \chi_{[\mathbf{0},\mathbf{b})}(\xi^\mu) \right\|_q, \tag{2.6}$$

where the L_q norm is taken with respect to $\mathbf{b}$ over the domain Ω_d. In order to emphasize the connection between discrepancy and numerical integration we use the notation

$$Q_m(\chi^d, \xi, q) := D_q(\xi), \quad 1 \le q \le \infty.$$

We note that the fact that the class χ^d is parametrized by $\mathbf{b} \in \Omega_d$ is a special important feature, which allows us to consider along with the worst case setting (2.5) the setting (2.6). We call the (2.6) setting in case $q < \infty$ – *the average case setting.*

3. A brief history of results on classical discrepancy

The first result on the lower bound for discrepancy was the following conjecture of van der Corput [8] and [9] for $d = 1$ formulated in 1935. Let $\xi^j \in [0, 1]$, $j = 1, 2, \dots$, then we have

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

This conjecture was proved by van Aardenne–Ehrenfest [47] in 1945 (see also [48]):

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

We now list some classical lower estimates of discrepancy of point sets in the unit cube Ω_d. Let us denote

$$D(m, d)_q := \inf_{\xi} D_q(\xi), \quad \xi = \{\xi_j\}_{j=1}^m, \quad 1 \le q \le \infty,$$

where $D_q(\xi)$ is defined in (2.6). In 1954 K. Roth [21] proved that

$$D(m, d)_2 \ge C(d)m^{-1}(\log m)^{(d-1)/2}. \tag{3.1}$$

In 1972 W. Schmidt [23] proved

$$D(m, 2)_\infty \ge Cm^{-1} \log m. \tag{3.2}$$

In 1977 W. Schmidt [24] proved

$$D(m, d)_q \ge C(d, q)m^{-1}(\log m)^{(d-1)/2}, \qquad 1 < q \le \infty. \tag{3.3}$$

In 1981 G. Halász [15] proved

$$D(m, d)_1 \ge C(d)m^{-1}(\log m)^{1/2}. \tag{3.4}$$

The following conjecture has been formulated in [2] as an excruciatingly difficult great open problem.

Conjecture 3.1. *We have for $d \ge 3$*

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

This problem is still open. Recently, D. Bilyk and M. Lacey [4] and D. Bilyk, M. Lacey, and A. Vagharshakyan [5] proved

$$D(m, d)_\infty \ge C(d)m^{-1}(\log m)^{(d-1)/2+\delta(d)}$$

with some positive $\delta(d)$.

For further historical discussion we refer the reader to surveys [3], [12], [31], and books [2], [18], [40].

4. Smooth discrepancy and numerical integration

We say that a univariate function f has smoothness 1 in L_1 if $\|\Delta_t f\|_1 \le C|t|$, where $\Delta_t f(x) := f(x) - f(x+t)$ is the first difference. In case $\|\Delta_t^r f\|_1 \le C|t|^r$, where $\Delta_t^r := (\Delta_t)^r$ is the rth difference operator, $r \in \mathbb{N}$, we say that f has smoothness r in L_1. In the above definitions the function class χ^d with $d = 1$ consists of characteristic functions, which have smoothness 1 in the L_1 norm. In numerical integration it is natural to study function classes with arbitrary smoothness r. There are different generalizations of the above concept of discrepancy to the case of *smooth discrepancy*. We discuss two of them here. In the definition of the first version of the r-discrepancy, $r \in \mathbb{N}$, (see [29], [40]) instead of the characteristic function (this corresponds to 1-discrepancy) we use the following function

$$B_r(\mathbf{x},\mathbf{y}) := \prod_{j=1}^{d} \left((r-1)!\right)^{-1}(y_j - x_j)_+^{r-1}, \quad \mathbf{x},\mathbf{y} \in \Omega_d, \qquad (a)_+ := \max(a,0).$$

In case $d = 1$ function $B_r(x,y)$ has smoothness r (on $\mathbb{R}_+$) in the L_1 norm.

Denote

$$\mathbf{B}^{r,d} := \{B_r(\mathbf{x},\mathbf{y}) : \mathbf{y} \in \Omega_d\}.$$

Definition 4.1. For a point set $\xi := \{\xi^\mu\}_{\mu=1}^m$ of cardinality m and weights $\Lambda := \{\lambda_\mu\}_{\mu=1}^m$ we define the r-discrepancy of the pair (ξ,Λ) by the formula

$$D_q^r(\xi,\Lambda) := \left\| \sum_{\mu=1}^{m} \lambda_\mu B_r(\xi^\mu,\mathbf{y}) - \prod_{j=1}^{d}(y_j^r/r!) \right\|_q, \quad 1 \le q \le \infty. \tag{4.1}$$

It is clear from the definition of $\Lambda_m(\mathbf{W},\xi)$ in (2.3) that

$$D_\infty^r(\xi,\Lambda) = \Lambda_m(\mathbf{B}^{r,d},\xi).$$

Consider the class $\dot{\mathbf{W}}_p^r$ consisting of the functions $f(\mathbf{x})$, which have an integral representation with the kernel $B_r(\mathbf{x},\mathbf{y})$,

$$\dot{\mathbf{W}}_p^r := \left\{ f \,:\, f(\mathbf{x}) = \int_{\Omega_d} B_r(\mathbf{x},\mathbf{y})\varphi(\mathbf{y})\mathrm{d}\mathbf{y}, \quad \|\varphi\|_p \le 1 \right\}.$$

In connection with the definition of the class $\dot{\mathbf{W}}_p^r$ we remark here that for the error of the cubature formula (ξ,Λ) with weights $\Lambda = (\lambda_1,\dots,\lambda_m)$ and knots $\xi = (\xi^1,\dots,\xi^m)$ the following relation holds with $p' := p/(p-1)$

$$\Lambda_m(\dot{\mathbf{W}}_p^r,\xi) = \left\| \sum_{\mu=1}^{m} \lambda_\mu B_r(\xi^\mu,\mathbf{y}) - \prod_{j=1}^{d}(y_j^r/r!) \right\|_{p'} = D_{p'}^r(\xi,\Lambda). \tag{4.2}$$

Thus, errors of numerical integration of classes $\dot{\mathbf{W}}_p^r$ are dual to the average errors of numerical integration of functions in classes $\mathbf{B}^{r,d}$.

We now consider classes of periodic functions (with period 2π in each variable) on $\mathbb{T}^d$, which usually are referred to as classes of functions with bounded mixed derivative. For $\mathbf{x} = (x_1,\dots,x_d)$ denote

$$F_r(\mathbf{x}) := \prod_{j=1}^{d} F_r(x_j), \quad F_r(x_j) := 1 + 2\sum_{k=1}^{\infty} k^{-r}\cos(kx_j - r\pi/2),$$

and

$$\mathbf{W}_p^r := \{f : f = F_r * \varphi := (2\pi)^{-d}\int_{\mathbb{T}^d} F_r(\mathbf{x}-\mathbf{y})\varphi(\mathbf{y})\mathrm{d}\mathbf{y}, \quad \|\varphi\|_p \le 1\}.$$

We note that in the univariate case for a 2π-periodic function f with $\int_0^{2\pi} f(x)dx = 0$ we have for $r \in \mathbb{N}$

$$f(x) = (2\pi)^{-1} \int_0^{2\pi} F_r(x-y) f^{(r)}(y)\mathrm{d}y.$$

For $f \in \mathbf{W}^r_p$ we denote $f^{(r)} := \varphi$ where φ is such that $f = F_r * \varphi$. In the case of $r \in \mathbb{N}$ the class $\mathbf{W}^r_p$ 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. We refer the reader to [40, Ch. 4], and [12, Ch. 3], for detailed discussion of periodic classes of functions with bounded mixed derivative. It turns out that errors of numerical integration of classes χ^d, defined on Ω_d, and classes $\mathbf{W}^1_1$, defined on $\mathbb{T}^d$, are very close. In order to formulate this, we introduce a notation. For a point set $\xi = \{\xi^j\}_{j=1}^m \subset \Omega_d$ denote $2\pi\xi := \{2\pi\xi^j\}_{j=1}^m \subset \mathbb{T}^d$.

Proposition 4.2. *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*

$$C_1(d)\Lambda_m(\chi^d,\xi) \le \Lambda_m(\mathbf{W}^1_1, 2\pi\xi) \le C_2(d)\Lambda_m(\chi^d,\xi). \tag{4.3}$$

The reader can find the proof of Proposition 4.2 in [31]. The following theorem, which extends (in a ceratin sense) Proposition 4.2 to the case of arbitrary smoothness $r \in \mathbb{N}$, is from [31] (see also [40, p. 250]). The definition of $\kappa_m(\mathbf{W})$, which we use in Theorem 4.3, is given in (2.4).

Theorem 4.3. *Let $1 \le p \le \infty$. Then for $r \in \mathbb{N}$*

$$\kappa_m(\dot{\mathbf{W}}^r_p(\Omega_d)) \asymp \kappa_m(\mathbf{W}^r_p). \tag{4.4}$$

We now proceed to the second version of smooth discrepancy – the r-smooth discrepancy. The classical definition of discrepancy $D_\infty(\xi)$ (see (2.5)) of a set ξ of points $\{\xi^1,\dots,\xi^m\} \subset [0,1)^d$, which is sometimes called *anchored discrepancy*, is equivalent within multiplicative constants that may only depend on d to the following definition

$$D^1(\xi) := \sup_{B\in\mathcal{B}} \left| \mathrm{vol}(B) - \frac{1}{m}\sum_{\mu=1}^m \chi_B(\xi^\mu) \right|, \tag{4.5}$$

where for $B = [\mathbf{a},\mathbf{b}) \in \mathcal{B}$ we denote $\chi_B(\mathbf{x}) := \prod_{j=1}^d \chi_{[a_j,b_j)}(x_j)$. Moreover, we consider the following optimized version of $D^1(\xi)$

$$D^{1,o}(\xi) := \inf_{\lambda_1,\dots,\lambda_m} \sup_{B\in\mathcal{B}} \left| \mathrm{vol}(B) - \sum_{\mu=1}^m \lambda_\mu \chi_B(\xi^\mu) \right|. \tag{4.6}$$

In the definition of $D^1(\xi)$ and $D^{1,o}(\xi)$ – the 1-smooth discrepancy – we use as a building block the univariate characteristic function. In numerical integration L_1-smoothness of a function plays an important role. A characteristic function of an interval has smoothness 1 in the L_1 norm. This is why we call the corresponding discrepancy characteristics the 1-smooth discrepancy. In the definition of $D^2(\xi)$, $D^{2,o}(\xi)$, $D^2(\xi,V)$, and $D^{2,o}(\xi,V)$ (see below and [42]) we use the hat function $h_{[-u,u)}(x) = u - |x|$ for $|x| \le u$ and $h_{[-u,u)}(x) = 0$ for $|x| \ge u$ instead of the characteristic function $\chi_{[-u/2,u/2)}(x)$. Function $h_{[-u,u)}(x)$ has smoothness 2 in L_1. This fact gives the corresponding name. Note that

$$h_{[-u,u)}(x) = \chi_{[-u/2,u/2)}(x) * \chi_{[-u/2,u/2)}(x),$$

where

$$f(x) * g(x) := \int_{\mathbb{R}} f(x-y)g(y)\mathrm{d}y.$$

Now, for $r = 1, 2, 3, \dots$ we inductively define

$$h^1(x,u) := \chi_{[-u/2,u/2)}(x), \qquad h^2(x,u) := h_{[-u,u)}(x),$$
$$h^r(x,u) := h^{r-1}(x,u) * h^1(x,u), \qquad r = 3, 4, \dots.$$

Then $h^r(x,u)$ has smoothness r in L_1 and has support $(-ru/2, ru/2)$. Represent a box $B \in \mathcal{B}$ in the form

$$B = \prod_{j=1}^{d} [x_j^0 - ru_j/2, x_j^0 + ru_j/2)$$

and define

$$h_B^r(\mathbf{x}) := h^r(\mathbf{x}, \mathbf{x}^0, \mathbf{u}) := \prod_{j=1}^{d} h^r(x_j - x_j^0, u_j).$$

We note that in the above definition of $h^r(\mathbf{x}, \mathbf{x}^0, \mathbf{u})$ parameters $\mathbf{x}^0$ and $\mathbf{u}$ play different roles. Parameter $\mathbf{x}^0$ gives a shift, while parameter $\mathbf{u}$ controls the size and the shape of the box B.

In [42] we modified definitions (4.5) and (4.6), replacing the characteristic function χ_B by a smoother (r-smooth) hat function h_B^r.

Definition 4.4. The r-smooth discrepancy, $r \in \mathbb{N}$, is defined as

$$D^r(\xi) := \sup_{B \in \mathcal{B}} \left| \int h_B^r(\mathbf{x}) \mathrm{d}\mathbf{x} - \frac{1}{m} \sum_{\mu=1}^{m} h_B^r(\xi^\mu) \right| \tag{4.7}$$

and its optimized version is defined as

$$D^{r,o}(\xi) := \inf_{\lambda_1,\dots,\lambda_m} \sup_{B \in \mathcal{B}} \left| \int h_B^r(\mathbf{x}) \mathrm{d}\mathbf{x} - \sum_{\mu=1}^{m} \lambda_\mu h_B^r(\xi^\mu) \right|. \tag{4.8}$$

Note that the known concept of r-discrepancy (see, for instance, [29], [31], and above in this section) is close to the concept of r-smooth discrepancy.

It is more convenient for us to consider the average setting in the periodic case. For a function $f \in L_1(\mathbb{R}^d)$ with a compact support we define its periodization $\tilde{f}$ as follows

$$\tilde{f}(\mathbf{x}) := \sum_{\mathbf{m} \in \mathbb{Z}^d} f(\mathbf{m} + \mathbf{x}).$$

For each $\mathbf{z} \in [0,1)^d$ and $\mathbf{u} \in (0, \frac{1}{2}]^d$ consider a periodization of the function $h^r(\mathbf{x}, \mathbf{z}, \mathbf{u})$ in $\mathbf{x}$ with period 1 in each variable $\tilde{h}^r(\mathbf{x}, \mathbf{z}, \mathbf{u})$. Consider the class of periodic r-smooth hat functions

$$\mathbf{H}^{r,d} := \{\tilde{h}^r(\mathbf{x}, \mathbf{z}, \mathbf{u}) : \mathbf{z} \in [0,1)^d; \mathbf{u} \in (0, 1/2]^d\}.$$

The following definition is from [36].

Definition 4.5. Define the *periodic r-smooth discrepancy* for a pair (ξ, Λ) as follows

$$\tilde{D}_\infty^r(\xi, \Lambda) := \sup_{\mathbf{z} \in [0,1)^d; \mathbf{u} \in (0,1/2]^d} \left| \int_{[0,1)^d} \tilde{h}^r(\mathbf{x}, \mathbf{z}, \mathbf{u}) \mathrm{d}\mathbf{x} - \sum_{\mu=1}^{m} \lambda_\mu \tilde{h}^r(\xi^\mu, \mathbf{z}, \mathbf{u}) \right|. \tag{4.9}$$

The corresponding optimized version of the periodic r-smooth discrepancy is defined as

$$\tilde{D}_\infty^{r,o}(m,d) := \inf_{\xi,\Lambda} \tilde{D}_\infty^r(\xi, \Lambda). \tag{4.10}$$

It is clear that

$$\tilde{D}_\infty^r(\xi, \Lambda) := \Lambda_m(\mathbf{H}^{r,d}, \xi).$$

Definition 4.6. For $\mathbf{p} = (p_1, p_2)$, $1 \le p_1, p_2 \le \infty$, define the corresponding *periodic r-smooth $L_\mathbf{p}$-discrepancy*, which also can be called *Weyl r-smooth $L_\mathbf{p}$-discrepancy* ([17], [49]), as follows

$$\widetilde{D}^r_{p_1,p_2}(\xi,\Lambda) := \left\| \left\| \int_{[0,1)^d} \widetilde{h}^r(\mathbf{x},\mathbf{z},\mathbf{u})\mathrm{d}\mathbf{x} - \sum_{\mu=1}^m \lambda_\mu \widetilde{h}^r(\xi^\mu,\mathbf{z},\mathbf{u}) \right\|_{p_1} \right\|_{p_2}, \tag{4.11}$$

where the L_{p_1} norm is taken with respect to $\mathbf{z}$ over the unit cube $[0,1)^d$ and the L_{p_2} norm is taken with respect to $\mathbf{u}$ over the cube $(0,1/2]^d$.

In the definition of $\widetilde{D}^r_{p_1,p_2}(\xi,\Lambda)$ parameters $\mathbf{z}$ and $\mathbf{u}$ play different roles. The most important parameter is $\mathbf{u}$ – it controls the size and the shape of supports of the corresponding hat functions. It seems like the most natural value for parameter p_2 is ∞. In this case we obtain bounds uniform with respect to the shape and the size of supports of hat functions.

5. Lower estimates for the smooth discrepancy

We begin with a presentation of results on the lower estimates for the r-discrepancy. As above for a point set $\xi := \{\xi^\mu\}_{\mu=1}^m$ of cardinality m and weights $\Lambda := \{\lambda_\mu\}_{\mu=1}^m$ we define the r-discrepancy $D^r_q(\xi,\Lambda)$ of the pair (ξ,Λ) by the formula (4.1). We denote

$$D^r_q(m,d) := \inf_\xi D^r_q(\xi,(1/m,\dots,1/m)),$$

where $D^r_q(\xi,\Lambda)$ is defined in (4.1) and also denote

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

It is clear that

$$D^{r,o}_q(m,d) \le D^r_q(m,d).$$

The first result on estimating the r-discrepancy was obtained in 1985 by V. A. Bykovskii [7]

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

This result is a generalization of Roth's result (3.1). The generalization of Schmidt's result (3.3) was obtained by the author in 1990 (see [28])

$$D^{r,o}_q(m,d) \ge C(r,d,q)m^{-r}(\log m)^{(d-1)/2}, \qquad 1 < q \le \infty. \tag{5.2}$$

In 1994 (see [30]) the author proved the lower bounds in the case of weights Λ satisfying an extra condition (2.2) and smoothness parameter $r \in \mathbb{N}$ being an even number.

Theorem 5.1. *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* (2.2)

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

we have for even integers r

$$D^r_\infty(\xi,\Lambda) \ge C(d,B,r)m^{-r}(\log m)^{d-1}$$

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

This result encouraged us to formulate the following generalization of Conjecture 3.1 (see [31]).

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

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

We point out that inequality (5.2) gives the same, in the sense of order, lower bound for all $1 < q < \infty$, while Conjecture 5.2 gives a stronger lower bound in case $q = \infty$.

We now proceed to the r-smooth $L_{\mathbf{p}}$-discrepancy (see Definitions 4.5 and 4.6 above). The first lower bound for such discrepancy was obtained in the case $\mathbf{p} = \infty$ under an extra condition (2.2) on the weights (see [36]) and under assumption that r is an even number. Here is the corresponding result from [36].

Theorem 5.3. *For any points $\xi^1, \dots, \xi^m \subset \Omega_d$ and weights $\Lambda = (\lambda_1, \dots, \lambda_m)$ satisfying condition* (2.2) *we have for even integers r*

$$\widetilde{D}^r_\infty(\xi, \Lambda) \ge C(d, B, r) m^{-r} (\log m)^{d-1}$$

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

Denote

$$\widetilde{D}^{r,o}_{\mathbf{p}}(m, d) := \inf_{\xi,\Lambda} \widetilde{D}^r_{\mathbf{p}}(\xi, \Lambda).$$

Theorem 5.3 supports the following conjecture.

Conjecture 5.4. *For all $d, r \in \mathbb{N}$ we have*

$$\widetilde{D}^{r,o}_\infty(m, d) \ge C(r, d) m^{-r} (\log m)^{d-1}.$$

We now proceed to the case $\mathbf{p} \ne \infty$. The following theorem is from [38].

Theorem 5.5. *Let $r \in \mathbb{N}$. Then for any (ξ, Λ) we have*

$$\widetilde{D}^r_{2,2}(\xi, \Lambda) \ge C(r, d) m^{-r} (\log m)^{(d-1)/2}, \qquad C(r, d) > 0.$$

Theorem 5.5 gives the following lower bound for $r \in \mathbb{N}$ and $\mathbf{p} \ge \mathbf{2}$

$$\widetilde{D}^{r,o}_{\mathbf{p}}(m, d) \ge C(r, d) m^{-r} (\log m)^{(d-1)/2}. \tag{5.3}$$

The lower bound (5.3) is different from the lower bound in Theorem 5.3. However, the following Proposition 5.6 (see [38]) shows that this bound is sharp in case $\mathbf{p} = \mathbf{2}$.

Proposition 5.6. *For $r \in \mathbb{N}$ there exists a cubature formula (ξ, Λ) such that*

$$\widetilde{D}^r_{2,\infty}(\xi, \Lambda) \le C(r, d) m^{-r} (\log m)^{(d-1)/2}, \qquad C(r, d) > 0.$$

Under a stronger assumption on r, namely, assuming that r is an even number, we obtained in [38] a stronger than (5.3) lower bound.

Theorem 5.7. *Let $r \in \mathbb{N}$ be an even number. Then for any (ξ, Λ) we have for $1 < p < \infty$*

$$\widetilde{D}^r_{p,1}(\xi, \Lambda) \ge C(r, d, p) m^{-r} (\log m)^{(d-1)/2}, \qquad C(r, d, p) > 0.$$

Theorem 5.7 gives that for even r for any $1 < p < \infty$

$$\widetilde{D}^{r,o}_{p,1}(m, d) \ge C(r, d, p) m^{-r} (\log m)^{(d-1)/2}, \qquad C(r, d, p) > 0.$$

The following result from [38] is an extension of Proposition 5.6.

Proposition 5.8. *For $r \in \mathbb{N}$ and $1 < p < \infty$ there exists a cubature formula (ξ, Λ) such that*

$$\widetilde{D}^r_{p,\infty}(\xi, \Lambda) \le C(r, p, d) m^{-r} (\log m)^{(d-1)/2}.$$

Proposition 5.8 shows that the above lower bound is sharp. Moreover, it shows that for r even we have for all $1 < p_1 < \infty$ and $1 \le p_2 \le \infty$

$$\widetilde{D}^{r,o}_{p_1,p_2}(m, d) \asymp m^{-r} (\log m)^{(d-1)/2}. \tag{5.4}$$

Comment. In some of the lower bounds reported in this section we are able to prove better results under an extra assumption that r is an even number. The reason of a distinction between odd and even r is in a special trick, which is used in the proof (see, for instance, [40, pp. 269–270]). This trick uses the fact that the Fourier coefficients $\hat{F}_r(\mathbf{k})$ of the function $F_r(\mathbf{x})$ are positive in the case of an even number $r \in \mathbb{N}$.

6. Fixed volume discrepancy

Along with $D^r(\xi)$ and $D^{r,o}(\xi)$ (see Definition 4.4 above) we consider a more refined quantity – the *r-smooth fixed volume discrepancy* – defined as follows.

Definition 6.1. Let $V \in (0,1]$. We define the r-smooth fixed volume discrepancy with equal weights as

$$D^r(\xi, V) := \sup_{B\in\mathcal{B}:\mathrm{vol}(B)=V} \left| \int h^r_B(\mathbf{x})\mathrm{d}\mathbf{x} - \frac{1}{m}\sum_{\mu=1}^{m} h^r_B(\xi^\mu) \right|. \tag{6.1}$$

The optimized version of the r-smooth fixed volume discrepancy is defined as follows

$$D^{r,o}(\xi, V) := \inf_{\lambda_1,\dots,\lambda_m} \sup_{B\in\mathcal{B}:\mathrm{vol}(B)=V} \left| \int h^r_B(\mathbf{x})\mathrm{d}\mathbf{x} - \sum_{\mu=1}^{m} \lambda_\mu h^r_B(\xi^\mu) \right|. \tag{6.2}$$

Clearly,

$$D^r(\xi) = \sup_{V\in(0,1]} D^r(\xi, V).$$

We begin with the case $d = 2$. It is well known that the Fibonacci cubature formulas are optimal in the sense of order for numerical integration of different kind of smoothness classes of functions of two variables (see [12], [29], [40]). We present a result from [42], which shows that the Fibonacci point set has good fixed volume discrepancy.

Let $\{b_n\}_{n=0}^{\infty}$, $b_0 = b_1 = 1$, $b_n = b_{n-1} + b_{n-2}$, $n \geq 2$, – be the Fibonacci numbers. Denote the nth *Fibonacci point set* by

$$\mathcal{F}_n := \{(\mu/b_n, \{\mu b_{n-1}/b_n\}),\ \mu = 1,\dots,b_n\}.$$

In this definition $\{a\}$ is the fractional part of the number a. The cardinality of the set $\mathcal{F}_n$ is equal to b_n. In [42] we proved the following upper bound.

Theorem 6.2. *Let $d = 2$, $r \geq 2$. There exists a constant $c(r) > 0$ such that for any $V \geq V_0 := c(r)/b_n$ we have for all $B \in \mathcal{B}$, $\mathrm{vol}(B) = V$*

$$\left| b_n^{-1} \sum_{\mu=1}^{b_n} h^r_B(\mu/b_n, \{\mu b_{n-1}/b_n\}) - \hat{h}^r_B(\mathbf{0}) \right| \leq C(r)\log(2V/V_0)/b_n^r. \tag{6.3}$$

Theorem 6.2 provides the following inequalities for the Fibonacci point sets $\mathcal{F}_n$ in case $r \geq 2$

$$D^{r,o}(\mathcal{F}_n, V) \leq D^r(\mathcal{F}_n, V) \leq C(r)(\log(2V/V_0))/b_n^r, \qquad V \geq V_0.$$

We now proceed to the case $d \geq 3$. It is well known that the *Frolov point sets* are very good for numerical integration of smoothness classes of functions of several variables (see [12], [13], [29], [31], [40], [43]). Theorem 6.4 below, which was proved in [42], shows that the Frolov point sets have good fixed volume discrepancy. Construction of the Frolov point sets is more involved than the construction of the Fibonacci point sets. We begin with a description of the Frolov point sets. The following lemma plays a fundamental role in the construction of such point sets (see [29, Ch. 4, §4] or [40, Ch. 6, §6.7] for its proof).

Lemma 6.3. *There exists a matrix A such that the lattice $L(\mathbf{m}) = A\mathbf{m}$*

$$L(\mathbf{m}) = \begin{bmatrix} L_1(\mathbf{m}) \\ \vdots \\ L_d(\mathbf{m}) \end{bmatrix},$$

where $\mathbf{m}$ is a (column) vector with integer coordinates, has the following properties

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

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 6.3. 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 with a compact support.

We call the *Frolov point set* the following set associated with the matrix A and parameter a

$$\mathcal{F}(a,A) := \left\{\left(\frac{(A^{-1})^T\mathbf{m}}{a}\right)\right\}_{\mathbf{m}\in\mathbb{Z}^d} \cap [0,1]^d =: \{z^\mu\}_{\mu=1}^N.$$

Clearly, the number $N = |\mathcal{F}(a,A)|$ of points of this set does not exceed $C(A)a^d$. The following Theorem 6.4 and its Corollary 6.5 are from [42].

Theorem 6.4. *Let $r \ge 2$. There exists a constant $c(d,A,r) > 0$ such that for any $V \ge V_0 := c(d,A,r)a^{-d}$ we have for all $B \in \mathcal{B}$, $\operatorname{vol}(B) = V$,*

$$|\Phi(a,A)(h^r_B) - \hat{h}^r_B(\mathbf{0})| \le C(d,A,r)a^{-rd}(\log(2V/V_0))^{d-1}. \tag{6.4}$$

Corollary 6.5. *For $r \ge 2$ there exists a constant $c(d,A,r) > 0$ such that for any $V \ge V_0 := c(d,A,r)a^{-d}$ we have*

$$D^{r,o}(\mathcal{F}(a,A),V) \le C(d,A,r)a^{-rd}(\log(2V/V_0))^{d-1}. \tag{6.5}$$

The following technical Lemma 6.6 played the main role in the proofs of Theorems 6.2 and 6.4. Lemma 6.6 might be of interest by itself. Consider

$$\sigma^r(v,\mathbf{u}) := \sum_{\|\mathbf{s}\|_1=v} \prod_{j=1}^d \min\left((2^{s_j}u_j)^{r/2}, \frac{1}{(2^{s_j}u_j)^{r/2}}\right), \quad v \in \mathbb{N}_0.$$

Denote

$$pr(\mathbf{u},d) := \prod_{j=1}^d u_j.$$

Lemma 6.6. *Let $v \in \mathbb{N}_0$ and $\mathbf{u} \in \mathbb{R}^d_+$. Then we have the following inequalities.*

(I) *Under condition $2^v pr(\mathbf{u},d) \ge 1$ we have*

$$\sigma^r(v,\mathbf{u}) \le C(d)\frac{\left(\log(2^{v+1}pr(\mathbf{u},d))\right)^{d-1}}{(2^v pr(\mathbf{u},d))^{r/2}}. \tag{6.6}$$

(II) *Under condition* $2^v pr(\mathbf{u},d)\le 1$ *we have*

$$\sigma^r(v,\mathbf{u})\le C(d)(2^v pr(\mathbf{u},d))^{r/2}\left(\log\frac{2}{2^v pr(\mathbf{u},d)}\right)^{d-1}. \tag{6.7}$$

In [36] we extended Theorem 6.4 and Corollary 6.5 to the periodic case (with period 1 in each variable). For that we need to modify the set $\mathcal{F}(a,A)$ and the cubature formula $\Phi(a,A)$. For $\mathbf{y}\in\mathbb{R}^d$ denote $\{\mathbf{y}\}:=(\{y_1\},\dots,\{y_d\})$, where for $y\in\mathbb{R}$ notation $\{y\}$ means the fractional part of y. For given a and A denote

$$\eta:=\{\eta^\mu\}_{\mu=1}^m:=\left\{\left(\frac{(A^{-1})^T\mathbf{m}}{a}\right)\right\}_{\mathbf{m}\in\mathbb{Z}^d}\cap[-1/2,3/2)^d$$

and

$$\xi:=\{\xi^\mu\}_{\mu=1}^m:=\{\{\eta^\mu\}\}_{\mu=1}^m. \tag{6.8}$$

Clearly, $m\le C(A)a^d$. Next, let $w(t)$ be infinitely differentiable on $\mathbb{R}$ function with the following properties

$$\operatorname{supp}(w)\subset(-1/2,3/2)\quad\text{and}\quad\sum_{k\in\mathbb{Z}}w(t+k)=1. \tag{6.9}$$

Denote $w(\mathbf{x}):=\prod_{j=1}^d w(x_j)$. Then for $f(\mathbf{x})$ defined on $[0,1)^d$ we consider the cubature formula

$$\Phi(a,A,w)(f):=\sum_{\mu=1}^m w_\mu f(\xi^\mu),\qquad w_\mu:=w(\eta^\mu).$$

In [36] we proved the following analogs of Theorem 6.4 and Corollary 6.5.

Theorem 6.7. *Let* $r\ge 2$. *There exists a constant* $c(d,A,r)>0$ *such that for any* $V\ge V_0:=c(d,A,r)a^{-d}$ *we have for all* $\mathbf{u}\in(0,1/2]^d$, $pr(\mathbf{u},d)=V$, *and* $\mathbf{z}\in[0,1)^d$

$$|\Phi(a,A,w)(\tilde h^r(\,\cdot\,,\mathbf{z},\mathbf{u}))-\hat{\tilde h}^r(\mathbf{0},\mathbf{z},\mathbf{u})|\le C(d,A,r,w)a^{-rd}(\log(2V/V_0))^{d-1}.$$

Definition 6.8. Let $V\in(0,1]$. We define the periodic r-smooth fixed volume discrepancy with equal weights as

$$\tilde D^r(\xi,V):=\sup_{B\in\mathcal{B}:\mathrm{vol}(B)=V}\left|\int\tilde h^r_B(\mathbf{x})\mathrm{d}\mathbf{x}-\frac{1}{m}\sum_{\mu=1}^m\tilde h^r_B(\xi^\mu)\right|. \tag{6.10}$$

The optimized version of the periodic r-smooth fixed volume discrepancy is defined as follows

$$\tilde D^{r,o}(\xi,V):=\inf_{\lambda_1,\dots,\lambda_m}\sup_{B\in\mathcal{B}:\mathrm{vol}(B)=V}\left|\int\tilde h^r_B(\mathbf{x})\mathrm{d}\mathbf{x}-\sum_{\mu=1}^m\lambda_\mu\tilde h^r_B(\xi^\mu)\right|. \tag{6.11}$$

Corollary 6.9. *For* $r\ge 2$ *there exists a constant* $c(d,A,r)>0$ *such that for any* $V\ge V_0:=c(d,A,r)a^{-d}$ *we have for the point set* ξ *defined by* (6.8)

$$\tilde D^{r,o}(\xi,V)\le C(d,A,r)a^{-rd}(\log(2V/V_0))^{d-1}.$$

In particular, Theorem 6.7 implies that the periodic r-smooth discrepancy (see Definition 4.5) satisfies the bound (for $r\in\mathbb{N}$, $r\ge 2$)

$$\tilde D^{r,o}_\infty(m,d)\le C(d,r)m^{-r}(\log m)^{d-1}. \tag{6.12}$$

Theorem 5.3 shows that the bound (6.12) cannot be improved for a natural class of weights $\lambda_1,\dots,\lambda_m$ used in the optimization procedure in the definition of $\tilde D^{r,o}_\infty(m,d)$, namely, for weights, satisfying condition (2.2)

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

7. Dispersion

We remind the definition of dispersion. Let $d \geq 2$ and $[0,1)^d$ be the d-dimensional unit cube. As above for $\mathbf{x}, \mathbf{y} \in [0,1)^d$ with $\mathbf{x} = (x_1, \dots, x_d)$ and $\mathbf{y} = (y_1, \dots, y_d)$ we write $\mathbf{x} < \mathbf{y}$ if this inequality holds coordinate-wise. For $\mathbf{x} < \mathbf{y}$ we write $[\mathbf{x}, \mathbf{y})$ for the axes-parallel box $[x_1, y_1) \times \cdots \times [x_d, y_d)$ and define

$$\mathcal{B} := \{[\mathbf{x}, \mathbf{y}) : \mathbf{x}, \mathbf{y} \in [0,1)^d, \mathbf{x} < \mathbf{y}\}.$$

For $N \geq 1$ let T be a set of points in $[0,1)^d$ of cardinality $|T| = N$. The volume of the largest empty (from points of T) axes-parallel box, which can be inscribed in $[0,1)^d$, is called the dispersion of T:

$$\operatorname{disp}(T) := \sup_{B \in \mathcal{B} : B \cap T = \emptyset} \operatorname{vol}(B).$$

An interesting extremal problem is to find (estimate) the minimal dispersion of point sets of fixed cardinality:

$$\operatorname{disp}^*(N,d) := \inf_{T \subset [0,1)^d, |T| = N} \operatorname{disp}(T).$$

It is known that

$$\operatorname{disp}^*(N,d) \leq C^*(d)/N. \tag{7.1}$$

Inequality (7.1) with $C^*(d) = 2^{d-1} \prod_{i=1}^{d-1} p_i$, where p_i denotes the ith prime number, was proved in [11] (see also [20]). The authors of [11] used the Halton–Hammersly set of N points (see [18]). Inequality (7.1) with $C^*(d) = 2^{7d+1}$ was proved in [1]. The authors of [1], following G. Larcher, used the (t,r,d)-nets (see [18], [19] for results on (t,r,d)-nets and Definition 8.2 below for the definition). We point out that both the Halton–Hammersly points and the (t,r,d)-nets provide constructive results in (7.1).

It was demonstrated in [42] how good upper bounds on fixed volume discrepancy can be used for proving good upper bounds for dispersion. This fact was one of the motivations for studying the fixed volume discrepancy. Theorem 7.1 below was derived from Theorem 6.2 (see [42]). The upper bound in Theorem 7.1 combined with the trivial lower bound shows that the Fibonacci point set provides optimal rate of decay for the dispersion.

Theorem 7.1. *There is an absolute constant C such that for all n we have*

$$\operatorname{disp}(\mathcal{F}_n) \leq C/b_n. \tag{7.2}$$

The following Theorem 7.2 was derived in [42] from Theorem 6.4.

Theorem 7.2. *Let A be a matrix from Lemma 6.3. There is a constant $C(d,A)$, which may only depend on A and d, such that for all a we have*

$$\operatorname{disp}(\mathcal{F}(a,A)) \leq C(A,d) a^{-d}. \tag{7.3}$$

Other proof of Theorem 7.2 was given by M. Ullrich [45]. His proof is based on a deep result on admissible matrices obtained by M. M. Skriganov [25].

We only discussed behavior of $\operatorname{disp}^*(N,d)$ as a function on N. Certainly, the problem of finding the right dependence of $\operatorname{disp}^*(N,d)$ on both parameters is an important and difficult problem. The reader can find further recent results in this direction in [16], [22], [26], [44], [45], and [46].

8. Universal discretization

In this section we demonstrate an application of results on dispersion from Section 7 to the problem of universal discretization. For a more detailed discussion of universality in approximation theory and in learning theory we refer the reader to [6], [12], [14], [27], [29], [31], [33], [40], [41]. We remind the discretization problem setting, which we plan to discuss (see [37] and [39]).

Marcinkiewicz problem. Let Ω be a compact subset of $\mathbb{R}^d$ with the probability measure μ. We say that a linear subspace X_N (usually N stands for the dimension of X_N) of the $L_q(\Omega) := L_q(\Omega, \mu)$, $1 \le q < \infty$, admits the Marcinkiewicz-type discretization theorem with parameters m and q if there exist a set $\{\xi^\nu \in \Omega, \nu = 1, \dots, m\}$ and two positive constants $C_j(d,q)$, $j = 1,2$, such that for any $f \in X_N$ we have

$$C_1(d,q)\|f\|_q^q \le \frac{1}{m}\sum_{\nu=1}^{m} |f(\xi^\nu)|^q \le C_2(d,q)\|f\|_q^q. \tag{8.1}$$

In the case $q = \infty$ we define L_∞ as the space of continuous on Ω functions and ask for

$$C_1(d)\|f\|_\infty \le \max_{1\le\nu\le m} |f(\xi^\nu)| \le \|f\|_\infty. \tag{8.2}$$

We will also use a brief way to express the above property: the $\mathcal{M}(m,q)$ theorem holds for a subspace X_N or $X_N \in \mathcal{M}(m,q)$. The reader can find a survey on recent results in this direction in the paper [10]. We only remind classical results on discretization of the L_p norms of trigonometric polynomials defined on $\mathbb{T}^d$. By Q we denote a finite subset of $\mathbb{Z}^d$, and $|Q|$ stands for the number of elements in Q. Let

$$\mathcal{T}(Q) := \left\{ f : f(\mathbf{x}) = \sum_{\mathbf{k}\in Q} c_{\mathbf{k}} e^{i(\mathbf{k},\mathbf{x})}, \quad c_{\mathbf{k}} \in \mathbb{C} \right\}.$$

Consider the case $Q = \Pi(\mathbf{N}) := [-N_1, N_1] \times \dots \times [-N_d, N_d]$, $N_j \in \mathbb{N}$ or $N_j = 0$, $j = 1, \dots, d$, $\mathbf{N} = (N_1, \dots, N_d)$. We set

$$P(\mathbf{N}) := \Big\{ \mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d : 0 \le n_j \le 2N_j, j = 1, \dots, d \Big\},$$

and

$$\mathbf{x}^{\mathbf{n}} := \left(\frac{2\pi n_1}{2N_1 + 1}, \dots, \frac{2\pi n_d}{2N_d + 1} \right), \qquad \mathbf{n} \in P(\mathbf{N}).$$

For any $f \in \mathcal{T}(\Pi(\mathbf{N}))$, one has

$$\|f\|_2^2 = \vartheta(\mathbf{N})^{-1} \sum_{\mathbf{n}\in P(\mathbf{N})} \big|f(\mathbf{x}^{\mathbf{n}})\big|^2,$$

where $\vartheta(\mathbf{N}) := \prod_{j=1}^d (2N_j + 1) = \dim \mathcal{T}(\Pi(\mathbf{N}))$. In particular, this implies that for any $\mathbf{N}$ one has

$$\mathcal{T}(\Pi(\mathbf{N})) \in \mathcal{M}(\vartheta(\mathbf{N}), 2). \tag{8.3}$$

In the case $1 < q < \infty$, the well-known Marcinkiewicz discretization theorem (for $d = 1$) is given as follows (see, for instance, [40, §§1.3.3 and 3.3.4]): for $f \in \mathcal{T}(\Pi(\mathbf{N}))$,

$$C_1(d,q)\|f\|_q^q \le \vartheta(\mathbf{N})^{-1} \sum_{\mathbf{n}\in P(\mathbf{N})} \big|f(\mathbf{x}^{\mathbf{n}})\big|^q \le C_2(d,q)\|f\|_q^q, \quad 1 < q < \infty.$$

This yields the following extension of (8.3):

$$\mathcal{T}(\Pi(\mathbf{N})) \in \mathcal{M}(\vartheta(\mathbf{N}), q), \quad 1 < q < \infty.$$

For $q = 1$ or $q = \infty$, one needs some adjustments. Let

$$P'(\mathbf{N}) := \Big\{ \mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d : 1 \le n_j \le 4N_j, j = 1, \dots, d \Big\}$$

and

$$\mathbf{x}(\mathbf{n}) := \left(\frac{\pi n_1}{2N_1}, \dots, \frac{\pi n_d}{2N_d} \right), \qquad \mathbf{n} \in P'(\mathbf{N}).$$

If $N_j = 0$, we let $x_j(\mathbf{n}) = 0$. Set $\overline{N} := \max(N, 1)$ and $\nu(\mathbf{N}) := \prod_{j=1}^d \overline{N}_j$. The following Marcinkiewicz-type discretization theorem (see [40, p. 102]): for $f \in \mathcal{T}(\Pi(\mathbf{N}))$

$$C_1(d,q)\|f\|_q^q \le \nu(4\mathbf{N})^{-1} \sum_{\mathbf{n}\in P'(\mathbf{N})} |f(\mathbf{x}(\mathbf{n}))|^q \le C_2(d,q)\|f\|_q^q, \quad 1 \le q \le \infty, \tag{8.4}$$

implies that

$$\mathcal{T}(\Pi(\mathbf{N})) \in \mathcal{M}(\nu(4\mathbf{N}), q), \quad 1 \le q \le \infty.$$

We note that $\nu(4\mathbf{N}) \le C(d) \dim \mathcal{T}(\Pi(\mathbf{N}))$.

Universal discretization problem. This problem is about finding (proving existence) of a set of points, which is good in the sense of the above Marcinkiewicz-type discretization for a collection of linear subspaces (see [41]). We formulate it in an explicit form. Let $\mathcal{X}_N := \{X_N^j\}_{j=1}^k$ be a collection of linear subspaces X_N^j of the $L_q(\Omega)$, $1 \le q \le \infty$. We say that a set $\{\xi^\nu \in \Omega, \nu = 1, \dots, m\}$ provides *universal discretization* for the collection $\mathcal{X}_N$ if, in the case $1 \le q < \infty$, there are two positive constants $C_i(d,q)$, $i = 1, 2$, such that for each $j \in [1,k]$ and any $f \in X_N^j$ we have

$$C_1(d,q)\|f\|_q^q \le \frac{1}{m} \sum_{\nu=1}^m |f(\xi^\nu)|^q \le C_2(d,q)\|f\|_q^q. \tag{8.5}$$

In the case $q = \infty$ for each $j \in [1,k]$ and any $f \in X_N^j$ we have

$$C_1(d)\|f\|_\infty \le \max_{1\le\nu\le m} |f(\xi^\nu)| \le \|f\|_\infty. \tag{8.6}$$

In [41] we studied the universal discretization for the collection of subspaces of trigonometric polynomials with frequencies from parallelepipeds (rectangles). For $\mathbf{s} \in \mathbb{N}_0^d$ define

$$R(\mathbf{s}) := \{\mathbf{k} \in \mathbb{Z}^d : |k_j| < 2^{s_j}, \quad j = 1, \dots, d\}.$$

Consider the collection $\mathcal{C}(n,d) := \{\mathcal{T}(R(\mathbf{s})), \|\mathbf{s}\|_1 = n\}$.

The following theorem was proved in [41].

Theorem 8.1. *Let a set T with cardinality $|T| = 2^r =: m$ have dispersion satisfying the bound $disp(T) < C(d)2^{-r}$ with some constant $C(d)$. Then there exists a constant $c(d) \in \mathbb{N}$ such that the set $2\pi T := \{2\pi\mathbf{x} : \mathbf{x} \in T\}$ provides the universal discretization in L_∞ for the collection $\mathcal{C}(n,d)$ with $n = r - c(d)$.*

Theorem 8.1 is a conditional result. As we discussed in Section 7 existence of sets with a property required in Theorem 8.1 is a non-trivial fact. In particular, the (t,r,d)-nets provide such existence. We now give a definition of the (t,r,d)-nets.

Definition 8.2. A (t,r,d)-net (in base 2) is a set T of 2^r points in $[0,1)^d$ such that each dyadic box $[(a_1-1)2^{-s_1}, a_1 2^{-s_1}) \times \cdots \times [(a_d-1)2^{-s_d}, a_d 2^{-s_d})$, $1 \le a_j \le 2^{s_j}$, $j = 1, \dots, d$, of volume 2^{t-r} contains exactly 2^t points of T.

We note that existence of (t,r,d)-nets is a very non-trivial problem. A construction of such nets for all d and $t \ge Cd$, where C is a positive absolute constant, $r \ge t$ is given in [19].

Theorem 8.1 in a combination with Theorems 7.1 and 7.2 guarantees that the appropriately chosen Fibonacci ($d = 2$) and Frolov (any $d \ge 2$) point sets provide universal discretization in L_∞ for the collection $\mathcal{C}(n,d)$.

The following Theorem 8.3 (see [41]) can be seen as an inverse to Theorem 8.1.

Theorem 8.3. *Assume that $T \subset [0,1)^d$ is such that the set $2\pi T$ provides universal discretization in L_∞ for the collection $\mathcal{C}(n,d)$ with a constant $C_1(d)$ (see (8.2)). Then there exists a positive constant $C(d)$ with the following property $disp(T) \le C(d)2^{-n}$.*

Arbitrary trigonometric polynomials. For $n \in \mathbb{N}$ denote $\Pi_n := \Pi(\mathbf{N}) \cap \mathbb{Z}^d$ with $\mathbf{N} = (2^{n-1} - 1, \dots, 2^{n-1} - 1)$, where $\Pi(\mathbf{N}) := [-N_1, N_1] \times \dots \times [-N_d, N_d]$. Then $|\Pi_n| = (2^n - 1)^d < 2^{dn}$. Let $v \in \mathbb{N}$ and $v \le |\Pi_n|$. Consider

$$\mathcal{S}(v, n) := \{Q \subset \Pi_n : |Q| = v\}.$$

Then it is easy to see that

$$|\mathcal{S}(v, n)| = \binom{|\Pi_n|}{v} < 2^{dnv}. \tag{8.7}$$

We are interested in solving the following problem of universal discretization. For a given $\mathcal{S}(v, n)$ and $q \in [1, \infty)$ find a condition on m such that there exists a set $\xi = \{\xi^\nu\}_{\nu=1}^m$ with the property: for any $Q \in \mathcal{S}(v, n)$ and each $f \in \mathcal{T}(Q)$ we have

$$C_1(q, d)\|f\|_q^q \le \frac{1}{m}\sum_{\nu=1}^m |f(\xi^\nu)|^q \le C_2(q, d)\|f\|_q^q. \tag{8.8}$$

We present results from [10] for $q = 2$ and $q = 1$.

Theorem 8.4. *There exist three positive constants $C_i(d)$, $i = 1, 2, 3$, such that for any $n, v \in \mathbb{N}$ and $v \le |\Pi_n|$ there is a set $\xi = \{\xi^\nu\}_{\nu=1}^m \subset \mathbb{T}^d$, with $m \le C_1(d)v^2 n$, which provides universal discretization in L_2 for the collection $\mathcal{S}(v, n)$: for any $f \in \cup_{Q \in \mathcal{S}(v,n)} \mathcal{T}(Q)$*

$$C_2(d)\|f\|_2^2 \le \frac{1}{m}\sum_{\nu=1}^m |f(\xi^\nu)|^2 \le C_3(d)\|f\|_2^2.$$

The classical Marcinkiewicz-type result for $\mathcal{T}(\Pi_n)$ (see, for instance, (8.3)) provides a universal set ξ with cardinality $m \le C(d)2^{dn}$. Thus, Theorem 8.4 gives a non-trivial result for v satisfying $v^2 n \le C(d)2^{dn}$. We now present a result for discretization of the L_1 norm.

Theorem 8.5. *There exist three positive constants $C_1(d)$, C_2, C_3, such that for any $n, v \in \mathbb{N}$ and $v \le |\Pi_n|$ there is a set $\xi = \{\xi^\nu\}_{\nu=1}^m \subset \mathbb{T}^d$, with $m \le C_1(d)v^2 n^{9/2}$, which provides universal discretization in L_1 for the collection $\mathcal{S}(v, n)$: for any $f \in \cup_{Q \in \mathcal{S}(v,n)} \mathcal{T}(Q)$*

$$C_2\|f\|_1 \le \frac{1}{m}\sum_{\nu=1}^m |f(\xi^\nu)| \le C_3\|f\|_1.$$

The classical Marcinkiewicz-type result for $\mathcal{T}(\Pi_n)$ (see (8.4) with $q = 1$) provides a universal set ξ with cardinality $m \le C(d)2^{dn}$. Thus, Theorem 8.5 gives a non-trivial result for v satisfying $v^2 n^{9/2} \le C(d)2^{dn}$.

9. Generalizations

In Section 4 we discussed numerical integration of functions from special classes $\mathbf{W}_p^r$ and $\dot{\mathbf{W}}_p^r$. Functions from both of these classes have integral representations with the kernel $F_r(\mathbf{x} - \mathbf{y})$ in case of the class $\mathbf{W}_p^r$ and with the kernel $B_r(\mathbf{x}, \mathbf{y})$ in case of the class $\dot{\mathbf{W}}_p^r$. In this section we present a generalization to the case, when a function class $\mathbf{W}_p^K$ is defined with a help of a general kernel $K(\mathbf{x}, \mathbf{y})$. As above for a function class $\mathbf{W}$ we have a concept of error of the cubature formula $\Lambda_m(\,\cdot\,, \xi)$

$$\Lambda_m(\mathbf{W}, \xi) := \sup_{f \in \mathbf{W}} \left| \int_\Omega f \mathrm{d}\mu - \Lambda_m(f, \xi) \right|. \tag{9.1}$$

If the class $\mathbf{W} = \{f(\mathbf{x}, \mathbf{y}) : \mathbf{y} \in Y\}$ is parametrized by a parameter $\mathbf{y} \in Y \subset \mathbb{R}^n$ with Y being a bounded measurable set, then we can consider a natural *average case setting.* For $\mathbf{p} = (p_1, \dots, p_n)$

define

$$\Lambda_m(\mathbf{W},\xi,\mathbf{p}) := \left\|\int_\Omega f(\,\cdot\,,\mathbf{y})\mathrm{d}\mu - \Lambda_m(f(\,\cdot\,,\mathbf{y}),\xi)\right\|_{\mathbf{p}}, \tag{9.2}$$

where the vector $L_{\mathbf{p}}$ norm is taken with respect to the Lebesgue measure on Y. We write

$$\Lambda_m(\mathbf{W},\xi,\infty) := \Lambda_m(\mathbf{W},\xi).$$

We are interested in dependence on m of the optimal errors of numerical integration with m knots

$$\kappa_m(\mathbf{W},\mathbf{p}) := \inf_{\lambda_1,\dots,\lambda_m;\xi^1,\dots,\xi^m} \Lambda_m(\mathbf{W},\xi,\mathbf{p})$$

for different classes $\mathbf{W}$. We have $\kappa_m(\mathbf{W},\infty) = \kappa_m(\mathbf{W})$ (see (2.4).

We now present a rather general setting of this problem. Let $1 \le q \le \infty$. We define a set $\mathcal{K}_q$ of kernels possessing the following properties. Let $K(\mathbf{x},\mathbf{y})$ be a measurable function on $\Omega^1 \times \Omega^2$. We assume that for any $\mathbf{x} \in \Omega^1$ we have $K(\mathbf{x},\cdot\,) \in L_q(\Omega^2)$; for any $\mathbf{y} \in \Omega^2$ the $K(\,\cdot\,,\mathbf{y})$ is integrable over Ω^1 and $\int_{\Omega^1} K(\mathbf{x},\cdot\,)\mathrm{d}\mathbf{x} \in L_q(\Omega^2)$. For $1 \le p \le \infty$ and a kernel $K \in \mathcal{K}_{p'}$, $p' := p/(p-1)$, we define the class

$$\mathbf{W}_p^K := \left\{ f : f(\mathbf{x}) = \int_{\Omega^2} K(\mathbf{x},\mathbf{y})\varphi(\mathbf{y})\mathrm{d}\mathbf{y}, \quad \|\varphi\|_{L_p(\Omega^2)} \le 1 \right\}. \tag{9.3}$$

Then each $f \in \mathbf{W}_p^K$ is integrable on Ω^1 (by Fubini's theorem) and defined at each point of Ω^1. We denote for convenience

$$J_K(\mathbf{y}) := \int_{\Omega^1} K(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x}.$$

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

$$\Lambda_m(\mathbf{W}_p^K,\xi) = \sup_{\|\varphi\|_{L_p(\Omega^2)}\le 1} \left| \int_{\Omega^2} \Big(J_K(\mathbf{y}) - \sum_{\mu=1}^m \lambda_\mu K(\xi^\mu,\mathbf{y})\Big)\varphi(\mathbf{y})\mathrm{d}\mathbf{y} \right|$$

$$= \left\| J_K(\,\cdot\,) - \sum_{\mu=1}^m \lambda_\mu K(\xi^\mu,\cdot\,) \right\|_{L_{p'}(\Omega^2)}. \tag{9.4}$$

Consider a problem of numerical integration of functions $K(\mathbf{x},\mathbf{y})$, $\mathbf{y} \in \Omega^2$, with respect to $\mathbf{x}$, $K \in \mathcal{K}_q$, in other words a problem of numerical integration of functions from the function class $\mathbf{K} := \{K(\mathbf{x},\mathbf{y}) : \mathbf{y} \in \Omega^2\}$:

$$\int_{\Omega^1} K(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x} - \sum_{\mu=1}^m \lambda_\mu K(\xi^\mu,\mathbf{y}).$$

Definition 9.1. (K,q)-discrepancy of a set of knots $\xi = \{\xi^1,\dots,\xi^m\}$ and a set of weights $\Lambda = \{\lambda_1,\dots,\lambda_m\}$ (a pair (ξ,Λ)) is

$$D(\xi,\Lambda,K,q) := \Lambda_m(\mathbf{K},\xi,q) = \left\| \int_{\Omega^1} K(\mathbf{x},\mathbf{y})\mathrm{d}\mathbf{x} - \sum_{\mu=1}^m \lambda_\mu K(\xi^\mu,\mathbf{y}) \right\|_{L_q(\Omega^2)}.$$

In a special case $\Lambda_m(\,\cdot\,,\xi) = Q_m(\,\cdot\,,\xi)$ we write $D(\xi,Q,K,q)$. The above definition of the (K,q)-discrepancy and relation (9.4) imply right a way the following relation

$$D(\xi,\Lambda,K,p') = \Lambda_m(\mathbf{W}_p^K,\xi). \tag{9.5}$$

Relation (9.5) shows that numerical integration in the class $\mathbf{W}_p^K$ and the (K,q)-discrepancy are tied by the duality principle. In a special case, when $K(\mathbf{x},\mathbf{y}) = B_r(\mathbf{x},\mathbf{y})$ the (K,q)-discrepancy $D(\xi,\Lambda,B_r,q)$ of a set of knots $\xi^1,\dots,\xi^m$ and a set of weights $\lambda_1,\dots,\lambda_m$ coincides with the r-discrepancy $D_q^r(\xi,\Lambda)$ of the pair (ξ,Λ) (see (4.1)).

Let us consider a special case, when $K(\mathbf{x},\mathbf{y}) = F(\mathbf{x}-\mathbf{y})$, $\Omega^1 = \Omega^2 = [0,1)^d$ and we deal with 1-periodic in each variable functions. Associate with a cubature formula (ξ,Λ) and the function F the following function

$$g_{\xi,\Lambda,F}(\mathbf{x}) := \sum_{\mathbf{k}} \Lambda(\xi,\mathbf{k})\hat{F}(\mathbf{k})e^{2\pi i(\mathbf{k},\mathbf{x})} - \hat{F}(\mathbf{0}),$$

where

$$\Lambda(\xi,\mathbf{k}) := \Lambda_m(e^{2\pi i(\mathbf{k},\mathbf{x})},\xi).$$

Then for the quantity $\Lambda_m(\mathbf{W}^F_p,\xi)$ we have ($p' := p/(p-1)$)

$$\begin{aligned}\Lambda_m(\mathbf{W}^F_p,\xi) &= \sup_{f\in\mathbf{W}^F_p} |\Lambda_m(f,\xi) - \hat{f}(\mathbf{0})| \\ &= \sup_{\|\varphi\|_p\le 1} \left|\Lambda_m(F(\mathbf{x})*\varphi(\mathbf{x}),\xi) - \hat{F}(\mathbf{0})\hat{\varphi}(\mathbf{0})\right| \\ &= \sup_{\|\varphi\|_p\le 1} \left|\langle g_{\xi,\Lambda,F}(-\mathbf{y}),\overline{\varphi(\mathbf{y})}\rangle\right| = \|g_{\xi,\Lambda,F}\|_{p'}. \end{aligned} \tag{9.6}$$

Thus, the worst case error of a cubature formula $\Lambda_m(\,\cdot\,,\xi)$ on the class $\mathbf{W}^F_p$ coincides with the $L_{p'}$ norm of the function $g_{\xi,\Lambda,F}$.

Let us discuss a special case of function F, which is very important in numerical integration (see, for instance, [12], [29], [31], and [40]). Let for $r>0$

$$F_{r,\alpha}(x) := 1 + 2\sum_{k=1}^{\infty} k^{-r}\cos(2\pi kx - \alpha\pi/2). \tag{9.7}$$

For $\mathbf{x} = (x_1,\dots,x_d)$, $\alpha = (\alpha_1,\dots,\alpha_d)$ denote

$$F_{r,\alpha}(\mathbf{x}) := \prod_{j=1}^{d} F_{r,\alpha_j}(x_j)$$

and

$$\mathbf{W}^r_{p,\alpha} := \mathbf{W}^{F_{r,\alpha}}_p = \{f : f(\mathbf{x}) = (F_{r,\alpha}*\varphi)(\mathbf{x}), \quad \|\varphi\|_p \le 1\},$$

where

$$(F_{r,\alpha}*\varphi)(\mathbf{x}) := \int_{[0,1)^d} F_{r,\alpha}(\mathbf{x}-\mathbf{y})\varphi(\mathbf{y})\mathrm{d}\mathbf{y}.$$

It is easy to see that

$$\|g_{\xi,\Lambda,F_{r,\alpha}}\|_2 = \left(\sum_{\mathbf{k}\ne\mathbf{0}} |\Lambda(\xi,\mathbf{k})|^2 \left(\prod_{j=1}^{d}(\max(|k_j|,1))^{-r}\right)^2 + |\Lambda(\xi,\mathbf{0}) - 1|^2\right)^{1/2}. \tag{9.8}$$

The above quantity in the case $r=1$ was introduced in [50] under the name *diaphony.* In case of generic r it was called *generalized diaphony* and was studied in [17]. Relation (9.6) shows that generalized diaphony is closely related to numerical integration of the class $\mathbf{W}^r_{2,\alpha}$. Following this analogy, we can call the quantity $\|g_{\xi,\Lambda,F_{r,\alpha}}\|_q$ the (r,q)-diaphony of the pair (ξ,Λ) (the cubature formula (ξ,Λ)). Behavior of $\kappa_m(\mathbf{W}^r_{p,\alpha})$ is well studied (see, for instance, [12] and [40]). By (9.6) results on $\kappa_m(\mathbf{W}^r_{p,\alpha})$ provide estimates on

$$\inf_{\xi,\Lambda} \|g_{\xi,\Lambda,F_{r,\alpha}}\|_{p'}.$$

For completeness we cite some known results on the lower bounds for $\kappa_m(\mathbf{W}^r_{p,\alpha})$. The reader can find these and other results with a historical discussion in [40, Ch. 6] and in [12, Ch. 8].

Theorem 9.2. *The following lower estimate is valid for any cubature formula (ξ, Λ) with m knots $(r > 1/p)$*

$$\Lambda_m(\mathbf{W}^r_{p,\alpha}, \xi) \geq C(r,d,p)m^{-r}(\log m)^{\frac{d-1}{2}}, \qquad 1 \leq p < \infty.$$

The rate of decay $m^{-r}(\log m)^{\frac{d-1}{2}}$ in the lower bound in Theorem 9.2 does not depend on p. Therefore, the larger the $p < \infty$ the stronger the lower bound. It turns out that in the case $p = 1$ one can improve the corresponding lower bound under certain restrictions on the weights of the cubature formula. We obtained the lower estimates for the quantities

$$\kappa^B_m(\mathbf{W}) := \inf_{\Lambda_m(\,\cdot\,,\xi)\in Q(B,m)} \Lambda_m(\mathbf{W}, \xi),$$

where $Q(B,m)$ is the collection of cubature formulas satisfying condition (2.2). We proved the following relation.

Theorem 9.3. *Let $r > 1$. Then*

$$\kappa^B_m(\mathbf{W}^r_{1,0}) \geq C(r,B,d)m^{-r}(\log m)^{d-1}, \qquad C(r,B,d) > 0.$$

The case $p = \infty$ is excluded in Theorem 9.2. There is no nontrivial general lower estimates in this case. We give one conditional result in this direction (see, for instance, [40, p. 271]).

Theorem 9.4. *Let the cubature formula (ξ, Λ) be such that the inequality*

$$\Lambda_m(\mathbf{W}^r_{p,\alpha}, \xi) \leq C_1(p,r,d)m^{-r}(\log m)^{(d-1)/2}, \qquad r > 1/p,$$

holds for some $1 < p < \infty$. Then there exists a constant $C_2(p,r,d) > 0$ such that

$$\Lambda_m(\mathbf{W}^r_{\infty,\alpha}, \xi) \geq C_2(p,r,d)m^{-r}(\log m)^{(d-1)/2}.$$

There are two big open problems in this area. We formulate them as conjectures.

Conjecture 9.5. *For any $d \geq 2$ and any $r \geq 1$ we have*

$$\kappa_m(\mathbf{W}^r_{1,\alpha}) \geq C(r,d)m^{-r}(\log m)^{d-1}.$$

Conjecture 9.6. *For any $d \geq 2$ and any $r > 0$ we have*

$$\kappa_m(\mathbf{W}^r_{\infty,\alpha}) \geq C(r,d)m^{-r}(\log m)^{(d-1)/2}.$$

We note that by Theorem 4.3 and (4.2) Conjecture 9.5 implies Conjecture 3.1 and Conjecture 9.6 implies for any cubature formula (ξ, Λ)

$$D^r_1(\xi, \Lambda) \geq C(r,d)m^{-r}(\log m)^{(d-1)/2}. \tag{9.9}$$

Remark 9.7. In the case $d = 2$, $r = 1$, and $\alpha = (1,1)$ Conjecture 9.5 holds.

Remark 9.7 follows from an analog of the Schmidt's bound (3.2) and Proposition 4.2. We discuss this in more detail. D. Bilyk and I observed that a slight modification of the proof of (3.2) from [3] gives the following lower bound. For any cubature formula (ξ, Λ) we have

$$\Lambda_m(\chi^2, \xi) \geq C_1 m^{-1}\log m. \tag{9.10}$$

Therefore, by Proposition 4.2 for any cubature formula (ξ, Λ), satisfying an extra condition $\sum_j \lambda_j = 1$, we have for $\mathbf{W}^1_1 = \mathbf{W}^1_{1,(1,1)}$

$$\Lambda_m(\mathbf{W}^1_1, \xi) \geq C_2 m^{-1}\log m. \tag{9.11}$$

Further, it is well known (see [40, p. 269]) and easy to check, that for a function class $\mathbf{W}$ of periodic functions, satisfying the condition: $1 \in \mathbf{W}$ and for $f \in \mathbf{W}$ we have $\frac{1}{2}(f - \hat{f}(\mathbf{0})) \in \mathbf{W}$, the inequality holds

$$\inf_{\Lambda} \Lambda_m(\mathbf{W}, \xi) \geq \frac{1}{4} \inf_{\Lambda:\sum_j \lambda_j = 1} \Lambda_m(\mathbf{W}, \xi). \tag{9.12}$$

Clearly, $\mathbf{W}_1^1$ satisfies the above condition on a class $\mathbf{W}$. Combining (9.10)–(9.12) we obtain for $d = 2$

$$\kappa_m(\mathbf{W}_1^1) \geq Cm^{-1} \log m.$$

10. Numerical integration without smoothness assumptions

In the previous sections we discussed numerical integration for classes of functions under certain conditions on smoothness. Parameter r controlled the smoothness. The above results show that the numerical integration characteristics decay with the rate $m^{-r}(\log m)^{c(d)}$, which substantially depends on smoothness r. The larger the smoothness – the faster the error decay. In this section we discuss the case, when we do not impose any of the smoothness assumptions. Surprisingly, even in such a situation we can guarantee some rate of decay. Results discussed in this section apply in a very general setting presented in Section 9. We present here results from [38]. The following result is proved in [35] (see also [31] for previous results). For the theory of greedy algorithms we refer the reader to [34]. Consider a dictionary

$$\mathcal{D} := \{K(\mathbf{x}, \cdot\,), \mathbf{x} \in \Omega^1\}$$

and define a Banach space $X(K,q)$ as the $L_q(\Omega^2)$-closure of span of $\mathcal{D}$.

Theorem 10.1. *Let* $\mathbf{W}_p^K$ *be a class of functions defined above in Section 9. Assume that* $K \in \mathcal{K}_{p'}$ *satisfies the condition*

$$\|K(\mathbf{x}, \cdot\,)\|_{L_{p'}(\Omega^2)} \leq 1, \quad \mathbf{x} \in \Omega^1, \quad |\Omega^1| = 1,$$

and $J_K \in X(K,p')$*. Then for any* m *there exists (provided by an appropriate greedy algorithm) a cubature formula* $Q_m(\,\cdot\,,\xi)$ *such that*

$$Q_m(\mathbf{W}_p^K, \xi) \leq C(p-1)^{-1/2} m^{-1/2}, \quad 1 < p \leq 2.$$

As a direct corollary of Theorem 10.1 and relation (9.5) we obtain the following result about the (K,q) – *discrepancy*.

Theorem 10.2. *Assume that* $K \in \mathcal{K}_q$ *satisfies the condition*

$$\|K(\mathbf{x}, \cdot\,)\|_{L_q(\Omega^2)} \leq 1, \quad \mathbf{x} \in \Omega^1, \quad |\Omega^1| = 1,$$

and $J_K \in X(K,q)$*. Then for any* m *there exists (provided by an appropriate greedy algorithm) a cubature formula* $Q_m(\,\cdot\,,\xi)$ *such that*

$$D(\xi, Q, K, q) \leq Cq^{1/2} m^{-1/2}, \quad 2 \leq q < \infty.$$

Remark 10.3. In Theorems 10.1 and 10.2 we impose the restriction $1 < p \leq 2$ or the dual one $2 \leq q < \infty$. The proof of Theorems 10.1 and 10.2 from [35] also works in the case $2 < p < \infty$ or $1 < q < 2$ and gives

$$Q_m(\mathbf{W}_p^K, \xi) \leq Cm^{-1/p}, \quad 2 < p < \infty,$$

$$D(\xi, Q, K, q) \leq Cm^{\frac{1}{q}-1}, \quad 1 < q < 2.$$

Let us discuss a special case $K(\mathbf{x},\mathbf{y}) = F(\mathbf{x}-\mathbf{y})$, $\Omega^1 = \Omega^2 = [0,1)^d$ and 1-periodic in each variable functions. Then, as in Section 9, we associate with a cubature formula (ξ, Λ) and a function F the function $g_{\xi,\Lambda,F}(\mathbf{x})$. The following Proposition is proved in [31].

Proposition 10.4. *Let* $1 < p < \infty$ *and* $\|F\|_{p'} \leq 1$*. Then the kernel* $K(\mathbf{x},\mathbf{y}) = F(\mathbf{x}-\mathbf{y})$ *satisfies the assumptions of Theorem 10.1.*

Proposition 10.4, Theorem 10.1, Remark 10.3, and relation (9.6) imply

Theorem 10.5. *Let $1 < p < \infty$ and $\|F\|_p \le 1$. Then there exists a set ξ of m points such that*

$$\|g_{\xi,Q,F}(\mathbf{x})\|_p \le Cp^{1/2}m^{-1/2}, \quad 2 \le p < \infty,$$

$$\|g_{\xi,Q,F}(\mathbf{x})\|_p \le Cm^{\frac{1}{p}-1}, \quad 1 < p < 2.$$

Here is a corollary of Theorem 10.2 and Proposition 10.4. Let $E \subset [0,1)^d$ be a measurable set. Consider $F(\mathbf{x}) := \widetilde{\chi}_E(\mathbf{x})$.

Theorem 10.6. *For any $p \in [2,\infty)$ there exists a set of m points ξ such that*

$$Q_m(\{\widetilde{\chi}_E(\mathbf{x}-\mathbf{z}), \mathbf{z} \in [0,1)^d\}, \xi, p) \le Cp^{1/2}m^{-1/2}.$$

We note that there are interesting results on the behavior of $Q_m(\{\chi_E(\mathbf{x}-\mathbf{z}), \mathbf{z} \in [0,1)^d\}, \xi, \infty)$ under assumption that E is a convex set (see [2]). Theorem 10.6 shows that for $p < \infty$ we do not need any assumptions on the geometry of E in order to get the upper bound $\ll m^{-1/2}$ for the discrepancy.

The proof of the above Theorems 10.1–10.6 is constructive (see [35]), it is based on the greedy algorithms. The use of greedy-type algorithms is an important new ingredient in numerical integration. For completeness, in order to give the reader an idea about a greedy-type algorithm, we briefly formulate the related result from the theory of greedy approximation. We remind some notations from the theory of greedy approximation in Banach spaces. The reader can find a systematic presentation of this theory in [34, Ch. 6]. 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 less than or equal to one ($\|g\| \le 1$) and the closure of span of $\mathcal{D}$ coincides with X.

For an element $f \in X$ we denote by F_f a norming (peak) functional for f:

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

The existence of such a functional is guaranteed by the Hahn–Banach theorem.

We proceed to the Incremental Greedy Algorithm (see [32] and [34, Ch. 6]). Let $\epsilon = \{\epsilon_n\}_{n=1}^{\infty}$, $\epsilon_n > 0$, $n = 1, 2, \dots$. For a Banach space X and a dictionary $\mathcal{D}$ define the following algorithm IA(ϵ) := IA($\epsilon, X, \mathcal{D}$).

Incremental Algorithm with schedule ϵ (IA($\epsilon, X, \mathcal{D}$)). Denote $f_0^{i,\epsilon} := f$ and $G_0^{i,\epsilon} := 0$. Then, for each $m \ge 1$ we have the following inductive definition.

(1) $\varphi_m^{i,\epsilon} \in \mathcal{D}$ is any element satisfying

$$F_{f_{m-1}^{i,\epsilon}}(\varphi_m^{i,\epsilon} - f) \ge -\epsilon_m.$$

(2) Define

$$G_m^{i,\epsilon} := (1 - 1/m)G_{m-1}^{i,\epsilon} + \varphi_m^{i,\epsilon}/m.$$

(3) Let

$$f_m^{i,\epsilon} := f - G_m^{i,\epsilon}.$$

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

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

It is well known that in the case $X = L_p$, $1 \le p < \infty$ we have

$$\rho(u) \le \begin{cases} u^p/p & \text{if} \quad 1 \le p \le 2, \\ (p-1)u^2/2 & \text{if} \quad 2 \le p < \infty. \end{cases} \tag{10.1}$$

Denote by $A_1(\mathcal{D}) := A_1(\mathcal{D}, X)$ the closure in X of the convex hull of $\mathcal{D}$. Proof of Theorem 10.1 and Remark 10.3 is based on the following theorem proved in [32] (see also [34, Ch. 6]).

Theorem 10.7. *Let X be a Banach space with modulus of smoothness $\rho(u) \le \gamma u^q$, $1 < q \le 2$. Set*

$$\epsilon_n := \beta\gamma^{1/q} n^{-1/p}, \qquad p := \frac{q}{q-1}, \quad n = 1, 2, \dots .$$

Then, for every $f \in A_1(\mathcal{D})$ we have

$$\|f_m^{i,\epsilon}\| \le C(\beta)\gamma^{1/q} m^{-1/p}, \qquad m = 1, 2 \dots .$$

Acknowledgements

The author is grateful to the referees for many useful suggestions.

References

[1] C. Aistleitner, A. Hinrichs, and D. Rudolf. On the size of the largest empty box amidst a point set. *Discrete Appl. Math.*, 230:146–150, 2017.

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

[3] D. Bilyk. Roth's Orthogonal Function Method in Discrepancy Theory and Some New Connections. In *Panorama of Discrepancy Theory. Lecture Notes in Mathematics. 2017*, pages 71–158. Springer, 2014.

[4] D. Bilyk and M. Lacey. On the Small Ball Inequality in three dimensions. *Duke Math. J.*, 143:81–115, 2008.

[5] D. Bilyk, M. Lacey, and A. Vagharshakyan. On the Small Ball Inequality in all dimensions. *J. Funct. Anal.*, 254:2470–2502, 2008.

[6] P. Binev, A. Cohen, W. Dahmen, R. DeVore, and V. N. Temlyakov. Universal algorithms for learning theory. Part I: piecewise constant functions. *J. Mach. Learn. Theory*, 6:1297–1321, 2005.

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

[8] J. G. van der Corput. Verteilungsfunktionen.I. *Proc. Akad. Wet. Amsterdam*, 38:813–821, 1935.

[9] J. G. van der Corput. Verteilungsfunktionen.II. *Proc. Akad. Wet. Amsterdam*, 38:1058–1066, 1935.

[10] F. Dai, A. Prymak, V. N. Temlyakov, and S. Tikhonov. Integral norm discretization and related problems. `https://arxiv.org/abs/1807.01353v1`, 2018.

[11] A. Dumitrescu and M. Jiang. On the largest empty axis-parallel box amidst n points. *Algorithmica*, 66:225–248, 2013.

[12] D. Dũng, V. N. Temlyakov, and T. Ullrich. Hyperbolic Cross Approximation. `https://arxiv.org/abs/1601.03978v2`, 2016.

[13] K. K. Frolov. Upper bounds on the error of quadrature formulas on classes of functions. *Dokl. Akad. Nauk SSSR*, 231:818–821, 1976. English transl. in *Sov. Math., Dokl.* 17, 1976.

[14] L. Györfy, M. Kohler, A. Krzyzak, and H. Walk. *A distribution-free theory of nonparametric regression.* Springer, 2002.

[15] G. Halász. On Roth's method in the theory of irregularities of points distributions. In *Recent Progress in Analytic Number Theory. Vol. 2*, pages 79–94. Academic Press Inc., 1981.

[16] D. Krieg. On the Dispersion of Sparse Grids. `https://arxiv.org/abs/1709.02983`, 2017.

[17] V. F. Lev. The exact order of generalized diaphony and multidimensional numerical integration. *J. Aust. Math. Soc., Ser. A*, 66:1–17, 1999.

[18] J. Matousek. *Geometric Discrepancy.* Springer, 1999.

[19] H. Niederreiter and C. Xing. Low-discrepancy sequences and global function fields with many rational places. *Finite Fields Appl.*, 2:241–273, 1996.

[20] G. Rote and F. Tichy. Quasi-Monte Carlo methods and the dispersion of point sequences. *Math. Comput. Modelling*, 23:9–23, 1996.

[21] K. F. Roth. On irregularities of distribution. *Mathematica*, 1:73–79, 1954.

[22] D. Rudolf. An Upper Bound of the Minimal Dispersion via Delta Covers. In *Contemporary Computational Mathematics - A Celebration of the 80th Birthday of Ian Sloan*, pages 1099–1108. Springer, 2018.

[23] W. M. Schmidt. Irregularities of distribution.VII. *Acta Arith.*, 21:45–50, 1972.

[24] W. M. Schmidt. Irregularities of distribution. X. In *Number Theory and Algebra*, pages 311–329. Academic Press Inc., 1977.

[25] M. M. Skriganov. Constructions of uniform distributions in terms of geometry of numbers. *Algebra Anal.*, 6:200–230, 1994.

[26] J. Sosnovec. A note on minimal dispersion of point sets in the unit cube. *Eur. J. Comb.*, 69:255–259, 2018.

[27] V. N. Temlyakov. Approximation by elements of a finite-dimensional subspace of functions from various Sobolev or Nikol'skii spaces. *Mat. Zametki*, 43:770–786, 1988. English transl. in *Math. Notes* 43, 1988.

[28] V. N. Temlyakov. On a way of obtaining lower estimates for the errors of quadrature formulas. *Mat. Sbornik*, 181:1403–1413, 1990. English transl. in *Math. USSR Sbornik* 71, 1992.

[29] V. N. Temlyakov. *Approximation of periodic functions.* Computational Mathematics and Analysis Series. Nova Science Publishers, 1993.

[30] V. N. Temlyakov. On error estimates for cubature formulas. *Tr. Mat. Inst. Steklova*, 207:326–338, 1994. English transl. in *Proc. Steklov Inst. Math.* 6:299–309, 1995.

[31] V. N. Temlyakov. Cubature formulas, discrepancy, and nonlinear approximation. *J. Complexity*, 19:352–391, 2003.

[32] V. N. Temlyakov. Greedy-Type Approximation in Banach Spaces and Applications. *Constr. Approx.*, 21:257–292, 2005.

[33] V. N. Temlyakov. On universal estimators in learning theory. *Tr. Mat. Inst. Steklova*, 255:256–272, 2006. English transl. in *Proc. Steklov Inst. Math.* 255:244–259, 2006.

[34] V. N. Temlyakov. *Greedy approximation.* Cambridge University Press, 2011.

[35] V. N. Temlyakov. Incremental Greedy Algorithm and Its Applications in Numerical Integration. In *Monte Carlo and Quasi-Monte Carlo Methods*, pages 557–570. Springer, 2016.

[36] V. N. Temlyakov. Fixed volume discrepancy in the periodic case. `https://arxiv.org/abs/1710.11499v1`, 2017.

[37] V. N. Temlyakov. The Marcinkiewicz-type discretization theorems for the hyperbolic cross polynomials. *Jaen J. Approx.*, 9:37–63, 2017.

[38] V. N. Temlyakov. Remarks on numerical integration, discrepancy, and diaphony. `https://arxiv.org/abs/1711.07017v1`, 2017.

[39] V. N. Temlyakov. The Marcinkiewicz-type discretization theorems. *Constr. Approx.*, 48:337–369, 2018.

[40] V. N. Temlyakov. *Multivariate approximation.* Cambridge University Press, 2018.

[41] V. N. Temlyakov. Universal discretization. *J. Complexity*, 47:97–109, 2018.

[42] V. N. Temlyakov. Smooth fixed volume discrepancy, dispersion, and related problems. *J. Approximation Theory*, 237:113–134, 2019.

[43] M. Ullrich. On "Upper error bounds for quadrature formulas on function classes" by K. K. Frolov. In *Monte Carlo and Quasi-Monte Carlo Methods, MCQMC*, pages 571–582. Springer, 2016.

[44] M. Ullrich. A lower bound for the dispersion on the torus. *Math. Comput. Simulation*, 143:186–190, 2018.

[45] M. Ullrich. A note on the dispersion of admissible lattices. *Discrete Appl. Math.*, 257:385–387, 2019.

[46] M. Ullrich and J. Vybiral. An upper bound on the minimal dispersion. *J. Complexity*, 45:120–126, 2018.

[47] T. van Aardenne-Ehrenfest. Proof of the impossibility of a just distribution of an infinite sequence of points over an interval. *Proc. Akad. Wet. Amsterdam*, 48:266–271, 1945.

[48] T. van Aardenne-Ehrenfest. On the impossibility of a just distribution. *Proc. Akad. Wet. Amsterdam*, 52:734–739, 1949.

[49] H. Weyl. Über die Gleichverteilung von Zahlen mod. Eins. *Math. Ann.*, 77:313–352, 1916.

[50] P. Zinterhof. Über einige Abschätzungen bei der Approximation von Funktionen mit Gleichverteilungsmethoden. *Österr. Akad. Wiss., Math.-Naturw. Kl., S.-Ber., Abt. II*, 185:121–132, 1976.