Technische Universität Chemnitz
Fakultät für Mathematik
Professur für Angewandte Analysis

Masterarbeit

zur Erlangung des akademischen Grades

Master of Science

vorgelegt von

Moritz Moeller

Gelfand numbers and best m-term trigonometric approximation for weighted mixed Wiener classes in L_2

Betreuer: Prof. Dr. Tino Ullrich
Gutachter: Dr. Martin Schäfer

Chemnitz, 23.02.23

Abstract

The best m-term approximation has since its first formulation by Stechkin in 1955 been a rather theoretical subject of study in approximation theory. Recently however Jahn, T.Ullrich and Voigtlaender have found some practical application for it by using it in a new bound on the sampling numbers. One important class of spaces where this bound can give an improvement over existing ones are mixed weighted Wiener spaces. Motivated by this a new bound for the best m-term approximation in these spaces will be developed in this thesis. This is achieved by using techniques from hyperbolic cross approximation. In addition a general optimality bound in form of a bound for the Gelfand numbers of these spaces will also be provided.

Contents

1 Introduction **1**

2 Setting and basics **3**
2.1 Setting . . . 3
2.2 Gelfand numbers . . . 4
2.3 Best m-term approximation . . . 6
2.4 Sampling numbers . . . 8
2.5 Existing results for s-numbers . . . 9

3 Gelfand numbers of diagonal operators **11**
3.1 Diagonal operators between ℓ_p and ℓ_2 . . . 11
3.2 Diagonal operators between $\ell_p(\mathbb{Z}^d)$ and $\ell_2(\mathbb{Z}^d)$ . . . 14

4 Best m-term trigonometric approximation of mixed weighted Wiener classes **21**
4.1 Weighted Wiener classes with quasi norms . . . 21
4.2 Results for $p > 1$ . . . 25
4.3 Basis pursuit denoising . . . 26
4.4 Numerical experiments . . . 28

5 Sampling numbers of mixed weighted Wiener spaces in L_2 **31**
5.1 Linear and non-linear sampling numbers . . . 31

6 Conclusion **33**

Bibliography **35**

1 Introduction

The periodic weighted Wiener classes with mixed weights $\mathcal{A}_p^{\alpha,d}$ and their embeddings into Lebesgue spaces have recently been studied in [15, 16, 18], where different upper and lower bounds were shown for different (quasi) s-numbers. V.D. Nguyen, V.K. Nguyen and Sickel examined the behaviour of some s-numbers of the embedding from $\mathcal{A}_1^{\alpha,d}$ into L_2 in their new paper [18] and showed bounds on the approximation, Bernstein, Kolmogorov and Weyl numbers. The authors, however, did study neither the Gelfand numbers nor the best m-term approximation of these embeddings.

This thesis seeks to remedy this fact by proving an asymptotic bound for Gelfand numbers of the identical embedding into L_2. This will be achieved via decomposing the error into different parts and estimating the finite-dimensional ones with results about Gelfand numbers of operators between some finite-dimensional ℓ_p spaces.

Theorem 1.1. *For $n, d \in \mathbb{N}$, $0 < p \leq 1$ and $\alpha > 0$ it holds*

$$c_n(D_\alpha \colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) \asymp n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha}$$

where $\lambda = \frac{1}{p} - \frac{1}{2}$, and

$$D_\alpha(x_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}^d} = \Big(\prod_{j=1}^{d}(1+|k_j|)^{-\alpha} x_{\mathbf{k}}\Big)_{\mathbf{k}\in\mathbb{Z}^d}.$$

One motivation to study Gelfand numbers is, that they give a lower bound on the non-linear sampling numbers or in other words, they bound the error of the best non-linear reconstruction of a function from linear samples from below.

Another asymptotic quantity of interest is the best m-term trigonometric approximation of $\mathcal{A}_p^{\alpha,d}$. Since Wiener classes, in general, are smoothness classes, the Fourier coefficients of functions from them decay fast. This makes them a prime target for best trigonometric m-term approximation, since functions in the unit ball of these spaces can not have large Fourier coefficients with high absolute value. To study these, greedy approximation results on hyperbolic crosses from [29] were employed. The results for the best m-term approximation then allow for the formulation of some basis pursuit denoising bounds that were supplemented with numerical experiments.

The best m-term approximation has for a long time been a purely theoretical subject of study. However, very recently Jahn, T. Ullrich and Voigtlaender have shown in [14] that the sampling numbers measured in L_2 can be bounded by a sum of the best m-term approximation and the best trigonometric approximation width measured in L_∞ .

Proposition 1.2 ([14, Theorem 3.1]). *For any $d \in \mathbb{N}$ and any quasi-normed function space $\mathcal{F} \hookrightarrow L_\infty(\mathbb{T}^d)$ it holds for $n, M \in \mathbb{N}$*

$$\varrho_{[nd\log(n)^3\log(M)]}\big(\mathcal{F}\big)_2 \lesssim \sigma_n\big(\mathcal{F}\big)_\infty + E_{[-M,M]^d\cap\mathbb{Z}^d}\big(\mathcal{F}\big)_\infty \tag{1.1}$$

And while the second term can be controlled in the setting of mixed weighted Wiener classes (or even completely prevented, see [5]) the bounds for the best m-term approximation in general are not known. While the authors showed a bound for $\alpha > \frac{1}{2}$ they did not give an estimate for smaller α.

Proposition 1.3 ([14, Lemma 4.3 (i)]). *Let $\alpha > \frac{1}{2}$ then for all $n \in \mathbb{N}$ it holds*

$$\sigma_n\big(\mathcal{A}_1^{\alpha,d}\big)_\infty \lesssim n^{-(\alpha+\frac{1}{2})}\log(n)^{(d-1)\alpha+\frac{1}{2}}. \tag{1.2}$$

The technique used to prove this result cannot be expanded to $0 < \alpha \leq \frac{1}{2}$ since the embedding $\mathcal{A}_2^{\alpha,d} \to L_\infty(\mathbb{T}^d)$ does not hold in this regime. Therefore, the second main objective of this thesis was to develop a new asymptotic bound for the best m-term approximation of mixed weighted Wiener classes for all $\alpha > 0$.

Theorem 1.4. *Let $n, d \in \mathbb{N}$, $0 < p \leq 1$ and $\alpha > 0$, then it holds*

$$n^{-(\alpha+\lambda)}\log(n)^{(d-1)\alpha} \lesssim \sigma_n\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_\infty \lesssim n^{-(\alpha+\lambda)}\log(n)^{(d-1)\alpha+\frac{1}{2}} \tag{1.3}$$

where $\lambda = \frac{1}{p} - \frac{1}{2}$.

This result can then be applied via Proposition 1.2 to achieve the last goal of this thesis, using the above results and their improvements to establish bounds on the sampling numbers of mixed weighted Wiener spaces using the new result from [14] as well as existing bounds in terms of Gelfand and Kolmogorov numbers.

Notation

As usual $\mathbb{N}$ denotes the natural numbers, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $\mathbb{Z}$ denotes the integers, $\mathbb{R}$ the real numbers and $\mathbb{R}_+$ the non-negative real numbers and $\mathbb{C}$ the complex numbers. If not indicated otherwise $\log(\cdot)$ denotes the natural logarithm of its argument, the positive part of $a \in \mathbb{R}$ is defined as $(a)_+ = \max(a, 0)$. $\mathbb{C}^n$ denotes the complex n-space. Vectors and matrices are usually typesetted boldface for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ the inner product $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}\mathbf{y}$ is defined as usual. For any (quasi) normed spaced X its (quasi) norm is denoted by $\|\cdot\|_X$ when $0 < p \leq \infty$ and $X = \ell_p$ or $X = L_p$ this is instead abbreviated by $\|\cdot\|_p$ where ℓ_p denotes the usual space of p-sumable sequences. The space $\ell_p(\mathbb{Z}^d)$ simply denotes its d-dimensional version. For $r > 0$ and a metric space X we denote byB_r^X the sphere with radius r measured in the metric of the space X. For two sequences $\mathbf{a} = (a_n)_{n\in\mathbb{N}_0}$ and $\mathbf{b} = (b_n)_{n\in\mathbb{N}_0}$ we write $\mathbf{a} \lesssim \mathbf{b}$ if there is a constant $C > 0$ such that for all $n \in \mathbb{N}_0$ it holds $a_n \leq Cb_n$. If both $\mathbf{a} \lesssim \mathbf{b}$ and $\mathbf{a} \gtrsim \mathbf{b}$ hold true we write $\mathbf{a} \asymp \mathbf{b}$. $\mathbb{T}$ denotes the quotient space $\mathbb{R}/(2\pi\mathbb{Z})$. The constant λ is defined as $\lambda = \frac{1}{p} - \frac{1}{2}$. For two operators S, T we denote their composition by $S \circ T$.

2 Setting and basics

2.1 Setting

Definition 2.1. *For $\alpha > 0$ and $0 < p < \infty$ we define the norm*

$$\|f\|_{\mathcal{A}_p^\alpha(\mathbb{T}^d)} \coloneqq \Big(\sum_{\mathbf{k}\in\mathbb{Z}^d} \prod_{i=1}^{d} (1+|k_i|)^{\alpha p} |\hat{f}(\mathbf{k})|^p \Big)^{\frac{1}{p}},$$

with

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

The corresponding space

$$\mathcal{A}_p^\alpha(\mathbb{T}^d) \coloneqq \Big\{ f \in L_1(\mathbb{T}^d) \Big| \|f\|_{\mathcal{A}_p^\alpha(\mathbb{T}^d)} < \infty \Big\}, \tag{2.1}$$

is called weighted Wiener space with mixed weights. The space $\mathcal{A}_p^\alpha(\mathbb{T}^d)$ is often simply denoted by $\mathcal{A}_p^{\alpha,d}$ and $\mathcal{A}_p^\alpha(\mathbb{T}^1)$ by $\mathcal{A}_p^\alpha$. In the case $p = \infty$ the usual modifications are made and the space is called Korobov space (see e.g. [8, Section 3.3]).

In the literature these spaces are often referred to as $\mathcal{A}_{\text{mix}}^s$ where s takes the place of α. These spaces are usually called (mixed) weighted Wiener spaces (see [16]), weighted mixed Wiener spaces (see [14]) or weighted Wiener algebras with mixed smoothness (see [18]) or some combination of these. The $\mathcal{A}_p^\alpha(\mathbb{T}^1)$ spaces do not truly have mixed weights since there is no product for $d = 1$.

Remark 2.2. *While the space $\mathcal{A}_p^{\alpha,d}$ is a function space it also can be embedded into ℓ_p quite naturally for all $p > 0$ by the following operator*

$$A_\alpha f = \Big(\prod_{i=1}^{d} (1+|k_i|)^\alpha \hat{f}(\mathbf{k}) \Big)_{\mathbf{k}\in\mathbb{Z}^d}, \quad \|A_\alpha \colon \mathcal{A}_p^{\alpha,d} \to \ell_p\| = 1.$$

This property will be very useful since there are already some results for Gelfand numbers between ℓ_p spaces.

Definition 2.3. *A diagonal operator D_α between two d-dimensional multi-indexed Lebesgue sequence spaces is an operator of the form*

$$D_\alpha(x_{\mathbf{k}})_{\mathbf{k}\in\mathbb{Z}^d} = \Big(\prod_{j=1}^{d} (1+|k_j|)^{-\alpha} x_{\mathbf{k}} \Big)_{\mathbf{k}\in I}$$

where $I \subset \mathbb{Z}^d$. In particular if $d = 1$ the product vanishes and one simply gets

$$D_\alpha (x_k)_{k\in\mathbb{Z}} = ((1 + |k|)^{-\alpha} x_k)_{k\in I}.$$

2.2 Gelfand numbers

In the following Gelfand numbers and some of their basic properties are presented. They were introduced under a similar name (n-th diameter/width in the sense of Gelfand) by Triebel in 1970 in the paper [31] based on work by Solomyak and Tikhomirov [26] and Tikhomirov [30]. The n-th Gelfand number gives a lower bound for the error of the best non-linear reconstruction of a function from n linear samples, see Lemma 2.11. For more details or further references see [8, 21].

Definition 2.4. *Let X, Y be quasi-Banach spaces and $T\colon X \to Y$ linear and continuous. The Gelfand numbers $c_n(T)$ are defined as*

$$c_n(T) = \inf \Big\{ \sup_{x\in B^X\cap M} \|Tx\|_Y \colon M \subset X \textit{ linear subspace with } \operatorname{codim} M < n \Big\}, \quad n \in \mathbb{N}.$$

We say that $M \subset X$ is of codimension m if there exist m linearly independent functionals λ_i such that

$$M = \{x \in X \colon\ \lambda_i(x) = 0,\ i = 1, \ldots , m\}.$$

However, this definition of codimension does not make sense in all quasi-Banach spaces, there the dimension of a subspace is viewed as a purely algebraic notion and as mentioned in [7, Definition 2.1.], the codimension of a subspace can for that purpose be alternatively defined as the dimension of the quotient space, see [24, Sect. 1.40].

Gelfand numbers are part of a larger class of so called of s-numbers (see, eg. in [21, Definition 2.2.1]). These share the following properties.

Definition 2.5. *Let X, Z be quasi Banach-spaces and Y be a p-Banach space for $p \leq 1$, let $S, T \in \mathcal{L}(X, Y)$ and $R \in \mathcal{L}(Y, Z)$. A mapping $s\colon T \to (s_n(T))_{n=1}^{\infty}$ with the following properties*

(S1) $\|T|\mathcal{L}(X, Y)\| = s_1(T) \geq s_2(T) \geq \ldots \geq 0$

(S2) *for all $n_1, n_2 \in \mathbb{N}$ it holds*

$$s_{n_1+n_2-1}(R \circ S) \leq s_{n_1}(R)s_{n_2}(S)$$

(S3) *for all $n_1, n_2 \in \mathbb{N}$ it holds*

$$s^p_{n_1+n_2-1}(S + T) \leq s^p_{n_1}(S) + s^p_{n_2}(T)$$

(S4) *If rank $T < n$ then $s_n(T) = 0$ and $s_n(id\colon \ell_2^n \to \ell_2^n) = 1$*

is called s-function. And $s_n(T)$ is called the n-th s-number of the operator T.

Proposition 2.6. *The Gelfand numbers form an s-function.*

Proof. To see ***(S1)*** we observe that the only space $M \subset X$ with codim < 1 is the space X itself. The monotonicity

$$c_{n+1}(T) \leq c_n(T)$$

is obvious. To see ***(S2)*** we choose $\varepsilon > 0$ and subspaces $M_1 \subset X$ with codim $M_1 < n_1$ and $M_2 \subset Y$ with codim $M_2 < n_2$ such that

$$\|Tx\|_Y \leq (1+\varepsilon)c_{n_1}(T)\|x\|_X, \ x \in M_1$$

$$\|Ry\|_Z \leq (1+\varepsilon)c_{n_2}(R)\|y\|_Y, \ y \in M_2.$$

Put $M = M_1 \cap T^{-1}(M_2)$ then for $x \in M$

$$\begin{aligned} \|(R \circ T)x\|_Z &\leq (1+\varepsilon)c_{n_2}(R)\|Tx\|_Y \\ &\leq (1+\varepsilon)^2 c_{n_2}(R)c_{n_1}(T)\|x\|_X. \end{aligned}$$

Clearly

$$M_1 = \{x \in X : \ f_i(x) = 0, \ i = 1, \dots, n_1 - 1\}$$
$$M_2 = \{x \in X : \ g_i(x) = 0, \ i = 1, \dots, n_2 - 1\}$$

$$\begin{aligned} \Longrightarrow T^{-1}(M_2) &= \{x \in X : \ Tx \in M_2\} \\ &= \{x \in X : \ g_i(Tx) = 0, \ i = 1, \dots, n_2 - 1\} \end{aligned}$$

$\Longrightarrow$ $M_1 \cap T^{-1}(M_2)$ has codimension less than $n_1 + n_2 - 1$.

(S3) goes similar. Choose $\varepsilon > 0$ and $M_1 \subset X$ with codim $M_1 < n_1$ and $M_2 \subset X$ with codim $M_2 < n_2$ such that

$$\|Sx\|_Y \leq (1+\varepsilon)c_{n_1}(S)\|x\|_X$$

$$\|Tx\|_Y \leq (1+\varepsilon)c_{n_2}(T)\|x\|_X.$$

Take $M_1 \cap M_2 = M$ and find for $x \in M$

$$\begin{aligned} \|(S+T)x\|_Y^p &\leq \|Sx\|_Y^P + \|Tx\|_y^p \\ &\leq (1+\varepsilon)^p(c_{n_1}^p(S) + c_{n_2}^p(T))\|x\|_X^p \\ \Longrightarrow c_{n_1+n_2-1}^p(S+T) &\leq c_{n_1}^p(S) + c_{n_2}^p(T). \end{aligned}$$

By the same reasoning as above we see that $M \subset X$ has codimension smaller than $n_1 + n_2 - 1$. For ***(S4)*** we notice that if $\operatorname{rank} T < n$ then $M = \ker T$ has codimension $< n$. Therefore

$$\sup_{x \in M \cap B^X} \|Tx\|_Y = 0$$

and hence $c_n(T) = 0$. What remains is $c_n(\mathrm{id}\colon \ell_2^n \to \ell_2^n) = 1$ which follows from the more general result [32, Lemma 4.3]. ■

2.3 Best m-term approximation

Next up be want to consider a different metric for approximation. The best m-term approximation. To be precise the best m-term trigonometric approximation, since in this thesis we only use the trigonometric system as a dictionary. This quantity was first devised by Stechkin in 1955 in the paper [27] and since then has been a very theoretical subject of study especially in the Russian literature.

Definition 2.7. *Let X be d-dimensional quasi-Banach function spaces and $f \in X$. The best n-term approximation of f is defined as*

$$\sigma_n(f)_X := \inf_{s \in \Sigma_n} \|f - s\|_X,$$

where Σ_n is the set of all trigonometric polynomials of the form

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} a_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}}$$

with at most n non-zero coefficients $a_{\mathbf{k}}$.

The best n-term approximation of a function class $\mathbf{W}$ is defined accordingly, as

$$\sigma_n(\mathbf{W})_X := \sup_{f \in \mathbf{W}} \sigma_n(f)_X.$$

Because we can apply it with Proposition 1.2 we are in particular interested in

$$\sigma_n\Big(B_1\big(\mathcal{A}_p^\alpha(\mathbb{T}^d)\big)\Big)_\infty,$$

the best trigonometric m-term approximation of a weighted Wiener class in L_∞. This approximation quantity is already different to Gelfand numbers in that it studies functions and function classes instead of operators between them. They do however have some similarities.

Lemma 2.8. *The best m-term approximation is not an s-function, but some properties still hold in a similar fashion. Let X, Y, Z be function classes that contain the trigonometric polynomials such that Y has an embedding into Z and X has an embedding into Y.*

(L) *for all $n \in \mathbb{N}$ and any $\alpha \in \mathbb{R}$ it holds*

$$\sigma_n(\alpha f)_X = |\alpha| \sigma_n(f)_X$$

(S1) $\|f\| = \sigma_0(f)_X \geq \sigma_1(f)_X \geq \ldots \geq 0$

(S2) *for all $n_1, n_2 \in \mathbb{N}$ it holds*

$$\sigma_{n_1+n_2-1}(X)_Z \leq \sigma_{n_1}(X)_Y \sigma_{n_2}(Y)_Z$$

(S3) *for all $n_1, n_2 \in \mathbb{N}$ and f, g from a p-Banach space it holds*

$$\sigma^p_{n_1+n_2-1}(f+g)_X \leq \sigma^p_{n_1}(f)_X + \sigma^p_{n_2}(g)_X$$

(S4) *if $f \in \Sigma_n$ then $\sigma_n(f)_X = 0$*

Proof. **(L)** and **(S1)** follow immediately from the definition.

Instead of just approximating a function f from X in Z one could first approximate the function in Y with n_1 terms and then $f - s_1$ in Z with n_2 terms since $f - s_1$ is in Y. However this second approximation error has a different scaling, namely the norm of the first approximation error or in other words the best n_1-term approximation of X in Y. Now applying **(L)** yields **(S2)**.

(S3) follows from the p-triangle inequality and **(S4)** is trivial. ■

Note that the property **(L)** is in stark contrast to the setting of Gelfand numbers. It also means that it only makes sense to study the best m-term approximation of bounded function classes e.g. the unit ball of a function space.

A useful tool in nonlinear approximation is the so-called Stechkin Lemma (see, e.g. [8, 28]).

Lemma 2.9 (Stechkin Lemma). *Let $0 < p < q \leq \infty$. Then it holds*

$$\sigma_n(\mathbf{x})_q \leq (n+1)^{\frac{1}{q}-\frac{1}{p}} \|\mathbf{x}\|_p \tag{2.2}$$

for all $\mathbf{x} \in \ell_p$, where $\sigma_n(\mathbf{x})_q$ of a sequence is defined as the q-norm of the sequence without its n largest entries.

Proof. Let $\mathbf{x}^*$ be the non-increasing reordering of $\mathbf{x}$ then it holds

$$\begin{aligned}
\sum_{i=n+1}^{\infty} |x_i^*|^q &\leq |x_{n+1}^*|^{q-p} \sum_{i=n+1}^{\infty} |x_i^*|^p \\
&\leq \Big(\frac{1}{n+1} \sum_{i=1}^{n+1} |x_i^*|^p\Big)^{\frac{q-p}{p}} \|\mathbf{x}\|_p^p \\
&\leq (n+1)^{-\frac{q-p}{p}} \|\mathbf{x}\|_p^{q-p} \|\mathbf{x}\|_p^p \\
&\leq (n+1)^{-\frac{q-p}{p}} \|\mathbf{x}\|_p^q,
\end{aligned} \tag{2.3}$$

and since $\sigma_n(\mathbf{x})_q = \big(\sum_{i=n+1}^{\infty} |x_i^*|^q\big)^{\frac{1}{q}}$ we get the assertion. ■

2.4 Sampling numbers

The third important quantity discussed in this thesis will be the sampling numbers, these are the worst case error of the best recovery operator using standard information.

Definition 2.10. *Let $X(\Omega)$ be a quasi-Banach function space and $\mathcal{F}$ be a quasi-normed space continuously embedded into X. Then the n-th sampling number is defined as*

$$\varrho_n(\mathcal{F})_X = \inf_{t_1 \ldots t_n \in \Omega} \inf_{R\colon \mathbb{C}^n \to X} \sup_{\|f\|_{\mathcal{F}} \leq 1} \big\| f - R\big(f(t_1) \ldots f(t_n)\big) \big\|_X \tag{2.4}$$

Between the Gelfand and sampling numbers the following relation holds.

Lemma 2.11 ([14][Lemma B.1])**.** *Let $\mathcal{F}$ be a Banach space continuously embedded into the Banach space X by the identity $T\colon \mathcal{F} \to X$. Then for all $n \in \mathbb{N}_0$ it holds*

$$c_n(T\colon \mathcal{F} \to X) \leq \varrho_n(\mathcal{F})_X. \tag{2.5}$$

Proof. Let $\varepsilon > 0$. There exist $t_1 \ldots t_n \in \Omega$ and $R\colon \mathbb{C}^n \to X$ such that

$$\sup_{f \in B^1_{\mathcal{F}}} \|f - R\big(f(t_1) \ldots f(t_n)\big)\|_X \leq (1+\varepsilon)\varrho_n(\mathcal{F})_X.$$

Chose now L as the following $L \coloneqq \big\{f \in \mathcal{F}\colon f(t_1) = \cdots = f(t_n) = 0\big\}$. This subspace $L \subset \mathcal{F}$ has codim $(L) \leq n$. Now it holds

$$\begin{aligned}
c_n(T\colon \mathcal{F} \to X) &\leq \sup_{f \in B^1_{\mathcal{F}} \cap L} \big\|f\big\|_X \\
&= \frac{1}{2} \sup_{f \in B^1_{\mathcal{F}} \cap L} \big\|f - (-f)\big\|_X \\
&= \frac{1}{2} \sup_{f \in B^1_{\mathcal{F}} \cap L} \big\|f - R\big(f(t_1) \ldots f(t_n)\big) - \Big(- f - R\big(- f(t_1) \ldots - f(t_n)\big)\Big) \big\|_X \\
&\leq \Big(\frac{1}{2} \sup_{f \in B^1_{\mathcal{F}} \cap L} \big\|f - R\big(f(t_1) \ldots f(t_n)\big)\big\| + \big\|(-f) - R\big(f(t_1) \ldots f(t_n)\big)\big\|_X\Big) \\
&\leq (1+\varepsilon)\varrho_n(\mathcal{F})_X.
\end{aligned}$$

Taking the limit $\varepsilon \to 0$ yields the assertion. ∎

There also exist a second kind of sampling numbers, the linear sampling numbers. These are similar to the (non-linear) ones above but only allow linear reconstruction operators.

Definition 2.12. *Let $X(\Omega)$ be a quasi-Banach function space and $\mathcal{F}$ be a quasi-normed space continuously embedded into X. Then the n-th linear sampling number is defined as*

$$\varrho_n^{\mathrm{lin}}(\mathcal{F})_X = \inf_{t_1 \ldots t_n \in \Omega} \inf_{R\colon \mathbb{C}^n \to X} \sup_{\|f\|_{\mathcal{F}} \leq 1} \big\| f - R\big(f(t_1) \ldots f(t_n)\big) \big\|_X \tag{2.6}$$

when $R\colon \mathbb{C}^n \to X$ is a linear map.

For the linear sampling numbers a similar relation to the one from Lemma 2.11 holds.

Lemma 2.13. *Let $\mathcal{F}$ be a Banach space continuously embedded into the Banach space X by the identity $T\colon \mathcal{F} \to X$. Then for all $n \in \mathbb{N}_0$ it holds*

$$d_n(T\colon \mathcal{F} \to X) \le \varrho_n^{\mathrm{lin}}(\mathcal{F})_X. \tag{2.7}$$

2.5 Existing results for s-numbers

In this section some existing results from the literature will be discussed. The first two of them give a bound on the Gelfand numbers of the identities between ℓ_p^N and ℓ_q^N. This is useful since a lot of results for Gelfand numbers can be proven by segmenting operators into smaller finite blocks and then using finite results in finite-dimensional ℓ_p spaces.

Definition 2.14. *The space ℓ_p^N is defined for $0 < p < \infty$ as*

$$\ell_p^N(\mathbb{R}) \coloneqq \Big\{(x_n)_{n=1}^N \subset \mathbb{R} \Big| \|\mathbf{x}\|_{\ell_p^N(\mathbb{R})} \coloneqq \Big(\sum_{n=1}^N |x_n|^p\Big)^{\frac{1}{p}} < \infty\Big\},$$

with the usual modifications for $p = \infty$.

First up is an upper bound shown by Vybrial in 2008.

Proposition 2.15 ([33, Lemma 4.9]). *For $0 < p \le 1$, $q = 2$ the following upper bound holds*

$$c_m(id : \ell_p^N \to \ell_q^N) \lesssim \Big(\frac{1 + \log(\frac{N}{m})}{m}\Big)^{\frac{1}{p} - \frac{1}{q}}.$$

This bound was later expanded to the case where $p < q \le 2$ and $p \le 1$ by Foucart, Pajor, Rauhut and T.Ullrich in 2010. These authors also showed a corresponding lower bound in the same paper.

Proposition 2.16 ([9, Theorem 1.1]). *For $0 < p \le 1$, $p < q \le 2$ and $m < N$ the following lower bound holds*

$$c_m(id : \ell_p^N \to \ell_q^N) \gtrsim \Big(\frac{1 + \log(\frac{N}{m})}{m}\Big)^{\frac{1}{p} - \frac{1}{q}}.$$

Another result for Gelfand numbers was given by Buchmann [4, Korollar 7.6] and Vybiral [33, Lemma 4.7] based on work by Gluskin [10] and is a two sided bound for the regime where $1 < p < 2$.

Proposition 2.17. *For $1 < p < 2$ and $m < N$ the following bound holds*

$$c_m(id : \ell_p^N \to \ell_q^N) \asymp m^{-\frac{1}{2}} N^{1-\frac{1}{p}}.$$

All of these results will be employed throughout this thesis to construct asymptotic bounds on Gelfand numbers of diagonal operators.

Another very different result with relation to the topic of this thesis comes in the form of a recent paper by V.D. Nguyen, V.K. Nguyen and Sickel where the authors studied a wide variety of (quasi) s-numbers of the embedding from $\mathcal{A}_1^{\alpha,d}$ into L_2. They obtained the following bounds.

Proposition 2.18 (see [18]). *For the approximation numbers a_n and Kolmogorov numbers d_n it holds for all $\alpha > 0$*

$$a_n\big(id\colon \mathcal{A}_1^{\alpha,d} \to L_2(\mathbb{T}^d)\big) \asymp d_n\big(id : \mathcal{A}_1^{\alpha,d} \to L_2(\mathbb{T}^d)\big) \asymp n^{-\alpha} \log(n)^{\alpha(d-1)}. \tag{2.8}$$

For the Bernstein numbers b_n and Weyl numbers x_n it holds for all $\alpha > 0$

$$b_n\big(id\colon \mathcal{A}_1^{\alpha,d} \to L_2(\mathbb{T}^d)\big) \asymp x_n\big(id : \mathcal{A}_1^{\alpha,d} \to L_2(\mathbb{T}^d)\big) \asymp n^{-(\alpha+\frac{1}{2})} \log(n)^{\alpha(d-1)}. \tag{2.9}$$

Note that Bernstein numbers do not fulfil the definition of s-numbers used in this thesis. They do however satisfy the older definition of s-numbers that can be found in [20] and are usually referred to as quasi s-numbers.

3 Gelfand numbers of diagonal operators

3.1 Diagonal operators between ℓ_p and ℓ_2

Before examining Wiener classes with mixed weights we will look at the regime where $d = 1$ and $0 < p \leq 1$. The space $\mathcal{A}_p^\alpha(\mathbb{T}^1)$ does not have the same product weights as the ones in higher dimensions and it behaves slightly differently. This situation could also be treated with results like [4, Satz 7.1] (see, e.g. [12, Proposition 6] for an English version).

Using the same commutative diagram as in Section 3.2 we can reformulate the question for the Gelfand numbers of the identical embedding $\mathcal{A}_p^\alpha$ to $L_2(\mathbb{T})$ as the question for the Gelfand numbers of the diagonal operator from ℓ_p to ℓ_2. To get such a bound, we use the results for finite-dimensional ℓ_p spaces from the previous chapter and a decomposition of the diagonal operator.

The decomposition method

The following Theorem employs the decomposition method to construct a bound on the Gelfand numbers of an operator with infinite entries. This is done by segmenting it into a sum of Gelfand numbers of finitely supported operators (and one with infinite support that can be estimated by its norm). It is similar to [32, Theorem 2.11] for entropy numbers.

Theorem 3.1. *Let $0 < p \leq 1$, $\alpha > 0$ and define the diagonal operator*

$$D_\alpha \big(x_k\big)_{k\in\mathbb{Z}} = \big((|k| + 1)^{-\alpha} x_k\big)_{k\in\mathbb{Z}},$$

mapping from $\ell_p(\mathbb{Z})$ to $\ell_2(\mathbb{Z})$. Then we have for $n \in \mathbb{N}$

$$c_n(D_\alpha \colon \ell_p(\mathbb{Z}) \to \ell_2(\mathbb{Z})) \asymp n^{-(\alpha+\lambda)},$$

where

$$\lambda = \frac{1}{p} - \frac{1}{2}.$$

Proof. Step 1. Lower bound: Fix $N \in \mathbb{N}$

$$\begin{tikzcd}
\ell_p^N \arrow[r, "\mathrm{id}"] \arrow[d, "A_\alpha"'] & \ell_2^N \\
\ell_p \arrow[r, "D_\alpha"'] & \ell_2 \arrow[u, "P_d"']
\end{tikzcd}$$

$$A_\alpha\colon (x_{-N},\dots,x_0,\dots,x_N) \to (\dots,0,(N+1)^\alpha x_{-N},\dots,x_0,\dots,(N+1)^\alpha x_N,0,\dots),$$
$$P_d\colon (x_k)_{k\in\mathbb{Z}} \to (x_{-N},\dots,x_0,\dots,x_N),$$
$$\implies c_n(\mathrm{id}\colon \ell_p^N \to \ell_2^N) \le \|A_\alpha\| c_n(D_\alpha\colon \ell_p \to \ell_2)\|P_N\|.$$

From Proposition 2.16 we know that for all $n \le d$

$$c_n(\mathrm{id}\colon \ell_p^N \to \ell_q^N) \gtrsim n^{-\lambda}.$$

Moreover, $\|A_\alpha\colon \ell_p^N \to \ell_p\| \le N^\alpha$ and $\|P_N\colon \ell_2 \to \ell_2^N\| = 1$. Therefore, we get for $n = N$ (note that since here we work in $\mathbb{Z}$ instead of $\mathbb{N}_0$ this still satisfies the conditions of Proposition 2.16)

$$n^{-\lambda} \lesssim n^\alpha c_n(D_\alpha\colon \ell_p \to \ell_2).$$
$$\implies c_n(D_\alpha\colon \ell_p \to \ell_2) \ge n^{-(\lambda+\alpha)}.$$

Step 2. Upper bound: Fix $n \in \mathbb{N}$ and assume without loss of generality that $n = 2^m$, $N = 2n+1$. Let us further decompose $D_\alpha = D_{\alpha,n} + D_\alpha^n$, where

$$D_{\alpha,n}\colon (x_k)_{k\in\mathbb{Z}} \to \big((|k|+1)^{-\alpha}x_k\big)_{-n\le k\le n}$$
$$D_\alpha^n\colon (x_k)_k \to \Big(\big((|k|+1)^{-\alpha}x_k\big)_{k=-n}^{-\infty}, \underbrace{0,\dots,0}_{2n-1}, \big((|k|+1)^{-\alpha}x_k\big)_{k=n}^{\infty}\Big).$$

Now we split $D_{\alpha,d}$ into dyadic blocks

$$\begin{aligned}
\Delta_0\colon &\quad (x_k)_{k\in\mathbb{Z}} \to (\dots,0,x_0,0,\dots)\\
\Delta_1\colon &\quad (x_k)_{k\in\mathbb{Z}} \to \Big(\dots,0,\frac{x_1}{2^\alpha},\frac{x_2}{3^\alpha},0,\dots\Big)\\
\Delta_{-1}\colon &\quad (x_k)_{k\in\mathbb{Z}} \to \Big(\dots,0,\frac{x_{-2}}{3^\alpha},\frac{x_{-1}}{2^\alpha},0,\dots\Big)\\
\Delta_2\colon &\quad (x_k)_{k\in\mathbb{Z}} \to \Big(\dots,0,\frac{x_3}{4^\alpha},\frac{x_4}{5^\alpha},0,\dots\Big)\\
\Delta_{-2}\colon &\quad (x_k)_{k\in\mathbb{Z}} \to \Big(\dots,0,\frac{x_{-4}}{5^\alpha},\frac{x_{-3}}{4^\alpha},0,\dots\Big)\\
&\vdots\\
\Delta_m\colon &\quad (x_k)_{k\in\mathbb{Z}} \to \Big(\dots,0,\frac{x_{2^{m-1}+1}}{(2^{m-1}+2)^\alpha},\dots,\frac{x_{2^m}}{(2^m+1)^\alpha},0,\dots\Big)\\
\Delta_{-m}\colon &\quad (x_k)_{k\in\mathbb{Z}} \to \Big(\dots,0,\frac{x_{-2^m}}{(2^m+1)^\alpha},\dots,\frac{x_{-2^{m-1}-1}}{(2^{m-1}+2)^\alpha},0,\dots\Big).
\end{aligned}$$

We use the subadditivity of the Gelfand numbers, see Proposition 2.6 and Definition 2.5, (S4) and obtain for $N = \sum_{k=-m}^{m} n_k$

$$c_N^p(D_{\alpha,n}\colon \ell_p \to \ell_2) \le \sum_{k=-m}^{m} c_{n_k}^p(\Delta_k\colon \ell_p \to \ell_2).$$

We can now further decompose D_α as follows

$$D_\alpha = \sum_{j=-m}^{m} \Delta_j + \sum_{j=m+1}^{L} \big(\Delta_j + \Delta_{-j}\big) + D_\alpha^{2^L+1},$$

where L is chosen later. Now choose

- $n_j = 2^{j-1}2^{(m-j)\eta}$ with $\eta < 1$, for $j = 1\dots,m$,
- $n_j = 2^{j-1}2^{(m-j)\delta}$ with $\delta > 1$ (chosen later), for $j = m+1,\dots,L$.

Clearly, we have

$$\sum_{j=1}^{L} n_j = \sum_{j=1}^{m} n_j + \sum_{j=m+1}^{L} n_j$$

and

$$\sum_{j=1}^{m} n_j = \sum_{j=1}^{m} 2^{j-1}2^{(m-j)\eta} \asymp 2^{m\eta}\sum_{j=1}^{m} 2^{j(1-\eta)} \lesssim 2^m$$

because of $\eta < 1$. Further, due to $\delta > 1$ we have

$$\sum_{j=m+1}^{L} n_j = \sum_{j=m+1}^{L} 2^{j-1}2^{(m-j)\delta} = 2^{m\delta}\sum_{j=m+1}^{L} 2^{j(1-\delta)} \lesssim 2^m,\ \delta > 1.$$

We will need that

$$2^{j-1}2^{(m-j)\delta} > j = \log 2^{j-1},\ j = m+1,\dots,L.$$

Hence, we get by the subadditivity of Gelfand numbers a sum that is reminiscent of Maiorov's discretisation technique [17]

$$c_{cN}^p(D_\alpha\colon \ell_p \to \ell_2) \le \sum_{j=-m}^{m} c_{n_{|j|}}^p(\Delta_j\colon \ell_p \to \ell_2) + 2\sum_{j=m+1}^{L} c_{n_j}^p(\Delta_j\colon \ell_p \to \ell_2) + 2c_1^r(D_\alpha^{2^L+1}\colon \ell_p \to \ell_2)$$

We estimate piece by piece. For $n_j \ge 2^{j-1}$ it obviously holds

$$\sum_{j=-m}^{m} c_{n_{|j|}}^r(\Delta_j\colon \ell_p \to \ell_2) \asymp c_{n_j}(\mathrm{id}\colon \ell_p^{2^{j-1}} \to \ell_2^{2^{j-1}})2^{-j\alpha} = 0$$

Proposition 2.15 now gives

$$\begin{aligned}
2\sum_{j=m+1}^{L} c_{n_j}^p(\Delta_j\colon \ell_p \to \ell_2) &= 2\sum_{j=m+1}^{L} c_{n_j}^p(\Delta_j\colon \ell_p \to \ell_2) \\
&\lesssim \sum_{j=m+1}^{L} \left(\frac{\log\left(\frac{2^{j-1}}{2^{j-1}2^{(m-j)\delta}}\right)}{2^{j-1}2^{(m-j)\delta}} \right)^{\lambda r} 2^{-j\alpha p} \\
&\asymp \sum_{j=m+1}^{L} [(j-m)\delta 2^{(j-m)\delta} 2^{-j}]^{\lambda r} 2^{-j\alpha p} \\
&= 2^{-m(\alpha+\lambda)r} \sum_{j=m+1}^{L} [(j-m)\delta 2^{(j-m)\delta} 2^{m-j}]^{\lambda r} 2^{(m-j)\alpha p} \\
&\asymp 2^{-m(\lambda+\alpha)r} \sum_{j=m+1}^{L} \left[(j-m)\delta 2^{-(j-m)[\alpha-(\delta-1)\lambda]}\right]^p.
\end{aligned}$$

The sum converges if $\alpha > (\delta-1)\lambda$ hence, we choose δ close to 1 but larger than 1.

What remains is

$$c_1(D_\alpha^{2^L+1}\colon \ell_p \to \ell_2) \le \|D_\alpha^{2^L+1}\colon \ell_p \to \ell_2\| \lesssim 2^{-L\alpha} \overset{!}{\lesssim} 2^{-m(\alpha+\lambda)}.$$

How do we choose the parameters?

- L such that $2^{-L\alpha} < 2^{-m(\alpha+\lambda)}$, $L = C(\alpha,\lambda)m$
- $\delta > 1$ small enough such that $\alpha > (\delta-1)\lambda$

This implies

$$c_{C(\alpha,\lambda)2^m}(D_\alpha\colon \ell_p \to \ell_2) \lesssim 2^{-m(\alpha+\lambda)}.$$

To obtain

$$c_n(D_\alpha\colon \ell_p \to \ell_2) \lesssim n^{-(\alpha+\lambda)}$$

we use a monotonicity argument for

$$C(\alpha,\lambda)2^m \le n \le C(\alpha,\lambda)2^{m+1}$$

which gives

$$c_n(D_\alpha\colon \ell_p \to \ell_2) \le c_{C(\alpha,\lambda)2^m} \lesssim 2^{-m(\alpha+\lambda)} \lesssim n^{-(\alpha+\lambda)}.$$

∎

3.2 Diagonal operators between $\ell_p(\mathbb{Z}^d)$ and $\ell_2(\mathbb{Z}^d)$

Recall the motivation from Section 3.1, where we studied diagonal operators between ℓ_p and ℓ_2. This will now be extended to the regime where $d > 1$ and $0 < p \le 2$. This setting is qualitatively different from the one dimensional one since we now have to deal with mixed (product) weights. Therefore, results like [4, Satz 7.1] do not apply here. Consider for $0 < p \leq 2$ the embedding of a space $\mathcal{A}_p^\alpha(\mathbb{T}^d)$ in $L_2(\mathbb{T}^d)$

$$\|f\|_{\mathcal{A}_p^{\alpha,d}} = \sum_{\mathbf{k}\in\mathbb{Z}^d} \prod_{i=1}^d (1+|k_i|)^\alpha |\hat{f}(\mathbf{k})|, \; \alpha > 0.$$

We will use the following commutative diagram to characterize this embedding via a diagonal operator.

$$\begin{tikzcd} \mathcal{A}_p^\alpha(\mathbb{T}^d) \arrow[r, "\mathrm{id}"] \arrow[d, "A_\alpha"'] & L_2(\mathbb{T}^d) \\ \ell_p(\mathbb{Z}^d) \arrow[r, "D_\alpha"'] & \ell_2(\mathbb{Z}^d) \arrow[u, "B"'] \end{tikzcd}$$

$$A_\alpha f = \big(\prod_{i=1}^d (1+|k_i|)^\alpha \hat{f}(\mathbf{k})\big)_{\mathbf{k}\in\mathbb{Z}^d}, \quad \|A\| = 1$$

$$B(x_\mathbf{k})_{\mathbf{k}\in\mathbb{Z}^d} = \sum_{\mathbf{k}\in\mathbb{Z}^d} x_\mathbf{k} e^{i\mathbf{k}\mathbf{x}}, \quad \|B\| = 1$$

$$D_\alpha(x_\mathbf{k})_{\mathbf{k}\in\mathbb{Z}^d} = \Big(\prod_{i=1}^d (1+|k_i|)^{-\alpha} x_\mathbf{k}\Big)_{\mathbf{k}\in\mathbb{Z}^d}$$

Since the operators A_α and B are invertible we get by Definition 2.5 (S2)

$$c_n(\mathrm{id}\colon \mathcal{A}_p^\alpha(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) = c_n(D_\alpha\colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)).$$

We can therefore study the Gelfand numbers of D_α to get them for the embedding $\mathrm{id}\colon \mathcal{A}^\alpha(\mathbb{T}^d) \to L_2(\mathbb{T}^d)$. This is done in the following Theorem that expands Theorem 3.1 to the regime where $d > 1$ and provides the bound for [14][Lemma 4.2 (i)]. Its proof uses the same general strategy but adopts it to accommodate the mixed weights. Therefore, the blocks in this proof will not be cubes but instead hyperbolic layers. It incorporates some ideas from the proof of [7, Proposition 7.1].

Theorem 3.2. *Let $n, d \in \mathbb{N}$ where $d > 1$ such that $\log(n) > \log(\log_2(n))2(d-1)$ and $0 < p \leq 1$. Let further $\alpha > 0$ with $d \leq \log(n)^{(d-1)\alpha}$ then it holds*

$$c_n(D_\alpha\colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) \asymp n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha}$$

where

$$\lambda = \frac{1}{p} - \frac{1}{2}.$$

Proof. Step 1. Upper bound: Assume without loss of generality that $n = 2^m$ with $m \in \mathbb{N}_0$. First note that D_α is radial-symmetric and can be split into 2^d operators D_{α_c} that each have support on one of the 2^d quadrants of $\mathbb{Z}^d$. Let us now further decompose D_{α_c} (w.l.o.g. this is the operator on the first quadrant) into blocks.

First segment $\mathbb{N}_0 = \bigcup_{j=0}^{\infty} N_j$ with $N_j = \{2^j - 1, \dots, 2^{j+1} - 2\}$, it holds $\#N_j = 2^j$. This then immediately induces a segmentation of $\mathbb{N}^d$ into cuboids $\square_{\mathbf{x}} = \prod_{l=1}^{d} N_{x_l}$, each of these blocks contains $\#\square_{\mathbf{x}} = 2^{|\mathbf{x}|}$ points.

Now one can group blocks on the same hyperbolic layer together

$$\square_j := \Big\{ \mathbf{n} \in \mathbb{N}_0^d \;\Big|\; \exists \mathbf{x} \in \mathbb{N}_0^d \colon \mathbf{n} \in \square_{\mathbf{x}}, \#\square_{\mathbf{x}} = 2^j \Big\}, \tag{3.1}$$

where the contents of every $\square_{\mathbf{x}}$ appear in exactly one $\square_j$, $j \in \mathbb{N}_0$ meaning $\mathbb{N}_0^d = \bigcup_{j=0}^{\infty} \square_j$.

This now allows us to decompose $D_{\alpha_c} = \sum_{j=0}^{\infty} \Delta_j$ where Δ_j is just D_{α_c} restricted to $\square_j$.

To continue we need to know how many points are in each hyperbolic layer $\square_j$. First consider, that $\square_j$ can be decomposed into a number of $\square_{\mathbf{x}}$ that each contain 2^j points. Since these sets are products of dyadic intervals their quantity is just the number of possibilities to distribute j to d different dimensions. In total

$$C_j := \#\square_j = 2^j \binom{j+d-1}{j} \asymp 2^j j^{d-1}.$$

The above decomposition together with the subadditivity of Gelfand numbers now gives for $r = \min\{1, p\}$

$$c^r_{n2^d}(D_\alpha) \le 2^d \Big(\sum_{j=0}^{L} c^r_{n_j}(\Delta_j) + \sum_{j=L+1}^{M} c^r_{n_j}(\Delta_j) + c^r_1\big(\sum_{j=M+1}^{\infty} \Delta_j \big) \Big), \tag{3.2}$$

where

$$L = \lfloor m - (d-1)\log_2 m \rfloor \tag{3.3}$$

and

$$M = \lfloor m(1 + \frac{\lambda}{\alpha}) - (d-1)\log_2 m \rfloor. \tag{3.4}$$

For the first sum in (3.2) choose $n_j = 2C_j$ (i.e. for $j = 0, \dots, m$). This then gives

$$\sum_{j=0}^{L} c^r_{n_j}(\Delta_j \colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) = 0 \tag{3.5}$$

and

$$\sum_{j=0}^{L} n_j \asymp 2\sum_{j=0}^{L} 2^j j^{d-1} \le 2\sum_{j=0}^{L} 2^j m^{d-1} \asymp 2^L m^{d-1} \asymp 2^m. \tag{3.6}$$

In the second sum in (3.2) choose $n_j = 2^j 2^{(L-j)\eta} j^{d-1}$ where $j = L+1, \dots, M$, $\eta > 1$ will be chosen later. It now holds

$$\sum_{j=L+1}^{M} n_j \asymp 2^{L\eta} \sum_{j=L+1}^{M} 2^{j(1-\eta)} j^{d-1} \asymp 2^L L^{d-1} \leq 2^L m^{d-1} = 2^m. \tag{3.7}$$

To estimate the second sum Proposition 2.15 is used

$$\begin{aligned}
\sum_{j=L+1}^{M} c_{n_j}^r(\Delta_j \colon \ell_p \to \ell_2) &\lesssim \sum_{j=L+1}^{M} c_{n_j}^r(\mathrm{id} \colon \ell_p^{C_j} \to \ell_2^{C_j}) 2^{-j\alpha r} \\
&\lesssim \sum_{j=L+1}^{M} \left(\frac{\log\left(\frac{2^j j^{d-1}}{2^j 2^{(L-j)\eta} j^{d-1}}\right)}{2^j 2^{(L-j)\eta} j^{d-1}} \right)^{\lambda r} 2^{-j\alpha r} \\
&\asymp \sum_{j=L+1}^{M} [(j-L)\eta 2^{(j-L)\eta} 2^{-j} j^{1-d}]^{\lambda r} 2^{-j\alpha r} \\
&= 2^{-L(\alpha+\lambda)r} \sum_{j=L+1}^{M} [(j-L)\eta 2^{(j-L)\eta} 2^{L-j} j^{1-d}]^{\lambda r} 2^{(L-j)\alpha r} \\
&\asymp 2^{-L(\lambda+\alpha)r} \sum_{j=L+1}^{M} \left[(j-L)\eta 2^{-(j-L)[\alpha-(\eta-1)\lambda]} j^{-\lambda(d-1)} \right]^r \\
&\asymp 2^{-(m-(d-1)\log_2 m)(\lambda+\alpha)r} (m-(d-1)\log_2 m)^{-\lambda(d-1)r} \\
&\asymp 2^{-m(\lambda+\alpha)r} m^{(d-1)(\lambda+\alpha)r} m^{-\lambda(d-1)r} \\
&= 2^{-m(\lambda+\alpha)r} m^{(d-1)\alpha r},
\end{aligned} \tag{3.8}$$

where the fact that $m > 2(d-1)\log_2 m$ was used in the second to last line. The sum only converges if $\alpha > (\eta-1)\lambda$, so η has to be chosen accordingly.

What remains is

$$\begin{aligned}
c_1^r\Big(\sum_{j=M+1}^{\infty} \Delta_j \colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)\Big) &\leq \| \sum_{j=M+1}^{\infty} \Delta_j \colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d) \|^r \\
&\lesssim 2^{-M\alpha r} \\
&\asymp 2^{-(m(1+\frac{\lambda}{\alpha})-(d-1)\log_2 m)\alpha r} \\
&= 2^{-m(\alpha+\lambda)r} m^{(d-1)\alpha r}.
\end{aligned} \tag{3.9}$$

We can therefore estimate (3.2) by (3.5), (3.8) and (3.9) as follows

$$\begin{aligned}
c_{2^m 2^d}^r\big(D_\alpha \colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)\big) &\leq 2^d\big(0 + 2^{-m(\lambda+\alpha)r} m^{(d-1)\alpha r} + 2^{-m(\lambda+\alpha)r} m^{(d-1)\alpha r}\big) \\
&\asymp 2^{-m(\lambda+\alpha)r} m^{(d-1)\alpha r}
\end{aligned} \tag{3.10}$$

Since $\sum_{j=0}^{M} n_j \lesssim 2^m$ by (3.6) and (3.7). After taking the r-th root this gives

$$c_{dC(\alpha,\lambda)2^m}(D_\alpha\colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) \lesssim 2^{-m(\alpha+\lambda)} m^{(d-1)\alpha}.$$

To obtain

$$c_{n2^d}(D_\alpha\colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) \lesssim n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha}$$

we use a monotonicity argument for

$$C(\alpha,\lambda)2^m m^{(d-1)\alpha} \leq n \leq C(\alpha,\lambda)2^{m+1} m^{(d-1)\alpha}$$

which yields

$$c_{n2^d}(D_\alpha\colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) \leq c_{2^d C(\alpha,\lambda)2^m} \lesssim 2^{-m(\alpha+\lambda)} m^{(d-1)\alpha} \lesssim n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha}.$$

Step 2. Lower bound: Let n, m, d, L be as above.
From Theorem 2.16 we know that

$$\begin{aligned}
c_{n2^d}(D_\alpha\colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) &\geq 2^d c_{2^L L^{d-1}}(\Delta_L\colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) \\
&\gtrsim c_{2^L L^{d-1}}(\mathrm{id}\colon \ell_p^{C_L} \to \ell_2^{C_L}) 2^{-L\alpha} \\
&\gtrsim \left(\frac{\log\left(\frac{2^L L^{d-1}}{2^{L-1}L^{d-1}}\right)}{2^{L-1}L^{d-1}} \right)^{\lambda} 2^{-L\alpha} \\
&\geq \left(2^{-L} L^{1-d}\right)^{\lambda} 2^{-L\alpha} \\
&= 2^{-L(\alpha+\lambda)} L^{-\lambda(d-1)} \\
&\geq 2^{-m(\alpha+\lambda)} m^{(d-1)(\alpha+\lambda)} m^{-(d-1)\lambda} \\
&= 2^{-m(\alpha+\lambda)} m^{(d-1)\alpha}.
\end{aligned} \tag{3.11}$$

And then again argue with monotonicity to get this for all n.

■

This result can be extended to $p \leq 1 < q \leq 2$ by using the extended upper bound from [9] instead of Lemma 2.15 for finite-dimensional ℓ_p spaces and to $p \leq 2$ by the following result, that employs Proposition 2.17. Also note that for $d = 1$ the one-dimensional result, Theorem 3.1 immediately follows.

Corollary 3.3. *Let $n, d \in \mathbb{N}$ such that $1 < p \leq 2$ and further $\alpha > \frac{p-1}{p}$, then it holds*

$$c_n(D_\alpha\colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) \asymp n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha}$$

where

$$\lambda = \frac{1}{p} - \frac{1}{2}.$$

Proof. Follow the proof of Theorem 3.2 until just after (3.6). Now one can choose $n_j = (j-L)^{-2} 2^L L^{d-1}$ instead. This still yields

$$\sum_{j=L+1}^{M} n_j = 2^L L^{d-1} \sum_{j=L+1}^{M} (j-L)^{-2} \asymp 2^L L^{d-1} \le 2^L m^{d-1} = 2^m. \tag{3.12}$$

We can then use Proposition 2.17 to bound the second sum

$$\begin{aligned}
\sum_{j=L+1}^{M} c_{n_j}(\Delta_j \colon \ell_p \to \ell_2) &\asymp \sum_{j=L+1}^{M} c_{n_j}(\mathrm{id} \colon \ell_p^{C_j} \to \ell_2^{C_j}) 2^{-j\alpha} \\
&\asymp \sum_{j=L+1}^{M} n_j^{-\frac{1}{2}} C_j^{1-\frac{1}{p}} 2^{-j\alpha} \\
&\asymp \sum_{j=L+1}^{M} (j-L) 2^{-\frac{L}{2}} L^{-\frac{d-1}{2}} j^{(d-1)\frac{p-1}{p}} 2^{j\frac{p-1}{p}} 2^{-j\alpha} \qquad (3.13) \\
&\asymp 2^{-L(\frac{1}{2}+\alpha-1+\frac{1}{p})} L^{-(\frac{1}{2}-1+\frac{1}{p})(d-1)} \\
&\asymp 2^{-(m-(d-1)\log_2 m)(\lambda+\alpha)} (m-(d-1)\log_2 m)^{-\lambda(d-1)} \\
&\asymp 2^{-m(\lambda+\alpha)} m^{(d-1)(\lambda+\alpha)} m^{-\lambda(d-1)} \\
&= 2^{-m(\lambda+\alpha)} m^{(d-1)\alpha}.
\end{aligned}$$

This only works if $\alpha > \frac{p-1}{p}$, see section 4.2 for further comments. ■

Remark 3.4. *Theorem 3.2 and Corollary 3.3 together with the considerations at the beginning of the section now give us for* $\alpha > \left(\frac{p-1}{p}\right)_+$ *and* $0 < p \le 2$

$$c_n(id \colon \mathcal{A}_p^\alpha(\mathbb{T}^d) \to L_2(\mathbb{T}^d)) = c_n(D_\alpha \colon \ell_p(\mathbb{Z}^d) \to \ell_2(\mathbb{Z}^d)) \asymp n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha},$$

the asymptotic behaviour of the identical embedding of $\mathcal{A}_p^\alpha(\mathbb{T}^d)$ *into* $L_2(\mathbb{T}^d)$ *where*

$$\lambda = \frac{1}{p} - \frac{1}{2}.$$

This successfully concludes the study of the Gelfand numbers of weighted Wiener classes with mixed weights by giving a sharp bound on their asymptotic behaviour.

4 Best m-term trigonometric approximation of mixed weighted Wiener classes

4.1 Weighted Wiener classes with quasi norms

To get a result for the best m-term approximation for arbitrary α we will split the error estimation into two parts. This can be done thanks to Lemma 2.8. We use Stechkin's Lemma to approximate the error from $\mathcal{A}_p^{\alpha,d}$ in $\mathcal{A}_1^{\alpha,d}$ and the treat the error of $\mathcal{A}_1^{\alpha,d}$ in L_q with Theorem 4.2.

To show this Theorem we first need an auxiliary greedy approximation results for hyperbolic crosses due to Temlyakov.

Lemma 4.1 ([29, Theorem 2.6]). *For every trigonometric polynomial with frequencies of maximal degree N, there exist constructive greedy-type approximation methods, which provide m-term polynomials Q_m with the following properties. For $2 \leq q < \infty$*

$$\left\|f - Q_m\right\|_q \lesssim m^{-\frac{1}{2}} \|f\|_{\mathcal{A}_1^\alpha}. \tag{4.1}$$

And for $q = \infty$

$$\left\|f - Q_m\right\|_\infty \lesssim m^{-\frac{1}{2}} \log(N)^{\frac{1}{2}} \|f\|_{\mathcal{A}_1^\alpha}. \tag{4.2}$$

The proof of this Theorem is based on the ideas from Theorem 3.2 combined with techniques from [6, Theorem 6.1] and employs results from [29, Theorem 2.5/2.6]. In particular, we are interested in the cases $q = 2$ and $q = \infty$, where the latter one can be applied to Proposition 1.2.

Theorem 4.2. *Let $n, d \in \mathbb{N}$ such that $p = 1$ and $2 \leq q \leq \infty$ let further $\alpha > 0$ then it holds*

$$n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha} \lesssim \sigma_n\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_q \lesssim n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha+\mu} \tag{4.3}$$

where

$$\lambda = \frac{1}{2},$$

and $\mu = \frac{1}{2}$ if both $q = \infty$ and $d > 1$ or 0 otherwise.

Proof. Again assume $n = 2^m$, using the same idea and notation as in (the proof of) Theorem 3.2 we can restrict and decompose $f \in \mathcal{A}_p^{\alpha,d}$ as follows

$$f(\mathbf{x}) = \sum_{i=0}^{\infty} f_i(\mathbf{x}),$$

where $f_j = \sum_{\mathbf{k}\in\square_j} \hat{f}(\mathbf{k})e^{i\mathbf{x}\mathbf{k}}$.

We now again want to split the problem into three blocks, the first finite one where we use a lot of our "approximation budget" per frequency to get a precise result, the second finite one where we start allowing some error, and an outer infinite one where the decay in α helps us bound the error

$$\|f-P\|_q \leq \sum_{k=0}^{L} \|f_k - f_k\|_q + \sum_{k=L+1}^{M} \|f_k - P_k\|_q + \sum_{k=M+1}^{\infty} \|f_k\|_q =: S_1 + S_2 + S_3. \quad (4.4)$$

For $j = 0 \ldots L$ we chose $P_k = f_k$, then obviously $S_1 = 0$.

To estimate S_2 now put $m_k = (k-L)^{-2}2^L L^{d-1}$ for all $k = L+1 \ldots M$ and choose, by Lemma 4.1 $P_k \in \Sigma_{m_k}$ such that

$$\|f_k - P_k\|_q \lesssim m_k^{-\lambda} m^{\mu} 2^{-(k\alpha)}. \quad (4.5)$$

This is possible because of (4.1), or (4.2) for $d = \infty$, where the norm $\|f_k\|_{\mathcal{A}_1^\alpha}$ is bounded by $2^{-(k\alpha)}$ since it has support only on the k-th hyperbolic layer, see (3.1). For $d = 1$ instead use [29, Theorem 2.3] to avoid the term $m^{\frac{1}{2}}$ in the case $q = \infty$. Choose now L as in (3.3) and M as in (3.4), then $P := \sum_{k=0}^{M} P_k$ is a linear combination of at most

$$\begin{aligned} \sum_{k=0}^{L} C_k + \sum_{k=L+1}^{M} m_k &\lesssim \sum_{k=0}^{L} 2^k k^{d-1} + \sum_{k=L+1}^{M} (k-L)^{-2} 2^L L^{d-1} \\ &\lesssim 2^L L^{d-1} \\ &\leq 2^L m^{d-1} \\ &\lesssim 2^m \end{aligned} \quad (4.6)$$

trigonometric terms.

Now we can estimate S_2 by using (4.5)

$$\begin{aligned} S_2 &\lesssim \sum_{k=L+1}^{M} m_k^{-\lambda} 2^{-k\alpha} m^{\mu} \\ &= \sum_{k=L+1}^{M} (k-L)^{2\lambda} 2^{-L\lambda} L^{-\lambda(d-1)} 2^{-k\alpha} m^{\mu} \\ &\lesssim 2^{-L(\alpha+\lambda)} L^{-\lambda(d-1)} m^{\mu} \\ &\asymp 2^{-(m-(d-1)\log_2 m)(\alpha+\lambda)} (m-(d-1)\log_2 m)^{-\lambda(d-1)} m^{\mu} \\ &\asymp 2^{-m(\alpha+\lambda)} m^{(d-1)(\alpha+\lambda)} m^{-\lambda(d-1)} m^{\mu} \\ &= 2^{-m(\alpha+\lambda)} m^{(d-1)\alpha+\mu}. \end{aligned} \quad (4.7)$$

To estimate S_3 we use the fact that

$$\|f_k\|_q \leq \|f_k\|_\infty \leq \sum_j |\hat{f}_k(j)| = \|\hat{f}_k(j)\|_{\ell_1} \lesssim 2^{-k\alpha},$$

and therefore

$$\begin{aligned} S_3 &\lesssim \sum_{k=M+1}^{\infty} 2^{-k\alpha} \\ &\leq 2^{-M\alpha} \\ &= 2^{-(m(1+\frac{\lambda}{\alpha})-(d-1)\log_2(m))\alpha} \\ &= 2^{-m(\alpha+\lambda)} m^{(d-1)\alpha}. \end{aligned} \tag{4.8}$$

Combining now (4.7) and (4.8) we can estimate (4.4) as follows

$$\begin{aligned} \sigma_{2^{m+d}}(f)_q &\leq \|f - P\|_q \\ &\leq S_0 + S_1 + S_2 \\ &\lesssim 0 + 2^{-m(\alpha+\lambda)} m^{(d-1)\alpha} + 2^{-m(\alpha+\lambda)} m^{(d-1)\alpha} \\ &\lesssim 2^{-m(\alpha+\lambda)} m^{\alpha} \end{aligned} \tag{4.9}$$

Now arguing with monotonicity as in the results before we obtain the assertion for all n.

To get a lower bound we will use a specific function for $n = 2^m m^{d-1}$, this also works for p other than $p = 1$. Choose

$$f(\mathbf{x}) = C n^{-(\alpha+\frac{1}{p})} \log(n)^{\alpha(d-1)} \sum_{\mathbf{k}\in\square_m} e^{i\mathbf{k}\mathbf{x}}. \tag{4.10}$$

This function has $\mathcal{A}_p^{\alpha,d}$-norm as follows

$$\begin{aligned} \|f\|_{\mathcal{A}_p^{\alpha,d}} &\asymp n^{-(\alpha+\frac{1}{p})} \log(n)^{\alpha(d-1)} \Big(\sum_{\mathbf{k}\in\square_m} \big(\prod_{j=1}^{d} (1+|k_j|)^{\alpha} \big)^p \Big)^{\frac{1}{p}} \\ &\asymp n^{-(\alpha+\frac{1}{p})} \log(n)^{\alpha(d-1)} \Big(\sum_{\mathbf{k}\in\square_m} 2^{\alpha p \|\mathbf{k}\|_1} \Big)^{\frac{1}{p}} \\ &\asymp n^{-(\alpha+\frac{1}{p})} \log(n)^{\alpha(d-1)} \Big(\sum_{\mathbf{k}\in\square_m} 2^{\alpha p m} \Big)^{\frac{1}{p}} \\ &\asymp n^{-(\alpha+\frac{1}{p})} \log(n)^{\alpha(d-1)} \Big(2^m m^{d-1} 2^{\alpha p m} \Big)^{\frac{1}{p}} \\ &\asymp 1. \end{aligned} \tag{4.11}$$

This implies that $f \in B_C\big(\mathcal{A}_p^{\alpha,d}\big)$ for properly chosen constant C.

Now define $g = C_2 n^{-\frac{1}{2}} \sum_{k\in\square_m\setminus K_m} e^{i\mathbf{k}\mathbf{x}}$ and $h = \sum_{k\in K_m} a_{\mathbf{k}} e^{i\mathbf{k}\mathbf{x}}$ where $K_m \subset \square_m$ is a frequency set with at most $\frac{n}{2}$ elements. For these functions we get

$$\|g\|_{L_2} = C_2 n^{-\frac{1}{2}} \big(|\square_m \setminus K_m|\big)^{\frac{1}{2}} \lesssim n^{-\frac{1}{2}} n^{\frac{1}{2}} = 1 \tag{4.12}$$

and

$$\begin{aligned}\langle f-h, g\rangle &= \langle f, g\rangle \\ &= CC_2 n^{-(\alpha+\frac{1}{p}+\frac{1}{2})}\log(n)^{\alpha(d-1)} \sum_{k\in\square_m\setminus K_m} 1 \\ &\geq CC_2 n^{-(\alpha+\frac{1}{p}+\frac{1}{2})}\log(n)^{\alpha(d-1)}\big(|\square_m| - |K_m|\big) \\ &\gtrsim n^{-(\alpha+\frac{1}{p}-\frac{1}{2})}\log(n)^{\alpha(d-1)}.\end{aligned} \tag{4.13}$$

Using the Cauchy-Schwarz inequality and (4.12) yields

$$\langle f-h, g\rangle \leq \|f-h\|_{L_2}\|g\|_{L_2} \leq \|f-h\|_2. \tag{4.14}$$

Applying now (4.13) and (4.14) gives

$$\begin{aligned}\sigma_{n2^d}\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_q &\gtrsim \inf_{h\in\Sigma_n} \|f-h\|_2 \\ &\geq \inf_{h\in\Sigma_n} \langle f-h, g\rangle \\ &\gtrsim n^{-(\alpha+\frac{1}{p}-\frac{1}{2})}\log(n)^{\alpha(d-1)}\end{aligned} \tag{4.15}$$

where indeed every $h \in \Sigma_{\frac{n}{2}}$ can be chosen as above and g accordingly. The factor $\frac{1}{2}$ can be ignored since we are investigating the asymptotic behaviour. ■

This Theorem is very useful for $p = 1$ but for smaller p we do not gain anything. To remedy this, we can employ the Stechkin Lemma to gain a factor when changing from $\mathcal{A}_p^{\alpha,d}$ to $\mathcal{A}_1^{\alpha,d}$.

Corollary 4.3. *Let $n, d \in \mathbb{N}$ such that $0 < p \leq 1$ and $2 \leq q \leq \infty$ let further $\alpha > 0$ then it holds*

$$n^{-(\alpha+\lambda)}\log(n)^{(d-1)\alpha} \lesssim \sigma_n\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_q \lesssim n^{-(\alpha+\lambda)}\log(n)^{(d-1)\alpha+\mu} \tag{4.16}$$

where

$$\lambda = \frac{1}{p} - \frac{1}{2},$$

and $\mu = \frac{1}{2}$ if both $q = \infty$ and $d > 1$ or 0 otherwise.

Proof. We use Lemma 2.8 to simply split the approximation into two parts and apply the Stechkin Lemma 2.9 to $\mathcal{A}_p^\alpha$ yielding

$$\sigma_{2n2^d}\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_q \leq \sigma_{n2^d}\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_{\mathcal{A}_1^{\alpha,d}}\sigma_{n2^d}\Big(B_1\big(\mathcal{A}_1^{\alpha,d}\big)\Big)_q \leq n^{-\frac{1}{p}+1}\sigma_{n2^d}\Big(B_1\big(\mathcal{A}_1^{\alpha,d}\big)\Big)_q$$

and then apply the upper bound from Theorem 4.2 to obtain

$$\sigma_{2n2^d}\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_q \lesssim n^{-(\alpha+\lambda)}\log(n)^{(d-1)\alpha+\mu}. \tag{4.17}$$

The lower bound can be shown as in Theorem 4.2. ■

This result indeed expands Proposition 1.3 to the regime where α maybe be less than one half. In addition, it works for a wider variety of p and q while also giving a lower bound that is sharp up to a logarithmic factor.

4.2 Results for $p > 1$

In the previous section results for the best m-term approximation of mixed weighted Wiener classes with $p \leq 1$ were shown, in particular a result for $\mathcal{A}_1^{\alpha,d}$. For $p = 2$ the weighted Wiener classes coincide with other important spaces and therefore the following result was already shown in 1989 by Belinskii.

Theorem 4.4 ([1, Theorem 2] and [2, Theorem 11.1.6])**.** *For $\alpha > \frac{1}{2}$ and $n \in \mathbb{N}_0$ it holds*

$$n^{-\alpha}\log(n)^{(d-1)\alpha} \lesssim \sigma_n\Big(B_1\big(\mathcal{A}_2^{\alpha,d}\big)\Big)_\infty \lesssim n^{-\alpha}\log(n)^{(d-1)\alpha+\frac{1}{2}}. \tag{4.18}$$

This result again only holds for $\alpha > \frac{1}{2}$ and unlike before there is no clear path to loosening the restrictions on α. To get results for $p > 1$ one could now just use Stechkin's Lemma 2.9 but then the restriction on α would be retained. Instead, one can try to modify the proof from Theorem 4.2 to lower the requirements on α. This is achieved by estimating the $\mathcal{A}_1$-norm of f on the hyperbolic layers in terms of the $\mathcal{A}_p$-norm by abusing the fact that these layers have finite support.

Theorem 4.5. *Let $n, d \in \mathbb{N}$ such that $1 < p \leq q$ and $2 \leq q \leq \infty$ let further $\alpha > 1 - \frac{1}{p}$ then it holds*

$$n^{-(\alpha+\lambda)}\log(n)^{(d-1)\alpha} \lesssim \sigma_{n2^d}\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_q \lesssim n^{-(\alpha+\lambda)}\log(n)^{(d-1)\alpha+\mu} \tag{4.19}$$

where

$$\lambda = \frac{1}{p} - \frac{1}{2},$$

and $\mu = \frac{1}{2}$ if $q = \infty$ and $d > 1$ or 0 otherwise.

Proof. Follow the proof of Theorem 4.2 until (4.5). There [29, Theorem 2.6] is used with $\|f_k\|_{\mathcal{A}_1^\alpha} \leq 2^{-(k\alpha)}$. However, in this setting now only the $\mathcal{A}_p^\alpha$-norm of f is bounded by one instead of the $\mathcal{A}_1^\alpha$-norm. To fix this recall that the support of f_k is limited to the k-th hyperbolic layer and therefore at most $2^k k^{d-1}$. This now gives, by Remark 2.2 and after considering that a finite sequence with fixed ℓ_p-norm has largest possible ℓ_1-norm when all its entries are equal, the following

$$\begin{aligned} \|f_k\|_{\mathcal{A}_1^\alpha} &\leq 2^k k^{d-1} \left(2^k k^{d-1}\right)^{-\frac{1}{p}} \\ &\leq \|f_k\|_{\mathcal{A}_p^\alpha} 2^{k(1-\frac{1}{p})} k^{(d-1)(1-\frac{1}{p})} \\ &\leq 2^{-(k\alpha)} 2^{k(1-\frac{1}{p})} k^{(d-1)(1-\frac{1}{p})}. \end{aligned} \tag{4.20}$$

To continue the estimation of S_2 as before $\alpha > 1 - \frac{1}{p}$ is needed, so that the sum below collapses to $k = L + 1$.

$$\begin{aligned} S_1 &\lesssim \sum_{k=L+1}^{M} d m_k^{-\frac{1}{2}} 2^{-(k\alpha)} 2^{k(1-\frac{1}{p})} k^{(d-1)(1-\frac{1}{p})} m^\mu \\ &= \sum_{k=L+1}^{M} d(k-L) 2^{-L\frac{1}{2}} L^{-\frac{1}{2}(d-1)} 2^{-k(\alpha-1+\frac{1}{p})} k^{(d-1)(1-\frac{1}{p})} m^\mu \\ &\lesssim 2^{-L(\frac{1}{2}+\alpha-1+\frac{1}{p})} L^{-(\frac{1}{2}-1+\frac{1}{p})(d-1)} m^\mu \\ &\asymp 2^{-(m-(d-1)\log_2 m)(\lambda+\alpha)} (m-(d-1)\log_2 m)^{-\lambda(d-1)} m^\mu \\ &\asymp 2^{-m(\lambda+\alpha)} m^{(d-1)(\frac{1}{2}+\alpha)} m^{-\lambda(d-1)} m^\mu \\ &= 2^{-m(\lambda+\alpha)} m^{(d-1)\alpha+\mu}. \end{aligned} \tag{4.21}$$

Note that here λ has a different value than in Theorem 4.2.

The estimation of S_3 and the lower bound now work as before. ■

This indeed recovers the original result for $1 < p < 2$. While the bound on α might seem arbitrary, even if one were to use interpolation techniques between $\mathcal{A}_1$ and $\mathcal{A}_2$ one would only achieve the same restrictions. Indeed, $\alpha > \frac{p-1}{p}$ seems to be the natural barrier.

4.3 Basis pursuit denoising

An application of the best n-term approximation results above will be the following bound for basis pursuit denoising (BPD) in weighted Wiener classes. The idea behind BPD is, given a function f to guarantee that there is a good sparse approximation with trigonometric polynomials (this is where the best n-term approximation comes in) and then employ compressed sensing techniques to recover this (unknown) trigonometric polynomial with just a few random points where the difference between the polynomial and f is treated as noise (hence denoising). Examples of this strategy can be found in [23, 22].

Let $\mathbf{X} = (x_1 \ldots x_n)$ be a sequence of N sampling points drawn i.i.d. from $\mathbb{T}^d$. Then the N by m matrix

$$(F_{\mathbf{X}})_{j,k} = e^{ikx_j}, \tag{4.22}$$

would satisfy $(F_{\mathbf{X}}\mathbf{c})_j = f(x_j)$ if f were a trigonometric polynomial of degree at most m and $\mathbf{c}$ its vector of Fourier coefficients. However, we now want to allow more general $f \in \mathcal{A}_p^{\alpha,d}$. By Corollary 4.3 we know that there is an n-sparse trigonometric polynomial f_n with $\|f - f_n\|_q \leq Cn^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha+\mu}$. We now want to recover this f_n with sparse approximation but can only sample f (since we do not know f_n, only that it exists). We therefore put $y_i = f(x_i)$ for $i = 1 \ldots N$.

Proposition 4.6 ([22, Theorem 3.2]). *Let $\mathbf{c}^*$ be the solution to the minimisation problem*

$$\min \|\mathbf{c}\|_1 \quad \text{subject to } \|F_{\mathbf{X}}\mathbf{c} - y\|_2 \leq \nu, \tag{4.23}$$

then it satisfies the bound

$$\|\mathbf{c} - \mathbf{c}^*\|_2 \leq C_1 \frac{\nu}{\sqrt{N}} \tag{4.24}$$

with probability at least $1 - \varepsilon$ if $\|f_n - f\|_2 \leq \nu$ and if the isometry constants of the Matrix satisfy $\delta_{3n} + 3\delta_{4n} < 2$. This is true when the number of samples taken is at least

$$\frac{N}{\log N} \geq C_0 \, n \, d \log(n)^3 \log(\varepsilon^{-1}).$$

If we choose N equal to this bound and set ν to the bound from Corollary 4.3 with $q = 2$ we get from (4.24)

$$\|\mathbf{c} - \mathbf{c}^*\|_2 \leq C_2 \frac{n^{-(\alpha+\frac{1}{p})} \log(n)^{(d-1)\alpha-2}}{\sqrt{d \log(\varepsilon^{-1})}}. \tag{4.25}$$

This is achieved with an oversampling factor that is logarithmic in n and linear in d. In other words the curse of dimensionality does not affect the number of required samples.

It can however still be slightly improved by employing a more recent result by Haviv and Regev based on work by Bourgain [3].

Proposition 4.7 ([11, Theorem 3.7]). *Let $\mathbf{M} \in \mathbb{C}^{n^d \times n^d}$ be a unitary matrix with $\|\mathbf{M}\|_\infty \leq O(n^{-\frac{1}{2}})$. For sufficiently small $\varepsilon > 0$ and*

$$N = C \, n \, d \log(n)^3 \, \varepsilon^{-4}$$

a matrix $\mathbf{A} \in \mathbb{C}^{N \times n^d}$ with independently uniformly chosen rows from $\mathbf{M}$ multiplied by $\sqrt{\frac{n^d}{N}}$ has restricted isometry constants $\delta_n < \varepsilon$ with high probability.

This result allows to reduce the number of sampling points in Proposition 4.6 by a factor of $\log(n)$.

This is done by starting with the Fourier matrix with n^d frequencies and points and then subsampling F from this.

4.4 Numerical experiments

To illustrate the BPD results from the previous section we can compute the approximation of e.g. a trigonometric monomial with high degree via ℓ_1 and ℓ_2 minimisation. In this context BP or ℓ_1 minimisation refers (4.23) while ℓ_2 minimisation refers to the same problem, where instead of $\|\mathbf{c}\|_1$ the term $\|\mathbf{c}\|_2$ is minimised. The error is always measured in the L_2-norm.

The following example is the approximation of $\exp(2\pi i\, 50)$ (scaled in such a way that it is in the 1-ball of $\mathcal{A}_1^1$) with 120 frequencies and 200 iterations from a "fista" (fast iterative soft thresholding algorithm) solver, which is a common ℓ_1 minimiser.

Figure 1: Approximation of monome

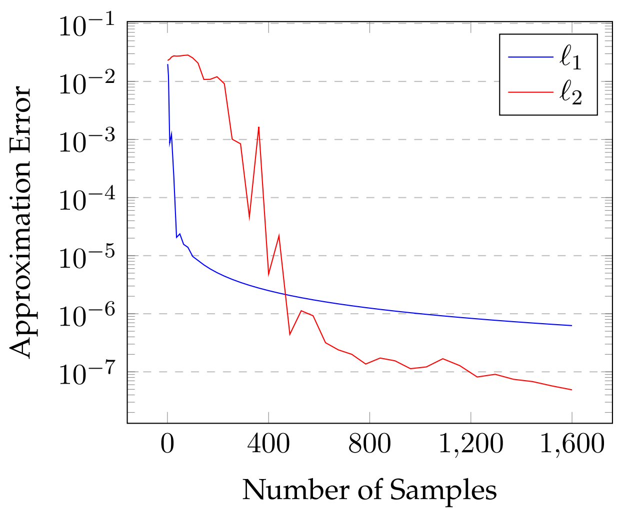
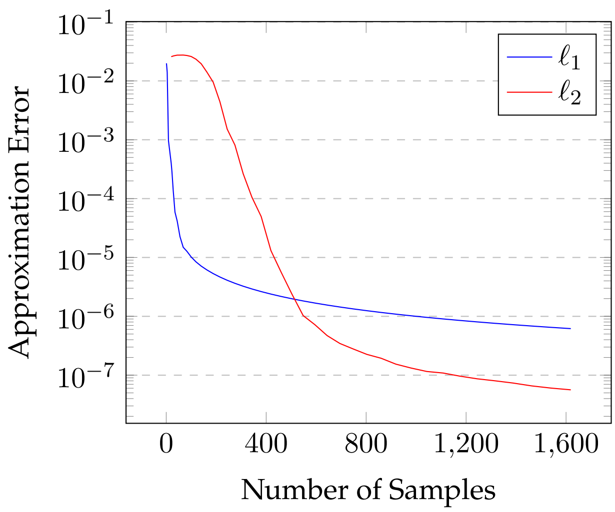


Figure 2: smoothed Approximation



Indeed, we see that error ℓ_1 obtained with BP becomes smaller a lot sooner than the ℓ_2 error. However, for large n the least squares error decays faster by a multiplicative constant factor. The ℓ_1 minimisation also seems to be more stable compared to the ℓ_2 one as can be seen in Figure 1. The second graphic shows the smoothed average over 100 calculations.

The second numerical experiment was done in 3 dimensions with a 3-dimensional monomial and "only" 20 frequencies in each dimension (so 8000 in total). Here the difference between the ℓ_1 and ℓ_2 minimisation becomes even more pronounced. Where the ℓ_1 minimisation has good decay almost immediately the error of the ℓ_2 minimisation

Figure 3: Approximation in 3 Dimensions

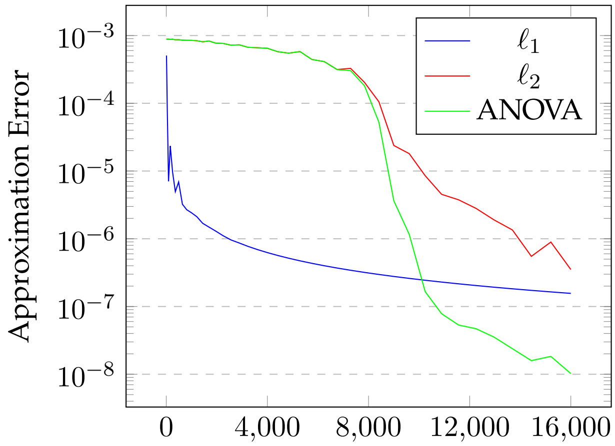
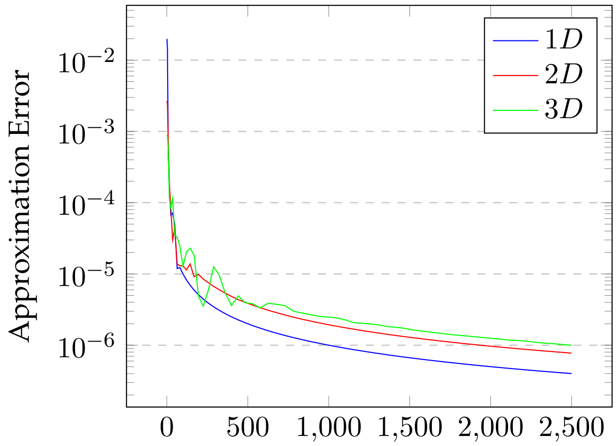


Figure 4: Dimension Comparison



only drops significantly after the system becomes overdetermined. This illustrates very well how the number of samples suffer from the curse of dimensionality in this regime while in the ℓ_1 case they do not. This picture was only smoothed over 20 runs since the runtime of this program does not behave as nicely as the number of required samples (Note that sophisticated approximation algorithms like ANOVA have better runtime then ℓ_1 or ℓ_2 minimisation for higher dimensions) and instead suffers from the curse of dimensionality even for the ℓ_1 case. For very large numbers of samples the ℓ_2 minimisation again closes the gap and the ANOVA (for details regarding ANOVA see [25]) approximation even overtakes it (as does ℓ_2, eventually) however this only happens for overdetermined systems.

Figure 4 shows the comparison between different dimensional problems for monomes with similar $\mathcal{A}_1^1$ norms. As one can see the number samples required to reach some error is indeed only influenced in a linear fashion by the dimension d. E.g. in one dimension it takes about 1000 sampling points for the error to reach 10^{-6} in two dimensions about 2000 and in three 2500. One reason why the gaps are not uniform might be that different regulariasation terms where used. The optimal terms could only be computed numerically here, since the above bounds are only asymptotic.

5 Sampling numbers of mixed weighted Wiener spaces in L_2

5.1 Linear and non-linear sampling numbers

As mentioned in the introduction the Kolmogorov numbers of mixed weighted Wiener space have already been studied in [18] and Proposition 2.18 states that they behave like

$$d_n\big(\mathrm{id}\colon \mathcal{A}_1^{\alpha,d} \to L_2(\mathbb{T}^d)\big) \asymp n^{-\alpha} \log(n)^{\alpha(d-1)}. \tag{5.1}$$

This gives a lower bound on the linear sampling numbers via Lemma 2.13.

$$\varrho_n^{\mathrm{lin}}\big(\mathcal{A}_1^{\alpha,d}\big)_2 \gtrsim n^{-\alpha} \log(n)^{\alpha(d-1)}. \tag{5.2}$$

Similar to the linear sampling numbers the non-linear ones can also be bounded from below. The results from Section 3 in particular Remark 3.4 give by Lemma 2.11 the following bound.

Theorem 5.1. *For $1 \le p \le 2$ and $\alpha > \frac{p-1}{p}$ it holds*

$$\varrho_n(\mathcal{A}_p^{\alpha,d})_2 \ge c_n(id\colon \mathcal{A}_p^{\alpha}(\mathbb{T}^d) \to L_q(\mathbb{T}^d)) \asymp n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha}. \tag{5.3}$$

In particular this implies

$$\varrho_n\big(\mathcal{A}_1^{\alpha,d}\big)_2 \gtrsim n^{-(\alpha+\frac{1}{2})} \log(n)^{(d-1)\alpha}. \tag{5.4}$$

This lower bound is better by one half in the main rate than the one for linear sampling numbers.

Now using the new Results from [14] we can even get an upper bound on the non-linear sampling numbers. This is done by using the improved results about best m-term approximation i.e. Corollary 4.3 and Theorem 4.5.

Proposition 5.2. *For $1 \le p \le 2$ and $\alpha > \frac{p-1}{p}$ it holds*

$$\varrho_n\big(\mathcal{A}_p^{\alpha,d}\big)_2 \lesssim n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha+3(\alpha+\lambda)+\frac{1}{2}} \tag{5.5}$$

where

$$\lambda = \frac{1}{p} - \frac{1}{2}$$

in particular for $p = 1$ it holds

$$\varrho_n\big(\mathcal{A}_1^{\alpha,d}\big)_2 \lesssim n^{-(\alpha+\frac{1}{2})} \log(n)^{(d-1)\alpha+3(\alpha+\frac{1}{2})+\frac{1}{2}} \tag{5.6}$$

Proof. Employ [14, Theorem 3.2] and follow the idea of [14, Corollary 4.4] but allow for general p. Choose $M = \lfloor n^{\frac{\alpha+\lambda}{\alpha}} \rfloor$ instead. ■

This means there is only a dimension-independent gap in the order of the log of the size $3(\alpha + \frac{1}{2})$ between the upper and lower bound of the (non-linear) sampling numbers of weighted Wiener classes with mixed weights.

In addition, even the upper bound on the non-linear sampling numbers is smaller by a factor of one half in the main rate then the lower bound of the linear sampling numbers i.e.

$$\frac{\varrho_n\big(\mathcal{A}_1^{\alpha,d}\big)_2}{\varrho_n^{\mathrm{lin}}\big(\mathcal{A}_1^{\alpha,d}\big)_2} \lesssim n^{-\frac{1}{2}} \log(n)^{3(\alpha+\frac{1}{2})}. \tag{5.7}$$

6 Conclusion

As mentioned in the introduction the main point of this thesis was to expand Proposition 1.3 for the best m-term approximation for weighted mixed Wiener classes from [14] to the regime where α may be less than one half. As well as to complement Proposition 2.18 from [18] with an asymptotic bound for the Gelfand numbers in these spaces.

Indeed, Theorem 3.2 shows that the Gelfand numbers have the same behaviour as the Weyl and Bernstein numbers in $\mathcal{A}_1^{\alpha,d}$

$$b_n\big(\mathrm{id}\colon \mathcal{A}_1^{\alpha,d} \to L_2\big) \asymp x_n\big(\mathrm{id}\colon \mathcal{A}_1^{\alpha,d} \to L_2\big) \asymp c_n\big(\mathrm{id}\colon \mathcal{A}_1^{\alpha,d} \to L_2\big) \asymp n^{-(\alpha+\frac{1}{2})} \log(n)^{\alpha(d-1)}.$$

Where L_2 is understood to be restricted to the torus $L_2(\mathbb{T}^d)$. In particular the Gelfand numbers decay faster than the Kolmogorov numbers by one half in the main rate.

However, Theorem 3.2 is not restricted to the case where $p = 1$. Instead, as stated in Remark 3.4 it holds for all $0 < p \leq 2$.

On the topic of the best m-term approximation for weighted Wiener classes with mixed weights the restriction on α was bypassed completely for $p \leq 1$, by showing the bound directly instead of embedding it into some L_2 Sobolev space of periodic functions with bounded mixed derivative, as was done in [14]. This direct result was obtained by using results from hyperbolic cross approximation [29] and combining them with decomposition methods also used for other quasi s-numbers. Therefore instead of having the bound (1.2)

$$\sigma_n\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_\infty \leq n^{-(\alpha+\frac{1}{2})} \log(n)^{(d-1)\alpha+\frac{1}{2}}, \tag{6.1}$$

for only $\alpha > \frac{1}{2}$ we get from Theorem 4.2 the bound

$$n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha} \lesssim \sigma_n\Big(B_1\big(\mathcal{A}_p^{\alpha,d}\big)\Big)_q \lesssim n^{-(\alpha+\lambda)} \log(n)^{(d-1)\alpha+\mu} \tag{6.2}$$

for all $\alpha > 0$, $p = 1$ and $2 \leq q \leq \infty$. In addition, this Theorem also provides a lower bound that is sharp up to a logarithmic factor.

Corollary 4.3 employs the Stechkin Lemma to expand this result to the quasi-Banach setting $0 < p \leq 1$ and Theorem 4.5 modifies the original proof to $p \geq 1$ for $\alpha > \frac{p-1}{p}$.

Both of these bounds can be useful tools when working with mixed weighted Wiener spaces as can be seen in [14]. The best m-term approximation result is used to show error estimates, while the Gelfand bound is used as an optimality statement to contrast this with, since the Gelfand numbers are the worst case the error of the best non-linear reconstruction of a function from linear samples. This was demonstrated in Section 5 were these bounds were used to get a lower bound on the linear sampling numbers and a (in the main rate) tight asymptotic bound on the non-linear sampling numbers by employing results from [14]. In particular, it was shown

$$n^{-(\alpha+\frac{1}{2})} \log(n)^{(d-1)\alpha} \lesssim \varrho_n\big(\mathcal{A}_1^{\alpha,d}\big)_2 \lesssim n^{-(\alpha+\frac{1}{2})} \log(n)^{(d-1)\alpha+3(\alpha+\frac{1}{2})+\frac{1}{2}}, \tag{6.3}$$

while for the linear sampling numbers only the lower bound

$$\varrho_n^{\mathrm{lin}}\big(\mathcal{A}_1^{\alpha,d}\big)_2 \gtrsim n^{-\alpha} \log(n)^{\alpha(d-1)} \tag{6.4}$$

holds. This shows that there exists a gap of at least one half in the main rate between linear and non-linear reconstruction in mixed weighted Wiener classes. This can be shown because the Kolmogorov numbers decay slower than the Gelfand numbers and the best m-term approximation.

Bibliography

[1] E. S. Belinskii. Approximation of functions of several variables by trigonometric polynomials with given number of harmonics, and estimates of ϵ- entropy. *Anal. Math.*, 15(2):67–74, 1989.

[2] E. S. Belinskii and R. M. Trigub. *Fourier Analysis and Approximation of Functions*. Springer Dordrecht, 2004.

[3] J. Bourgain. *An improved estimate in the restricted isometry problem*. In Springer International Publishing, 2014, pages 65–70.

[4] N. Buchmann. *Fehlerabschätzungen von Näherungslösungen unendlicher Gleichungssysteme durch Gelfandzahlen von Tensorproduktoperatoren*. Dissertation, University of Oldenburg, 1999.

[5] F. Dai and V. N. Temlyakov. Universal discretization and sparse sampling recovery. *preprint*, 2023.

[6] R. DeVore and V. N. Temlyakov. Nonlinear approximation by trigonometric sums. *Journal of Fourier Analysis and Applications*, 2:29–48, 1995.

[7] S. Dirksen and T. Ullrich. Gelfand numbers related to structured sparsity and Besov space embeddings with small mixed smoothness. *Journal of Complexity*, 48:69–102, 2018.

[8] D. Dũng, V. N. Temlyakov, and T. Ullrich. *Hyperbolic Cross Approximation*. Advanced Courses in Mathematics - CRM Barcelona. Birkhäuser/Springer, 2018.

[9] S. Foucart, A. Pajor, H. Rauhut, and T. Ullrich. The Gelfand widths of ℓ_p-balls for $0 < p \leq 1$. *Journal of Complexity*, 26(6):629–640, 2010.

[10] E. D. Gluskin. Нормы случайных матриц и поперечники конечномерных множеств. *Matematicheskii Sbornik*, 120(2):180, 1983. English translation: E. D. Gluskin. Norms of random matrices and widths of finite-dimensional sets. *Mathematics of the USSR-Sbornik*, 48(1):173, 1984.

[11] I. Haviv and O. Regev. *The Restricted Isometry Property of subsampled Fourier matrices*. In Springer International Publishing, 2017, pages 163–179.

[12] A. Hinrichs, J. Prochno, and M. Sonnleitner. Random sections of ℓ_p-ellipsoids, optimal recovery and Gelfand numbers of diagonal operators. *preprint*, 2022.

[13] T. Jahn and T. Ullrich. On the optimal constants in the two-sided Stechkin inequalities. *Journal of Approximation Theory*, 269:105607, 2021.

[14] T. Jahn, T. Ullrich, and F. Voigtlaender. Sampling numbers of smoothness classes via ℓ^1-minimization. *preprint*, 2023.

[15] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate periodic functions by trigonometric polynomials based on rank-1 lattice sampling. *Journal of Complexity*, 31(4):543–576, 2015.

[16] Y. Kolomoitsev, T. Lomako, and S. Tikhonov. Sparse grid approximation in weighted Wiener spaces. *preprint*, 2021.

[17] V. E. Maiorov. Дискретизация задачи о поперечниках. *Uspekhi Mat. Nauk*, 30:179–180, 6(186), 1975.

[18] V. D. Nguyen, V. K. Nguyen, and W. Sickel. s-Numbers of embeddings of weighted Wiener algebras. *Journal of Approximation Theory*, 279, 2022.

[19] E. Novak and H. Wozniakowski. *Tractability of multivariate problems. Vol. 1: Linear information*. EMS Tracts in Mathematics, 2008, pages xii+384.

[20] A. Pietsch. s-Numbers of operators in Banach spaces. eng. *Studia Mathematica*, 51(3):201–223, 1974.

[21] A. Pietsch. *Eigenvalues and s-Numbers*. Akademische Verlagsgesellschaft Gest & Portig K.-G., 1987.

[22] H. Rauhut. Stability results for random sampling of sparse trigonometric polynomials. *IEEE Transactions on Information Theory*, 54(12):5661–5670, 2008.

[23] H. Rauhut and R. Ward. Sparse legendre expansions via ℓ_1 minimization. *Journal of Approximation Theory*, 164, 2012.

[24] W. Rudin. *Functional Analysis*. McGraw-Hill, 1991.

[25] P. Schröter. *Hochdimensionale Approximation mit teilweise periodischen Randbedingungen*. Master's thesis, Faculty of Mathematics, TU Chemnitz, 2023.

[26] M. Z. Solomyak and V. M. Tikhomirov. О геометрических характеристиках вложения классов W_p^α в C. *Izv. Vyssh. Uchebn. Zaved. Mat.*, 10:76–82, 1967.

[27] S. B. Stechkin. Об абсолютной сходимости ортогональных рядов. *Dokl. Akad. Nauk SSSR*, 102(1):37–40, 1955.

[28] V. N. Temlyakov. *Approximations of functions with bounded mixed derivative*, volume 178. Trudy Mat. Inst. Steklov., 1986, pages 3–113.

[29] V. N. Temlyakov. Constructive sparse trigonometric approximation and other problems for functions with mixed smoothness. *Sbornik: Mathematics*, 206, 2015.

[30] V. M. Tikhomirov. Одно замечание об n-мерных поперечниках множеств в банаховых пространствах. *Uspekhi Mat. Nauk*, 20:227–230, 1(121), 1965.

[31] H. J. Triebel. Interpolationseigenschaften von Entropie und Durchmesseridealen kompakter Operatoren. *Studia Mathematica*, 34:89–107, 1970.

[32] T. Ullrich. Lecture notes: metric entropy and n-widths in approximation theory.

[33] J. Vybíral. Widths of embeddings in function spaces. *Journal of Complexity*, 24(4):545–570, 2008.

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Masterarbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe.

Die vorliegende Arbeit ist frei von Plagiaten. Alle Ausführungen, die wörtlich oder inhaltlich aus anderen Schriften entnommen sind, habe ich als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch bei keinem anderen Prüfer als Prüfungsleistung eingereicht und ist auch noch nicht veröffentlicht.

....

Moritz Moeller