Industrial Mathematics Institute

1998:06

Non-linearity versus linearity in ridge approximation

K.I. Oskolkov

IMI
Preprint Series

Department of Mathematics
University of South Carolina

Non-linearity versus linearity in ridge approximation

K.I. Oskolkov

Abstract

The goal is to compare *free (non-linear)*, *equispaced* ridge and *algebraic polynomial* approximations $\mathcal{R}^{\rm fr}_N[f]$, $\mathcal{R}^{\rm eq}_N[f]$, $\mathcal{E}_N[f]$ of individual functions $f(\mathbf{x})$ in the norm of $\mathcal{L}^2(\mathbb{B}^2)$, $\mathbb{B}^2$ – the unit disc $|\mathbf{x}|\le 1$ on the plane $\mathbb{R}^2$. By definition

$$\mathcal{R}^{\rm fr}_N[f] := \inf_{R\in\mathcal{W}^{\rm fr}_N}\|f-R\|, \quad \mathcal{R}^{\rm eq}_N[f] := \min_{R\in\mathcal{W}^{\rm eq}_N}\|f-R\|, \quad \mathcal{E}_N[f] := \min_{P\in\mathcal{P}^2_{N-1}}\|f-P\|.$$

Here, $\mathcal{W}^{\rm fr}_N$ denotes the set of all N-term linear combinations of planar wave functions $R(\mathbf{x}) = \sum_1^N W_j(\mathbf{x}\cdot\boldsymbol{\theta}_j)$ of arbitrary profiles $W_j(x)$, $x\in\mathbb{R}^1$ and directions $\{\boldsymbol{\theta}_j\}_1^N$; $\mathcal{W}^{\rm eq}_N$ – the subset of $\mathcal{W}^{\rm fr}_N$ corresponding to N equispaced directions, and $\mathcal{P}^2_{N-1} := \mathrm{Span}\,\{x_1^k x_2^l\}_{k+l<N}$. One has $\mathcal{R}^{\rm fr}_N[f]\le\mathcal{R}^{\rm eq}_N[f]\le\mathcal{E}_N[f]$.

The central question is: when $\mathcal{R}^{\rm fr}_N[f] = o\left(\mathcal{R}^{\rm eq}_N[f]\right)$, $N\to\infty$, i. e. when non-linear approximation $\mathcal{R}^{\rm fr}$ is more efficient than linear $\mathcal{R}^{\rm eq}$? It is proved that this is the case for harmonic functions: $\forall\varepsilon>0\ \exists c_\varepsilon>0$ such that if $\Delta f(\mathbf{x})=0$, $|\mathbf{x}|<1$, $f\in\mathcal{L}^2(\mathbb{B}^2)$, then

$$\mathcal{R}^{\rm fr}_N[f]\le c_\varepsilon\left(\mathcal{R}^{\rm eq}_N[f]\exp(-N^\varepsilon)+\mathcal{R}^{\rm eq}_{N^{2-3\varepsilon}}[f]\right).$$

On the other hand, $\mathcal{R}^{\rm fr}_N[f]\ge\frac{1}{c}\mathcal{R}^{\rm eq}_{N^2}[f]$. Thus, for $f=f_{\rm harm}$, $\mathcal{R}^{\rm fr}_N[f]$ is "almost square times better" than $\mathcal{R}^{\rm eq}_N[f]$. However, these ultra-convergence rates are achieved at the expense of *collapse* of wave vectors.

On the contrary, non-linearity in $\mathcal{R}^{\rm fr}$ does not bring any essential gain in approximation orders, say, for all radial functions. If $f(\mathbf{x})=f(|\mathbf{x}|)$, then $\mathcal{E}_{2N}[f]\ge\mathcal{R}^{\rm eq}_N[f]\ge\sqrt{\frac{2}{3}}\mathcal{E}_{2N}(f)$ and $\mathcal{R}^{\rm fr}_N[f]\ge\sup_{\varepsilon>0}\sqrt{\frac{\varepsilon}{3(1+\varepsilon)}}\,\mathcal{R}^{\rm eq}_{(1+\varepsilon)N}[f]$.

These problems are elaborated via Fourier – Chebyshev analysis in $\mathbb{B}^2$ and arising *duality* between ridge approximation and optimization of *quadrature formulas*, in the sense of Kolmogorov – Nikol'skii [1], on classes of trigonometric polynomials.

1 Introduction

We consider here a special case of the general problem of ridge approximation. First of all, we restrict ourselves to the case of (complex valued) functions of two real variables $f(\mathbf{x}) = f(x_1, x_2)$, supported in the unit disc $\mathbb{B}^2 := \{\mathbf{x} : |\mathbf{x}| := \sqrt{x_1^2 + x_2^2} \le 1\}$ on the real Euclidean plane $\mathbb{R}^2$. Further, we assume that $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$, and focus on the approximation problem exclusively in the norm of the Hilbert space $\mathcal{L}^2(\mathbb{B}^2)$,

$$\|f(\mathbf{x}),\ \mathcal{L}^2(\mathbb{B}^2)\| := \left(\iint_{\mathbb{B}^2} |f(\mathbf{x})|^2 \, \mu(d\mathbf{x}) \right)^{\frac{1}{2}},$$

where $\mu(d\mathbf{x}) := \frac{dx_1 dx_2}{\pi}$ denotes the normalized Lebesgue measure on $\mathbb{B}^2$.

Let us introduce some other notations. $\mathbf{x} \cdot \mathbf{y}$ will denote the usual inner product of vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$; $\mathcal{S}^1$ - the unit circle $|\mathbf{x}| = 1$; $\boldsymbol{\theta} = \boldsymbol{\theta}(\vartheta) := \langle \cos\vartheta, \sin\vartheta \rangle$, $\vartheta \in [0, 2\pi)$ – polar parametrization of $\mathcal{S}^1$. Further, for $N = 1, 2, \ldots$, we will apply the vector notations $\vec{\vartheta} := \{\vartheta_j\}_1^N \in \mathbb{R}^N$ for N element sets of directional angles; $\boldsymbol{\theta}_j := \langle \cos\vartheta_j, \sin\vartheta_j \rangle$, $\vec{\boldsymbol{\theta}} = \vec{\boldsymbol{\theta}}(\vec{\vartheta}) := \{\boldsymbol{\theta}_j\}_1^N$.

Let us consider the following sets $\mathcal{W}(\vec{\vartheta})$, $\mathcal{W}_N^{\mathrm{eq}}$, $\mathcal{W}_N^{\mathrm{fr}}$ of *ridge functions* – N-terms linear combinations of *planar waves*:

$$\begin{aligned}
\mathcal{W}(\vec{\vartheta}) &:= \left\{ R(\mathbf{x}) = \sum_1^N W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j) \right\},\ \vec{\vartheta} \in \mathbb{R}^N : \quad \mathcal{W}_N^{\mathrm{fr}} := \bigcup_{\vec{\vartheta} \in \mathbb{R}^N} \mathcal{W}(\vec{\vartheta}), \\
\mathcal{W}_N^{\mathrm{eq}} &:= \left\{ R(\mathbf{x}) = \sum_1^N W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j) : \quad \vartheta_j = \frac{\pi j}{N},\ j = 0, \ldots, N-1 \right\}.
\end{aligned}$$

In the above definitions, $W_j(x)$, $x \in \mathbb{R}^1$ are single-variate functions (profiles of waves). Clearly, in the definition of $\mathcal{W}(\vec{\vartheta})$ we can confine the components ϑ_j of $\vec{\vartheta}$ to the interval $[0, \pi)$, and consider only non-degenerate $\vec{\vartheta}$, i. e. the case when ϑ_j are pairwise non-congruent $\bmod \pi$.

Thus, $\mathcal{W}(\vec{\vartheta})$, $\mathcal{W}_N^{\mathrm{fr}}$, $\mathcal{W}_N^{\mathrm{eq}}$ consist of N-term linear combinations of planar waves of arbitrary profiles; $\mathcal{W}(\vec{\vartheta})$ corresponds to a fixed set of directional angles $\vec{\vartheta} \in \mathbb{R}^N$; $\mathcal{W}_N^{\mathrm{fr}}$ – the widest collection of all functions of such type, $\mathcal{W}_N^{\mathrm{eq}}$ – the particular case of $\mathcal{W}_N^{\vec{\vartheta}}$ with N *equispaced* wave vectors.

Our goal is to study, for a fixed function $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$, the extremal problems $\mathcal{R}(\vec{\vartheta})$, $\mathcal{R}^{\mathrm{fr}}$, $\mathcal{R}^{\mathrm{eq}}$ associated with the following quantities:

$$\begin{aligned}
\mathcal{R}[f, \vec{\vartheta}] &:= \inf_{R \in \mathcal{W}(\vec{\vartheta})} \|f - R,\ \mathcal{L}^2(\mathbb{B}^2)\|,\ \vec{\vartheta} \in \mathbb{R}^N; \quad \mathcal{R}_N^{\mathrm{fr}}[f] := \inf_{R \in \mathcal{W}_N^{\mathrm{fr}}} \|f - R,\ \mathcal{L}^2(\mathbb{B}^2)\|, \\
\mathcal{R}_N^{\mathrm{eq}}[f] &:= \min_{R \in \mathcal{W}_N^{\mathrm{eq}}} \|f - R,\ \mathcal{L}^2(\mathbb{B}^2)\|.
\end{aligned}$$

Obviously, $\mathcal{R}_N^{\rm fr}[f] = \inf_{\vec{\vartheta}\in\mathrm{I\!R}^N} \mathcal{R}[f,\vec{\vartheta}] \le \mathcal{R}_N^{\rm eq}[f]$.

Of particular interest is to clarify *structural* (geometric, differential, etc.) conditions on the given $f(\mathbf{x})$ when freedom in selection of wave vectors $\{\boldsymbol{\theta}_j\}_1^N$ in the problem $\mathcal{R}^{\rm fr}$ brings an essential advantage in approximation rates over those associated with $\mathcal{R}^{\rm eq}$. Quantitatively, this advantage is expressed by order relations $\mathcal{R}_N^{\rm fr}[f] = o(\mathcal{R}_N^{\rm eq}[f])$, $N\to\infty$. A partial answer is given below in theorem 3 (see also corollary 1) regarding two important types of functions: radials and harmonics.

A very essential difference between the problems $\mathcal{R}^{\rm fr}$ and $\mathcal{R}^{\rm eq}$ cosists in *non-linearity* of $\mathcal{R}^{\rm fr}$. The latter is associated with a complete freedom in the choice of wave vectors $\{\boldsymbol{\theta}_j\}_1^N$, that are allowed to be selected optimally for a given function $f(\mathbf{x})$. On the contrary, for each fixed $\vec{\vartheta}\in\mathrm{I\!R}^N$ the problem $\mathcal{R}[f,\vec{\vartheta}]$ is linear and the solution is provided by the orthogonal projection in $\mathcal{L}^2(\mathrm{I\!B}^2)$ onto the corresponding subspace $\mathcal{L}^2(\mathrm{I\!B}^2)^{\vec{\vartheta}}$, cf. also theorem 4 below.

Further, *non-existence and non-uniqueness* of the element of best ridge approximation are quite typical for the problem $\mathcal{R}_N^{\rm fr}$, if $N\ge 2$. This can be seen from the following.

1) If $j\ge 1$, $W(x)$, $|x|\le 1$ – a smooth single variate function and $f(\mathbf{x},\vartheta) := \left(\frac{\partial}{\partial\vartheta}\right)^{j-1} W(\mathbf{x}\cdot\boldsymbol{\theta})$, then for each fixed ϑ one has $\mathcal{R}_j^{\rm fr}[f]=0$. This follows from consideration of the angular derivative as limit of divided differences.

This simple observation directs to a natural completion of $\mathcal{W}(\vec{\vartheta})$ and $\mathcal{W}_N^{\rm fr}$ by the following sets of *collapsed ridge functions*:

$$\overline{\mathcal{W}}_N(\vec{\vartheta}) := \left\{ R(\mathbf{x}) = \sum_j \left(\sum_{\nu=1}^{N_j} \left(\frac{\partial}{\partial\vartheta}\right)^{\nu-1} W_{j,\nu}(\mathbf{x}\cdot\boldsymbol{\theta}) \right)\Bigg|_{\vartheta=\vartheta_j} : \sum_j N_j = N \right\}, \quad \vec{\vartheta} = \{\vartheta_j\}$$

(obviously, the class $\overline{\mathcal{W}}_N(\vec{\vartheta})$ is non-trivial only if $\dim\vec{\vartheta}\le N$). Respectively, the extremal problem $\mathcal{R}^{\rm fr}$ can be "sliced" as follows:

$$\overline{\mathcal{R}}_N[f,\vec{\vartheta}] := \inf_{R\in\overline{\mathcal{W}}_N(\vec{\vartheta})} \|f(\mathbf{x}) - R(\mathbf{x})\|; \quad \overline{\mathcal{R}}_{M,N}[f] := \inf_{\vec{\vartheta}\in\mathrm{I\!R}^M} \overline{\mathcal{R}}_N[f,\vec{\vartheta}]; \quad \mathcal{R}_N^{\rm fr}[f] = \min_{1\le M\le N} \overline{\mathcal{R}}_{M,N}[f].$$

A particular case is approximation by *completely collapsed ridge functions*:

$$\mathcal{R}_N^{\rm fr}[f] \le \mathcal{R}_N^{\rm col}[f,\vartheta] := \inf_{\{W_j(x)\}_{j=1}^N} \left\| f(\mathbf{x}) - \sum_{j=1}^N \left(\frac{\partial}{\partial\vartheta}\right)^{j-1} W_j(\mathbf{x}\cdot\boldsymbol{\theta}),\ \mathcal{L}^2(\mathrm{I\!B}^2) \right\|, \tag{1}$$

which is in a sense the direct counterpart of equispaced ridge approximation.

2) Denote, respectively, $\mathcal{P}_N^1 := \mathrm{Span}\,\{x^k\}_{k\le N}$ and $\mathcal{P}_N^2 := \mathrm{Span}\,\{x_1^k x_2^l\}_{k+l\le N}$ the subspaces of algebraic polynomials of degree N in one and two real variables. If the components ϑ_j of $\vec{\vartheta}\in\mathrm{I\!R}^N$

are pairwise non-congruent modπ, then (cf. e. g. [2] and theorem 1 below) every polynomial $P(\mathbf{x}) \in \mathcal{P}^2_{N-1}$ can be represented as a linear combination of planar wave polynomials of degree $N-1$

$$P(\mathbf{x}) = \sum_{j=1}^{N} P_j(\mathbf{x}\cdot\boldsymbol{\theta}_j),\ P_j(x) \in \mathcal{P}^1_{N-1}, \quad \text{or} \quad \mathcal{P}^2_{N-1} \subset \mathcal{W}(\vec{\vartheta}), \quad \mathcal{R}[P,\vec{\vartheta}] = 0. \tag{2}$$

Thus, the element of best ridge approximation is not unique for all algebraic polynomials. Further, the classical quantities – best algebraic polynomial approximations

$$\mathcal{E}_N[f] := \min_{P\in\mathcal{P}^2_{N-1}} \|f - P, \mathcal{L}^2(\mathbb{B}^2)\|$$

majorize ridge approximations *for every non-degenerate* $\vec{\vartheta} \in \mathbb{R}^N$

$$\mathcal{R}[f,\vec{\vartheta}] \le \mathcal{E}_N[f]. \tag{3}$$

The solution of the problem of ridge approximation of the given function $f(\mathbf{x})$ depends upon the *Chebyshev orthogonal momenta* $a_n(f,\vartheta)$ generated by Fourier analysis in $\mathbb{B}^2$:

$$a_n(f,\vartheta) := \int_{\mathbb{B}^2} f(\mathbf{x})u_n(\mathbf{x}\cdot\boldsymbol{\theta})\,\mu(d\mathbf{x}), \quad u_n(x) := \frac{\sin{(n+1)\arccos x}}{\sqrt{1-x^2}}, \quad n = 0,\,1,\,\ldots. \tag{4}$$

By means of the latter, the problem $\mathcal{R}_N(f)$ is split into an infinite series of Kolmogorov – Nikol'skii type problems, cf. [1], concerning *optimal quadrature formulas for recovery of linear functionals*

$$\mathcal{F}_n(f)[T] := \int_0^{2\pi} a_n(f,\vartheta)T(\vartheta)\,\mu(d\vartheta), \quad n = 0,\,1,\,\ldots.$$

In the case of a general function $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$, momenta $a_n(f,\vartheta)$ are trigonometric polynomials of n-th order, satisfying $a_n(f,\vartheta+\pi) \equiv (-1)^n a_n(f,\vartheta)$. Let us denote $\mathcal{T}^{\pm}_n$ the whole subspace of trigonometric polynomials possessing this property, i. e. $\mathcal{T}^{\pm}_n := \mathrm{Span}\,\{e^{im\vartheta}\}_{|m|\le n(2)}$; here and below we use the notation $|m| \le n(2)$ for the set of integers m with $|m| \le n$ and $m \equiv n \pmod 2$. Further, denote $\mu(d\vartheta) := \frac{d\vartheta}{2\pi}$ the normalized Lebesgue measure on $\mathcal{S}^1$, $\|T,\mathcal{L}^2_{2\pi}\| := \left(\int_0^{2\pi} |T(\vartheta)|^2\,\mu(d\vartheta)\right)^{\frac{1}{2}}$, and let

$$\mathbb{B}(\mathcal{T}^{\pm}_n) := \left\{T \in \mathcal{T}^{\pm}_n :\ \|T,\mathcal{L}^2_{2\pi}\| \le 1\right\}, \quad \mathbb{B}(\mathcal{P}^1_n) := \{P(z) \in \mathcal{P}^1_n :\ \|P(e^{i\vartheta}),\ \mathcal{L}^2_{2\pi}\| \le 1\}.$$

A quadrature formula $\sigma(\vec{\vartheta},\vec{w})[T]$ with the *nodes* $\vec{\vartheta} := \{\vartheta_j\}_1^N \in \mathbb{R}^N$ and *weights* $\vec{w} := \{w_j\}_1^N \in C^N$ is a point-values functional

$$\sigma(\vec{\vartheta},\vec{w})[T] := \sum_{j=1}^{N} w_j T(\vartheta_j),$$

and the following quantities are typical for Kolmogorov – Nikol'skii setting, cf. [1] of the problem on optimization of quadrature formulas for recovery of linear functionals:

$$\mathcal{Q}_n[a,\vec{\vartheta}] := \inf_{\vec{w}\in C^N} \sup_{T\in\mathbb{B}(\mathcal{T}_n^{\pm})} \left| \int_0^{2\pi} a(f,\vartheta)T(\vartheta)\,\mu(d\vartheta) - \sigma(\vec{\vartheta},\vec{w})[T] \right|, \quad \vec{\vartheta}\in\mathbb{R}^N;$$

$$\mathcal{Q}^{\mathrm{eq}}_{n,N}[a] := \mathcal{Q}_n[a,\vec{\vartheta}],\ \vartheta_j = \frac{\pi j}{N}; \quad \mathcal{Q}^{\mathrm{opt}}_{n,N}[a] := \inf_{\vec{\vartheta}\in\mathbb{R}^N} \mathcal{Q}_n[a,\vec{\vartheta}]. \tag{5}$$

In the above, $a = a(\vartheta)$ is a fixed trigonometric polynomial, $a \in \mathcal{T}_n^{\pm}$.

Theorem 1 *Let* $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$, $\vec{\vartheta} \in \mathbb{R}^N$ *where the coordinates* ϑ_j *are pairwise non-congruent* $\mathrm{mod}\pi$. *Then*

$$\mathcal{R}[f,\vec{\vartheta}] = \sqrt{\sum_{n=N}^{\infty} (n+1)\left(\mathcal{Q}_n[a_n(f),\vec{\vartheta}]\right)^2}, \quad \mathcal{R}^{\mathrm{fr}}_N[f] = \inf_{\vec{\vartheta}\in\mathbb{R}^N} \sqrt{\sum_{n=N}^{\infty} (n+1)\left(\mathcal{Q}_n[a_n(f),\vec{\vartheta}]\right)^2}. \tag{6}$$

In particular,

$$\mathcal{R}^{\mathrm{fr}}_N[f] \ \geq \ \sqrt{\sum_{n=N}^{\infty} (n+1)\left(\mathcal{Q}^{\mathrm{opt}}_{n,N}[a_n(f)]\right)^2} \tag{7}$$

and

$$\mathcal{P}^2_{N-1} \subset \mathcal{W}^{\vec{\vartheta}}_N, \quad \mathcal{R}[f,\vec{\vartheta}] \leq \mathcal{E}_N[f]. \tag{8}$$

Like in free ridge approximation, in Kolmogorov – Nikol'skii problem we should admit *collapsed* quadrature formulas, involving linear combinations of point values of linear differential operators of total degree $\leq N-1$. A particular case is represented by *completely collapsed* quadrature formulas $\sigma^{\mathrm{col}}(P,\vartheta)[T] := P\left(\frac{d}{d\vartheta}\right)T(\vartheta)$, $P \in \mathcal{P}^1_{N-1}$. A proper version of the quantities $\mathcal{Q}_n[a,\vec{\vartheta}]$ answering this case is

$$\mathcal{Q}^{\mathrm{col}}_{n,N}[a,\vartheta] := \inf_{P\in\mathcal{P}^1_{N-1}} \sup_{T\in\mathbb{B}(\mathcal{T}_n^{\pm})} \left| \int_0^{2\pi} a(f,\varphi)T(\varphi)\,\mu(d\varphi) - \sigma^{\mathrm{col}}(P,\vartheta)[T] \right|,$$

and according to (6) free ridge approximations can be estimated from above as follows:

$$\mathcal{R}^{\mathrm{fr}}_N[f] \leq \mathcal{R}^{\mathrm{col}}_N[f] = \inf_{\vartheta} \sqrt{\sum_{n=N}^{\infty} (n+1)\left(\mathcal{Q}^{\mathrm{col}}_{n,N}[a_n(f),\vartheta]\right)^2} \tag{9}$$

Momenta $a_n(f,\vartheta)$, cf. (4), are especially simple for radial or harmonic functions in $\mathbb{B}^2$, i. e. when $f(\mathbf{x}) = f(|\mathbf{x}|)$ or, respectively, $\Delta f(\mathbf{x}) = 0$, $|\mathbf{x}| < 1$ (in the sequel, we will apply the self explanatory notations $f = f_{\rm rad}$, $f = f_{\rm harm}$ in these cases). For $f = f_{\rm rad}$ the n-th momentum is a constant, $a_n(f, \vartheta) = \alpha_n$ (and all while for $f = f_{\rm harm}$ it is a monomial of the highest frequency, $a_n(f, \vartheta) = \beta_n e^{in\vartheta} + \gamma_n e^{-in\vartheta}$ (cf. lemma 1 below)

It is clear from theorem 1 what special cases of Kolmogorov – Nikol'skii problems should be solved. In the case of $f = f_{\rm rad}$, we need to solve the problem concerning $\mathcal{Q}^{\rm opt}_{n,N}[1]$ – optimal recovery of the averages $\int_0^{2\pi} T(\vartheta)\, \mu(d\vartheta)$ of $T \in \mathcal{T}_n^{\pm}$ (the problem is non-trivial only for even n, $n \geq N$). For $f = f_{\rm harm}$, we need the recovery of $\int_0^{2\pi} T(\vartheta) \left(\beta e^{in\vartheta} + \gamma e^{-in\vartheta}\right) \mu(d\vartheta)$, i. e. linear combinations of the senior Fourier coefficients $\hat{T}(\pm n) := \int_0^{2\pi} T(\vartheta) e^{\pm in\vartheta}\, \mu(d\vartheta)$.

Seemingly, the problems concerning $\mathcal{Q}^{\rm opt}_{n,N}[1]$ and $\mathcal{Q}^{\rm opt}_{n,N}[e^{\pm in\vartheta}]$ are of a quite analogous nature simply because all coefficients of polynomials in $\mathbb{B}(\mathcal{T}_n^{\pm})$ possess "equal rights". Moreover, the problem concerning recovery of $\hat{T}(\pm n)$ can be also reformulated, cf. (50) below, as that of complex polynomials $P(z) \in \mathbb{B}(\mathcal{P}_n^1)$ in the center $z = 0$ of the disc $|z| \leq 1$ via their values $P(z_j)$ on the circumference $\mathcal{S}^1 = \{z :\ |z| = 1\}$:

$$\mathcal{Q}^{\rm opt}_{n,N}[e^{\pm in\vartheta}] = \mathcal{Q}^{\rm opt}_{n,N}[1, \mathbb{B}(\mathcal{P}_n^1)] := \inf_{\{z_j\}_1^N \in \mathcal{S}^1,\ \{w_j\}_1^N}\ \sup_{P \in \mathbb{B}(\mathcal{P}_n^1)} \left| P(0) - \sum_{j=1}^N w_j P(z_j) \right|. \tag{10}$$

Thus, the following conjectures are plausible:
1) optimal nodal points $\vec{\vartheta} = \{\vartheta_j\}_1^N$ should be "uniformly distributed";
2) if the nodal deficiency is essential, i. e. the ratio $\frac{N}{n}$ is small, then it is impossible to recover Fourier coefficients of all polynomials in $\mathbb{B}(\mathcal{T}_n^{\pm})$ with a small error: neither of the quantities $\mathcal{Q}^{\rm opt}_{n,N}[1]$, $\mathcal{Q}^{\rm opt}_{n,N}[e^{\pm in\vartheta}]$ can be small.

However, these conjectures fail to be true in the part concerning $\mathcal{Q}^{\rm opt}_{n,N}[e^{\pm in\vartheta}]$ (see Theorem 2 below). *Recovery of the senior Fourier coefficient (or the value $P(0)$ of algebraic polynomials $P(z) \in \mathbb{B}(\mathcal{P}_n^1)$, cf. (10)) with a small global error on the class $\mathbb{B}(\mathcal{T}_n^{\pm})$ is possible even if the sampling number N is much smaller than n, and it is rather the ratio $\frac{N}{\sqrt{n}}$ that governs this effect.*

Theorem 2 *Let n, N be positive integers, n even, $n \geq 2N$. Then*

$$\sqrt{\frac{1}{2}\left(1 - \frac{2N}{n+1}\right)} \leq \mathcal{Q}^{\rm opt}_{n,N}[1] \leq \sqrt{2\left(1 - \frac{2N}{n+2}\right)}. \tag{11}$$

Furthermore, the following relations hold true for $n \geq N \geq 5$

$$e^{-\frac{2N^2}{n}} \leq \mathcal{Q}^{\rm opt}_{n,N}[e^{\pm in\vartheta}] \leq \mathcal{Q}^{\rm col}_{n,N}[e^{\pm in\vartheta}, 0] \leq \min\left(1,\ \sqrt{2n}e^{-\frac{N}{\sqrt{n}}}\right). \tag{12}$$

The meaning of the next statement is the following.
1) Solutions of the problem $\mathcal{R}^{\text{eq}}$ (equispaced ridge approximation) for $f = f_{\text{rad}}$ and $f = f_{\text{harm}}$ are qualitatively and quantitatively the same, and in essence coincide with $\mathcal{E}$.
2) For $f = f_{\text{rad}}$, freedom in the choice of directions, in particular, the effect of collapse is not associated with any essential gain in approximation orders of $\mathcal{R}^{\text{fr}}$ compared with $\mathcal{R}^{\text{eq}}$ or $\mathcal{E}$.
3) On the contrary, for $f = f_{\text{harm}}$ free ridge approximation $\mathcal{R}^{\text{fr}}$ is "almost square times better" than $\mathcal{R}^{\text{eq}}$, due to the effect of collapse of wave vectors.

Theorem 3 *The following relations hold true*

$$\mathcal{R}_N^{\text{eq}}[f] = \sqrt{2\sum_{q=1}^{\infty} \frac{(\mathcal{E}_{2qN}[f])^2}{4q^2-1}}\,, \qquad \sqrt{\frac{2}{3}}\,\mathcal{E}_{2N}[f] \le \mathcal{R}_N^{\text{eq}}[f] \le \mathcal{E}_{2N}[f]\,, \quad f = f_{\text{rad}}\,;$$

$$\sqrt{\frac{1}{3}}\,\mathcal{E}_{N+1}[f] \;\le\; \mathcal{R}_N^{\text{eq}}[f] \le \mathcal{E}_N[f]\,, \quad f = f_{\text{harm}}\,; \tag{13}$$

$$\mathcal{R}_N^{\text{fr}}[f] \ge \sup_{M\ge N} \sqrt{\frac{M-N}{2M}}\,\mathcal{E}_{2M}[f] \ge \sup_{M\ge N} \sqrt{\frac{M-N}{2M}}\,\mathcal{R}_M^{\text{eq}}[f]\,, \quad f = f_{\text{rad}}\,; \tag{14}$$

$$e^{-8}\,\mathcal{E}_{N^2}[f] \le \mathcal{R}_N^{\text{fr}}[f] \le \mathcal{R}_N^{\text{col}}[f] \le \min_{M\ge N}\left(\sqrt{2M}e^{-\frac{N}{\sqrt{M}}}\mathcal{E}_N[f] + \mathcal{E}_{M+1}[f]\right)\,, \quad N \ge 5,\; f = f_{\text{harm}}\,. \tag{15}$$

Corollary 1 *If $f = f_{\text{rad}}$ then*

$$\mathcal{R}_N^{\text{fr}}[f] \ge \sup_{\varepsilon>0} \sqrt{\frac{\varepsilon}{2(1+\varepsilon)}}\,\mathcal{E}_{2(1+\varepsilon)N}[f] \ge \sup_{\varepsilon>0} \sqrt{\frac{\varepsilon}{3(1+\varepsilon)}}\,\mathcal{R}_{(1+\varepsilon)N}^{\text{eq}}[f]\,.$$

If $f = f_{\text{harm}}$ then

(i) $\forall\varepsilon > 0\ \exists c_\varepsilon > 0 :\ \mathcal{R}_N^{\text{fr}}[f] \le c_\varepsilon\,(\mathcal{R}_N^{\text{eq}}[f]\exp(-N^\varepsilon) + \mathcal{R}_{N^{2-3\varepsilon}}^{\text{eq}}[f])\,;$

(ii) $\exists\delta > 0 :\ \mathcal{E}_{N^{2-\delta}}[f] = o\,(\mathcal{E}_N[f]) \Longrightarrow \mathcal{R}_N^{\text{fr}}[f] = o\,(\mathcal{R}_N^{\text{eq}}[f])\,;$

(iii) $\exists\alpha > 0 :\ \mathcal{R}_N^{\text{eq}}[f] = O(N^{-\alpha}) \Longrightarrow \forall\varepsilon > 0 :\ \mathcal{R}_N^{\text{fr}}[f] = O(N^{-2\alpha+\varepsilon}),\ N \to \infty.$

2 Proofs

2.1 Chebyshev – Fourier analysis in $\mathcal{L}^2(\mathbb{B}^2)$

For a given non-negative integer n, let $D_n(\vartheta)$ denote the Dirichlet kernel for the subspace $\mathcal{T}_n^{\pm}$:

$$D_n(\vartheta) := \sum_{|m|\leq n(2)} e^{im\vartheta} = \frac{\sin{(n+1)\vartheta}}{\sin\vartheta}.$$

Obviously,

$$T(\vartheta) = [T * D_n](\vartheta) := \int_0^{2\pi} T(\varphi) D_n(\varphi - \vartheta)\, \mu(d\vartheta), \quad T \in \mathcal{T}_n^{\pm}. \tag{16}$$

Further, let as above $u_n(x)$ denote the n-th Chebyshev polynomial of the second kind, i. e. $u_n(x) = D_n(\arccos x) = \frac{\sin{(n+1)}\arccos x}{\sqrt{1-x^2}}, \quad |x| \leq 1.$

Chebyshev – Fourier analysis in $\mathcal{L}^2(\mathbb{B}^2)$ consists in the orthogonal expansion

$$f(\mathbf{x}) \overset{\mathcal{L}^2(\mathbb{B}^2)}{=} \int_0^{2\pi} \left(\sum_{n=0}^{\infty} (n+1) a_n(f,\vartheta) u_n(\mathbf{x}\cdot\boldsymbol{\theta}) \right) \mu(d\vartheta), \tag{17}$$

cf. [2] – [4]. For each fixed $\boldsymbol{\theta} \in \mathcal{S}^1$, the corresponding planar wave Chebyshev polynomial $u_n(\mathbf{x}\cdot\boldsymbol{\theta})$ is orthogonal in $\mathcal{L}^2(\mathbb{B}^2)$ to all polynomials of degree $\leq n-1$:

$$\int_{\mathbb{B}^2} u_n(\mathbf{x}\cdot\boldsymbol{\theta}) P(\mathbf{x})\, \mu(d\mathbf{x}) = 0, \quad \forall P(\mathbf{x}) \in \mathcal{P}_{n-1}^2. \tag{18}$$

As mentioned above, the ortogonal momenta, or *Fourier – Chebyshev coefficients* $a_n(f,\vartheta)$, are trigonometric polynomials $a_n(f) \in \mathcal{T}_n^{\pm}$. The *Parceval identity* is given by

$$\left\| f,\, \mathcal{L}^2(\mathbb{B}^2) \right\|^2 = \sum_{n=0}^{\infty} (n+1) \left\| a_n(f),\, \mathcal{L}_{2\pi}^2 \right\|^2 = \frac{1}{2\pi} \sum_{n=0}^{\infty} (n+1) \int_0^{2\pi} |a_n(f,\vartheta)|^2\, d\vartheta. \tag{19}$$

Furthermore, if $\{a_n(\vartheta)\}_{n=0}^{\infty}$ is a sequence of trigonometric polynomials satisfying the conditions

$$a_n \in \mathcal{T}_n^{\pm}, \quad \sum_{n=0}^{\infty} (n+1) \left\| a_n,\, \mathcal{L}_{2\pi}^2 \right\|^2 < \infty,$$

then (*Plancherel's theorem*) there exists a function $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$, unique up to a set of Lebesgue measure 0, such that $a_n(f,\vartheta) \equiv a_n(\vartheta),\ n = 0, 1, \ldots$.

Orthogonal projection $\mathrm{Proj}_N[f](\mathbf{x})$ in $\mathcal{L}^2(\mathbb{B}^2)$ of a function $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$ onto the subspace of algebraic polynomials of degree $N-1$ is given by the partial sum of first N terms of the expansion (17),

$$\mathrm{Proj}_N[f](\mathbf{x}) = \int_0^{2\pi} \left(\sum_{n=0}^{N-1} (n+1) a_n(f, \vartheta) u_n(\mathbf{x} \cdot \boldsymbol{\theta}) \right) \mu(d\vartheta), \quad N = 1, 2, \ldots,$$

and in particular

$$\mathcal{E}_N[f] = \sqrt{\sum_{n=N}^{\infty} (n+1) \left\| a_n(f),\ \mathcal{L}^2_{2\pi} \right\|^2}. \tag{20}$$

2.2 Momenta of radial, harmonic and planar wave functions

Lemma 1 1) *If $f(\mathbf{x}) = g(|\mathbf{x}|^2)$, where $g(x) \in \mathcal{L}^2(0,1)$ and $g(x) \overset{\mathcal{L}^2(0,1)}{=} \sum_{\nu=0}^{\infty} \grave{g}_\nu l_\nu(x)$ is the Fourier – Legendre expansion of $g(x)$, then*

$$a_{2\nu+1}(f) = 0; \quad a_{2\nu}(f) = \frac{\grave{g}_\nu}{\sqrt{2\nu+1}}, \quad \nu = 0, 1, \ldots . \tag{21}$$

2) *If $f(r\boldsymbol{\varphi}) = \hat{f}(0) + \sum_{n=1}^{\infty} r^n \left(\hat{f}(-n) e^{-in\varphi} + \hat{f}(n) e^{in\varphi} \right), \quad 0 \le r < 1$ is the standard representation of a harmonic function $f = f_{\mathrm{harm}} \in \mathcal{L}^2(\mathbb{B}^2)$ in the polar coordinates $\mathbf{x} = r\boldsymbol{\varphi}$, then*

$$a_0(f) = \hat{f}(0), \quad a_n(f, \vartheta) = \frac{\hat{f}(-n) e^{-in\vartheta} + \hat{f}(n) e^{in\vartheta}}{n+1}, \quad n = 1, 2, \ldots . \tag{22}$$

3) *Let $\omega(x) := \frac{2}{\pi}\sqrt{1-x^2}$, $|x| \le 1$, $W(x) \in \mathcal{L}^2_\omega(-1,1)$ and $W(x) \overset{\mathcal{L}^2_\omega(-1,1)}{=} \sum_{n=0}^{\infty} \check{W}_n u_n(x)$ – the Fourier – Chebyshev expansion of W, $\boldsymbol{\varphi} \in \mathcal{S}^1$ – a fixed unit vector. Then $W(\mathbf{x} \cdot \boldsymbol{\varphi}) \in \mathcal{L}^2(\mathbb{B}^2)$ and*

$$a_n\left(W(\mathbf{x} \cdot \boldsymbol{\varphi}), \vartheta\right) = \frac{\check{W}_n}{n+1} D_n(\vartheta - \varphi) = \frac{\check{W}_n \sin(n+1)(\vartheta - \varphi)}{(n+1)\sin(\vartheta - \varphi)}, \quad n = 1, 2, \ldots . \tag{23}$$

Proof. For the proof of (21), see e.g. [5]. Let us also note the following relations between Legendre and Chebyshev polynomials of the second kind:

$$\int_0^{2\pi} u_{2\nu+1}(\mathbf{x} \cdot \boldsymbol{\theta})\, \mu(d\vartheta) = 0, \quad \int_0^{2\pi} u_{2\nu}(\mathbf{x} \cdot \boldsymbol{\theta})\, \mu(d\vartheta) = \frac{l_\nu(|\mathbf{x}|^2)}{\sqrt{2\nu+1}}, \quad \nu = 0, 1, \ldots .$$

For the proof of claim 2), let us note that the terms $p_m^{\pm}(\mathbf{x}) := r^m e^{\pm im\varphi}$ are harmonic algebraic polynomials, $p_m^{\pm}(\mathbf{x}) \in \mathcal{P}_m^2$. By (18), $\int_{\mathbb{B}^2} u_n(\mathbf{x}\cdot\boldsymbol{\theta}) p_m^{\pm}(\mathbf{x})\,\mu(d\mathbf{x}) = 0$ for all $m < n$. The same relations are valid for $m > n$. Indeed, for each fixed $|\mathbf{x}| = r \geq 0$, $u_n(\mathbf{x}\cdot\boldsymbol{\theta}) = u_n(r\cos(\vartheta - \varphi))$ is a trigonometric polynomial in φ, of degree n, so that $\int_0^{2\pi} u_n(r\cos(\vartheta-\varphi))e^{\pm im\varphi}\,d\varphi = 0$, if $m > n$. Thus, for the proof of (22) we need to consider only the case $m = n$.

Let $T_n(x) := \cos(n \arccos x)$, $|x| \leq 1$ denote the n-th Chebyshev polynomial of the first kind. We have $u_0(x) = T_0(x)$, $u_n(x) = 2\sum_{0\leq m\leq n(2)} T_m(x)$ for $n \geq 1$, and $T_n(x) = 2^{n-1}x^n + q(x)$, where $q(x) \in \mathcal{P}_{n-2}^1$. Thus $u_n(x) = 2^{n-1}x^n + q_1(x)$, $q_1(x) \in \mathcal{P}_{n-2}^1$ and for fixed $n \geq 1$, $r > 0$ and ϑ, $u_n(r\cos(\vartheta-\varphi)) = 2^n(r\cos(\vartheta-\varphi))^n + t(e^{i\varphi}) = 2r^n\cos n(\vartheta-\varphi) + t_1(e^{i\varphi})$ where $t,\, t_1 \in \mathcal{T}_{n-2}^{\pm}$. Consequently, $\int_0^{2\pi} u_n(r\cos(\vartheta-\varphi))e^{\pm in\varphi}\,\mu(d\varphi) = r^n e^{\pm in\vartheta}$ and

$$\int_{\mathbb{B}^2} p_n^{\pm}(\mathbf{x}) u_n(\mathbf{x}\cdot\boldsymbol{\theta})\,\mu(d\mathbf{x}) = 2\int_0^1 r\left(\int_0^{2\pi} r^n e^{\pm in\varphi} u_n(r\cos(\vartheta-\varphi))\,\mu(d\varphi)\right)dr = 2e^{\pm in\vartheta}\int_0^1 r^{2n+1}\,dr,$$

whence the relations (22) follow. Let us also note that for $f = f_{\text{harm}} \in \mathcal{L}^2(\mathbb{B}^2)$ [1]

$$\|f;\, \mathcal{L}^2(\mathbb{B}^2)\|^2 = \sum_{n=-\infty}^{\infty} \frac{|\hat{f}(n)|^2}{|n|+1}.$$

Further, for a fixed vector $\mathbf{x}$, $u_n(\mathbf{x}\cdot\boldsymbol{\theta})$ is a trigonometric polynomial in ϑ, and $u_n(\mathbf{x}\cdot\boldsymbol{\theta}) \in \mathcal{T}_n^{\pm}$. Thus, by (16)

$$\int_0^{2\pi} \frac{\check{W}_n}{n+1} D_n(\vartheta-\varphi) u_n(\mathbf{x}\cdot\boldsymbol{\theta})\,\mu(d\vartheta) = \frac{\check{W}_n}{n+1} u_n(\mathbf{x}\cdot\boldsymbol{\varphi}),$$

and (23) follow from (17).

2.3 Proof of theorem 1

Now let us consider a ridge function $R(\mathbf{x}) = \sum_{j=1}^N W_j(\mathbf{x}\cdot\boldsymbol{\theta}_j) \in \mathcal{L}^2(\mathbb{B}^2)$. Then the momenta of $R(\mathbf{x})$ are linear combinations of shifted Dirichlet kernels:

$$a_n(R,\vartheta) = \frac{1}{n+1}\sum_{j=1}^N \check{W}_{j,n} D_n(\vartheta - \vartheta_j), \tag{24}$$

where $\check{W}_{j,n}$ denotes the n-th Fourier – Chebyshev coefficient of the function $W_j(x) \in \mathcal{L}_\omega^2(-1,1)$, cf. (23). Let us upply vector notations:

$$\vec{\check{W}}_n := \{\check{W}_{j,n}\}_{j=1}^N \in C^N, \quad |\vec{\check{W}}_n|^2 := \sum_{j=1}^N |\check{W}_{j,n}|^2, \quad \vec{U}\cdot\vec{V} = \sum_{j=1}^N U_j V_j;$$

[1] Note that the condition $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$ does not guarantee the existence of the boundary values $f(\boldsymbol{\theta})$.

let also $\mathcal{D}_n(\vec{\vartheta})$ denote the $N\times N$ symmetric matrix $\{\mathcal{D}_n(\vartheta_j-\vartheta_k)\}_{j,k=1}^N$, and $\mathcal{D}'_n(\vec{\vartheta}) := \mathcal{D}_n(\vec{\vartheta})-D_n(0)\mathcal{I} = \mathcal{D}_n(\vec{\vartheta}) - (n+1)\mathcal{I}$, where $\mathcal{I}$ is the N-th identity matrix; z^* - the conjugate of a complex number z.

Lemma 2 *Let n be a positive integer, N – a natural number, $\vec{\vartheta} = \{\vartheta_j\}_1^N \in \mathbb{R}^N$, where ϑ_j are pairwise distinct* $\mathrm{mod}\pi$. *Further, let $a(\vartheta)$ be a fixed polynomial of the class $\mathcal{T}_n^{\pm}$, $\vec{a}(\vec{\vartheta}) := \{a(\vartheta_j)\}_1^N$. Then:*

$$1) \qquad \mathcal{Q}_n[a,\vec{\vartheta}] = \min_{\vec{w}\in C^N} \left\| a(\vartheta) - \sum_{j=1}^N w_j D_n(\vartheta-\vartheta_j),\, \mathcal{L}^2(\mathbb{B}^2)\right\|. \tag{25}$$

2) *Denote $\vec{w} = \{w_j\}_1^N$ the vector of optimal weights, or, which is the same, the minimizer of the extremal problem on the right of (25). Then $\vec{w}$ satisfies the following system of N linear equations:*

$$\sum_{j=1}^N w_j D_n(\vartheta_k-\vartheta_j) = a(\vartheta_k),\ k=1,\dots,N, \quad \textit{or} \quad \mathcal{D}_n(\vec{\vartheta})\vec{w} = \vec{a}(\vec{\vartheta}). \tag{26}$$

$$3) \quad \mathrm{rank}\ \mathcal{D}_n(\vec{\vartheta}) = \dim \mathrm{Span}\,\{\mathcal{D}_n(\vartheta-\vartheta_j)\}_{j=1}^N = \min\,(N, n+1). \tag{27}$$

$$4) \quad \mathcal{Q}_n[a,\vec{\vartheta}] = 0,\ n\le N-1; \quad \mathcal{Q}_n[a,\vec{\vartheta}] = \sqrt{\|a,\mathcal{L}^2_{2\pi}\|^2 - \vec{w}\cdot\vec{a}^*(\vec{\vartheta})},\ n\ge N, \tag{28}$$

where $\vec{w}$ is the vector of optimal weights.

$$5) \qquad \sup_n \|\mathcal{D}'_n(\vec{\vartheta})\|_{l^2\mapsto l^2} = C(\vec{\vartheta}) < \infty, \tag{29}$$

i. e., the $l^2\mapsto l^2$-norm of the matrix $\mathcal{D}'_n(\vec{\vartheta})$ is uniformly bounded in n.

Proof. First of all,

$$\left\| a(\vartheta) - \sum_{j=1}^N w_j D_n(\vartheta-\vartheta_j)\right\| = \sup_{T\in\mathbb{B}(\mathcal{T}_n^{\pm})} \left| \int_0^{2\pi} a(\vartheta)T(\vartheta)\,\mu(d\vartheta) - \sum_{j=1}^N w_j T(\vartheta_j)\right|.$$

This relation is a corollary of (16):

$$\int_0^{2\pi} a(\vartheta)T(\vartheta)\,\mu(d\vartheta) - \sum_{j=1}^N w_j T(\vartheta_j) = \int_0^{2\pi}\left(a(\vartheta) - \sum_{j=1}^N w_j D_n(\vartheta-\vartheta_j)\right) T(\vartheta)\,\mu(d\vartheta).$$

and the resonance case of Cauchy inequality. Thus, (25) follows from the definition of $\mathcal{Q}_n[a,\vec{\vartheta}]$.

Next, (26) follows from (25), because

$$\int_0^{2\pi} D_n(\vartheta-\vartheta_j)D_n(\vartheta-\vartheta_k)\,\mu(d\vartheta)=\mathcal{D}_n(\vartheta_j-\vartheta_k),\quad \int_0^{2\pi} a(\vartheta)\mathcal{D}_n(\vartheta-\vartheta_k)\,\mu(d\vartheta)=a(\vartheta_k).$$

The system (26) is consistent for *every* polynomial $a(\vartheta)\in\mathcal{T}_n^{\pm}$. Since $\dim\mathcal{T}_n^{\pm}=n+1$, by Lagrange interpolation over $\vec{\vartheta}$, from here we conclude that

$$\operatorname{rank}\mathcal{D}_n(\vec{\vartheta})=\dim\{\mathcal{D}_n(\vec{\vartheta})\vec{w}:\ \vec{w}\in C^N\}=\dim\{\vec{a}(\vec{\vartheta}):\ a(\vartheta)\in\mathcal{T}_n^{\pm}\}=\min(N,n+1).$$

Further, $\dim\operatorname{Span}\{\mathcal{D}_n(\vartheta-\vartheta_j)\}_{j=1}^N=\operatorname{rank}\mathcal{D}_n(\vec{\vartheta})$, because $\mathcal{D}_n(\vec{\vartheta})$ is the Gramm matrix of the system $\{\mathcal{D}_n(\vartheta-\vartheta_j)\}_{j=1}^N$, which completes the proof of (27).

Since $\operatorname{Span}\{\mathcal{D}_n(\vartheta-\vartheta_j)\}_{j=1}^N=\mathcal{T}_n^{\pm}$ for $n\le N-1$, the equalities $\mathcal{Q}_n(a,\vec{\vartheta})=0,\ n\le N-1$ follow from (27).

Remark. Linear independence of the system $\{\mathcal{D}_n(\vartheta-\vartheta_j)\}_{j=1}^N$ for $n\ge N-1$ and (27) is not a new result, cf. e.g. [2].

The equality

$$\left(\mathcal{Q}_n[a,\vec{\vartheta})\right]^2=\min_{\vec{w}\in C^N}\left\|a(\vartheta)-\sum_{j=1}^N w_jD_n(\vartheta-\vartheta_j),\,\mathcal{L}^2(\mathrm{I\!B}^2)\right\|^2=\|a,\mathcal{L}^2_{2\pi}\|^2-\vec{w}\cdot\vec{a}^*(\vec{\vartheta})$$

is a corollary of (26). Finally, the entries of the matrix $\mathcal{D}_n'(\vec{\vartheta})$ are majorized by $\max_{j\ne k}|\csc(\vartheta_j-\vartheta_k)|$, which implies (29).

Theorem 1 follows from the Parceval identity (19), claim 3) of Lemma 1, (23) and the definition of $\mathcal{Q}_n[a,\vec{\vartheta}]$. Indeed, by (24) the n-th momentum of a ridge function $R(\mathbf{x})=\sum_{j=1}^N W_j(\mathbf{x}\cdot\boldsymbol{\theta}_j)$ is a linear combination of shifted Dirichlet kernels:

$$a_n(R,\vartheta)=\sum_{j=1}^N w_{j,n}D_n(\vartheta-\vartheta_j),\quad w_{j,n}:=\frac{\check{W}_{j,n}}{n+1},\quad j=1,\dots,N. \tag{30}$$

Therefore, the selection of optimal profiles $W_j(x)$ for a given $f(\mathbf{x})\in\mathcal{L}^2(\mathrm{I\!B}^2)$ and a non-degenerate set of directional angles $\vec{\vartheta}=\{\vartheta_j\}_{j=1}^N\in\mathrm{I\!R}^N$ is performed in the following three steps.

Step 1. Find the point values of the Chebyshev momenta

$$a_n(f,\vartheta_j)=\int_{\mathrm{I\!B}^2} f(\mathbf{x})u_n(\mathbf{x}\cdot\boldsymbol{\theta}_j)\,\mu(d\mathbf{x}),\quad j=1,\dots,N,\ n=0,1,\dots.$$

Step 2. For each fixed $n = 0, 1, \ldots$, solve the system of N linear equations

$$\sum_{j=1}^{N} D_n(\vartheta_j - \vartheta_k) w_{j,n} = a_n(f, \vartheta_k), \quad k = 1, \ldots, N.$$

with respect to the unknown $\vec{w}_n = \{w_{j,n}\}_{j=1}^{N}$. Note, that all these systems are cosistent. However, the solution is not unique, if $n \leq N-2$ ("low frequences"), and is unique for all $n \geq N-1$ ("high frequences").

Step 3. Let

$$W_j(x) := \sum_{n=0}^{\infty} (n+1) w_{j,n} u_n(x), \quad j = 1, \ldots, N.$$

Due to non-uniqueness on Step 2, optimal profiles $W_j(x)$ are always non-unique, if $N \geq 2$. However, it is easy to see that these profiles are "unique up to low frequences" – in the orthogonal complement ${}^{\perp}\mathcal{P}^2_{N-1} := \mathcal{L}^2(\mathbb{B}^2) \ominus \mathcal{P}^2_{N-1}$ of the subspace of algebraic polynomials $\mathcal{P}^2_{N-1}$ within $\mathcal{L}^2(\mathbb{B}^2)$. Thus, the set of optimal profiles $\{W_j(\mathbf{x})\}_1^N = \{W_j(f, \vec{\vartheta}, \mathbf{x}, \vec{\vartheta})\}_1^N$, i. e. the minimizer in the problem $\mathcal{R}_N(f - \mathrm{Proj}_N[f], \vec{\vartheta})$, is uniquelly defined.

In the next statement we apply the notation $\vec{W}(x) = \{W_j(x)\}_1^N$ for a set of N univariate functions and let

$$\mathcal{E}_M[\vec{W}] := \sqrt{\sum_{j=1}^{N} \left(\mathcal{E}_M[W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j)]\right)^2} = \sqrt{\sum_{n=M}^{\infty} |\vec{W}_n|^2}, \quad M = 1, 2, \ldots,$$

Theorem 4 *Assume that the components ϑ_j of $\vec{\vartheta} \in \mathbb{R}^N$ are pairwise distinct* $\bmod \pi$*, and let $R(\mathbf{x}) = \sum_{j=1}^{N} W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j)$. Then*

$$\mathcal{E}_M[\vec{W}] = \left(1 + O_{\vec{\vartheta}}\left(\frac{1}{M}\right)\right) \mathcal{E}_M[R], \quad M \to \infty; \quad \mathcal{E}_M[\vec{W}] \leq C(\vec{\vartheta}) \mathcal{E}_M[R], \quad M \geq N-1, \tag{31}$$

where the constant in $O_{\vec{\vartheta}}$ and $C(\vec{\vartheta})$ depend only upon $\vec{\vartheta}$;
Further, the operator

$$\vec{W}_N: \quad f(\mathbf{x}) \mapsto \vec{W}_N(f) = \{W_j(f, x)\}_{j=1}^{N} := \arg\min \left\| f(\mathbf{x}) - \sum_{j=1}^{N} W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j), \mathcal{L}^2(\mathbb{B}^2) \right\|$$

is well-defined, linear and bounded from ${}^{\perp}\mathcal{P}^2_{N-1}$ into $\vec{\mathcal{L}}_{\omega,N}$, where $\omega(x) = \frac{2}{\pi}\sqrt{1-x^2}$ and

$$\vec{\mathcal{L}}_{\omega,N} := \left\{ \vec{W}(x) = \{W_j(x)\}_{j=1}^{N} : \quad \left\| \vec{W}, \vec{\mathcal{L}}_{\omega,N} \right\| = \sum_{j=1}^{N} \|W_j(x), \mathcal{L}^2_{\omega}(-1,1)\| < \infty \right\}.$$

Proof. According to (20) and (30),

$$(\mathcal{E}_M[R])^2 = \sum_{n=M}^{\infty} (n+1)\|a_n(R), \mathcal{L}^2_{2\pi}\|^2 = \sum_{n=M}^{\infty} \frac{\mathcal{D}_n(\vec{\vartheta})}{n+1} \vec{\check{W}}_n \cdot \vec{\check{W}}_n^*$$

$$= \sum_{n=M}^{\infty} \left(|\vec{\check{W}}_n|^2 - \frac{\mathcal{D}'_n(\vec{\vartheta})}{n+1} \vec{\check{W}}_n \cdot \vec{\check{W}}_n^* \right) = \left(\mathcal{E}_M[\vec{W}]\right)^2 - \sum_{n=M}^{\infty} \frac{\mathcal{D}'_n(\vec{\vartheta})}{n+1} \vec{\check{W}}_n \cdot \vec{\check{W}}_n^*, \tag{32}$$

and making use of (29), we further have

$$\left| \sum_{n=M}^{\infty} \frac{\mathcal{D}'_n(\vec{\vartheta})}{n+1} \vec{\check{W}}_n \cdot \vec{\check{W}}_n^* \right| \le \frac{C(\vec{\vartheta})}{M+1} \sum_{n=M}^{\infty} |\vec{\check{W}}_n|^2 = \frac{C(\vec{\vartheta})}{M+1} \left(\mathcal{E}_M[\vec{W}]\right)^2,$$

and the asymptotic formula in (31) follows. This also implies the estimate $\mathcal{E}_M[\vec{W}] \le C(\vec{\vartheta})\mathcal{E}_M[R]$ for all sufficiently large $M \ge M_0(\vec{\vartheta})$. To prove that the same estimate is true for all remaining M, i. e. $N-1 \le M < M_0(\vec{\vartheta})$, we note that, according to (27), for $n \ge M \ge N-1$, all matrices $\mathcal{D}_n(\vec{\vartheta})$ are *strictly* positive definite, and thus (cf. (32))

$$\sum_{M \le n < M_0(\vec{\vartheta})} |\vec{\check{W}}_n|^2 \le C'(\vec{\vartheta}) \sum_{M \le n < M_0(\vec{\vartheta})} \frac{\mathcal{D}_n(\vec{\vartheta})}{n+1} \vec{\check{W}}_n \cdot \vec{\check{W}}_n^*.$$

The proof of claim 2) is analogous, and we omit the details.

Remark. If $R(\mathbf{x}) = \sum_{j=1}^N W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j)$, then obviously $\|R\| \le \sum_{j=1}^N \|W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j)\|$. It is also true that if $R(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$, then all summands $W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j)$ are in $\mathcal{L}^2(\mathbb{B}^2)$, too. However, it is easy to see that for $N \ge 2$ it is impossible to estimate the norms of the $W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j)$ via that of R. Indeed, the set of ridge functions of the class $\mathcal{W}(\vec{\vartheta})$ contains "kernels" of the type $0 \equiv \sum_{j=1}^N P_j(\mathbf{x} \cdot \boldsymbol{\theta}_j)$, where $\{P_j(x)\}$ are non-trivial collections of single variate polynomials of degree $N-2$.

Thus, (31) is a correct form of inverse type estimates of the planar wave components $W_j(\mathbf{x} \cdot \boldsymbol{\theta}_j)$ via their sum R [2].

2.4 Equispaced quadrature formulas and ridge approximation

In this section, we consider equispaced quadrature formulas and prove the relations (13) of theorem 3 concerning equispaced ridge approximations of f_{rad} and f_{harm}.

[2]It is an interesting open problem whether estimates of such type are true for the functional norms other than $\mathcal{L}^2(\mathbb{B}^2)$.

We will explicitly solve the series of optimization problems $\mathcal{Q}^{\rm eq}_{n,N}[a]$, cf. also [2] and [6]. In this section, we consider only equispaced nodes $\vartheta_j = \frac{\pi j}{N}$.

Let us consider the *spectral matrix*, whose entries are Fourier coefficients of the Chebyshev momenta $a_n(f,\vartheta)$ of the given function $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$:

$$\hat{\mathcal{A}}(f) := \Big\{ \{\hat{a}_{m,n}(f)\}_{|m|\le n(2)} \Big\}_{n=0}^{\infty}, \quad a_n(f,\vartheta) = \sum_{|m|\le n(2)} \hat{a}_{m,n}(f) e^{im\vartheta}. \tag{33}$$

Fix n and denote $\hat{\mathbf{a}} = \hat{\mathbf{a}}_n(f) = \langle \hat{a}_{m,n} \rangle_{|m|\le n(2)}$ the nth column of $\hat{\mathcal{A}}(f)$; $\quad |\hat{\mathbf{a}}| := \sqrt{\sum_{|m|\le n(2)} |\hat{a}_{m,n}|^2} = \|a_n(f), \mathcal{L}^2_{2\pi}\|$. Optimization of quadrature formulas for recovery of $\int_0^{2\pi} a(\vartheta) T(\vartheta)\, d\vartheta$ via equispaced nodal data is dual to the following type of approximation of the vector $\hat{\mathbf{a}}$:

$$\mathcal{E}^{\rm dif}_N[\hat{\mathbf{a}}] := \min\{ |\hat{\mathbf{a}} - \mathbf{c}| : \ c_l = c_m, \ l \equiv m(2N), \ |l|, |m| \le n(2) \}; \quad \mathbf{c}(\hat{\mathbf{a}}, N) := \arg \mathcal{E}^{\rm dif}_N(\hat{\mathbf{a}}).$$

Geometrically, this problem is solved by orthogonal projection in l^2 of $\hat{\mathbf{a}}$ onto the subspace of vectors $\mathbf{c}$ whose coordinates c_m, $|m| \le n(2)$ are constant along arithmetical progressions mod$2N$:

$$c_m(\hat{\mathbf{a}}, N) = \frac{1}{\delta(m,n,N)} \sum_{l\equiv m(2N)} \hat{a}_l; \quad \delta(m,n,N) := \left[\frac{n+m}{2N}\right] + \left[\frac{n-m}{2N}\right] + 1, \quad |m| \le n(2), \tag{34}$$

($[x]$ is the integral part of $x \in \mathbb{R}^1$). The operator $\mathbf{C}_{N,n}: \quad \hat{\mathbf{a}}_n \mapsto \mathbf{c}(\hat{\mathbf{a}}, N)$ *diffracts* the coordinates of $\hat{\mathbf{a}}_n$. It is easy to see that

$$\mathbf{c}(\hat{\mathbf{a}}, N) = \hat{\mathbf{a}}, \ n \le N-1; \quad \hat{\mathbf{a}} - \mathbf{c}(\hat{\mathbf{a}}, N) \perp \mathbf{c}(\hat{\mathbf{a}}, N) \tag{35}$$

Lemma 3 1) *Let (cf. (34))*

$$B(a,N,\vartheta) := \sum_{m\in(-N,N], |m|\le n(2)} c_m(\hat{\mathbf{a}}, N)\, e^{im\vartheta}.$$

Then the set of optimal weights $\vec{w}(a,n,N) = \{w_j\}_1^N$ and the optimal error of the equisapced quadrature are given by the relations

$$w_j = \frac{B(a,N,\vartheta_j)}{N}, \ j = 1, \ldots, N; \quad \mathcal{Q}^{\rm opt}[a, \vec{\vartheta}] = \mathcal{E}^{\rm dif}_N[\hat{\mathbf{a}}] = \sqrt{|\hat{\mathbf{a}}|^2 - |\mathbf{c}(\hat{\mathbf{a}}, N)|^2}. \tag{36}$$

2) *For $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$*

$$\mathcal{R}^{\rm eq}_N[f] = \sqrt{\sum_{n=N}^{\infty} (n+1) \left(\mathcal{E}^{\rm dif}_N[\hat{\mathbf{a}}_n]\right)^2}, \tag{37}$$

and if $f(\mathbf{x}) \in^{\perp} \mathcal{P}^2_{N-1}$, *then the set of optimal profiles (high frequences)* $\vec{W}(x) = \{W_j(x)\}_1^N = \arg\min \left\| f(\mathbf{x}) - \sum_{j=1}^N W_j(\mathbf{x}\cdot\boldsymbol{\theta}_j)\right\|$ *is determined by*

$$W_j(x) = \sum_{n=N}^{\infty} \frac{n+1}{N} B(a_n(f), N, \vartheta_j) u_n(x)\,. \tag{38}$$

Proof. It is not hard to see that

$$\sum_{j=1}^{N} e^{im\vartheta_j} \mathcal{D}_n(\vartheta_k - \vartheta_j) = N\delta(m,n,N)e^{im\vartheta_k}; \quad \frac{1}{N}\sum_{j=1}^{N} B(a,N,\vartheta_j)\,\mathcal{D}_n(\vartheta_k-\vartheta_j) = a(\vartheta_k),$$

so that the equality $\vec{w} = \frac{1}{N}\vec{B}$ for the optimal weights follows from (26). The equality $\mathcal{Q}^{\mathrm{opt}}[a,\vec{\vartheta}] = \mathcal{E}_N^{\mathrm{dif}}[\hat{\mathbf{a}}]$ follows from (28) and (35) in an anlogous way, and we omit the details. Relations (37) and (38) are corollaries of (6), (36) and (30).

Now we can finish the proof of the relations (13) in theorem 3.

According to (21), momenta $a_n(f_{\mathrm{rad}})$ are constants, and $a_{2\nu+1}(f_{\mathrm{rad}}) = 0$. Thus, for even n

$$\begin{aligned}
&\hat{a}_{m,n} = 0,\ m \neq 0;\ \hat{a}_{0,n} = \alpha_n; \quad \delta(0,n,N) = 2\left[\frac{n}{2N}\right] + 1; \quad B(\vartheta) = \frac{\alpha_n}{\delta(0,n,N)} \\
&\left(\mathcal{E}_N^{\mathrm{dif}}[\hat{\mathbf{a}}_n]\right)^2 = |\alpha_n|^2\left(1 - \frac{1}{\delta(0,n,N)}\right) = |\alpha_n|^2 \frac{2\left[\frac{n}{2N}\right]}{2\left[\frac{n}{2N}\right]+1} = |\alpha_n|^2\frac{2q}{2q+1}\,, \\
&n = 2(qN+m),\ m = 0,1,\dots,N-1,\ q = 1,2,\dots\,.
\end{aligned} \tag{39}$$

Let $\varepsilon_n := \left(\mathcal{E}_n[f]\right)^2$, $\omega_n := (n+1)|\alpha_n(f)|^2$; note that $\omega_n = 0$ for odd n. Then, by (20) $\varepsilon_N = \sum_{n=N}^{\infty} \delta_n$ and according to (37) we have

$$\begin{aligned}
\left(\mathcal{R}_N^{\mathrm{eq}}[f]\right)^2 &= \sum_{n=N}^{\infty}(n+1)\left(\mathcal{E}_N^{\mathrm{dif}}[\hat{\mathbf{a}}_n]\right)^2 = \sum_{n=N}^{\infty}\omega_n \frac{2\left[\frac{n}{2N}\right]}{2\left[\frac{n}{2N}\right]+1} \\
&= \sum_{q=1}^{\infty}\frac{2q}{2q+1}\sum_{n=2qN}^{2(q+1)N-1}\omega_n = \sum_{q=1}^{\infty}\frac{2q}{2q+1}\left(\varepsilon_{2qN} - \varepsilon_{2(q+1)N}\right) = 2\sum_{q=1}^{\infty}\frac{\varepsilon_{2qN}}{4q^2-1},
\end{aligned}$$

whence (13) for f_{rad} follows.

It is also easy to find the minimizer $W(x) := \arg\min \left\| f_{\mathrm{rad}} - \sum_{j=1}^{N} W(\mathbf{x}\cdot\boldsymbol{\theta}_j)\right\|$: if $f_{\mathrm{rad}} = g(|\mathbf{x}|^2)$ and $g(x) \overset{\mathcal{L}^2(0,1)}{=} \sum_{n=0}^{\infty} \grave{g}_n\, l_n(x)$ is the Fourier–Legendre expansion of g, then

$$W(x) \overset{\mathcal{L}^2_\omega(-1,1)}{=} \frac{1}{N}\sum_{n=0}^{\infty} \frac{\grave{g}_n}{\sqrt{2n+1}\left(2\left[\frac{n}{N}\right]+1\right)} u_{2n}(x), \quad \omega(x) = \frac{2}{\pi}\sqrt{1-x^2}.$$

Further, according to (22),$a_n(f_{\mathrm{harm}},\vartheta) := \beta e^{-in\vartheta} + \gamma e^{in\vartheta}$. In this case, $\hat{a}(-n) = \beta$, $\hat{a}(n) = \gamma$; $\hat{a}(l) = 0$, $|l| < n(2)$, $\delta(\pm n, n, N) = \left[\frac{n}{N}\right] + 1$,

$$\left(\mathcal{E}_N^{\mathrm{dif}}[\hat{\mathbf{a}}_n]\right)^2 = \begin{cases} (|\beta|^2 + |\gamma|^2)\frac{q}{q+1} & \text{if} \quad n = qN+m, \quad m = 1, \ldots, N-1 \\ |\beta|^2 + |\gamma|^2 - \frac{|\beta+\gamma|^2}{q+1} & \text{if} \qquad n = qN, \qquad q = 1, 2, \ldots; \end{cases}$$

In particular, $\left(\mathcal{E}_N^{\mathrm{dif}}[\hat{\mathbf{a}}_n]\right)^2 \geq \frac{1}{3}(|\beta|^2 + |\gamma|^2)$ for $n \geq N+1$. After these calculations, the estimates (13) for $\mathcal{R}_N^{\mathrm{eq}}[f_{\mathrm{harm}}]$ are proved like it was done above for $\mathcal{R}_N^{\mathrm{eq}}[f_{\mathrm{rad}}]$. We omit the details.

2.5 Proof of theorem 2

Recovery of integrals. Let us start from the lower estimate of the quantities $\mathcal{Q}^{\mathrm{opt}}[1]$ in (11).

The idea is that *linear combinations* $\sum_{j=1}^{N} w_j D_{n-1}(\vartheta - \vartheta_j)$ *of a small number of shifted Dirichlet kernels of high order are always fast oscillating, and thus cannot approximate slow polynomials* $a(\vartheta)$*, say,* $\equiv 1$. Let us re-write such a linear combination as follows:

$$\sum_{j=1}^{N} w_j \frac{\sin n(\vartheta - \vartheta_j)}{\sin(\vartheta - \vartheta_j)} = F(\vartheta)\sin n\vartheta - G(\vartheta)\cos n\vartheta = H(\vartheta)\sin(n\vartheta - \Phi(\vartheta)), \tag{40}$$

where

$$F(\vartheta) := \sum_{j=1}^{N} \frac{w_j \cos n\vartheta_j}{\sin(\vartheta - \vartheta_j)}, \quad G(\vartheta) := \sum_{j=1}^{N} \frac{w_j \sin n\vartheta_j}{\sin(\vartheta - \vartheta_j)}, \quad \Phi(\vartheta) := \arctan\frac{G(\vartheta)}{F(\vartheta)},$$

and $H(\vartheta) = \sqrt{F^2(\vartheta) + G^2(\vartheta)}$. Let us consider the following sets

$$\begin{aligned} &\mathcal{E}_- := \{\vartheta :\ \vartheta \in [0, 2\pi),\ \sin(n\vartheta - \Phi(\vartheta)) \leq 0\}, \quad \mathcal{E}_+ := [0, 2\pi) \setminus \mathcal{E}_-, \\ &\mathcal{F}_+ := \{\varphi :\ \varphi = n\vartheta - \Phi(\vartheta),\ \vartheta \in \mathcal{E}_+\}, \quad \mathcal{G}_- := \{\varphi :\ \sin\varphi \leq 0,\ \varphi \in [0, 2\pi n)\}, \end{aligned}$$

and prove that the following estimates hold for Lebesgue measures:

$$|\mathrm{meas}\,\mathcal{E}_{\pm} - \pi| \leq \frac{2\pi N}{n} = \frac{2\pi N}{n}. \tag{41}$$

These estimates can be interpreted as that $\Phi(\vartheta)$ is a " slow" perturbation of the function $n\vartheta$ if n is essentially larger than N. Further, it is enough to prove just one estimate $\mathrm{meas}\,\mathcal{E}_- \geq \pi - \frac{2\pi N}{n}$, because $\mathcal{E}_+ \bigcup \mathcal{E}_- = [0, 2\pi)$, $\mathcal{E}_+ \bigcap \mathcal{E}_- = \emptyset$, and the converse relations $\mathrm{meas}\,\mathcal{E}_+ \geq \pi - \frac{2\pi N}{n}$, $\mathrm{meas}\,\mathcal{E}_- \leq \pi + \frac{2\pi N}{n}$ follow by symmetry.

Let $N(t)$ denote the Banach indicatrix of the function $\Phi(\vartheta)$, i. e.

$$N(t) :=^{\#} \{\vartheta \in [0, 2\pi) : \ \Phi(\vartheta) = t\}, \quad |t| \leq \frac{\pi}{2}.$$

By the definition of $\Phi(\vartheta)$, $N(t)$ equals the number of solutions $\vartheta \in [0, 2\pi)$ of the equation $G(\vartheta) = (\tan t) H(\vartheta)$. With a possible exception of one value of t, $N(t) \leq 2N - 1$, because a non-trivial trigonometric polynomial of degree $N - 1$ cannot have more than $2N - 1$ zeros on the period. Consequently, the period can be represented as a union of $M \leq 2N - 1$ disjoint intervals $\bigcup_{k=1}^{M} I_k$ so that $\Phi(\vartheta)$ is monotone and absolutely continuous on each of I_k. Taking also into account that $|\Phi(\vartheta)| \leq \frac{\pi}{2}$ we obtain:

$$\mathrm{meas}\,\mathcal{F}_+ \geq n\,\mathrm{meas}\,\mathcal{E}_+ - M\pi \geq n\,\mathrm{meas}\,\mathcal{E}_+ - 2\pi N + \pi.$$

On the other hand,

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

so that $\mathrm{meas}\,\mathcal{F}_+ \leq \pi n + \pi$. Comparing these estimates we see that $\pi n + \pi \geq n\mathrm{meas}\,\mathcal{E}_+ - 2\pi N + \pi$, or $\mathrm{meas}\,\mathcal{E}_+ \leq \pi + \frac{2\pi N}{n}$. As mentioned above, this implies (41).

It follows that

$$\int_0^{2\pi} \left(1 - \sum_{j=1}^{N} w_j \mathcal{D}_{n-1}(\vartheta - \vartheta_j)\right)^2 \mu(d\vartheta) = \int_0^{2\pi} [1 - H(\vartheta)\sin(n\vartheta - \Phi(\vartheta))]^2 \mu(d\vartheta)$$
$$\geq \int_{\mathcal{E}_-} \mu(d\vartheta) = \frac{\mathrm{meas}\,\mathcal{E}_-}{2\pi} \geq \frac{1}{2}\left(1 - \frac{2N}{n}\right),$$

and since the lower bound on the right is independent from the selection of $\vec{w}$ and $\vec{\vartheta}$, this completes the proof of the lower estimate of $\mathcal{Q}^{\mathrm{opt}}[1]$ in (11).

The lower estimates of $\mathcal{R}_N^{\mathrm{fr}}[f_{\mathrm{rad}}]$ in (14) follow from (7), lower estimates of $\mathcal{Q}_{n,N}^{\mathrm{opt}}[1]$ in (11) and the comparison result (13) of $\mathcal{R}_N^{\mathrm{eq}}[f_{\mathrm{rad}}]$ and $\mathcal{E}_{2N}[f_{\mathrm{rad}}]$:

$$\left(\mathcal{R}_N^{\mathrm{fr}}[f_{\mathrm{rad}}]\right)^2 \geq \sum_{m=N}^{\infty} (2m+1)|\alpha_{2m}|^2 \left(\mathcal{Q}_{2m,N}^{\mathrm{opt}}[1]\right)^2 \geq \frac{1}{2} \sum_{m=N}^{\infty} (2m+1)|\alpha_{2m}|^2 \left(1 - \frac{N}{m}\right)$$
$$\geq \sup_{M>N} \frac{M-N}{2M} \sum_{m=M}^{\infty} (2m+1)|\alpha_{2m}|^2 = \sup_{M>N} \frac{M-N}{2M} \left(\mathcal{E}_{2M}[f_{\mathrm{rad}}]\right)^2 \geq \sup_{M>N} \frac{M-N}{2M} \left(\mathcal{R}_M^{\mathrm{eq}}[f_{\mathrm{rad}}]\right)^2.$$

Let us prove the upper estimates of $\mathcal{Q}_{n,N}^{\mathrm{opt}}[1]$ in (11) (the suggested method is not related with ridge approximation problem because the nodes depend on n). Let $\mu := \frac{n}{2}+1$, $l := \mu - N$, $\vartheta_j = \vartheta_j^{(n)} := \frac{\pi j}{\mu}$, $j = 0, \pm 1, \ldots$, and consider the "incomplete" quadrature formula of rectangles, with the nodes $\{\vartheta_j\}$, $j = 1, \ldots, N$ and weights $w_j := \frac{1}{\mu}$.

The idea of such formula belongs to E.A. Rakhmanov (personal communication). Let us extend this formula by adding l extra nodes $\vartheta_j := \frac{\pi j}{\mu}$, $j = N+1, N+2, \ldots, \mu$. Since

$$\frac{1}{\mu} \sum_{j=1}^{\mu} D_n(\vartheta - \vartheta_j) = \frac{1}{\mu} \sum_{j=1}^{\mu} \sum_{|m| \leq \mu-1} e^{i2m(\vartheta - \vartheta_j)} = \sum_{|m| \leq \mu-1} e^{2im\vartheta} \frac{1}{\mu} \sum_{j=1}^{\mu} e^{-\frac{2\pi imj}{\mu}} \equiv 1,$$

the extended (complete) formula of rectangles is exact for all polynomials in $\mathcal{T}_n^{\pm}$. Thus

$$1 - \frac{1}{\mu} \sum_{j=1}^{N} D_n(\vartheta - \vartheta_j) = \frac{1}{\mu} \sum_{j=N+1}^{\mu} D_n(\vartheta - \vartheta_j).$$

Further, we have

$$\sum_{j=N+1}^{\mu} D_n(\vartheta - \vartheta_j) = \sum_{|m| \leq \mu-1} e^{-2im\vartheta} \sum_{j=N+1}^{\mu} e^{2ij\vartheta_m} = \sum_{|m| \leq \mu-1} e^{-2im\vartheta}\, u_m \mathcal{D}_{l-1}(\vartheta_m),$$

where $u_m = u_{m,l}$ are unimodular complex factors, i. e. $|u_m| = 1$, so that

$$\left\| \sum_{j=N+1}^{\mu} D_n(\vartheta - \vartheta_j),\ \mathcal{L}_{2\pi}^2 \right\|^2 = \sum_{|m| \leq \mu-1} D_{l-1}^2(\vartheta_m) = 2 \sum_{m=0}^{\mu-1} D_{l-1}^2(\vartheta_m) - D_{l-1}^2(0)$$
$$= 2\frac{\mu}{2\pi} \int_0^{2\pi} \mathcal{D}_{l-1}^2(\vartheta)\, d\vartheta - l^2 = 2\mu l - l^2.$$

It follows that

$$\mathcal{Q}^{\mathrm{opt}}_{n,N}[1] \le \frac{1}{\mu}\left\| \sum_{j=N+1}^{\mu} \mathcal{D}_n(\vartheta-\vartheta_j),\ \mathcal{L}^2_{2\pi}\right\| \le \sqrt{\frac{l}{\pi\mu}} = \sqrt{2\left(1-\frac{2N}{n+2}\right)},$$

which completes the proof of the upper estimate in (11).

Collapsed quadratures and ridge functions. Now let us pass to optimization of collapsed quadratures and ridge functions, cf. (1) and (9).

Lemma 4 1) *Let* $a(\vartheta) = \sum_{|m|\le n(2)} \hat{a}_m e^{im\vartheta} \in \mathcal{T}^{\pm}_n$. *Then*

$$\mathcal{Q}^{\mathrm{col}}_{n,N}[a,\varphi] = \min_{P\in\mathcal{P}^1_{N-1}} \sqrt{\sum_{|m|\le n(2)} |\hat{a}_m - P(m)e^{-im\varphi}|^2}. \tag{42}$$

In particular, for $\varphi = 0$ *the problem of optimal collapsed quadrature formula is equivalent to the least squares discrete algebraic approximation of the data sequence* $\{\hat{a}_m\}_{|m|\le n(2)}$*:*

$$\mathcal{Q}^{\mathrm{col}}_{n,N}[a,0] = \min_{P\in\mathcal{P}^1_{N-1}} \sqrt{\sum_{|m|\le n(2)} |\hat{a}_m - P(m)|^2}. \tag{43}$$

2) *Let* $\{W_j(x)\}_{j=1}^N$ $|x|\le 1$ *be an arbitrary set of sufficiently smooth single variate functions,* $W_j(x) = \sum_{n=0}^{\infty} \check{W}_{j,n} u_n(x)$ *– the Fourier – Chebyshev expansion of the function* $W_j(x)$. *Then*

$$a_n\left(\sum_{j=1}^{N}\left(\frac{i\partial}{\partial\varphi}\right)^{j-1} W_j(\mathbf{x}\cdot\boldsymbol{\varphi}),\ \vartheta\right) = \frac{1}{n+1}\sum_{|m|\le n(2)} P(m)e^{im(\vartheta-\varphi)} \tag{44}$$

where $P(x) := \sum_{j=1}^{N} \check{W}_{j,n} x^{j-1}$.

3) *Let* $f(\mathbf{x}) \in \mathcal{L}^2(\mathbb{B}^2)$ *and* $a_n(f,\vartheta) = \sum_{|m|\le n(2)} \hat{a}_{m,n}(f)e^{im\vartheta}$ *be the Chebyshev momenta of* f. *Then*

$$\mathcal{R}^{\mathrm{col}}_N[f,0] = \sqrt{\sum_{n=N}^{\infty}(n+1)\min_{P\in\mathcal{P}^1_{N-1}}\sum_{|m|\le n(2)} |\hat{a}_{m,n}(f) - P(m)|^2}. \tag{45}$$

Poof. 1) Fix polynomials $P \in \mathcal{P}^1_{n-1}$ and $T \in \mathcal{T}^{\pm}_n$. Then

$$P\left(\frac{id}{d\varphi}\right)T(\varphi) = \int_0^{2\pi} T(\vartheta)P\left(\frac{i\partial}{\partial\varphi}\right)D_n(\varphi-\vartheta)\,\mu(d\vartheta),$$

$$\int_0^{2\pi} T(\vartheta)a(\vartheta)\,\mu(d\vartheta) - P\left(\frac{id}{d\varphi}\right)T(\varphi) = \int_0^{2\pi} T(\vartheta)\left(a(\vartheta) - \sum_{|m|\le n(2)} P(m)e^{im(\vartheta-\varphi)}\right)\,\mu(d\vartheta);$$

$$\sup_{T\in \mathbb{B}(\mathcal{T}_n^{\pm})}\left|\int_0^{2\pi} T(\vartheta)a(\vartheta)\,\mu(d\vartheta) - P\left(\frac{id}{d\varphi}\right)T(\varphi)\right| = \left\| a(\vartheta) - \sum_{|m|\le n(2)} P(m)e^{im(\vartheta-\varphi)},\ \mathcal{L}^2_{2\pi}\right\|,$$

whence (42) follows by Parceval identity and minimization in $P\in\mathcal{P}^1_{N-1}$.

2) Let us apply termwise differentiation in the angular variable φ to the expansion

$$W(\mathbf{x}\cdot\boldsymbol{\varphi}) = \int_0^{2\pi}\left(\sum_{n=0}^{\infty}\check{W}_n\mathcal{D}_n(\vartheta-\varphi)u_n(\mathbf{x}\cdot\boldsymbol{\theta})\right)\,\mu(d\vartheta).$$

Then

$$\left(\frac{i\partial}{\partial\varphi}\right)^{j-1} W_j(\mathbf{x}\cdot\boldsymbol{\varphi}) = \int_0^{2\pi}\left(\sum_{n=0}^{\infty}\check{W}_{j,n}u_n(\mathbf{x}\cdot\boldsymbol{\theta})\left(\frac{i\partial}{\partial\varphi}\right)^{j-1}\mathcal{D}_n(\vartheta-\varphi)\right)\,\mu(d\vartheta)$$

$$= \int_0^{2\pi}\sum_{n=0}^{\infty}\check{W}_{j,n}u_n(\mathbf{x}\cdot\boldsymbol{\theta})\left(\sum_{|m|\le n(2)} m^{j-1}e^{im(\vartheta-\varphi)}\right)\,\mu(d\vartheta),$$

whence (44) follows by addition in j.

3) (45)is an immediadte corollary of (43) and (9).

Lemma 5 *For each set of N points $Z=\{z_j\}_1^N$, $|z_j|=1$ and each natural number m there exists a polynomial $P(z)=P_Z(z)$ satisfying*

$$P(z)\in\mathcal{P}^1_{mN},\quad P(0)=1,\ P(z_j)=0,\ j=1,\dots,N;\quad \max_{|z|\le 1}|P(z)|\le e^{\frac{2N}{m}}. \tag{46}$$

In particular, for $n\ge N$ and fixed complex numbers β, γ

$$\mathcal{Q}^{\rm opt}_{n,N}[e^{\pm in\vartheta}] = \mathcal{Q}^{\rm opt}_N[1,\mathbb{B}(\mathcal{P}^1_n)] \ge e^{-\frac{4N^2}{n}},\quad \mathcal{Q}^{\rm opt}_{n,N}[\beta e^{in\vartheta}+\gamma e^{-in\vartheta}] \ge \sqrt{|\beta|^2+|\gamma|^2}\,e^{-\frac{8N^2}{n}}. \tag{47}$$

Lemma 6 *Let $n\ge N$ and β, γ be fixed complex numbers. Then*[3]

$$\mathcal{Q}^{\rm col}_{n,N}[\beta e^{in\vartheta}+\gamma e^{-in\vartheta},0] = \min_{P\in\mathcal{P}^1_{N-1}}\sqrt{|\beta-P(-n)|^2+\sum_{|m|<n(2)}|P(m)|^2+|\gamma-P(n)|^2}, \tag{48}$$

[3]The problem concerning the structure of "minimizer" in $\mathcal{Q}^{\rm opt}_{n,N}[e^{\pm in\vartheta}]$ remains open. In particular, it seems interesting to clarify when the collapsed nodes are exactly optimal, i. e. $\mathcal{Q}^{\rm opt}_{n,N}[e^{\pm in\vartheta}] = \mathcal{Q}^{\rm col}_{n,N}[e^{\pm in\vartheta}]$.

and further

$$\mathcal{Q}^{\rm col}_{n,N}[\beta e^{in\vartheta}+\gamma e^{-in\vartheta},0] \le \sqrt{|\beta|^2+|\gamma|^2}\,\min\left(1,\sqrt{2n}e^{-\frac{N}{\sqrt{n}}}\right), \quad n\ge N\ge 5. \tag{49}$$

Proof of lemma 5. Let us make use of the known solution of the following problem posed by G. Hálász. Find

$$\kappa_m := \min_{p\in\mathcal{P}^{(1,0)}_m}\max_{|z|\le 1}|p(z)|, \quad \text{where} \quad \mathcal{P}^{(1,0)}_m := \{p(z)\in\mathcal{P}^1_m,\ p(0)=1,\ p(1)=0\}$$

and the extremal polynomial for which the min is attained. The exact solution was found in [8]: $\gamma_m = \left(\sec\frac{\pi}{2(m+1)}\right)^{m+1}$, and it is interesting to note that a properly scaled Chebyshev polynomial of the first kind is extremal. For our purpose, a simplified version, namely, the estimate $\kappa_m \le 1+\frac{2}{m}$ is sufficient. The latter was proved by H.L. Motgomery (see [7], Ch. 5). Let us take $P(z) := \prod_{j=1}^N p(zz_j^{-1})$, where $p(z)\in\mathcal{P}^{(1,0)}_m$ is the mth Hálász' extremal polynomial. Then obviously $P(z)\in\mathcal{P}^1_{mN}$, $P(0)=1,\ P(z_j)=0,\ j=1,\dots,N;\quad \max_{|z|\le 1}|P(z)| \le (\kappa_m)^N \le \left(1+\frac{2}{m}\right)^N \le e^{\frac{2N}{m}}$, which proves (46).

A polynomial $T(\vartheta)\in \mathbb{B}(\mathcal{T}^{\pm}_n)$ can be represented as $T(\vartheta)=e^{-in\vartheta}P(e^{2i\vartheta}),\ P\in\mathbb{B}(\mathcal{P}^1_n)$, and respectively

$$\int_0^{2\pi} T(\vartheta)e^{-in\vartheta}\,\mu(d\vartheta) - \sum_{j=1}^N w_jT(\vartheta_j) = P(0) - \sum_{j=1}^N \left(w_j e^{-in\vartheta_j}\right)P(z_j), \quad z_j := e^{2i\vartheta_j}, \tag{50}$$

whence the equalities $\mathcal{Q}^{\rm opt}_{n,N}[e^{\pm in\vartheta}] = \mathcal{Q}^{\rm opt}_N[1,\mathbb{B}(\mathcal{P}^1_n)]$ and $\mathcal{Q}^{\rm col}_{n,N}[e^{\pm in\vartheta}] = \mathcal{Q}^{\rm col}_N[1,\mathbb{B}(\mathcal{P}^1_n)]$ easily follow.

The lower estimates in (47) follow from (46). Indeed, let $m := \left[\frac{n}{N}\right]$ and for a given set of nodal points $Z=\{z_j\}_1^N,\ |z_j|=1$ consider the polynomial $\Pi(z) := e^{-\frac{2N}{m}}P_Z(z)$ where $P_Z(z)$ satisfies (**??**). Then $\Pi\in\mathcal{P}^1_{mN}\subset\mathcal{P}^1_n,\ \Pi(z_j)=0$ and $\Pi\in\mathbb{B}(\mathcal{P}^1_n)$, because $|\Pi(z)|\le 1,\ |z|\le 1$. Consequently, for each quadrature with the nodes on the set Z and arbitrary weights we have $\Pi(0)-\sum_1^N w_j\Pi(z_j) = \Pi(0) \ge e^{-\frac{2N}{m}} \ge e^{-\frac{4N^2}{n}},\ n\ge N$, which completes the proof of the estimate $\mathcal{Q}^{\rm opt}_N[1,\mathbb{B}(\mathcal{P}^1_n)] \ge e^{-\frac{4N^2}{n}}$. Further, the lower estimate for $\mathcal{Q}^{\rm opt}_{n,N}[\beta e^{in\vartheta}+\gamma e^{-in\vartheta}]$ follows by splitting polynomials $T(\vartheta)\in\mathcal{T}^{\pm}_n$ into parts $T_+(\vartheta) := \sum_{0\le m\le n(2)}\hat{T}(m)e^{im\vartheta},\ T_-(\vartheta) := T(\vartheta)-T_+(\vartheta)$ and subsequent reduction to the lower estimate of $\mathcal{Q}^{\rm opt}_N\left[1,\mathbb{B}\left(\mathcal{P}^1_{\frac{n}{2}}\right)\right]$. We omit the details.

Proof of lemma 6. First of all, (48) is a particular case of (43). Further, let $T_M(x)=\cos(M\arccos x),\quad x\in$

$[-1,1]$ denote the Chebyshev polynomials of the first kind, $M=N-1$ or $M=N-2$, and

$$P_{n,M}(x) := \frac{T_M\left(\left(1+\frac{2}{n}\right)\frac{x}{n}\right)}{T_M\left(1+\frac{2}{n}\right)}, \quad P_{n,N}(x,\beta,\gamma) := \frac{\gamma-(-1)^N\beta}{2}P_{n,N-1}(x)+\frac{\gamma+(-1)^N\beta}{2}P_{n,N-2}(x).$$

These polynomials are in $\mathcal{P}^1_{N-1}$, and since $P_{n,M}(1)=1,\ P_{n,M}(-1)=(-1)^M$, we see that $P_{n,N}(-n,\beta,\gamma)=\beta,\ P_{n,N}(n,\beta,\gamma)=\gamma$. For $|m|\le n-2$ we have $\left|T_M\left(\left(1+\frac{2}{n}\right)\frac{m}{n}\right)\right|\le 1$, so that (49) follows from the estimates

$$\begin{aligned}
&\sum_{|m|<n(2)} |P_{n,M}(m)|^2 \le \frac{n-1}{T_M^2\left(1+\frac{2}{n}\right)} \le 4(n-1)e^{-\frac{4M}{\sqrt{n+1}}},\\
&\sum_{|m|<n(2)} |P_{n,N}(m,\beta,\gamma)|^2 \le 2(n-1)(|\beta|^2+|\gamma|^2)e^{-\frac{4(N-2)}{\sqrt{n+1}}} \le 2n(|\beta|^2+|\gamma|^2)e^{-\frac{2N}{\sqrt{n}}}, \quad N\ge 5.
\end{aligned}$$

In the above, we made use of the well-known estimates of Chebyshev polynomials $T_M(x)$ for $x>1$:

$$T_M(x) = \frac{(x+\sqrt{x^2-1})^M+(x-\sqrt{x^2-1})^M}{2} \ge \frac{1}{2}\exp\left(M\int_1^x \frac{dy}{\sqrt{y^2-1}}\right) \ge \frac{1}{2}\exp\left(2M\sqrt{\frac{x-1}{x+1}}\right).$$

Relations (47) and (49) imply the claim (12) of theorem 2. They also imply (15) of theorem 3. Indeed, let $\delta_n := \sqrt{|\beta_n(f)|^2+|\gamma_n(f)|^2}$. Then according to (7), (20) and (47) we have

$$\begin{aligned}
&\left(\mathcal{R}^{\mathrm{fr}}_N[f]\right)^2 \ge \sum_{n=N}^{\infty}(n+1)\left(\mathcal{Q}^{\mathrm{opt}}_{n,N}[\beta_n e^{in\vartheta}+\gamma_n e^{-in\vartheta}]\right)^2 \ge \sum_{n=N}^{\infty}(n+1)\delta_n^2 e^{-\frac{16N^2}{n}}\\
&\ge e^{-16}\sum_{n=N^2}^{\infty}(n+1)\delta_n^2 = e^{-16}\left(\mathcal{E}_{N^2}[f]\right)^2,
\end{aligned}$$

whence the lower estimate of $\mathcal{R}^{\mathrm{fr}}_N[f_{\mathrm{harm}}]$ in (15) follows. Finally, for $M\ge N\ge 5$

$$\begin{aligned}
&\left(\mathcal{R}^{\mathrm{col}}_N[f,0]\right)^2 = \sum_{n=N}^{\infty}(n+1)\left(\mathcal{Q}^{\mathrm{col}}_{n,N}[\beta_n e^{in\vartheta}+\gamma_n e^{-in\vartheta},0]\right)^2 \le \sum_{n=N}^{\infty}(n+1)\delta_n^2\min\left(1,2ne^{-\frac{2N}{\sqrt{n}}}\right)\\
&\le 2Me^{-\frac{2N}{\sqrt{M}}}\sum_{n=N}^{M}(n+1)\delta_n^2+\sum_{n=M+1}^{\infty}(n+1)\delta_n^2 \le 2Me^{-\frac{2N}{\sqrt{M}}}\left(\mathcal{E}_N[f]\right)^2+\left(\mathcal{E}_{M+1}[f]\right)^2.
\end{aligned}$$

which finishes the proof of theorem 3.

2.6 Comments

It is hard to indicate the primary source in literature of the general problem of ridge approximation. For the author, such a source was [3]. Problems of ridge approximation naturally appear in applications, such as Radon transformations and local tomography [4], [2], and in geometry [9].

Recently, a considerable attention was tributed to *constrained ridge approximation*, when restrictions of different kinds are imposed on the wave profiles. This includes a version of the so-called neural networks approximations when the profiles $W_j(x)$ are assumed to be piecewise constant functions or some more smooth splines, cf. e. g. [10].

It should be noted that a wide circle of problems remains open, concerning free and constrained ridge approximation in functional metrics other than $\mathcal{L}^2$, in particular, uniform metric $\mathcal{L}^\infty$. Very promising and difficult seem to be generalizations to functions of more than two variables, see [11], [12].

Estimates of free ridge approximation, in particular, comparison of $\mathcal{R}_N^{\mathrm{fr}}[f]$ with $\mathcal{E}_N[f]$ and $\mathcal{R}_N^{\mathrm{eq}}[f]$, were discussed by D.L. Donoho and I.M. Johnstone [6]. In [6], cf. p. 73, a conjecture was made that *equispaced wave vectors are best for* f_{rad} *and* f_{harm} (for brevity, conjecture(*eq*) in the sequel) in the problem $\mathcal{R}_N^{\mathrm{fr}}$ posed in the weighted space $\mathcal{L}_w^2(\mathbb{R}^2)$ with the Gaussian weight $w(\mathbf{x}) := e^{-\pi|\mathbf{x}|^2}$:

$$\mathcal{R}_N^{\mathrm{fr}}\left[f, \mathcal{L}_w^2(\mathbb{R}^2)\right] := \inf_{R\in\mathcal{W}_N^{\mathrm{fr}}} \sqrt{\iint_{\mathbb{R}^2} |f(\mathbf{x}) - R(\mathbf{x})|^2 w(\mathbf{x})\, d\mathbf{x}}.$$

The strong form of conjecture(*eq*) is that exact equalities $\mathcal{R}_N^{\mathrm{fr}}\left[f,\ \mathcal{L}_w^2(\mathbb{R}^2)\right] = \mathcal{R}_N^{\mathrm{eq}}\left[f,\ \mathcal{L}_w^2(\mathbb{R}^2)\right]$ hold for all f_{rad} and f_{harm}.

A weaker version of conjecture(*eq*), concerning orders of ridge approximation of f_{rad} in the metric $\mathcal{L}^2(\mathbb{B}^2)$, is confirmed by theorem 3 of the present paper (cf. (13) and (14), corollary 1 and [5]).

On the contrary, (15) implies that for f_{harm} and approximation in the metric of $\mathcal{L}^2(\mathbb{B}^2)$ the conjecture(*eq*) principally fails to be true.

(15) apparently represents a new effect in non-linear ridge approximation. The complete freedom in the choice of N wave vectors does bring an essential gain in the orders of approximation. Harmonic functions represent a wide set for which this effect takes place.

Upper bounds of $\mathcal{R}_N^{\mathrm{fr}}(f)$ on Sobolev classes were considered by V.E. Maiorov [13], [12] and V.N. Temlyakov [14]. In particular, in the recent preprint [12] Maiorov considers the upper bounds of free ridge approximation on classes $W_2^{r,d}$ in the unit ball $\mathbb{B}^d$, $d \geq 2$ of d-dimensional Euclidean space $\mathbb{R}^{\mathrm{d}}$:

$$\operatorname{dist}\left(W_2^{r,d},\ \mathcal{W}_N;\ \mathcal{L}^2(\mathbb{B}^d)\right) := \sup_{f\in W_2^{r,d}} \mathcal{R}_N^{\mathrm{fr}}[f]; \quad W_2^{r,d} := \left\{f(\mathbf{x}) : \quad \max_{\rho\leq r} \left\|\mathcal{D}^\rho f;\ \mathcal{L}^2(\mathbb{B}^d)\right\| \leq 1\right\}.$$

As an improvement and generalization of earlier results of the works [14] and [13], dealing with the case $d = 2$, the main result of [12] (Theorem 1) is the exact order estimate dist $\left(W_2^{r,d},\ \mathcal{W}_N,\ \mathcal{L}^2(\mathbb{B}^d)\right) \sim N^{-\frac{r}{d-1}}, \quad N \to \infty$.

Returning to the conjecture(eq), let us note that there are corresponding counterexamples to its' strong version for $f = f_{\text{rad}}$, too (cf. M.E. Davison and F.A. Grunbaum [2], p. 104). There exist radial polynomials $P(|\mathbf{x}|^2)$, $\deg P = 2, 3, \dots$ such that strict inequalities $\mathcal{R}_2^{\text{fr}}\left[P(|\mathbf{x}|^2)\right] < \mathcal{R}_2^{\text{eq}}\left[P(|\mathbf{x}|^2)\right]$ hold in the weighted spaces $\mathcal{L}^2_\omega(\mathbb{B}^2)$, where $\omega(\mathbf{x}) = (1 - |\mathbf{x}|^2)^\lambda$, $\mathbf{x} \in \mathbb{B}^2$ is a Gegenbauer weight. E. g., in the exact solution of the extremal problem $\mathcal{R}_2^{\text{fr}}\left[|\mathbf{x}|^4 - |\mathbf{x}|^2, \mathcal{L}^2(\mathbb{B}^2)\right]$, the angle between optimal directions $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ satisfies $\vartheta_2 - \vartheta_1 = \arccos\sqrt{\frac{3}{8}}$, i. e. $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ are not mutually perpendicular, cf. [5]. In fact, this peculiarity is equivalent to the strict inequality $\mathcal{Q}_{4,2}^{\text{opt}}[1] < \mathcal{Q}_{4,2}^{\text{eq}}[1]$ for optimal quadratures with just two nodes for the class $\mathbb{B}(\mathcal{T}_n^{\pm})$, or

$$\mathcal{Q}_2^{\text{opt}}[1, \mathbb{B}(\mathcal{T}_2)] < \mathcal{Q}_2^{\text{eq}}[1, \mathbb{B}(\mathcal{T}_2)]. \tag{51}$$

Here $\mathbb{B}(\mathcal{T}_n)$ denotes the $\mathcal{L}^2_{2\pi}$-unit ball in the subspace $\mathcal{T}_n$ of all trigonometric polynomials of degree n. It is easy to see that $\mathcal{Q}_N^{\text{opt}}[1, \mathbb{B}(\mathcal{T}_n)] = \mathcal{Q}_N^{\text{opt}}[1, \mathbb{B}(\mathcal{T}_{2n}^{\pm})] := \mathcal{Q}_{2n,N}^{\text{opt}}[1]$.

Kolmogorov – Nikol'skii problem has a long history, cf. [1],[15]-[17]. Starting from the sixties, an essential part of efforts was concentrated on the conjecture(eq) for quadrature formulas: *equispaced nodes and the formula of rectangles* $\frac{1}{N}\sum_{j=1}^{N} f\left(\frac{2\pi}{N}\right)$ *are optimal for recovery of integrals* $\int_0^{2\pi} f(\vartheta)\mu(d\vartheta)$ *on all periodic classes* $W^r\left(\mathcal{L}^p_{2\pi}\right)$, $1 \le p \le \infty$. That this conjecture is right for $p = \infty$ and all natural $r \ge 4$ (low smootheness cases $r = 1, 2, 3$ had been solved earlier) was proved by V.P. Motornyi [16]. Subsequently A.A. Zhensykbaev [17] extended this result of Motornyi on all $p \in [1, \infty)$. For large indices r of differentiability (in fact, for $r \ge 4$), one of the difficulties was in the *existence* of the optimal quadrature formula, in particular, the proof that the optimal nodes do not collapse.

To find the limits for validity of the conjecture(eq), the author [18], [19] considered modifications of the periodic classes $W^r\left(\mathcal{L}^p_{2\pi}\right)$:

$$\left\| P\left(\frac{d}{d\vartheta}\right) f(\vartheta),\ \mathcal{L}^p_{2\pi} \right\| \le 1,$$

where $P\left(\frac{d}{d\vartheta}\right)$ is a fixed differential operator. It turned out that the solution qualitatively depends on the spectrum of P. In particular, for classes of the type $\|f''(\vartheta) + \omega^2 f(\vartheta),\ \mathcal{L}^p_{2\pi}\| \le 1$ (oscillatory differential operators) the conjecture(eq) fails to be true, at least for some small initial values of N. Relation (51) provides another counter-example, for the class $\mathbb{B}(\mathcal{T}_2)$.

Lower bounds for the quantities $\mathcal{Q}_N^{\text{opt}}[1, \mathbb{B}(\mathcal{T}_n)]$ and their multivariate analogs were considered by V.N. Temlyakov [20]. Basing on results of B.S. Kashin [21] (cf. also [22]), it was proved in [20] that if $N \leq (1-\varepsilon)n$, where $\varepsilon > 0$ then $\mathcal{Q}_N^{\rm opt}[1, \mathbb{B}(\mathcal{T}_n)]$ are bounded below, i. e. $\mathcal{Q}_N^{\rm opt}[1, \mathbb{B}(\mathcal{T}_n)] \geq c_\varepsilon > 0$. Using the latter result, it was established in [5] that $\forall \varepsilon > 0$, $\exists c_\varepsilon > 0 :\ \mathcal{R}_N^{\rm fr}(f_{\rm rad}) \geq c_\varepsilon \mathcal{E}_{2(1+\varepsilon)N}(f_{\rm rad})$. (11) and (14) represent more explicit versions of these results.

Preliminary upper estimates of the righthand side of (48) based on Chebyshev polynomials $T_M(x)$ appeared in discussions with my colleagues at USC P. Petrushev, B. Popov and O. Trifonov. A subsequent improvement of the type (49) was later communicated to the author by I.I. Sharapudinov. Recently, using properties of discrete Chebyshev polynomials, Sharapudinov [23] proved that the condition $\frac{N}{\sqrt{n}} \to \infty$ is necessary and sufficient for

$$\mathcal{Q}_N^{\rm col}[1, \mathbb{B}(\mathcal{P}_n^1] = \min_{P \in \mathcal{P}_N^1} \sqrt{(1-P(0))^2 + \sum_{m=1}^{n} P^2(m)} \to 0.$$

(In view of this result, it seems likely that the factor $\sqrt{2n}$ on the right of (49) can be substituted by a constant.) That the condition $\frac{N}{\sqrt{n}} \to \infty$ is necessary, follows also from the lower estimate (47) of errors of quadrature formulas with free nodes.

2.7 Acknowledgements

The author is grateful to his colleagues at USC R. DeVore, R. Howard, P. Petrushev, B. Popov, V.N. Temlyakov, O. Trifonov for many useful discussions and valuable comments. V.N. Temlyakov indicated on the above mentioned results in [20]. I.I. Sharapudinov communicated with the author concerning discrete Chebyshev polynomials. The idea of using incomplete formulas of rectangles in upper estimates of $\mathcal{Q}_{n,N}^{\rm opt}[1, \mathbb{B}(\mathcal{T}_n)]$ belongs to E.A. Rakhmanov. The lower estimate (47) of $\mathcal{Q}_N^{\rm fr}[1, \mathbb{B}(\mathcal{P}_n^1]$ was established in a close collaboration with B.S. Kashin during his recent visit to USC (April – May 1998). The author also had a number of very useful e-mail communications with B. Bojanov and V. Totik concerning optimal quadrature formulas.

An especially warm gratitude is directed to S.M. Nikol'skii, who visited USC in October – November 1997. The author enjoyed several personal discussions with him of the results of the present paper.

The work was supported by the NSF Grant DMS-9706883.

References

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

[2] M.E. Davison snd F.A. Grunbaum. *Tomographic reconstruction with arbitrary directions.* Comm. Pure and Appl. Math., **34**(1981), pp. 77 – 120.

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

[4] A.G. Ramm, A.I. Katsevich. *The Radon transform and local tomography.* 1996, CRC Press, Inc., 485 pp. ISBN 0-8493-9492-9

[5] K.I. Oskolkov. *Ridge approximation, Chebyshev – Fourier analysis and optimal quadrature formulas.* Proc. Steklov Math. Inst., **219** (1997), pp. 269 – 285 (in Russian).

[6] D.L. Donoho and I.M. Johnstone. *Projection-based approximation and a duality with kernel methods.* The Annals of Statistics, **17** (1989), No. 1, pp. 58 – 106.

[7] H.L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis.* expository lectures from the NSF-CBMS Regional Conference held at Kansas State University, Manhattan, Kansas, May 22-25, 1990. (Regional conference series in mathematics, ISSN 0160-7642; no.84), AMS, 220 pp.

[8] M. Lachance, E.B. Saff and R.S. Varga. *Inequalities for polynomials with a prescribed zero.* Mathematische Zeitschrift, **168** (1979), pp. 105-116.

[9] H. Groemer. Geometric applications of Fourier series and spherical harmonics. (Encyclopedia of Mathematics and its applications, **61**). Cambridge University Press, 1996, 329 pp.

[10] 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.

[11] P.P. Petrushev. *Approximation by ridge functions and neural networks.* Research report 1997:18, Industrial mathematics initiative. USC, 41 pp. To appear in: SIAM Journal on Mathematical Analysis (1998).

[12] V.E. Majorov. *On best approximation by ridge functions.* Preprint, Department of Mathematics, Technion, Haifa, Israel, January 14, 1998. 27 pp.

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

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

[15] N.P. Korneichuk. Splines in approximation theory, "Nauka", 1984, 352 pp. (in Russian).

[16] V.P. Motornyi. *On the best quadrature formula* $\sum_{k=1}^{n} p_k f(x_k)$ *for certain classes of periodic differentiable functions.* – Izvestiia AN SSSR, Ser. matem., **38**, no. 3 (1974), pp. 583 – 614 (in Russian).

[17] A.A. Zhensykbaev. *Monosplines of minimal norm and optimal quadrature formulas.* – Uspekhi Matem. Nauk, **36**, no. 4 (1981), pp. 107 – 159 (in Russian).

[18] K.I. Oskolkov. *On the optimality of the quadrature formula with equidistant nodes on classes of periodic functions.* Soviet Math. Dokl., **20**(1979), No. 6, pp. 1211 – 1214.

[19] K.I. Oskolkov. *On optimal quadrature formulae on certain classes of periodic functions.* Appl. Math. Optim., **8**(1982), pp. 245 – 263.

[20] V.N. Temlyakov. *Approximation of periodic functions.* ISBN 1-56072-131-6, Nova Science Publishers, Inc. (Ch. 1, Theorem 1.3, and Ch. 2, Theorem 1.1).

[21] 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)).

[22] 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).

[23] I.I. Sharapudinov. *On a new application of Chebyshev polynomials orthogonal on equispaced nets.* Manuscript (1998), 6pp. Submitted to Mathematical Notes (in Russian).