INDUSTRIAL MATHEMATICS INSTITUTE

1997:09

Ridge approximation, Chebyshev-Fourier analysis and optimal quadrature formulas

K. Oskolkov

IMI
Preprint Series

Department of Mathematics
University of South Carolina

Ridge approximation, Chebyshev – Fourier analysis and optimal quadrature formulas

K.I. Oskolkov

Abstract

Free (non-linear) ridge L^2-approximation $\mathcal{NRA}_n(f)$, $n = 1, 2, \dots$, of a function $f(\mathbf{x}) = f(x_1, x_2)$ in the unit disc $\mathbb{B}^2$ is considered:

$$\left\| f - \sum_{1}^{n} F_j(\mathbf{x} \cdot \boldsymbol{\xi}_j);\ L^2(\mathbb{B}^2) \right\| \Longrightarrow \inf \text{ in } \{F_j(t)\}_1^n \text{ and } \{\boldsymbol{\xi}_j\}_1^n \subset \mathcal{S}^1,$$

where $\{F_j(t)\}_1^n$ denotes a set of n single variate functions. Geometrically, the $\mathcal{NRA}$-problem means approximation of f by a linear combination of n planar waves of arbitrary shapes F_j and directions of propagation (wave vectors) $\boldsymbol{\xi}_j$.

A duality relation is established between the $\mathcal{NRA}$ problem and that of optimal quadrature formulas, in the sense of Kolmogorov – Nikol'skii, for classes of trigonometric polynomials.

On the base of this duality and lower estimates of errors of quadrature formulas, it is proved that if $f(\mathbf{x})$ is radial, $f(\mathbf{x}) = f(|\mathbf{x}|)$, then algebraic polynomials in two variables provide "almost best" tool for ridge approximation:

$$\frac{1}{c_0}\mathcal{PA}_{3n}(f) \le \mathcal{NRA}_n(f) \le \mathcal{PA}_{n-1}(f), \quad n = 1, 2, \dots,$$

where c_0 is an absolute positive constant, and $\mathcal{PA}_n(f)$ denotes the n-th best algebraic polynomial approximation of f in $L^2(\mathbb{B}^2)$:

$$\mathcal{PA}_n(f) := \min_{p(\mathbf{x}) \in \mathcal{P}_n^2} \left\| f - p;\ L^2(\mathbb{B}^2) \right\|; \quad \mathcal{P}_n^2 := \text{Span}\left\{ x_1^k x_2^l \right\}_{k+l \le n}.$$

It is known that algebraic polynomials of degree n in two variables can be represented as linear combinations of $n+1$ planar wave polynomials. Radon – Fourier analysis via Chebyshev ridge polynomials is crucial in the proof.

Math. Subject Classification: AMS(MOS) 41A25, 41A63.

Keywords and Phrases: Ridge Approximation, Radon Transform, Chebyshev Polynomials, Optimal Quadrature Formulas.

Acknowledgements and Notes. The author was supported by the Office of the Naval Research Grant N00014-91-5-1076.

The author expresses his deep gratitude to R.A. DeVore, B.S. Kashin, P.P. Petrushev and V.N. Temlyakov for many useful discussions. The author also thanks V.E. Majorov for cooperation and providing the manuscript of the paper [4].

1 Introduction

Let n be a natural number, $\mathbb{R}$. Denote $\mathcal{R}_n$ the set of linear combinations of n *planar waves* on the real plane R^2, i. e.

$$R(\mathbf{x}) \in \mathcal{R}_n \Longleftrightarrow R(\mathbf{x}) = \sum_1^n F_j(\mathbf{x} \cdot \boldsymbol{\xi}_j),$$

where $F_j(t)$ are functions of a single real variable t, $\boldsymbol{\xi}_j$ are unit vectors, i.e. $\boldsymbol{\xi}_j \in \mathcal{S}^1$ (*wave vectors*), and $\mathbf{x} \cdot \boldsymbol{\xi}$ denotes the usual inner product of vectors $\mathbf{x}$, $\boldsymbol{\xi}$. Thus, functions from $\mathcal{R}_n$ are linear combinations of n planar waves, in general, of arbitrary shapes and directions of propagation.

Obviously, double trigonometric polynomials

$$T(\mathbf{x}) = \sum_1^n c_j e^{i\omega_j(\boldsymbol{\xi}_j \cdot \mathbf{x})}$$

are elements of $\mathcal{R}_n$ for every choice of frequencies ω_j and wave vectors $\boldsymbol{\xi}_j$.

The fact that *algebraic polynomials* in two variables of degree $n-1$ also belong to the set $\mathcal{R}_n$:

$$\mathcal{P}^2_{n-1} := \mathrm{Span}\left\{x_1^k x_2^l\right\}_{k+l\le n-1} \subset \mathcal{R}_n, \quad n = 1,\, 2,\, \ldots \tag{1}$$

is somewhat less obvious (cf. (16) below). Its significance in problems of *Radon inversion* and the so called *non-linear ridge approximation* (for brevity, $\mathcal{NRA}$ in the sequel). was demonstrated in [1], cf. also [2].

A particular case of the latter problem, answering the functions $f(\mathbf{x}) = f(x_1, x_2)$, $\mathbf{x} \in \mathbb{B}^2$ supported in the unit disc $\mathbb{B}^2 := \{\mathbf{x} : \ |\mathbf{x}| \le 1\}$, and the usual $L^2(\mathbb{B}^2)$ norm,

$$\left\| f;\ L^2(\mathbb{B}^2) \right\| := \sqrt{\int_{\mathbb{B}^2} |f(\mathbf{x})|^2 \, d\mathbf{x}}\,,$$

is formulated as follows. *Given a function* $f \in L^2(\mathbb{B}^2)$*, and a natural* n*, find the quantity*

$$\mathcal{NRA}_n(f) = \mathcal{NRA}_n\left(f;\ L^2(\mathbb{B}^2)\right) := \inf_{R \in \mathcal{R}_n} \left\| f - R;\ L^2(\mathbb{B}^2) \right\|,$$

and the corresponding minimizer $R^*(f) \in \mathcal{R}_n$*, if the latter exists.* Thus, geometrically the $\mathcal{NRA}$ problem consists in searching for linear combinations of n planar waves, of arbitrary shapes and directions, that are best fit to $f(\mathbf{x})$ in the sense of L^2 distance. It obviously follows from the classical K. Weierstrass' approximation theorem that $\mathcal{NRA}_n(f) \to 0, \quad n \to \infty,\ \forall f \in L^2(\mathbb{B}^2)$. Moreover, it follows directly from (1) that the following inequalities hold true

$$\mathcal{NRA}_n(f) \le \mathcal{PA}_{n-1}(f), \quad n = 1,\, 2,\, \ldots \tag{2}$$

where $\mathcal{PA}_n(f)$ denote the values of best algebraic polynomial approximations of f:

$$\mathcal{PA}_n(f) = \mathcal{PA}_n\left(f;\ L^2(\mathbb{B}^2)\right) := \min_{P \in \mathcal{P}_n^2} \left\| f - p;\ L^2(\mathbb{B}^2) \right\|, \quad n = 0,\, 1,\, \ldots .$$

It should be noted that even in the simplest case of the metric L^2, the extremal problem of $\mathcal{NRA}$ is of *highly non-linear nature.* This non-linearity dwells in the selection of the optimal set of wave vectors $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$, that are allowed to depend upon the given function $f(\mathbf{x})$. Neither of the basic mathematical questions of *existence, uniquenness* of the optimal linear combination $R_n^*(\mathbf{x}) = R_n^*(f, \mathbf{x}) \in \mathcal{R}_n$ of n planar waves can be solved in general terms. An a priori reason for the effect of non-existence can be seen in *non-compactness* of the class of all admissible single variate functions $\{F_j(t)\}$, generating planar waves $F_j(\mathbf{x} \cdot \boldsymbol{\xi}_j)$. As a result, some of the wave vectors, in approach to inf in the definition of $\mathcal{NRA}_n(f)$, can tend to *couple*, or asymptotically stick together. *The set* $\mathcal{R}_n$ *is not closed* for $n \ge 2$, since no restrictions are imposed on the distribution of the wave vectors $\boldsymbol{\xi}_j \in \mathcal{S}^1$. Indeed, if $F(t),\ t \in R^1$ is a differentiable function, $\boldsymbol{\xi} := \mathbf{e}_\vartheta = \langle \cos\vartheta, \sin\vartheta \rangle$ - a fixed unit vector, $\boldsymbol{\xi}^\perp := \langle -\sin\vartheta, \cos\vartheta \rangle$, then the function

$$\begin{aligned} f(\mathbf{x}) &= f_\vartheta(\mathbf{x}) := (\mathbf{x} \cdot \boldsymbol{\xi}^\perp)\, F'(\mathbf{x} \cdot \boldsymbol{\xi}) = \frac{\partial F(x_1 \cos\vartheta + x_2 \sin\vartheta)}{\partial \vartheta} \\ &= \lim_{\vartheta_1, \vartheta_2 \to \vartheta} \left(\frac{F(x_1 \cos\vartheta_1 + x_2 \sin\vartheta_1)}{\vartheta_2 - \vartheta_1} + \frac{F(x_1 \cos\vartheta_2 + x_2 \sin\vartheta_2)}{\vartheta_1 - \vartheta_2} \right) \end{aligned}$$

is a *limit* of a sequence of linear combinations of 2 planar waves, i. e belongs to the closure of $\mathcal{R}_2$. Consequently, for every function $f(\mathbf{x})$ of the type $f(\mathbf{x}) = (\mathbf{x} \cdot \boldsymbol{\xi}^\perp)\, F'(\mathbf{x} \cdot \boldsymbol{\xi})$ one has $\mathcal{NRA}_2(f) = 0$, but obviously in general $f_\vartheta(\mathbf{x}) \notin \mathcal{R}_2$. Further, consider the function $f(\mathbf{x}) := x_1 x_2$. For $\varphi \neq \frac{k\pi}{2},\ k \in \mathrm{Z}$ one obviously has

$$x_1 x_2 = \frac{(x_1 \cos\varphi + x_2 \sin\varphi)^2}{2\sin 2\varphi} - \frac{(x_1 \cos\varphi - x_2 \sin\varphi)^2}{2\sin 2\varphi},$$

so that again $\mathcal{NRA}_2(f) = 0$, the optimal combination of 2 planar waves *exists, but is essentially non-unique.*

Let us also note the example of the function

$$f(\mathbf{x}) := |\mathbf{x}|^2 - |\mathbf{x}|^4 = x_1^2 + x_2^2 - (x_1^2 + x_2^2)^2,$$

which is a *radial* algebraic polynomial of degree = 4. This function provides a warning against possible "naive" conjectures about the structure of the set of optimal wave vectors. Such an apparent conjecture "on the run", without thorough analysis, is that the *optimal wave vectors should be selected equidistributed on the circle* $\mathcal{S}^1$, simply because the function is radial. In the case $n = 2$, this would mean that the two optimal wave vectors should be mutually perpendicular, i. e. $\boldsymbol{\xi}_1 \cdot \boldsymbol{\xi}_2 = 0$. However, we will prove that the latter is not true: *for* $f(\mathbf{x}) := |\mathbf{x}|^2 - |\mathbf{x}|^4$ *and* $n = 2$*, the optimal wave vectors* $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ *are defined by the relation*

$$\boldsymbol{\xi}_1 \cdot \boldsymbol{\xi}_2 = \sqrt{\frac{3}{8}}. \tag{3}$$

The main goal of the present paper is to establish the following statement.

Theorem 1 *There exists an absolute positive constant* c_0 *such that nonlinear ridge- and polynomial approximations of every radial function* $f(\mathbf{x}) = f(|\mathbf{x}|)$ *are related by the following inequalities:*

$$\frac{1}{c_0}\mathcal{PA}_{3n}(f) \le \mathcal{NRA}_n(f) \le \mathcal{PA}_{n-1}(f), \quad n = 1, 2, \ldots . \tag{4}$$

As mentioned above, cf. (2), only the lower estimate $\mathcal{NRA}_n(f) \ge \frac{1}{c_0}\mathcal{PA}_{2n}(f)$ in this statement is new. The meaning of this statement is that *for each radial function* $f(|\mathbf{x}|)$ *that is "not too smooth", namely,* $\mathcal{PA}_n(f) = O\left(\mathcal{PA}_{3n}(f)\right),\ n \to \infty$*, orthogonal projections onto subspaces of algebraic polynomials represent the optimal in order and linear tool of ridge approximation.*

In connection with the latter corollary, recent results of V.E. Majorov [4] and V.N. Temlyakov [5] can be quoted, on estimates from below of *upper bounds* of $\mathcal{NRA}_n(f)$ on variants of Sobolev classes $W^r(\mathbb{B}^2)$. In somewhat loose words, the class $W^r(\mathbb{B}^2)$ in the papers [4] and [5] consists of functions $f(\mathbf{x})$ whose polynomial approximations satisfy the estimate $\mathcal{PA}_n(f) = O\left(n^{-r}\right),\ n \to \infty$, and *existence* of a function $f(\mathbf{x}) \in W^r(\mathbb{B}^2)$ is established for which

$$\mathcal{NRA}_n(f) \ge (n \ln n)^{-r}, \quad n = 2, 3, \ldots .$$

It follows from (4) that the factors $(\ln n)^{-r}$ in this result can be dropped: *for every radial function* whose polynomial approximations $\mathcal{PA}_n(f)$ are of exact order n^{-r}, the non-linear ridge approximations $\mathcal{NRA}_n(f)$ are of the same order of magnitude.

The proof of Theorem 1 relies upon Radon inversion formula and the corresponding Fourier – Chebyshev analysis in $\mathbb{B}^2$. In the next section we list the necessary facts (for more details, the reader may be referred to [1], [2] or [3]).

2 Radon inversion formula via Chebyshev – Fourier series

A general approach to the problem of ridge approximation can be seen from the following. On the first step, find an *integral representation* of the function $f(\mathbf{x})$ of the form

$$f(\mathbf{x}) \sim \frac{1}{2\pi} \int_{\mathcal{S}^1} F(\boldsymbol{\xi}, \mathbf{x} \cdot \boldsymbol{\xi})\, d\boldsymbol{\xi}, \tag{5}$$

where $F(\boldsymbol{\xi}, t) = F(f;\, \boldsymbol{\xi}, t)$. Then discretize the integral on the righthand side by a suitable *quadrature formula* (Riemannian sum)

$$\int_{\mathcal{S}^1} F(\boldsymbol{\xi}, \mathbf{x} \cdot \boldsymbol{\xi})\, d\boldsymbol{\xi} \sim \sum_1^n |\Delta \boldsymbol{\xi}_j| F(\boldsymbol{\xi}_j, \mathbf{x} \cdot \boldsymbol{\xi}_j). \tag{6}$$

The first step is accomplished by applying direct and inverse Radon transforms. For $f(\mathbf{x}) \in L^1(\mathbb{R}^2)$, denote $R(f;\, \boldsymbol{\xi}, t)$, $\boldsymbol{\xi} \in \mathcal{S}^1$, $t \in \mathbb{R}^1$ the *direct Radon transform*:

$$R(f;\, \boldsymbol{\xi}, t) := \int_{\mathbf{y} \cdot \boldsymbol{\xi} = t} f(\mathbf{y})\, m_1(d\mathbf{y}),$$

where $m_1(d\mathbf{y})$ stands for the 1d Lebesgue measure on the real line $\mathcal{R}^1$. Then each sufficiently smooth and rapidly decreasing function $f(\mathbf{x})$ can be reconstructed by applying to $R(f;\, \boldsymbol{\xi}, t)$ the *inverse Radon transform*:

$$f(\mathbf{x}) = \frac{1}{4\pi} \int_{\mathcal{S}^1} (\mathcal{H}\mathcal{D}) R(f;\, \boldsymbol{\xi}, \cdot)\Big|_{t = \mathbf{x} \cdot \boldsymbol{\xi}} d\boldsymbol{\xi}. \tag{7}$$

Here $(\mathcal{H}\mathcal{D})$ denotes the composition of commuting one-dimensional operators of Hilbert transform $\mathcal{H}$ and differentiation $\mathcal{D}$, i.e.

$$(\mathcal{H}\mathcal{D}) R(f;\, \boldsymbol{\xi}, \cdot)(t) := \frac{1}{2\pi} \frac{\partial}{\partial t} \int_{\mathcal{R}^1} R(f;\, \boldsymbol{\xi}, s) \cot \frac{t - s}{2}\, ds\,.$$

Thus, in capacity of $F(f;\,\boldsymbol{\xi},t)$ in the integral representation (6) one can take the function

$$F(f;\,\boldsymbol{\xi},t) := (\mathcal{H}\mathcal{D})R(f;\,\boldsymbol{\xi},\cdot)(t) = \frac{1}{4\pi}\frac{\partial}{\partial t}\int_{\mathcal{R}^1} R(f;\,\boldsymbol{\xi},s)\,\cot\frac{t-s}{2}\,ds\,.$$

Obviously, Radon inversion operator (7) is a composition of two operators of polarly different nature. *Singular* part $(\mathcal{H}\mathcal{D})$ is applied to the direct Radon transform $R(f;\,\boldsymbol{\xi},t)$ in the space variable t. After it and the substitution $t = \mathbf{x}\cdot\boldsymbol{\xi}$, the smoothing operator of *averaging* $\frac{1}{2\pi}\int_{\mathcal{S}^1}$ in the angular variable $\boldsymbol{\xi}$ is applied. Thus, the difficulty in the second step of construction of ridge approximation consists in search of a suitable quadrature formula (6) for the image of the singular operator $(\mathcal{H}\mathcal{D})$.

For functions $f(\mathbf{x})$ supported in $\mathbb{B}^2$, the direct Radon transforms $R(f;\,\boldsymbol{\xi},t)$ are obviously supported on the interval $t\in\mathbb{B}^1 := (-1,1)$. Restriction of the general Radon inversion operator (7) on the class of such functions naturally generates Fourier analysis where Chebyshev polynomials of the second kind

$$u_n(t) := \frac{1}{\sqrt{\pi}}\,\mathcal{D}_n(\arccos t), \quad \text{where } \mathcal{D}_n(\vartheta) := \frac{\sin{(n+1)\vartheta}}{\sin\vartheta}\,,\ n = 0,\,1,\,\ldots\,.$$

play the crucial role, cf. [1], [2]. These classical polynomials constitute a complete orthonormal system $\mathcal{U} = \{u_{n-1}(t)\}_{n=1}^{\infty}$ in $L^2_w(-1,1)$ with the weight $w(t) = 2\sqrt{1-t^2}$.

Moreover, these polynomials constitute the complete system of *eigen-functions* of the operator $(\mathcal{H}\mathcal{D}w)$ in the $1d$ spectral problem $(\mathcal{H}\mathcal{D})w(t)u(t) = \lambda u(t),\ t\in\mathbb{B}^1$.

After substitution $t = \mathbf{x}\cdot\boldsymbol{\xi},\ \mathbf{x}\in\mathbb{B}^2,\ \boldsymbol{\xi}\in\mathcal{S}^1,\ u_n(\mathbf{x}\cdot\boldsymbol{\xi})$ generate *a family of complete orthonormal systems of ridge polynomials* in $L^2(\mathbb{B}^2)$. The fundamental properties of this system

$$\mathcal{U}\mathcal{S}^1 := \{\{u_{n-1}(\mathbf{x}\cdot\boldsymbol{\xi})\}_{n=1}^{\infty}\}_{\boldsymbol{\xi}\in\mathcal{S}^1}$$

are expressed by the following

1. Orthogonality relation.

$$\int_{\mathbb{B}^2} u_n(\mathbf{x}\cdot\boldsymbol{\xi})P(\mathbf{x})\,d\mathbf{x} = 0 \quad \forall P(\mathbf{x})\in\mathcal{P}^2_{n-1},\ n = 1,\,2,\,\ldots\,,\ \forall\boldsymbol{\xi}\in\mathcal{S}^1, \tag{8}$$

or $u_n(\mathbf{x}\cdot\boldsymbol{\xi}) \perp \mathcal{P}^2_{n-1}$ in $L^2(\mathbb{B}^2)$.

A proof of this property can be carried out using Chebyshev's general ideas of polynomials of best approximation and symmetry of $\mathbb{B}^2$. First of all, not loosing generality we can take $\boldsymbol{\xi} = \langle 1,0\rangle$. Next, let $f(\mathbf{x}) = f(x_1)\in L^2(\mathbb{B}^2)$ be a single variate function in $L^2(\mathbb{B}^2)$, or $f(t)\in L^2_w(\mathbb{B}^1),\ w(t) = 2\sqrt{1-t^2}$.

Consider the problem of best approximation of this function by all polynomials in *two variables* $x_1 := t,\ x_2$, of the class $\mathcal{P}^2_{n-1}$, in $L^2(\mathbb{B}^2)$:

$$\|f(t) - P(\mathbf{x});\ L^2(\mathbb{B}^2)\| \Longrightarrow \min \quad \text{in } \ P(\mathbf{x}) \in \mathcal{P}^2_{n-1}.$$

Then using Jensen's inequality and symmetry of $\mathbb{B}^2$ it is not hard to see, that the minimizer P^* of this problem is indeed a *single variate polynomial*, $P^*(\mathbf{x}) = P^*(t) \in \mathcal{P}^1_{n-1}$. Obviously, we have also $f(t) - P^*(t) \perp \mathcal{P}^2_{n-1}$ in $L^2(\mathbb{B}^2)$ and in particular

$$\int_{\mathbb{B}^2} (f(t) - P^*(t))P(t)\, d\mathbf{x} = \int_{-1}^{1} (f(t) - P^*(t))P(t)w(t)\, dt = 0 \quad \forall P(t) \in \mathcal{P}^1_{n-1}.$$

If we take $f(t) := t^n$, we easily see that the corresponding mimimizer $t^n - P^*(t)$ is a multiple of the n-th Chebyshev polynomial $u_n(t)$, and (8) follows, because we have $t^n - P^*(t) \perp \mathcal{P}^2_{n-1}$ in $L^2(\mathbb{B}^2)$.

2. Inner products of Chebyshev ridge polynomials.

$$\int_{\mathbb{B}^2} u_n(\mathbf{x}\cdot\boldsymbol{\xi})u_n(\mathbf{x}\cdot\boldsymbol{\eta})\, d\mathbf{x} = \frac{u_n(\boldsymbol{\xi}\cdot\boldsymbol{\eta})}{u_n(1)} = \frac{\sqrt{\pi}}{n+1}\, u_n(\boldsymbol{\xi}\cdot\boldsymbol{\eta}), \quad \forall \boldsymbol{\xi},\boldsymbol{\eta} \in \mathcal{S}^1. \tag{9}$$

Furtermore, let $\mathcal{T}_n^{\pm}$ denote the subspace of trigonometric polynomials $a(\boldsymbol{\xi}),\ \boldsymbol{\xi} \in \mathcal{S}^1$, of degree $\deg a = n$ and satisfying $a(-\boldsymbol{\xi}) \equiv (-1)^n a(\boldsymbol{\xi}),\ \boldsymbol{\xi} \in \mathcal{S}^1$. Then

$$\frac{1}{2\pi} \int_{\mathcal{S}^1} \sqrt{\pi} u_n(\boldsymbol{\xi}\cdot\boldsymbol{\eta})\, a(\boldsymbol{\eta})\, d\boldsymbol{\eta} = a(\boldsymbol{\xi}), \quad a(\boldsymbol{\xi}) \in \mathcal{T}_n^{\pm},\ \boldsymbol{\xi} \in \mathcal{S}^1 \tag{10}$$

and in particular

$$\frac{1}{2\sqrt{\pi}} \int_{\mathcal{S}^1} u_n(\boldsymbol{\xi}_1\cdot\boldsymbol{\eta})\, u_n(\boldsymbol{\eta}\cdot\boldsymbol{\xi}_2)\, d\boldsymbol{\eta} = u_n(\boldsymbol{\xi}_1\cdot\boldsymbol{\xi}_2), \quad \boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \mathcal{S}^1$$

The latter two relations simply means that the convolution $\sqrt{\pi}u_n * a$ represents the identity operator on $\mathcal{T}_n^{\pm}$, i.e. $\sqrt{\pi}u_n$ is the Dirichlet kernel.

3. Integral representation.

Given a function $f(\mathbf{x}) \in L^2(\mathbb{B}^2)$, consider the following Chebyshev – Fourier coefficients, depending on $\boldsymbol{\xi} \in \mathcal{S}^1$ as a parameter:

$$a_n(f,\boldsymbol{\xi}) := \int_{\mathbb{B}^2} f(\mathbf{y})u_n(\mathbf{y}\cdot\boldsymbol{\xi})\, d\mathbf{y}\ ,\ n = 0,\, 1,\, \dots\ , \quad \boldsymbol{\xi} \in \mathcal{S}^1.$$

The following relation represents the integral form of Chebyshev ridge polynomial Fourier series, which is in fact *Radon inversion formula* (7) for $f(\mathbf{x})$ expressed via Chebyshev ridge polynomials:

$$f(\mathbf{x}) \overset{L^2(\mathbb{B}^2)}{=} \frac{1}{2\pi} \sum_{n=1}^{\infty} n \int_{\mathcal{S}^1} a_{n-1}(f,\boldsymbol{\xi}) u_{n-1}(\mathbf{x}\cdot\boldsymbol{\xi})\, d\boldsymbol{\xi}\,. \tag{11}$$

4. Integral form of Parceval's identity.
If $f(\mathbf{x}), g(\mathbf{x}) \in L^2(\mathbb{B}^2)$, then $a_n(f,\boldsymbol{\xi}), a_n(g,\boldsymbol{\xi}) \in \mathcal{T}_n^{\pm}$ and

$$\int_{\mathbb{B}^2} f(\mathbf{x}) g(\mathbf{x})\, d\mathbf{x} = \frac{1}{2\pi} \sum_{n=1}^{\infty} n \int_{\mathcal{S}^1} a_{n-1}(f,\boldsymbol{\xi}) a_{n-1}(g,\boldsymbol{\xi})\, d\boldsymbol{\xi}\,.$$

In particular, if $f(\mathbf{x}) \in L^2(\mathbb{B}^2)$, then

$$\| f(\mathbf{x}),\ L^2(\mathbb{B}^2)\|^2 = \frac{1}{2\pi} \sum_{n=1}^{\infty} n \| a_{n-1}(\boldsymbol{\xi}),\ L^2(\mathcal{S}^1)\|^2. \tag{12}$$

These relations easily follow from (9) and (10).
5. Plancherel's theorem.
If $\{a_n(\boldsymbol{\xi})\}_{n=0}^{\infty}$ is a sequence of trigonometric polynomials, satisfying the conditions

$$a_n(\boldsymbol{\xi}) \in \mathcal{T}_n^{\pm}\,, \quad \sum_{n=1}^{\infty} n\, \| a_{n-1}(\boldsymbol{\xi}),\ L^2(\mathcal{S}^1)\|^2 < \infty$$

then there exists a unique function $f(\mathbf{x}) \in L^2(\mathbb{B}^2)$ such that $a_n(f,\boldsymbol{\xi}) = a_n(\boldsymbol{\xi}),\ n = 0,\, 1,\, \ldots$.
6. Partial sums and orthogonal projections onto $\mathcal{P}_n^2$.
Let $n = 1,\, 2,\, \ldots$.' Then the finite partial sum of the Fourier expansion (11)

$$S_n(f,\ \mathbf{x}) := \frac{1}{2\pi} \sum_{m=1}^{n} m \int_{\mathcal{S}^1} a_{m-1}(f,\boldsymbol{\xi}) u_{m-1}(\mathbf{x}\cdot\boldsymbol{\xi})\, d\boldsymbol{\xi}$$

coincides with orthogonal projection of $f(\mathbf{x})$ onto the subspace of algebraic polynomials $\mathcal{P}_{n-1}^2$ in $L^2(\mathbb{B}^2)$:

$$\| f(\mathbf{x}) - S_n(f,\ \mathbf{x}),\ L^2(\mathbb{B}^2\| = \min_{P\in\mathcal{P}_{n-1}^2} \| f(\mathbf{x}) - P(\mathbf{x}),\ L^2(\mathbb{B}^2\| = \mathcal{P}\mathcal{A}_{n-1}(f).$$

Obviously, $S_n(f,\ \mathbf{x})$ is a linear integral operator:

$$S_n(f,\ \mathbf{x}) = \int_{\mathbb{B}^2} D_n(\mathbf{x},\mathbf{y}) f(\mathbf{y})\, d\mathbf{y}, \quad \text{where } D_n(\mathbf{x},\mathbf{y}) := \frac{1}{2\pi} \sum_{m=1}^{n} m \int_{\mathcal{S}^1} u_{m-1}(\mathbf{x}\cdot\boldsymbol{\xi}) u_{m-1}(\mathbf{y}\cdot\boldsymbol{\xi})\, d\boldsymbol{\xi}.$$

The functions $D_n(\mathbf{x}, \mathbf{y})$ are the corresponding *Chebyshev – Dirichlet kernels* .

The kernels $D_n(\mathbf{x}, \mathbf{y})$ are complicated, with complete absence of localization in the usual sense. They are *strongly oscillatory*, with large amplitudes of oscillations. It seems interesting to investigate these kernels more closely, as well as possibilities of other *summation methods* of the series (11). In particular, such investigation may be worthy for understanding of ridge approximation in L^p-metrics for $p \neq 2$.

7. Discretization.

For a fixed natural n, consider a set of n points $\boldsymbol{\Xi}_n = \boldsymbol{\Xi}_n(\varphi_n) = \{\boldsymbol{\xi}_k^n\}_{k=1}^n$ equidistributed on a semicircle:

$$\boldsymbol{\Xi}_n := \{\boldsymbol{\xi} = \mathbf{e}_\vartheta\}_{\vartheta \in \Theta_n}, \quad \text{where } \mathbf{e}_\vartheta = \langle \cos\vartheta, \sin\vartheta \rangle, \quad \Theta_n := \left\{ \frac{k\pi}{n} + \varphi_n \right\}_{k=1}^n ,$$

where φ_n is an arbitrary fixed real number. Then using the relations

$$\frac{1}{n} \sum_{k=1}^n e^{\frac{2\pi ijk}{n}} = \begin{cases} 1 & \text{if } \; j \equiv 0 \pmod{n}, \\ 0 & \text{if } \; j \not\equiv 0 \pmod{n} \end{cases} \tag{13}$$

it is easy to see that

$$\int_{\mathcal{S}^1} a(\boldsymbol{\xi})\, d\boldsymbol{\xi} = \frac{2\pi}{n} \sum_{\boldsymbol{\xi} \in \boldsymbol{\Xi}_n} a(\boldsymbol{\xi}), \quad \forall a(\boldsymbol{\xi}) \in \mathcal{T}_{2(n-1)}^{\pm}. \tag{14}$$

Next note that for fixed $\mathbf{x}$, $\mathbf{y}$, the product $u_{n-1}(\mathbf{x}\cdot\boldsymbol{\xi})u_{n-1}(\mathbf{y}\cdot\boldsymbol{\xi})$, as a function of $\boldsymbol{\xi} \in \mathcal{S}^1$, is a trigonometric polynomial of the class $\mathcal{T}_{2(n-1)}^{\pm}$. Therefore,

$$\int_{\mathcal{S}^1} u_{n-1}(\mathbf{x}\cdot\boldsymbol{\xi})u_{n-1}(\mathbf{y}\cdot\boldsymbol{\xi})\, d\boldsymbol{\xi} = \frac{2\pi}{n} \sum_{\boldsymbol{\xi} \in \boldsymbol{\Xi}_n} u_{n-1}(\mathbf{x}\cdot\boldsymbol{\xi})u_{n-1}(\mathbf{y}\cdot\boldsymbol{\xi}).$$

Cosequently, the integral Chebyshev ridge polynomial Fourier series (11) can be rewritten in discrete form as follows:

$$f(\mathbf{x}) \overset{L^2(\mathbb{B}^2)}{=} \sum_{n=1}^{\infty} \sum_{\boldsymbol{\xi} \in \boldsymbol{\Xi}_n} a_{n-1}(f, \boldsymbol{\xi}) u_{n-1}(\mathbf{x}\cdot\boldsymbol{\xi}), \quad a_n(f, \boldsymbol{\xi}) = \int_{\mathbb{B}^2} f(\mathbf{y}) u_n(\mathbf{y}\cdot\boldsymbol{\xi})\, d\mathbf{y}. \tag{15}$$

It follows that *for an arbitrary choice of φ_n in the definition of $\boldsymbol{\Xi}_n(\varphi_n)$ the corresponding, double indexed, discrete set of Chebyshev ridge polynomials*

$$\mathcal{US}^1(\Phi) := \left\{ \{u_{n-1}(\mathbf{x}\cdot\boldsymbol{\xi})\}_{\boldsymbol{\xi} \in \boldsymbol{\Xi}_n} \right\}_{n=1}^{\infty}, \quad (\Phi := \{\varphi_n\}_1^{\infty})$$

is a complete orthonormal system in $L^2(\mathbb{B}^2)$. The Parceval's identity answering such a system is given by

$$\int_{\mathbb{B}^2} f(\mathbf{x})g(\mathbf{x})\,d\mathbf{x} = \sum_{n=1}^{\infty}\sum_{\boldsymbol{\xi}\in\boldsymbol{\Xi}_n} a_{n-1}(f,\,\boldsymbol{\xi})a_{n-1}(g,\,\boldsymbol{\xi}), \qquad \left\|f,\ L^2(\mathbb{B}^2)\right\|^2 = \sum_{n=1}^{\infty}\sum_{\boldsymbol{\xi}\in\boldsymbol{\Xi}_n} |a_{n-1}(f,\,\boldsymbol{\xi})|^2.$$

However note, that such classical aspects of Fourier analysis as proper analogues of Riemann–Lebesgue theorem for systems $\mathcal{US}^1(\Phi)$ and functions $f(\mathbf{x}) \in L^p(\mathbb{B}^2)$ with $p < 2$ are by far not clarified yet.

The following discrete representation of the Dirichlet kernel $D_n(\mathbf{x},\mathbf{y})$ is also an easy corollary of (13):

$$D_n(\mathbf{x},\mathbf{y}) = \sum_{m=1}^{n}\sum_{k=1}^{n} \frac{m}{n}\, u_{m-1}(\mathbf{x}\cdot\boldsymbol{\xi}_n^k)u_{m-1}(\mathbf{y}\cdot\boldsymbol{\xi}_n^k)\,, \quad \boldsymbol{\xi}_n^k := \left\langle \cos\frac{k\pi}{n}, \sin\frac{k\pi}{n}\right\rangle .$$

This relation implies, in particular, that a general polynomial of two variables can be represented as linear combination of ridge polynomials of same degree. Indeed, if $P(\mathbf{x}) \in \mathcal{P}^2_{n-1}$, then

$$P(\mathbf{x}) = \sum_{k=1}^{n} P_k(\mathbf{x}\cdot\boldsymbol{\xi}_n^k), \quad P_k(t) = \sum_{m=1}^{n}\frac{m}{n}\left(\int_{\mathbb{B}^2} P(\mathbf{y})u_{m-1}(\mathbf{y}\cdot\boldsymbol{\xi}_n^k)\,d\mathbf{y}\right)u_{m-1}(t)\,, \tag{16}$$

and obviously $P_k(t) \in \mathcal{P}^1_{n-1}$.

However, for our direct goal - the proof of Theorem 1 - we will use the integral form (11) of Chebyshev ridge polynomial Fourier series.

8. A relation between Chebyshev and Legendre polynomials.

Let $\mathcal{L} := \{l_n(t)\}_{n=0}^{\infty}$ denote the system of Legendre polynomials orthonormal in $L^2(0,1)$. Then following relations are true:

$$\frac{1}{2\pi}\int_{\mathcal{S}^1} u_n(\mathbf{x}\cdot\boldsymbol{\xi})\,d\boldsymbol{\xi} = \sqrt{\frac{1}{\pi(n+1)}}\begin{cases} l_{\frac{n}{2}}(|\mathbf{x}|^2) & \text{for even } n \\ 0 & \text{for odd } n. \end{cases} \tag{17}$$

Indeed, Chebyshev polynomials $u_n(t)$ with odd indices n are odd functions, so the integral on the left is obviously 0. On the other hand, if n is even, say $n = 2m$, it is easy to see that the integral is a polynomial in $|\mathbf{x}|^2$ of the form $P(|\mathbf{x}|^2)$, where $P \in \mathcal{P}^1_m$. Due to the orthogonality relation (8), we also have $P(|\mathbf{x}|^2) \perp \mathcal{P}^2_{2m-1}$ in $L^2(\mathbb{B}^2)$, and in particular $P(|\mathbf{x}|^2) \perp \forall Q(|\mathbf{x}|^2)$ where $Q \in \mathcal{P}^1_{m-1}$. In polar coordinates,

$$0 = \int_{\mathbb{B}^2} P(|\mathbf{x}|^2)Q(|\mathbf{x}|^2)\,d\mathbf{x} = 2\pi\int_0^1 rP(r^2)Q(r^2)\,dr = \pi\int_0^1 P(t)Q(t)\,dt, \quad \forall Q \in \mathcal{P}^1_{m-1},$$

and consequently $P(t)$ is indeed a constant multiple of the m-th Legendre polynomial, i.e. $P(t) = \kappa_m l_m(t)$. A calculation of these constants is based on (9) and can be carried out for $n = 2m$ as follows:

$$\kappa_m^2 = \int_0^1 P^2(t)\,dt = 2\int_0^1 rP(r^2)\,dr = \frac{1}{\pi}\int_{\mathbb{B}^2} P^2(|\mathbf{x}|^2)\,d\mathbf{x} = \frac{1}{4\pi^3}\int_{\mathbb{B}^2}\left(\int_{\mathcal{S}^1} u_n(\mathbf{x}\cdot\boldsymbol{\xi})\,d\boldsymbol{\xi}\right)^2 d\mathbf{x}$$
$$= \frac{1}{4\pi^3}\int_{\mathcal{S}^1}\int_{\mathcal{S}^1}\left(\int_{\mathbb{B}^2} u_n(\mathbf{x}\cdot\boldsymbol{\xi})\,u_n(\mathbf{x}\cdot\boldsymbol{\eta})\,d\mathbf{x}\right) d\boldsymbol{\xi}d\boldsymbol{\eta} = \frac{1}{4\pi^3}\int_{\mathcal{S}^1}\int_{\mathcal{S}^1}\frac{u_n(\boldsymbol{\xi}\cdot\boldsymbol{\eta})}{u_n(1)}\,d\boldsymbol{\xi}d\boldsymbol{\eta}$$
$$= \frac{\sqrt{\pi}}{4\pi^3(n+1)}\int_{\mathcal{S}^1}\int_{\mathcal{S}^1} u_n(\boldsymbol{\xi}\cdot\boldsymbol{\eta})\,d\boldsymbol{\xi}d\boldsymbol{\eta} = \frac{1}{4\pi^3(n+1)}\int_0^{2\pi}\int_0^{2\pi} D_n(\vartheta-\varphi)\,d\vartheta d\varphi = \frac{1}{\pi(n+1)},$$

whence (17) follows.

9. Chebyshev ridge polynomial Fourier series of radial functions.

Let $f(\mathbf{x})$ is a radial function, $f(\mathbf{x}) = f(|\mathbf{x}|)$. Then it is easy to see that the corresponding Chebyshev ridge Fourier coefficients $a_n(f,\boldsymbol{\xi})$ in (11) are trigonometric polynomials of degree 0, i.e. simply constants: $a_n(f,\boldsymbol{\xi}) = a_n(f)$. Moreover, for odd n one has $a_n(f) = 0$, and thus

$$f(|\mathbf{x}|) \overset{L^2(\mathbb{B}^2)}{=} \frac{1}{2\pi}\sum_{m=0}^{\infty}(2m+1)a_{2m}(f)\int_{\mathcal{S}^1} u_{2m}(\mathbf{x}\cdot\boldsymbol{\xi})\,d\boldsymbol{\xi}. \tag{18}$$

Using (17) one can express $a_{2m}(f)$ via Fourier–Legendre coefficients of the function $f(\sqrt{t}),\ t\in(0,1)$:

$$a_{2m}(f) \equiv \frac{1}{2\pi}\int_{\mathcal{S}^1} a_{2m}(f)\,d\boldsymbol{\xi} = \int_{\mathbb{B}^2} f(|\mathbf{x}|)\left(\frac{1}{2\pi}\int_{\mathcal{S}^1} u_n(\mathbf{x}\cdot\boldsymbol{\xi})\,d\boldsymbol{\xi}\right) d\mathbf{x} \tag{19}$$
$$= \int_{\mathbb{B}^2} f(|\mathbf{x}|)\sqrt{\frac{1}{\pi(2m+1)}}\,l_m(|\mathbf{x}|^2)\,d\mathbf{x} = 2\pi\sqrt{\frac{1}{\pi(2m+1)}}\int_0^1 rf(r)\,l_m(r^2)\,dr$$
$$= \sqrt{\frac{\pi}{2m+1}}\int_0^1 f(\sqrt{t})\,l_m(t)\,dt.$$

Respectively, the partial sums $S_n(f,\mathbf{x})$ for radial functions $f(\mathbf{x}) = f(|\mathbf{x}|)$ can be rewritten as follows:

$$S_n\left(f,\sqrt{|\mathbf{x}|}\right) = \pi\sum_{2m\le n-1}\left(\int_0^1 f(\sqrt{t})\,l_m(t)\,dt\right) l_m(|\mathbf{x}|).$$

10. Chebyshev Fourier series of ridge functions.

Let $\boldsymbol{\eta}\in\mathcal{S}^1$ be a fixed wave vector, and $F(t)\in L^2_w(\mathbb{B}^1),\ w(t) = 2\sqrt{1-t^2}$ - a single variate function, with Chebyshev–Fourier expansion

$$F(t) \overset{L^2_w(\mathbb{B}^1)}{=} \sum_{m=0}^{\infty} \hat{F}(m)\, u_m(t)\,, \quad \text{where} \quad \hat{F}(m) = 2\int_{-1}^{1} F(t) u_m(t)\sqrt{1-t^2}\,dt, \tag{20}$$

and obviously

$$F(\mathbf{x}\cdot\boldsymbol{\eta}) \overset{L^2(\mathbb{B}^2)}{=} \sum_{m=0}^{\infty} \hat{F}(m)\, u_m(\mathbf{x}\cdot\boldsymbol{\eta})\,.$$

Therefore (cf. (10)),

$$a_m(F(\mathbf{x}\cdot\boldsymbol{\eta}),\boldsymbol{\xi}) = \hat{F}(m)\int_{\mathcal{S}^1} u_m(\mathbf{y}\cdot\boldsymbol{\eta})u_m(\mathbf{y}\cdot\boldsymbol{\xi})\,d\mathbf{y} = \hat{F}(m)\,\frac{u_m(\boldsymbol{\eta}\cdot\boldsymbol{\xi})}{u_m(1)} = \frac{\sqrt{\pi}\hat{F}(m)}{m+1}\,u_m(\boldsymbol{\eta}\cdot\boldsymbol{\xi}),$$

and if $R(\mathbf{x}) = \sum_1^n F_j(\mathbf{x}\cdot\boldsymbol{\xi}_j)$ is a function of the class $\mathcal{R}_n$, then

$$a_m(R,\boldsymbol{\xi}) = \frac{\sqrt{\pi}}{m+1}\sum_{j=1}^{n} \hat{F}_j(m)u_m(\boldsymbol{\eta}\cdot\boldsymbol{\xi}_j)\,, \quad m = 0,\, 1,\, \ldots\,. \tag{21}$$

3 $\mathcal{NRA}$ and optimal quadrature formulas for trigonometric polynomials

Denote $\mathcal{T}_m^{\pm}(L^2)$ the unit $L^2(\mathcal{S}^1)$-ball in the subspace of trigonometric polynomials $\mathcal{T}_m^{\pm}$:

$$\mathcal{T}_m^{\pm}(L^2) := \left\{ t(\boldsymbol{\xi}) \in \mathcal{T}_m^{\pm} : \quad \|t(\boldsymbol{\xi}),\ L^2(\mathcal{S}^1)\| \le 1 \right\},$$

and let $f(\mathbf{x}) \in L^2(\mathbb{B}^2)$. By duality, the Parceval's identity (12) can be rewritten as follows:

$$\left\| f(\mathbf{x}),\ L^2(\mathbb{B}^2) \right\|^2 = \sum_{m=0}^{\infty} (m+1) \sup_{t\in\mathcal{T}_m^{\pm}(L^2)} \left| \int_{\mathcal{S}^1} a_m(f,\boldsymbol{\xi}) t(\boldsymbol{\xi})\,d\boldsymbol{\xi} \right|^2. \tag{22}$$

The next simple statement contains a duality relation between ridge approximation and *errors of quadrature formulas* for computation of linear functionals $\int_{\mathcal{S}^1} a_m(f,\,\boldsymbol{\xi})d\boldsymbol{\xi}$ on $\mathcal{T}_m^{\pm}(L^2)$. The latter formulas correspond to the *nodes* $\{\boldsymbol{\xi}_j\}$ and *weights* $\hat{F}_j(m)$, generated by the given linear combination of ridge functions $R(\mathbf{x}) = \sum_1^n F_j(\mathbf{x}\cdot\boldsymbol{\xi}_j) \in \mathcal{R}_n$.

Lemma 1 *Let* $R(\mathbf{x}) = \sum_1^n F_j(\mathbf{x}\cdot\boldsymbol{\xi}_j)$¿ *Then*

$$\left\| f(\mathbf{x}) - R(\mathbf{x});\ L^2(\mathbb{B}^2)\right\|^2$$
$$= \frac{1}{2\pi}\sum_{m=0}^{\infty}(m+1)\sup_{t\in\mathcal{T}_m^{\pm}(L^2)}\left|\int_{\mathcal{S}^1} a_m(f,\boldsymbol{\xi})t(\boldsymbol{\xi})\,d\boldsymbol{\xi} - \frac{2\pi}{m+1}\sum_{j=1}^{n}\hat{F}_j(m)\,t(\boldsymbol{\xi}_j)\right|^2, \qquad (23)$$

where $\hat{F}_j(m)$ *denote the Fourier – Chebyshev coefficients of the function* $F_j(t)$*:*

$$\hat{F}_j(m) = 2\int_{-1}^{1} F_j(t)u_m(t)\sqrt{1-t^2}\,dt.$$

This statement is a corollary of (22), (21) and (10), because for $t(\boldsymbol{\xi}) \in \mathcal{T}_m^{\pm}(L^2)$

$$\int_{\mathcal{S}^1} a_m(R,\boldsymbol{\xi})t(\boldsymbol{\xi})\,d\boldsymbol{\xi} = \frac{1}{m+1}\sum_{j=1}^{n}\hat{F}_j(m)\int_{\mathcal{S}^1}\sqrt{\pi}\,u_m(\boldsymbol{\xi}_j\cdot\boldsymbol{\xi})t(\boldsymbol{\xi})\,d\boldsymbol{\xi} = \frac{2\pi}{m+1}\sum_{j=1}^{n}\hat{F}_j(m)\,t(\boldsymbol{\xi}_j)\,.$$

Note the for each $m = 0, 1, \dots,$ the expression

$$Q(m,n,\hat{R})(t) := \frac{2\pi}{m+1}\sum_{j=1}^{n}\hat{F}_j(m)\,t(\boldsymbol{\xi}_j)$$

on the right of (23) can be interpreted as a *quadrature formula* with n nodes $\{\boldsymbol{\xi}_j\}_1^n$ and weights

$$w_j := w_j(m,\hat{R}) := \frac{2\pi}{m+1}\hat{F}_j(m),\ j = 1, 2, \dots, n$$

for computation of the linear functional

$$A_m(f,t) := \int_{\mathcal{S}^1} a_m(f,\boldsymbol{\xi})t(\boldsymbol{\xi})\,d\boldsymbol{\xi} \sim Q(m,n,\hat{R})(t) := \sum_{j=1}^{n} w_j t(\boldsymbol{\xi}_j)$$

on the class of trigonometric polynomials $t \in \mathcal{T}_m^{\pm}(L^2)$. Moreover, the upper bound

$$\sup_{t\in\mathcal{T}_m^{\pm}(L^2)}\left|\int_{\mathcal{S}^1} a_m(f,\boldsymbol{\xi})t(\boldsymbol{\xi})\,d\boldsymbol{\xi} - \frac{2\pi}{m+1}\sum_{j=1}^{n}\hat{F}_j(m)\,t(\boldsymbol{\xi}_j)\right| = \sup_{t\in\mathcal{T}_m^{\pm}(L^2)}\left|A_m(f,t) - Q(m,n,\hat{R})\right|$$

on the right of (23) represents the value of the global *error* of the quadrature formula $Q(m,n,\hat{R})$ on this class.

At this point, it is convenient to introduce the following variant of the general notion of *optimal quadrature formulas*, due to A.N. Kolmogorov and S.M. Nikol'skii, cf. [6], adjusted to the special case of compact classes of trigonometric polynomials.

Definition. Let $m,\ n = 1, 2, \dots$. Denote $\mathcal{T}_m$ the subspace of trigonometric polynomials of degree m, and let for $1 \le p \le \infty$ $\mathcal{T}_m(L^p)$ be the L^p-unit ball in $\mathcal{T}_m$:

$$\mathcal{T}_m(L^p) := \text{Span } \left\{ e^{ik\vartheta} \right\}_{k=-n}^{n} , \quad \mathcal{T}_m(L^2) := \{ T(\vartheta) \in \mathcal{T}_m \ : \| T(\vartheta),\ L^p(0, 2\pi) \| \le 1 \} \ .$$

Then the quantity

$$\mathcal{Q}(m,n)(L^p) := \inf_{\{w_j\}_1^n, \{\vartheta_j\}_1^n} \left\{ \sup_{T \in \mathcal{T}_m(L^p)} \left| \int_0^{2\pi} T(\vartheta)\, d\vartheta - \sum_{j=1}^{n} w_j t(\boldsymbol{\xi}_j) \right| \right\}$$

is called *optimal quadrature error* with n nodes for the class $\mathcal{T}_m(L^p)$. If inf in this definition is attained for certain sets of nodes and weights $\Theta_n = \{\vartheta_j^*\}_1^n$, $W_n = \{w_j^*\}_1^n$, the corresponding quadrature formula $\sum_1^n w_j^* T(\vartheta_j^*)$ is called *optimal* for the class $\mathcal{T}_m(L^p)$.

This definition, (23) and (18) imply a *lower* eslimate for $\mathcal{NRA}$ of each fixed radial function.

Lemma 2 *Let $f(\mathbf{x}) = f(|\mathbf{x}|)$ be a radial function in $L^2(\mathbb{B}^2)$. Then the following estimates hold true:*

$$\mathcal{NRA}_n(f) \ge \sqrt{\frac{1}{2\pi} \sum_{m=0}^{\infty} (2m+1) |a_{2m}|^2 \left(\mathcal{Q}(m,n)(L^2) \right)^2} , \quad n = 1, 2, \dots , \tag{24}$$

Indeed, as mentioned above, in the case of radial functions, the corresponding Chebyshev – Fourier coefficients $a_m(f, \boldsymbol{\xi})$ are in fact constants, and the latter are non-zero only for even indices m. Thus, (23) for a radial function f can be rewritten as follows:

$$\left\| f(|\mathbf{x}|) - \sum_{j=1}^{n} F_j(\mathbf{x} \cdot \boldsymbol{\xi}_j);\ L^2(\mathbb{B}^2) \right\|^2$$

$$\ge \frac{1}{2\pi} \sum_{m=0}^{\infty} (2m+1) \sup_{t \in \mathcal{T}_{2m}^{\pm}(L^2)} \left| a_{2m}(f) \int_0^{2\pi} t(\vartheta)\, d\vartheta - \frac{2\pi}{2m+1} \sum_{j=0}^{N} \hat{F}_j(2m)\, t(\vartheta_j) \right|^2 . \tag{25}$$

A polynomial $t(\vartheta) \in \mathcal{T}_{2m}^{\pm}$ is of the form $t(\vartheta) = T(2\vartheta)$, where $T(\vartheta) \in \mathcal{T}_m$. Thus, the following estimate from below holds true for each term of the series on the righthand side:

$$\sup_{t \in \mathcal{T}_{2m}^{\pm}(L^2)} \left| a_{2m}(f) \int_0^{2\pi} t(\vartheta)\, d\vartheta - \frac{2\pi}{2m+1} \sum_{j=1}^{n} \hat{F}_j(2m)\, t(\vartheta_j) \right|^2 \ge |a_{2m}(f)|^2 \left(\mathcal{Q}(m,n)(L^2) \right)^2 ,$$

and (24) follows.

The following inequalities are obvious:

$$\begin{aligned} \mathcal{Q}(m,n)(L^p) &\leq \mathcal{Q}(m,n)(L^r),\ (p \geq q); \quad \mathcal{Q}(m,n)(L^p) \leq \mathcal{Q}(m_1,n)(L^p),\ (m_1 \geq m); \\ \mathcal{Q}(m,n)(L^p) &\leq \mathcal{Q}(m,n_1)(L^p),\ (n_1 \leq n). \end{aligned}$$

Since (cf. (13))

$$\int_0^{2\pi} T(\vartheta)\, d\vartheta = \frac{2\pi}{n} \sum_1^n T\left(\frac{2\pi j}{n}\right), \quad \forall T \in \mathcal{T}_m,\ n > m$$

(i.e., quadrature formula of rectangles is *exact*), one obviously has $\mathcal{Q}(m,n)(L^p) = 0$ if $n > m$.

To finish the proof of Theorem 1, now we need explicit estimates of $\mathcal{Q}(m,n)(L^2)$ from below.

These estimates are discussed in V.N. Temlyakov's monograph [7]. For our goal, the following particular case of Lemma 5.1, p. 125, and also Theorem 1.3, p. 31, from [7] is crucial. Part 1) of this assertion is a result of B.S.Kashin [8] (for a stronger result, involving the norm U of uniform convergence in $\mathcal{T}_m$, see also [9].)

Lemma 3 *Let $\varepsilon > 0$ be a fixed number. Then there exists a constant $C(\varepsilon) > 0$ such that:*
1) *in every subspace $\Psi \subset \mathcal{T}_m$ of dimension* $\dim \Psi \geq \varepsilon(2m+1)$ *there exists a polynomial $T \in \Psi$ with equivalent norms in all L^p, $1 \leq p \leq \infty$:*

$$\|t,\, L^1(0,2\pi)\| \geq C(\varepsilon)\, \|t,\, L^\infty(0,2\pi)\| > 0; \tag{26}$$

2) *if $n \leq (1-\varepsilon)m$, then the following estimates of $\mathcal{Q}(m,n)$ from below hold true*

$$\mathcal{Q}(n,m)(L^\infty) \geq C(\varepsilon). \tag{27}$$

For completeness sake, let us reproduce a deduction of (27) from (26), see [7], Lemma 5.1, p. 125. Given a set of n nodes $\Theta = \{\vartheta_j\}_1^n$, denote $\Psi := \Psi(\Theta)$ the subspace of all polynomials $T(\vartheta) \in \mathcal{T}_{\frac{m}{2}}$ which vanish at all nodal points, i.e. $T(\vartheta) = 0$, $\forall \vartheta \in \Theta$ (if nodes are multiple – all corresponding derivatives must also vanish). Clearly we have $\dim \Psi \geq 2\left[\frac{m}{2}\right] + 1 - n \geq m - n - 1 \geq \varepsilon m - 1 \geq \frac{\varepsilon}{4}(2m+1)$ for all sufficiently large m. Thus, according to (26), there exists a polynomial $t(\vartheta) \in \Psi$ such that $\|t,\, L^\infty(0,2\pi)\| = 1$ and $\|t,\, L^2(0,2\pi)\|^2 \geq C'(\varepsilon)$. Take $T(\vartheta) := |t^(\vartheta)|^2$. Then obviously $T \in \mathcal{T}_m$, $\|T,\, L^\infty(0,2\pi)\| = 1$, $T(\vartheta) = 0$, $\forall \vartheta \in \Theta$, so that every quadrature formula with the nodal points Θ provides zero result for this polynomial. On the other hand, one has $\int_0^{2\pi} T(\vartheta)\, d\vartheta = \|t,\, L^2(0,2\pi)\|^2 \geq C'(\varepsilon)$, which completes the deduction of (27).

It follows from (24) and (27) that for each $\varepsilon > 0$ there is a constant $C''(\varepsilon) > 0$ such that for all radial functions f

$$\mathcal{NRA}_n(f) \geq C''(\varepsilon)\sqrt{\sum_{2m>2(1+\varepsilon)n}^{\infty}(2m+1)|a_{2m}|^2} = C''(\varepsilon)\,\mathcal{PA}_{2(1+\varepsilon)n}\,, \quad n = 1,\, 2,\, \ldots . \tag{28}$$

Theorem 1 is a corollary of this relation, corresponding to $\varepsilon := \frac{1}{2}$.

4 Comments and open problems.

1. An interesting open problem is to elaborate approach to $\mathcal{NRA}$ in metrics of L^p for $p \neq 2$, in particular, for $p = \infty$. Here, it is natural to expect that an analogue of Theorem 1 is true in all L^p. Also, $\mathcal{NRA}$ of functions of higher number of variables $f(\mathbf{x}) = f(x_1,\, x_2,\, \ldots,\, x_d)$ is the field of big theoretical and applied interest.

2. Surprisingly little is known concerning classical problem of optimal quadrature formulas for classes of trigonometric polynomials $\mathcal{T}_n(L^p)$ in case of *deficiency* of nodes, i.e. for $n \leq m$. It seems to be interesting to find a direct and simpler proof of the lower estimate (27) of $\mathcal{Q}(m,n)$, avoiding reference to the very deep result (26) of Kashin.

Even the problem of *existence* of optimal quadrature formula seems to be open. Here, the main source of difficulty is in *non-compactness* of the class of admissible weights $\{w_j\}_1^n$. Thus, a priori it may be more profitable to measure at certain nodal points ϑ_j not only the point values of the polynomials $T(\vartheta_j)$, but also those of their derivatives $T^{(k)}(\vartheta_j)$ up to a certain order. The nature of this difficulty is quite analogous to that of existence of the element of best non-linear ridge approximation, discussed in the Introduction.

It is not hard to see that

$$\mathcal{Q}(m,n)(L^2) = \inf_{\{w_j\}_1^n,\,\{\vartheta_j\}_1^n}\left\|1 - \sum_{j=1}^{n}\frac{w_j}{\pi}\,\mathcal{D}_m(\vartheta - \vartheta_j);\ L^2(0,2\pi)\right\|, \quad \mathcal{D}_m(\vartheta) := \frac{\sin\left(m+\frac{1}{2}\right)\vartheta}{2\sin\frac{\vartheta}{2}}\,. \tag{29}$$

Moreover, for a fixed set $\Theta = \{\vartheta_j\}_1^n$ of pairwise distinct nodes, the corresponding optimal weights $W = W(\Theta) = \{w_j\}_1^n$ can be selected as a solution of the following system of linear equations:

$$\sum_{j=1}^{n}\frac{w_j}{\pi}\,\mathcal{D}_m(\vartheta_k - \vartheta_j) = 1, \quad k = 1,\, 2,\, \ldots\, n. \tag{30}$$

If $2m+1 \geq n$, the system of shifted Dirichlet kernels $\left\{\frac{1}{\pi}\mathcal{D}_m(\vartheta - \vartheta_j)\right\}_1^n$ is linearly independent. Thus the Gram matrix $\left[\frac{1}{\pi}\mathcal{D}_m(\vartheta_k - \vartheta_j)\right]_{k,j=1}^n$ is nondegenerate, the solution $W = W(\Theta)$ of (30) is unique, and the error of the corresponding formula with optimally chosen weights $= \sqrt{2\pi - \sum_1^n w_j}$.

3. Let us apply (29), (30) to the analysis of the simplest non-trivial case.

Lemma 4 *The optimal quadrature formula with two nodes $(\vartheta_1, \vartheta_2)$ for the class of trigonometric polynomials of second order $\mathcal{T}_2(L^2)$ exists. The optimal nodes satisfy the relation*

$$\vartheta_2 - \vartheta_1 = \pi - \arccos\frac{1}{4} < \pi, \tag{31}$$

i.e., they are not equidistant on the period $[0, 2\pi)$. One has $\mathcal{Q}(2,2)(L^2) = \sqrt{\frac{14\pi}{15}}$ and the optimal weights are given by $w_1 = w_2 = \frac{8\pi}{15}$.

Without loss of generality, we may assume that $\vartheta_1 = 0$. Let $\vartheta_2 := \vartheta \neq 0$, and consider the system (30), answering the case $n = m = p = 2$:

$$\begin{cases} w_1 D(0) + w_2 D(\vartheta) = \pi \\ w_1 D(\vartheta) + w_2 D(0) = \pi \end{cases} \quad \text{where} \quad \mathcal{D}(\vartheta) := \mathcal{D}_2(\vartheta) = \frac{\sin\frac{5}{2}\vartheta}{2\sin\frac{\vartheta}{2}}$$

Obviously, the weights are equal $w_1 = w_2 = w(\vartheta) = \frac{\pi}{\mathcal{D}(0)+\mathcal{D}(\vartheta)}$, and further, choosing ϑ optimally, one has

$$\mathcal{Q}^2(2,2)(L^2) = \min_{\vartheta}(2\pi - 2w(\vartheta)) = 2\pi\left(1 - \max_{\vartheta}\frac{1}{\mathcal{D}(0)+\mathcal{D}(\vartheta)}\right).$$

Thus, we need to find the minimizer ϑ in $\min_\vartheta \mathcal{D}(\vartheta)$. It is not hard to see that such ϑ can be found as the smallest positive solution of the equation $\mathcal{D}'(\vartheta) = -\sin\vartheta - 2\sin 2\vartheta = -\sin\vartheta(1 + 4\cos\vartheta) = 0$. Thus,

$$\vartheta = \pi - \arccos\frac{1}{4}, \quad \mathcal{D}(\vartheta) = \frac{1}{2} + \cos\vartheta + \cos 2\vartheta = -\frac{5}{8},$$
$$\mathcal{D}(0) + \mathcal{D}(\vartheta) = \frac{5}{2} - \frac{5}{8} = \frac{15}{8}, \quad w_1 = w_2 = \frac{8\pi}{15}, \quad \left(\mathcal{Q}(2,2)(L^2)\right)^2 = 2\pi - w_1 - w_2 = \frac{14\pi}{15},$$

which completes the proof.

4. Let us prove the result mentioned in the Introduction, see (3), concerning geometric peculiarity in $\mathcal{NRA}$ of the radial function $f(|\mathbf{x}|) = |\mathbf{x}|^4 - |\mathbf{x}|^2$.

Lemma 5 *The exact solution of the $\mathcal{NRA}_2$-problem*

$$\mathcal{NRA}_2(f) = \inf_{\boldsymbol{\xi}_0,\boldsymbol{\xi}_1;F_1,F_2} \left\| |\mathbf{x}|^4 - |\mathbf{x}|^2 + \frac{1}{6} - F_1(\mathbf{x}\cdot\boldsymbol{\xi}_1) - F_2(\mathbf{x}\cdot\boldsymbol{\xi}_2),\ L^2(\mathbb{B}^2)\right\|$$

is attained for $\boldsymbol{\xi}_0, \boldsymbol{\xi}_1$ satisfying

$$\boldsymbol{\xi}_1 \cdot \boldsymbol{\xi}_2 = \sqrt{\frac{3}{8}} \neq 0, \tag{32}$$

which means that the optimal wave vectors are not mutually perpendicular. The corresponding optimal functions $F_1(t)$, $F_2(t)$ are constant multiples of Chebyshev polynomial $u_4(t)$ of 4th degree, $F_1(t) = F_2(t) = \mathrm{const}\cdot u_4(t)$.

This statement is a corollary of the previous lemma and (23), (25). Indeed, we have

$$f(|\mathbf{x}|) := |\mathbf{x}|^4 - |\mathbf{x}|^2 + \frac{1}{6} = \mathrm{const}\cdot l_2(|\mathbf{x}|^2), \quad \text{where} \quad l_2(t) = 6\sqrt{5}\left(t^2 - t + \frac{1}{6}\right)$$

is Legendre polynomial of order 2 on $(0,1)$. By (17),

$$f(|\mathbf{x}|) = \mathrm{const}\cdot \int_{\mathcal{S}^1} u_4(\mathbf{x}\cdot\boldsymbol{\xi})\, d\boldsymbol{\xi},$$

so that the representation (18) reduces to *a single* non-zero term answering $m = 2$. The $\mathcal{NRA}_2$-problem for such function is reduced to minimization

$$\min_{\boldsymbol{\xi}_1,\boldsymbol{\xi}_2;w_1,w_2} \left\| \int_{\mathcal{S}^1} u_4(\mathbf{x}\cdot\boldsymbol{\xi})\, d\boldsymbol{\xi} - w_1 u_4(\mathbf{x}\cdot\boldsymbol{\xi}_1) - w_2 u_4(\mathbf{x}\cdot\boldsymbol{\xi}_2),\ L^2(\mathbb{B}^2)\right\|$$

and, further, according to (23) – to search of *a single* optimal quadrature formula with 2 nodes for trigonometric polynomials of the class $\mathcal{T}_4^{\pm}(L^2)$. The latter in its turn is equivalent to $\mathcal{Q}(2,2)(L^2)$-problem for trigonometric polynomials of 2nd degree, considered in Lemma 4. Finally, it is easy to see that the angle $\alpha = \arccos(\boldsymbol{\xi}_1\cdot\boldsymbol{\xi}_2)$ between optimal wave vectors for the $\mathcal{NRA}_2(f)$-problem is determined by

$$\alpha = \frac{\vartheta_2 - \vartheta_1}{2} = \frac{1}{2}\left(\pi - \arccos\frac{1}{4}\right),$$

whence (32) follows.

5. Let us note that, due to the general duality relation (23), ridge approximation of functions, other than radial, requires a progress in the problem of *optimal recovery of linear functionals* for trigonometric polynomials, of the general form

$$\mathcal{Q}(m,n,A,\ L^2) := \inf_{\{w_j\}_1^n,\{\vartheta_j\}_1^n} \left\{ \sup_{T\in\mathcal{T}_m(L^2)} \left| \int_0^{2\pi} A(\vartheta)T(\vartheta)\,d\vartheta - \sum_{j=1}^n w_j T(\vartheta_j) \right| \right\},$$

where $A(\vartheta)$ is a certain fixed trigonometric polynomial in $\mathcal{T}_m$. An interesting special class of such problems is *reconstruction of a harmonic* in polynomials of the class $\mathcal{T}_n(L^2)$:

$$\mathcal{Q}^{(l)}(m,n\ L^2) := \inf_{\{w_j\}_1^n,\{\vartheta_j\}_1^n} \left\{ \sup_{T\in\mathcal{T}_m(L^2)} \left| \int_0^{2\pi} e^{-ilt}T(\vartheta)\,d\vartheta - \sum_{j=1}^n w_j T(\vartheta_j) \right| \right\}, \quad |l| \le m.$$

It may be *conjectured* that, say, for $n \le \frac{m}{4}$ the quantities $\mathcal{Q}^{(l)}(m,n\ L^2)$ are bounded below by an absolute positive constant: $\mathcal{Q}^{(l)}(m,n\ L^2) \ge c_0 > 0$. The latter simply means that it is impossible, using quadrature formulas, to reconstruct the lth harmonic of all polynomials from $\mathcal{T}_n(L^2)$ with a small error, if measurements of point values are available at "too few" nodes. In contrast to the case of $l=0$, such a generalization of (27) does not seem to be directly deductable from Kashin's result (26).

The extreme case $l=m$ corresponds to $\mathcal{NRA}$ of *harmonic functions* in the open disc $\mathbb{B}^2$:

$$\Delta f(\mathbf{x}) := \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} = 0, \quad |\mathbf{x}| < 1.$$

It is not hard to see that in this case the polynomial Fourier coefficients $a_m(f,\boldsymbol{\xi})$ in (11) are monomials, $a_m(f,\boldsymbol{\xi}) = \alpha_m(f)\,\cos m(\boldsymbol{\xi}\cdot\boldsymbol{\xi}_n)$. In polar coordinates $\boldsymbol{\xi} = \mathbf{e}_\vartheta = \langle\cos\vartheta,\sin\vartheta\rangle$ one has

$$A_m(f,\vartheta) = a_m(f,\mathbf{e}_\vartheta) = \alpha_m(f)\,\cos m(\vartheta - \vartheta_m),$$

where $\alpha_m(f)$ are determined by Fourier coefficients of the boundary value $f(\boldsymbol{\xi})$, $\boldsymbol{\xi}\in\mathcal{S}^1$ of $f(\mathbf{x})$:

$$\alpha_n(f) = \frac{\sqrt{\pi}\,\rho_n(f)}{n+1}, \quad \text{where} \quad f(\mathbf{e}_\vartheta) := F(\vartheta) \sim \sum_{n=0}^{\infty} \rho_n(f)\cos n(\vartheta - \vartheta_n)\,.$$

5 Appendix. More on quadrature formulas

Let us provide an alternative proof of a variant of the estimate (27), not referring to (26):

$$\mathcal{Q}(m,n)(L^2) \ge \sqrt{\pi\left(1 - \frac{n+1}{m}\right)}, \quad n < m. \tag{33}$$

The idea is quite transparent: the sum of small number, say, $n \le (1-\varepsilon)m$ of shifted Dirichlet kernels $\mathcal{D}_m(\vartheta - \vartheta_j)$ in (29) is a "fast" oscillating function, and thus cannot approximate $f(\vartheta) := 1$ *even in measure*. We have

$$\sum_{j=1}^{n} \frac{w_j}{\pi} \mathcal{D}_m(\vartheta - \vartheta_j) = F(\vartheta)\cos m\vartheta + G(\vartheta)\sin m\vartheta = H(\vartheta)\cos(m\vartheta - \Phi(\vartheta)), \tag{34}$$

where

$$\begin{aligned} F(\vartheta) &:= -\frac{1}{2\pi}\sum_{j=1}^{n} w_j \left(\sin m\vartheta_j \cot\frac{\vartheta-\vartheta_j}{2} + \cos m\vartheta_j\right), \\ G(\vartheta) &:= \frac{1}{2\pi}\sum_{j=1}^{n} w_j \left(\cos m\vartheta_j \cot\frac{\vartheta-\vartheta_j}{2} + \sin m\vartheta_j\right), \end{aligned}$$

and $H(\vartheta) := \sqrt{F^2(\vartheta) + G^2(\vartheta)}$, $\Phi(\vartheta) := \arctan\frac{G(\vartheta)}{F(\vartheta)}$. The result will follow, if we prove that

$$\operatorname{meas}\mathcal{E}_- \ge \pi\left(1 - \frac{n+1}{m}\right), \quad \text{where} \quad \mathcal{E}_- := \{\vartheta : \ \cos(m\vartheta - \Phi(\vartheta)) \le 0, \ \vartheta \in [0, 2\pi)\}, \tag{35}$$

because

$$\int_0^{2\pi} |1 - H(\vartheta)\cos(m\vartheta - \Phi(\vartheta))|^2 \, d\vartheta \ge \int_{\mathcal{E}_-} 1 \cdot d\vartheta = \operatorname{meas}\mathcal{E}_-$$

Although the functions $F(\vartheta)$ and $G(\vartheta)$ can take on rather big values, they are *piecewise monotonic*. In the representation (34) the phase $\Phi(\vartheta)$ is bounded, $|\Phi(\vartheta)| \le \frac{\pi}{2}$, and what is essential, the total variation of this function satisfies the estimate

$$\operatorname{var}\{\Phi(\vartheta), [0, 2\pi)\} \le \pi n. \tag{36}$$

Indeed, for a fixed $t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ denote $N(t)$ the number of solutions $\vartheta \in [0, 2\pi)$ of the equation $\Phi(\vartheta) = t$. Thus $N(t)$ coincides with the number of solutions $\vartheta \in [0, 2\pi)$ of $G(\vartheta) = (\tan t)H(\vartheta)$, which equals the the number of solutions of $G(2\vartheta) = (\tan t)H(2\vartheta)$ in $\vartheta \in [0, \pi)$. Since $G(2\vartheta + \pi) \equiv G(2\vartheta)$, $H(2\vartheta + \pi) \equiv H(2\vartheta)$ we see that $2N(t) = M(t)$, where $M(t)$ is the number of solutions of the equation $G(2\vartheta) = (\tan t)H(2\vartheta)$ on $[0, 2\pi)$. After multiplication of both sides by $\omega(\vartheta) := \prod_{j=1}^{n} \sin\left(\vartheta - \frac{\vartheta_j}{2}\right)$, the latter equation transfers into $T(\vartheta) = (\tan t)S(\vartheta)$ where T, S are trigonometric polynomials of degree n. Thus, by Fundamental Theorem of Algebra, we have $M(t) \le 2n$, or $N(t) \le n$, and (36) follows.

Further, let

$$\mathcal{E}_+ := \{\vartheta : \cos(m\vartheta - \Phi(\vartheta)) > 0,\ \vartheta \in [0, 2\pi)\}, \quad \mathcal{F}_+ := \{\varphi : \varphi = m\vartheta - \Phi(\vartheta),\ \vartheta \in \mathcal{E}_+\},$$
$$\mathcal{G}_- := \{\varphi : \cos\varphi \le 0,\ \varphi \in [0, 2\pi m)\}.$$

Obviously,

$$\operatorname{meas}\mathcal{G}_- = \pi m, \quad \mathcal{G}_- \bigcap \mathcal{F}_+ = \emptyset, \quad \mathcal{G}_- \bigcup \mathcal{F}_+ \subset \left(-\frac{\pi}{2}, 2\pi m + \frac{\pi}{2}\right), \quad \operatorname{meas}\mathcal{G}_- + \operatorname{meas}\mathcal{F}_+ \le 2\pi m + \pi,$$

so that $\operatorname{meas}\mathcal{F}_+ \le \pi m + \pi$. On the other hand, $\operatorname{meas}\mathcal{F}_+ \ge m \operatorname{meas}\mathcal{E}_+ - \operatorname{var}\Phi \ge m \operatorname{meas}\mathcal{E}_+ - \pi n$, and consequently $\operatorname{meas}\mathcal{E}_+ \le \pi\left(\frac{n+1}{m} + 1\right)$. This implies (35), because by the definitions $\mathcal{E}_+ \bigcup \mathcal{E}_- = [0, 2\pi)$.

References

[1] B. Logan, L. Schepp. *Optimal reconstruction of a function from its projections.* Duke Mathematical Journal, **42**(1975), pp. 645 – 659.

[2] A.G. Ramm, A.I. Katsevich. *The Radon transform and local tomography.* 1996, CRC Press, Inc. 485 pp.

[3] R.A. DeVore, K.I. Oskolkov and P.P. Petrushev. *Approximation by feed-forward neural networks.* Annals of Numerical Mathematics, **4** (1997), pp. 261 – 287.

[4] V.E. Majorov. *On best approximation by ridge functions.* Preprint, Department of Mathematics, Technion, Haifa, Israel, 1997. 16 pp.

[5] V.N. Temlyakov. *On approximation by ridge functions.* Preprint, Department of Mathematics, University of South Carolina, 1996. 12 pp.

[6] S.M. Nikol'skii. Quadrature formulas, 2nd ed., "Nauka", Moscow, 1974 (in Russian). (Translation into English of 1st ed.: Hindustan, Delhi, 1964.)

[7] V.N. Temlyakov. *Approximation of periodic functions.* Preprint, Department of Mathematics, University of South Carolina, 1996. 272 pp. (Ch. 1, Theorem 1.3, and Ch. 2, Theorem 1.1).

[8] B.S. Kashin. *On certain properties of trigonometric polynomials with the uniform norm.* Trudy MIAN, **145** (1980), pp. 111 – 116; English transl. in: Proceedings of Steklov Math. Inst., issue **1** (1981)).

[9] B.S. Kashin. On certain properties of the space of trigonometric polynomials in connection with uniform convergence. Soobsch. AN Gruz. SSR,**93**(1979), pp. 281 – 284 (in Russian).