

Radial Time-Frequency Analysis and Embeddings of Radial Modulation Spaces

Holger Rauhut
University of Wrocław, Mathematical Institute
Pl. Grunwaldzki 2/4, 50–384 Wrocław, Poland
rauhut@ma.tum.de

February 8, 2005

Abstract

In this paper we construct frames of Gabor type for the space $L^2_{rad}(\mathbb{R}^d)$ of radial L^2 -functions, and more generally, for subspaces of modulation spaces consisting of radial distributions. Hereby, each frame element itself is a radial function. This construction is based on a generalization of the so called Feichtinger-Gröchenig theory – sometimes also called coorbit space theory – which was developed in an earlier article. We show that this new type of Gabor frames behaves better in linear and non-linear approximation in a certain sense than usual Gabor frames when approximating a radial function. Moreover, we derive new embedding theorems for coorbit spaces restricted to invariant vectors (functions) and apply them to modulation spaces of radial distributions. As a special case this result implies that the Feichtinger algebra $(S_0)_{rad}(\mathbb{R}^d) = M^1_{rad}(\mathbb{R}^d)$ restricted to radial functions is embedded into the Sobolev space $H^{(d-1)/2}_{rad}(\mathbb{R}^d)$. Moreover, for $d \geq 2$ the embedding $(S_0)_{rad}(\mathbb{R}^d) \hookrightarrow L^2_{rad}(\mathbb{R}^d)$ is compact.

2000 AMS subject classification: 42C40, 46E35, 41A46

Keywords: time-frequency analysis, radial functions, radial Gabor frames, modulation spaces, Feichtinger algebra, linear approximation, nonlinear approximation, compact embedding, entropy numbers

1 Introduction

Nowadays time-frequency analysis is a well-developed field with many applications in signal analysis and wireless communication [14]. The modulation spaces introduced by Feichtinger in the early 80's [7] play a fundamental role in this mathematical area [14].

The basic operators in time-frequency analysis on \mathbb{R}^d are the translation $T_x f(y) = f(y-x)$ and the modulation $M_\omega f(y) = e^{2\pi i y \cdot \omega} f(y)$. For a fixed window function $g \in L^2(\mathbb{R}^d)$ the short time Fourier transform (STFT) is given by $\text{STFT}_g f(x, \omega) = \langle f, M_\omega T_x g \rangle$. It is well-known that for a suitable choice of constants $a, b > 0$ and of $g \in L^2(\mathbb{R}^d)$ the family $\{M_{bj} T_{ak} g : j, k \in \mathbb{Z}^d\}$ forms a frame - a so called Gabor frame. In other words, the STFT admits discretizations. In particular, we can write an arbitrary L^2 -function as $f = \sum_{j, k \in \mathbb{Z}^d} \lambda_{j, k} M_{bj} T_{ak} g$. Furthermore, expansions of this type extend to modulation spaces [14].

In [19, 20, 21] we considered radial functions, or more generally functions which are invariant under the action of some subgroup of $O(d)$, and raised the question whether it is

possible to develop an adapted time-frequency analysis for such functions. Of course, one can apply all results that are valid for functions on \mathbb{R}^d . However, it seems natural that the additional information that the function under consideration is radial should allow one to gain some advantages. When starting investigations in this direction one immediately observes that the operators T_x and M_ω do not preserve radiality except for trivial cases. So the natural requirement is to replace these operators by some that preserve radiality.

In [19] we found a natural candidate for such operators in the more general context of square-integrable group representations. In particular, we developed the Hilbert space theory for the abstract continuous transform and applied it to the special case of the short time Fourier transform (STFT) of radial functions. In [20] we treated the corresponding discrete theory which lead to the construction of discrete (Banach) frames for spaces of invariant vectors (functions). To do this we generalized the well-established coorbit space theory originally developed by Feichtinger and Gröchenig [9, 10, 11, 13]. This discretization method does not only work on the Hilbert space level (i.e., for L^2) but also for more general Banach spaces, so-called coorbit spaces of invariant vectors, see Section 2.1. In the special case of time-frequency analysis they are precisely the modulation spaces. We remark that coorbit space theory was further generalized to the setting of an abstract continuous frame in [12].

In [20] we did not give examples. So in this paper we show in detail how to apply the general results in order to obtain radial Gabor frames and atomic decompositions for radial modulation spaces explicitly. In order to study the quality of approximation with radial Gabor frames we compare it to the one with usual Gabor frames. In particular, we investigate linear and non-linear approximation. We first derive results in the general context of coorbit spaces and apply them to approximation in radial modulation spaces later on. To some extent it turns out that the radial Gabor frames really perform better than the usual Gabor frames when approximating a radial function. This provides a further justification of the original motivation to consider an adapted analysis for radial functions.

Moreover, we present how embeddings of coorbit spaces of invariant elements can be studied by means of embeddings of certain sequence spaces. This result generalizes the work of Feichtinger and Gröchenig in [9]. The general embedding theorem indicates that restricting coorbit spaces to invariant vectors (functions) may enforce certain embeddings of coorbit spaces to become compact. In the case of the Besov and Triebel-Lizorkin spaces this phenomenon was observed earlier by Skrzypczak et al. in [17, 22, 23, 24].

We apply these results to derive new embedding results for subspaces of modulation spaces consisting of radial functions. Moreover, we determine the entropy numbers of certain embeddings of radial modulation spaces. As special case it turns out that the Feichtinger algebra $(S_0)_{rad}(\mathbb{R}^d) = M_{rad}^1(\mathbb{R}^d)$ restricted to radial functions is embedded into the Sobolev space $H_{rad}^{(d-1)/2}(\mathbb{R}^d)$. This is rather surprising because S_0 (not restricted to radial functions) is not a subspace of $H^{(d-1)/2}$. Moreover, for $d \geq 2$ the embedding $(S_0)_{rad}(\mathbb{R}^d) \hookrightarrow H_{rad}^s(\mathbb{R}^d)$ is compact if and only if $s < (d-1)/2$. In particular, $(S_0)_{rad}(\mathbb{R}^d)$ is compactly embedded into $L_{rad}^2(\mathbb{R}^d)$ for $d \geq 2$.

The paper is organized as follows. In Section 2 we introduce the necessary notation and background information. Moreover, we prove the general results about embeddings of abstract coorbit spaces. Also we provide statements about linear and non-linear approximation. In Section 3 we apply the general theory to time-frequency analysis of radial functions in \mathbb{R}^d . We explicitly construct the radial Gabor frames and state results about linear and non-linear approximation with these. Moreover, we prove the mentioned embedding theorems for subspaces of modulation spaces consisting of radial functions (distributions).

2 Coorbit spaces of invariant elements

2.1 Notation and Preliminaries

As announced we will derive our results first in the abstract setting of coorbit spaces and then apply it to modulation spaces of radial functions. To this end we need to introduce some notation from [10, 19, 20, 21].

Let \mathcal{G} be a locally compact group and \mathcal{A} be a compact automorphism group of \mathcal{G} , such that \mathcal{A} acts continuously on \mathcal{G} , i.e., the mapping $\mathcal{G} \times \mathcal{A} \rightarrow \mathcal{G}, (x, A) \mapsto Ax$ is continuous. We denote the left Haar measures on \mathcal{G} and \mathcal{A} by μ and ν , where ν is assumed to be normalized. However, we usually write dx and dA in integrals. The modular function on \mathcal{G} is denoted by Δ and the left and right translation operators on \mathcal{G} by $L_y F(x) = F(y^{-1}x)$ and $R_y F(x) = F(xy)$. Furthermore, we define two involutions by $F^\vee(x) = F(x^{-1})$ and $F^\nabla(x) = \overline{F(x^{-1})}$. The action of \mathcal{A} on functions on \mathcal{G} is denoted by $F_A(x) = F(A^{-1}x)$, $A \in \mathcal{A}$, and the action on measures $\tau \in M(\mathcal{G})$, the space of complex bounded Radon measures on \mathcal{G} (the dual space of $C_0(\mathcal{G})$), by $\tau_A(F) = \tau(F_{A^{-1}})$, $A \in \mathcal{A}$, $\tau \in M(\mathcal{G})$, $F \in C_0(\mathcal{G})$.

The functions (measures) which satisfy $F_A = F$ for all $A \in \mathcal{A}$ are called invariant (under \mathcal{A}). A standard argument shows that the Haar-measure μ and the modular function Δ are invariant under any compact automorphism group [21]. For a function (measure) space Y on \mathcal{G} we denote its subspace of invariant elements by $Y_{\mathcal{A}} := \{F \in Y, F_A = F \text{ for all } A \in \mathcal{A}\}$. An invariant function on \mathcal{G} can be interpreted as a function on $\mathcal{K} := \mathcal{A}(\mathcal{G})$ the space of all orbits of the form $\mathcal{A}x := \{Ax, A \in \mathcal{A}\}$, $x \in \mathcal{G}$. The orbit space \mathcal{K} becomes a hypergroup by inheriting the topology and the convolution structure of \mathcal{G} in a natural way [16, 19, 21]. For some positive measurable weight function m on \mathcal{G} we define the weighted space $L_m^p := \{F \text{ measurable}, Fm \in L^p\}$ with norm $\|F|L_m^p\| := \|Fm|L^p\|$ where the L^p -spaces on \mathcal{G} are defined as usual.

In this paper we will work with Banach spaces of functions on \mathcal{G} which will usually be denoted by Y . Similarly as in [13] we will make the following assumptions on Y .

1. Y is continuously embedded into $L_{loc}^1(\mathcal{G})$, the locally integrable functions on \mathcal{G} .
2. Y is solid, i.e., if $F \in L_{loc}^1(\mathcal{G})$, $G \in Y$ and $|F(x)| \leq |G(x)|$ a.e. then $F \in Y$ and $\|F|Y\| \leq \|G|Y\|$.
3. Y is invariant under left and right translations. We may hence define the two functions $u(x) := \|L_x|Y \rightarrow Y\|$ and $v(x) := \|R_{x^{-1}}|Y \rightarrow Y\|\Delta(x^{-1})$. Clearly, $u(xy) \leq u(x)u(y)$ and $v(xy) \leq v(x)v(y)$, i.e., u and v are submultiplicative. Additionally, we require that u and v are continuous. Under these assumptions it holds, see [10, 21]

$$Y * L_v^1 \subset Y, \quad \|F * G|Y\| \leq \|F|Y\| \|G|L_v^1\| \quad \text{for all } F \in Y, G \in L_v^1. \quad (2.1)$$

4. \mathcal{A} acts continuously on Y . Without loss of generality we may then even assume that \mathcal{A} acts isometrically on Y implying $u(Ax) = u(x)$ and $v(Ax) = v(x)$ for all $A \in \mathcal{A}$. (In case this is not true define an invariant norm on Y by $\|F|Y\|' := \int_{\mathcal{A}} \|F_A|Y\| dA$. Since \mathcal{A} acts continuously on Y this is an equivalent norm on Y .) Then $Y_{\mathcal{A}}$ is a closed non-trivial subspace of Y .

Given a submultiplicative weight w , another continuous weight function m is called w -moderate if $m(xyz) \leq w(x)m(y)w(z)$ for all $x, y, z \in \mathcal{G}$. Spaces L_m^p with invariant moderate weight function m are examples of spaces Y with the properties above.

We will always associate a weight function w to Y which is defined by

$$w(x) := \max\{u(x), u(x^{-1}), v(x), v(x^{-1})\Delta(x^{-1})\}. \quad (2.2)$$

Then as a consequence w is continuous, $w(xy) \leq w(x)w(y)$, $w(x) \geq 1$ and $w(Ax) = w(x)$ for all $A \in \mathcal{A}$ and $x \in \mathcal{G}$. Furthermore, by (2.1) it holds

$$Y * L_w^1 \subset Y, \quad \|F * G|Y\| \leq \|F|Y\| \|G|L_w^1\|. \quad (2.3)$$

We further assume that we have given a unitary, irreducible (strongly continuous) representation π of \mathcal{G} on some Hilbert space \mathcal{H} and some unitary (strongly continuous) representation σ of \mathcal{A} on the same Hilbert space \mathcal{H} (not necessarily irreducible) such that the following basic relation is satisfied (see also [19, 21]),

$$\pi(A(x))\sigma(A) = \sigma(A)\pi(x) \quad \text{for all } x \in \mathcal{G}, A \in \mathcal{A}. \quad (2.4)$$

In other words, we require that the representations $\pi_A := \pi \circ A$ are all unitarily equivalent to π and that the intertwining operators $\sigma(A)$ form a representation of \mathcal{A} .

For $f \in \mathcal{H}$ we denote $f_A = \sigma(A)f$ and $\mathcal{H}_{\mathcal{A}} := \{f \in \mathcal{H}, f_A = f \text{ for all } A \in \mathcal{A}\}$, the closed(!) subspace of invariant elements. We always assume that $\mathcal{H}_{\mathcal{A}}$ is not trivial. The wavelet transform or voice transform is defined by

$$V_g f(x) := \langle f, \pi(x)g \rangle.$$

It maps \mathcal{H} into $C^b(\mathcal{G})$, the space of bounded continuous functions on \mathcal{G} . With an element $g \in \mathcal{H}_{\mathcal{A}}$ we denote by \tilde{V}_g the restriction of V_g to $\mathcal{H}_{\mathcal{A}}$. We recall some facts from [19, 21].

- For $f, g \in \mathcal{H}_{\mathcal{A}}$ the function $\tilde{V}_g f$ is invariant under \mathcal{A} , i.e., \tilde{V}_g maps $\mathcal{H}_{\mathcal{A}}$ into $C_{\mathcal{A}}^b(\mathcal{G})$.
- For $x \in \mathcal{G}$ we define

$$\tilde{\pi}(x) := \int_{\mathcal{A}} \pi(Ax) dA \quad (2.5)$$

in a weak sense. This operator maps $\mathcal{H}_{\mathcal{A}}$ onto $\mathcal{H}_{\mathcal{A}}$ and depends only on the orbit of x under \mathcal{A} , i.e., $\tilde{\pi}(Bx) = \tilde{\pi}(x)$ for all $B \in \mathcal{A}$. Furthermore, it holds

$$\tilde{V}_g f(x) = \langle f, \tilde{\pi}(x)g \rangle_{\mathcal{H}_{\mathcal{A}}}. \quad (2.6)$$

- The operators $\tilde{\pi}(x)$ form an irreducible representation of the orbit hypergroup \mathcal{K} .

We further require that π is integrable which means that there exists a nonzero element $g \in \mathcal{H}$ such that $\int_{\mathcal{G}} |V_g g(x)| dx < \infty$. This implies that π is square-integrable, i.e., there exists $g \in \mathcal{H}$ such that $\int_{\mathcal{G}} |V_g f(x)|^2 dx < \infty$ for all $f \in \mathcal{H}$. Such a g (corresponding to the square-integrability condition) is called admissible. We list some further properties from [4] and [19] that hold under the square-integrability condition.

- There exists a positive, densely defined operator K such that the domain $\mathcal{D}(K)$ of K consists of all admissible vectors and the orthogonality relation

$$\int_{\mathcal{G}} V_{g_1} f_1(x) \overline{V_{g_2} f_2(x)} dx = \langle K g_2, K g_1 \rangle \langle f_1, f_2 \rangle$$

holds for all $f_1, f_2 \in \mathcal{H}, g_1, g_2 \in \mathcal{D}(K)$.

- The operator K commutes with the action of \mathcal{A} , i.e., $\sigma(A)K = K\sigma(A)$ for all $A \in \mathcal{A}$. Furthermore, $\mathcal{D}_{\mathcal{A}}(K) := \mathcal{D}(K) \cap \mathcal{H}_{\mathcal{A}}$ is dense in $\mathcal{H}_{\mathcal{A}}$ and K maps $\mathcal{D}_{\mathcal{A}}(K)$ into $\mathcal{H}_{\mathcal{A}}$.
- For $g \in \mathcal{D}_{\mathcal{A}}(K)$ with $\|Kg\| = 1$ we have the following inversion formula on $\mathcal{H}_{\mathcal{A}}$

$$f = \int_{\mathcal{K}} \tilde{V}_g f(y) \tilde{\pi}(y) g d\tilde{\mu}(y), \quad f \in \mathcal{H}_{\mathcal{A}} \quad (2.7)$$

where $\tilde{\mu}$ denotes the projection of the Haar measure μ onto the orbit space \mathcal{K} . The integral is understood in a weak sense.

In order to define the coorbit spaces we need to introduce the space of analyzing vectors. For some submultiplicative weight function w satisfying $w \geq 1$ it is defined by

$$\mathbb{A}_w := \{g \in \mathcal{H}, V_g g \in L_w^1(\mathcal{G})\}$$

and its subspace of invariant elements by $\mathbb{A}_w^{\mathcal{A}} := \mathbb{A}_w \cap \mathcal{H}_{\mathcal{A}}$. We only consider those weights for which $\mathbb{A}_w^{\mathcal{A}} \neq \{0\}$. It is clear that $\mathbb{A}_w^{\mathcal{A}} \subset \mathcal{D}(K)$. Now for some fixed non-zero vector $g \in \mathbb{A}_w^{\mathcal{A}}$ we define

$$\mathcal{H}_w^1 := \{f \in \mathcal{H}, V_g f \in L_w^1\}$$

with norm

$$\|f|_{\mathcal{H}_w^1}\| := \|V_g f|_{L_w^1}\|.$$

Its subspace of invariant elements is denoted by

$$(\mathcal{H}_w^1)_{\mathcal{A}} := \mathcal{H}_{\mathcal{A}} \cap \mathcal{H}_w^1 = \{f \in \mathcal{H}_{\mathcal{A}}, \tilde{V}_g f \in L_w^1\}.$$

It is shown in [20, 21] that $\tilde{\pi}(x)g \in \mathbb{A}_w^{\mathcal{A}}$ for all $g \in \mathbb{A}_w^{\mathcal{A}}, x \in \mathcal{G}$. As a consequence, both $\mathbb{A}_w^{\mathcal{A}}$ and $(\mathcal{H}_w^1)_{\mathcal{A}}$ are dense in $\mathcal{H}_{\mathcal{A}}$. As a reservoir for the general coorbit spaces we take the space $(\mathcal{H}_w^1)^{\top}$ of all bounded conjugate linear functionals on \mathcal{H}_w^1 (the anti-dual of \mathcal{H}_w^1) and its subspace $(\mathcal{H}_w^1)_{\mathcal{A}}^{\top}$ of invariant elements (the anti-dual of $(\mathcal{H}_w^1)_{\mathcal{A}}$), respectively. We may extend the voice transform onto $(\mathcal{H}_w^1)^{\top}$ by

$$V_g f(x) = f(\tilde{\pi}(x)g) = \langle f, \tilde{\pi}(x)g \rangle, \quad f \in (\mathcal{H}_w^1)^{\top}, g \in \mathbb{A}_w.$$

Now let Y be some function space on \mathcal{G} that satisfies our hypothesis and let w be the weight function defined by (2.2). Then for some fixed non-zero vector $g \in \mathbb{A}_w^{\mathcal{A}}$ the coorbit space associated to Y is defined by

$$\text{Co}Y := \{f \in (\mathcal{H}_w^1)^{\top}, V_g f \in Y\}$$

with natural norm

$$\|f|\text{Co}Y\| := \|V_g f|_Y\|.$$

Its subspace of invariant elements is denoted by

$$\text{Co}Y_{\mathcal{A}} := \text{Co}Y \cap (\mathcal{H}_w^1)_{\mathcal{A}}^{\top} = \{f \in (\mathcal{H}_w^1)_{\mathcal{A}}^{\top}, \tilde{V}_g f \in Y_{\mathcal{A}}\}.$$

It was shown in [10, 20, 21] that the coorbit spaces are Banach spaces whose definition does not depend on the particular choice of $g \in \mathbb{A}_w^{\mathcal{A}}$.

We will write $A \asymp B$ if there exist constants $C_1, C_2 > 0$ such that $C_1 A \leq B \leq C_2 B$ independently of other expressions on which A, B might depend.

2.2 Atomic decompositions and Banach frames in coorbit spaces of invariant elements

In [20, 21] atomic decompositions and Banach frames for coorbit spaces of invariant elements have been derived. In order to state these results we need to recall a definition from [20, 21].

Definition 2.1. Let $X = (x_i)_{i \in I} \subset \mathcal{G}$ be some family of points in \mathcal{G} indexed by some discrete index set I and let $V = V^{-1} = \mathcal{A}(V)$ be some relatively compact neighborhood of $e \in \mathcal{G}$.

- (a) X is called V - \mathcal{A} -dense if $\mathcal{G} = \bigcup_{i \in I} \mathcal{A}(x_i V)$.
- (b) X is called V - \mathcal{A} -separated if $\mathcal{A}(x_i V) \cap \mathcal{A}(x_j V) = \emptyset$ for all $i \neq j$.
- (c) X is called relatively separated with respect to \mathcal{A} if for all compact sets $W \subset \mathcal{G}$ there exists some constant C_W such that

$$\sup_{j \in I} \#\{i \in I, \mathcal{A}(x_i W) \cap \mathcal{A}(x_j W) \neq \emptyset\} \leq C_W < \infty.$$

- (d) X is called well-spread with respect to \mathcal{A} if it is both V - \mathcal{A} -dense (for some V) and relatively separated.

The existence of well-spread sets was shown in [21, 19]. Moreover, we have the following lemma relating separated and relatively separated families.

Lemma 2.1. *Let $X = (x_i)_{i \in I} \subset \mathcal{G}$. The following properties are equivalent.*

- $X = (x_i)_{i \in I}$ is relatively separated with respect to \mathcal{A} .
- For any compact set $K = K^{-1} = \mathcal{A}(K)$ there exists a finite partition of the index set $I = \bigcup_{r=1}^s I_r$ such that each family $(x_i)_{i \in I_r}$ is K - \mathcal{A} -separated.

Proof: This follows immediately from Lemma 2.9 in [8]. □

Let us now define the sequence spaces that will characterize the coorbit spaces $\text{Co}Y_{\mathcal{A}}$. For some well-spread set $X = (x_i)_{i \in I}$ let

$$Y_{\mathcal{A}}^{\flat} := Y_{\mathcal{A}}^{\flat}(X) := \{(\lambda_i)_{i \in I}, \sum_{i \in I} |\lambda_i| \chi_{\mathcal{A}(x_i U)} \in Y\},$$

$$Y_{\mathcal{A}}^{\sharp} := Y_{\mathcal{A}}^{\sharp}(X) := \{(\lambda_i)_{i \in I}, \sum_{i \in I} |\lambda_i| |\mathcal{A}(x_i U)|^{-1} \chi_{\mathcal{A}(x_i U)} \in Y\}$$

with natural norms

$$\|(\lambda_i)_{i \in I}\|_{Y_{\mathcal{A}}^{\flat}} := \left\| \sum_{i \in I} |\lambda_i| \chi_{\mathcal{A}(x_i U)} \right\|_Y,$$

$$\|(\lambda_i)_{i \in I}\|_{Y_{\mathcal{A}}^{\sharp}} := \left\| \sum_{i \in I} |\lambda_i| |\mathcal{A}(x_i U)|^{-1} \chi_{\mathcal{A}(x_i U)} \right\|_Y.$$

Hereby, $|\mathcal{A}(x_i U)|$ denotes the Haar measure and $\chi_{\mathcal{A}(x_i U)}$ the characteristic function of the set $\mathcal{A}(x_i U)$. If $Y = L_m^p(\mathcal{G})$, $1 \leq p \leq \infty$, with invariant moderate weight function m then $Y_{\mathcal{A}}^{\flat}(X) = \ell_{m_p}^p(I)$ and $Y_{\mathcal{A}}^{\sharp}(X) = \ell_{\nu_p}^p(I)$ (with equivalent norms) where

$$m_p(i) := m(x_i) |\mathcal{A}(x_i U)|^{\frac{1}{p}}, \quad \nu_p(i) := m(x_i) |\mathcal{A}(x_i U)|^{\frac{1}{p}-1}.$$

As another ingredient we need Wiener amalgam spaces as introduced by Feichtinger [5, 6]. For the definition we take a two-sided translation invariant solid BF-space Y and another two-sided invariant Banach space B of functions or measures on \mathcal{G} . Using a non-zero window function $k \in C_c(\mathcal{G})$ we define the control function

$$K(F, k, B)(x) := \|(L_x k)F|B\|, \quad x \in \mathcal{G} \quad (2.8)$$

where F is locally contained in B , in symbols $F \in B_{loc}$. The Wiener amalgam space $W(B, Y)$ is then defined by

$$W(B, Y) := \{F \in B_{loc}, K(F, k, B) \in Y\}$$

with norm

$$\|F|W(B, Y)\| := \|K(F, k, B)|Y\|.$$

It was shown in [6] that these spaces are two-sided invariant Banach spaces which do not depend on the particular choice of the window function k . Moreover, different functions k define equivalent norms. We will mainly need the spaces $W(L^\infty, Y)$ and $W(C_0, Y)$. Replacing the left translation L_x with the right translation R_x in the definition (2.8) of the control function leads to right Wiener amalgam spaces $W^R(B, Y)$. The subspace of functions which are invariant under \mathcal{A} is denoted by $W_{\mathcal{A}}(B, Y)$ (or $W_{\mathcal{A}}^R(B, Y)$ for right amalgams). If \mathcal{A} acts isometrically on B and Y then $W_{\mathcal{A}}$ is a closed subspace of $W_{\mathcal{A}}(B, Y)$. In order to derive Banach frames we need the 'better space' of analyzing vectors

$$\mathbb{B}_w := \{g \in \mathbb{A}_w, V_g g \in W^R(C_0, L_w^1)\}, \quad \mathbb{B}_w^{\mathcal{A}} := \{g \in \mathbb{A}_w^{\mathcal{A}}, \tilde{V}_g g \in W_{\mathcal{A}}^R(C_0, L_w^1)\}.$$

It was shown in [9, 21] that \mathbb{B}_w is dense in \mathcal{H} and that $\mathbb{B}_w^{\mathcal{A}}$ is dense in $\mathcal{H}_{\mathcal{A}}$.

Now we are ready to formulate the result concerning Banach frames and atomic decompositions of $\text{Co}Y_{\mathcal{A}}$, see [20, Theorem 7.3] or [21, Theorem 4.6.3].

Theorem 2.2. *Let $g \in \mathbb{B}_w^{\mathcal{A}} \setminus \{0\}$. Then there exists a relatively compact neighborhood $U = U^{-1} = \mathcal{A}U$ of $e \in \mathcal{G}$ such that for any family $X = (x_i)_{i \in I}$, which is U -dense and well-spread with respect to \mathcal{A} , the family $\{\tilde{\pi}(x_i)g\}_{i \in I}$ is an atomic decomposition for $\text{Co}Y_{\mathcal{A}}$. This means that*

- there exist elements $\{e_i\}_{i \in I}$ in $(\mathcal{H}_w^1)_{\mathcal{A}}$ such that $(\langle f, e_i \rangle)_{i \in I} \in Y_{\mathcal{A}}^{\natural}(X)$ for all $f \in \text{Co}Y_{\mathcal{A}}$ and

$$\|(\langle f, e_i \rangle)_{i \in I}|Y_{\mathcal{A}}^{\natural}\| \asymp \|f|_{\text{Co}Y_{\mathcal{A}}}\|;$$

- it holds $f = \sum_{i \in I} \langle f, e_i \rangle \tilde{\pi}(x_i)g$ for all $f \in \text{Co}Y_{\mathcal{A}}$ with norm convergence if the finite sequences are dense in $Y_{\mathcal{A}}^{\natural}$ and with w -* convergence in $(\mathcal{H}_w^1)_{\mathcal{A}}$ in general.

Moreover, $\{\tilde{\pi}(x_i)g\}_{i \in I}$ is a Banach frame for $\text{Co}Y_{\mathcal{A}}$, i.e.,

- $(\langle f, \tilde{\pi}(x_i)g \rangle)_{i \in I} \in Y_{\mathcal{A}}^{\flat}$ for all $f \in \text{Co}Y_{\mathcal{A}}$ and

$$\|(\langle f, \tilde{\pi}(x_i)g \rangle)_{i \in I}|Y_{\mathcal{A}}^{\flat}\| \asymp \|f|_{\text{Co}Y_{\mathcal{A}}}\|;$$

- there exists a bounded operator $\Omega : Y_{\mathcal{A}}^{\flat} \rightarrow \text{Co}Y_{\mathcal{A}}$ such that $\Omega((\langle f, \tilde{\pi}(x_i)g \rangle)_{i \in I}) = f$ for all $f \in \text{Co}Y_{\mathcal{A}}$.

This is a very general discretization theorem. Moreover, it characterizes invariant coorbit spaces by means of the sequence spaces $Y_{\mathcal{A}}^{\natural}$ and $Y_{\mathcal{A}}^{\flat}$. We remark that in [20, 21] we stated an explicit condition on the set U which uses a certain maximal function. This condition depends only on g and on w . This means that the theorem is valid 'uniformly' for all Y whose associated weight function w_Y defined in (2.2) is dominated by w .

We will apply Theorem 2.2 to the special case of radial Gabor frames in the next section. For an application to radial wavelet frames we refer to [21].

2.3 Embeddings of Coorbit Spaces

Next we study how embeddings of coorbit spaces and embeddings of sequence spaces are related. We start with an auxiliary result about interpolation of voice transforms.

Theorem 2.3. *Suppose Y and w are related as usual and let $g \in \mathbb{B}_w^{\mathcal{A}}$ with $\|Kg|\mathcal{H}\| = 1$. Then there exists some compact set K such that for any K - \mathcal{A} -separated family $X = (x_i)_{i \in I}$, there exists a linear bounded operator $S : Y_{\mathcal{A}}^{\flat}(X) \rightarrow \text{Co}Y_{\mathcal{A}}$ such that whenever $f = S(\lambda_i)_{i \in I}$ for $(\lambda_i)_{i \in I} \in Y_{\mathcal{A}}^{\flat}(X)$ then*

$$(V_g f(x_i))_{i \in I} = (\lambda_i)_{i \in I} \quad \text{and} \quad \|f|\text{Co}Y\| \leq C \|(\lambda_i)_{i \in I}|Y_{\mathcal{A}}^{\flat}(X)\|.$$

Proof: This Theorem is shown in completely the same way as Proposition 8.2 in [11]. In particular, this task consists in slightly adjusting the proof of Theorem 7.3 in [11]. We omit the details. \square

In other words the mapping $\text{Co}Y_{\mathcal{A}} \rightarrow Y_{\mathcal{A}}^{\flat}(X)$, $f \mapsto (V_g f(x_i))_{i \in I}$ is surjective if X is K - \mathcal{A} -separated for K large enough. Further, we need some statements about Wiener amalgam spaces of invariant elements.

Lemma 2.4. *Let $X = (x_i)$ be some well-spread set with respect to \mathcal{G} and $Q = Q^{-1} = \mathcal{A}(Q)$ be some neighborhood of $e \in \mathcal{G}$. Then a function F , which is invariant under \mathcal{A} , is contained in $W_{\mathcal{A}}(L^{\infty}, Y)$ if and only if $(\|\chi_{\mathcal{A}(x_i Q)} F\|_{\infty})_{i \in I} \in Y_{\mathcal{A}}^{\flat}(X)$ and there are constants $C_1, C_2 > 0$ such that*

$$C_1 \|F|W_{\mathcal{A}}(L^{\infty}, Y)\| \leq \|(\|\chi_{\mathcal{A}(x_i Q)} F\|_{\infty})_{i \in I}|Y_{\mathcal{A}}^{\flat}(X)\| \leq C_2 \|F|W_{\mathcal{A}}(L^{\infty}, Y)\|. \quad (2.9)$$

Proof: Let us assume that the characteristic function χ_Q is taken for the definition of $W_{\mathcal{A}}(L^{\infty}, Y)$.

Suppose $F \in W_{\mathcal{A}}(L^{\infty}, Y)$. We note that

$$\sum_{i \in I} \|F \chi_{\mathcal{A}(x_i Q)}\|_{\infty} \chi_{\mathcal{A}(x_i Q)}(x) = \sum_{i \in I_x} \sup_{y \in \mathcal{A}(x_i Q)} |F(y)| \chi_{\mathcal{A}(x_i Q)}(x)$$

where the sum runs over the finite index set $I_x = \{i \in I, x \in \mathcal{A}(x_i Q)\}$. Indeed, $\#I_x \leq N$ uniformly in x . This yields

$$\sum_{i \in I} \|F \chi_{\mathcal{A}(x_i Q)}\|_{\infty} \chi_{\mathcal{A}(x_i Q)}(x) \leq N \sup_{y \in \mathcal{A}(x Q^2)} |F(y)|.$$

By invariance of F under \mathcal{A} it holds

$$K(F, \chi_{Q^2}, L^{\infty})(x) = \sup_{y \in x Q^2} |F(y)| = \sup_{y \in \mathcal{A}(x Q^2)} |F(y)|.$$

Using the solidity of Y we obtain

$$\|(\|F \chi_{\mathcal{A}(x_i Q)}\|_{\infty})_{i \in I}|Y_{\mathcal{A}}^{\flat}\| \leq N \|K(F, \chi_{Q^2}, L^{\infty})|Y\| \leq C_1 \|F|W_{\mathcal{A}}(L^{\infty}, Y)\|.$$

Hereby, it is used that different window functions generate equivalent norms on $W(L^{\infty}, Y)$.

For the converse inequality, let

$$J_x := \{i \in I, \mathcal{A}(x_i Q) \cap \mathcal{A}(x Q) \neq \emptyset\}.$$

By the defining properties of a well-spread set J_x is finite and $\#J_x \leq N$ uniformly in x . Moreover, since the sets $\mathcal{A}(x_i Q)$ form a covering of \mathcal{G} it holds $\mathcal{A}(xQ) \subset \bigcup_{i \in J_x} \mathcal{A}(x_i Q)$. By invariance of F under \mathcal{A} we obtain

$$\begin{aligned} K(F, \chi_Q, L^\infty)(x) &= \sup_{y \in xQ} |F(y)| = \sup_{y \in \mathcal{A}(xQ)} |F(y)| \leq \sum_{i \in J_x} \sup_{y \in \mathcal{A}(x_i Q)} |F(y)| \\ &\leq \sum_{i \in I} \sup_{y \in \mathcal{A}(x_i Q)} |F(y)| \chi_{\mathcal{A}(x_i Q^2)}(x). \end{aligned}$$

The last inequality follows from the fact that $\mathcal{A}(xQ) \cap \mathcal{A}(x_i Q) \neq \emptyset$ implies $x \in \mathcal{A}(x_i Q^2)$. By solidity of Y this gives the lower estimate in (2.9). \square

Lemma 2.5. *Suppose Y, Z are two solid BF spaces on \mathcal{G} . Then $W_{\mathcal{A}}(L^\infty, Y) \subset W_{\mathcal{A}}(L^\infty, Z)$ if and only if $Y_{\mathcal{A}}^b(X) \subset Z_{\mathcal{A}}^b(X)$.*

Proof: First assume $Y_{\mathcal{A}}^b(X) \subset Z_{\mathcal{A}}^b(X)$ and let $F \in W_{\mathcal{A}}(L^\infty, Y)$. By Lemma 2.4 $(\|F \chi_{\mathcal{A}(x_i Q)}\|_\infty)_{i \in I} \in Y_{\mathcal{A}}^b(X) \subset Z_{\mathcal{A}}^b(X)$ and, thus, again by Lemma 2.4 F is contained in $W_{\mathcal{A}}(L^\infty, Z)$.

Conversely, suppose $W_{\mathcal{A}}(L^\infty, Y) \subset W_{\mathcal{A}}(L^\infty, Z)$ and let $(\lambda_i)_{i \in I} \in Y_{\mathcal{A}}^b(X)$. We claim that the function $F(x) = \sum_{i \in I} |\lambda_i| \chi_{\mathcal{A}(x_i Q)}(x)$ is contained in $W_{\mathcal{A}}(L^\infty, Y)$. Indeed,

$$\begin{aligned} K(F, \chi_Q, L^\infty)(x) &= \sup_{y \in xQ} \sum_{i \in I} |\lambda_i| \chi_{\mathcal{A}(x_i Q)}(y) \leq \sum_{i \in I} |\lambda_i| \sup_{y \in xQ} \chi_{\mathcal{A}(x_i Q)}(y) \\ &\leq \sum_{i \in I} |\lambda_i| \chi_{\mathcal{A}(x_i Q^2)}(x). \end{aligned}$$

The latter function is contained in Y by definition of $Y_{\mathcal{A}}^b(X)$, hence $F \in W_{\mathcal{A}}(L^\infty, Y)$. Since $W_{\mathcal{A}}(L^\infty, Z) \subset Z_{\mathcal{A}}$ we thus deduce $F \in Z_{\mathcal{A}}$ which is equivalent to $(\lambda_i)_{i \in I} \in Z_{\mathcal{A}}^b(X)$ by definition. Altogether we deduced $Y_{\mathcal{A}}^b(X) \subset Z_{\mathcal{A}}^b(X)$. \square

Corollary 2.6. *Suppose X_1 and X_2 are two well-spread sets and let Y, Z be solid BF-spaces. Then $Y_{\mathcal{A}}^b(X_1) \subset Z_{\mathcal{A}}^b(X_1)$ if and only if $Y_{\mathcal{A}}^b(X_2) \subset Z_{\mathcal{A}}^b(X_2)$. Hence, we may unambiguously write $Y_{\mathcal{A}}^b \subset Z_{\mathcal{A}}^b$.*

Proof: The assertion follows easily from Lemma 2.5 since $W_{\mathcal{A}}(L^\infty, Y)$ does not depend on the well-spread set X . \square

Now we are ready to prove the following characterization of coorbit spaces.

Theorem 2.7. *It holds $\text{Co}Y_{\mathcal{A}} \subset \text{Co}Z_{\mathcal{A}}$ if and only if $Y_{\mathcal{A}}^b \subset Z_{\mathcal{A}}^b$. In particular, two coorbit spaces coincide if and only if the corresponding sequence spaces coincide.*

Proof: Assume $Y_{\mathcal{A}}^b \subset Z_{\mathcal{A}}^b$. By Lemma 2.5 this implies $W_{\mathcal{A}}(L^\infty, Y) \subset W_{\mathcal{A}}(L^\infty, Z)$. It follows from Theorem 8.3 in [9] that $\text{Co}Y_{\mathcal{A}} = \text{Co}W_{\mathcal{A}}(L^\infty, Y)$. Since the implication $Y_{\mathcal{A}}^1 \subset Y_{\mathcal{A}}^2 \implies \text{Co}Y_{\mathcal{A}}^1 \subset \text{Co}Y_{\mathcal{A}}^2$ is trivial we conclude that $\text{Co}Y_{\mathcal{A}} \subset \text{Co}Z_{\mathcal{A}}$.

Conversely, let $\text{Co}Y_{\mathcal{A}} \subset \text{Co}Z_{\mathcal{A}}$. By Corollary 2.6 we may choose the well-spread set $X = (x_i)_{i \in I}$ according to Theorem 2.3. Now assume $Y_{\mathcal{A}}^b(X) \not\subset Z_{\mathcal{A}}^b(X)$, i.e., that there exists some $(\lambda_i)_{i \in I} \in Y_{\mathcal{A}}^b(X) \setminus Z_{\mathcal{A}}^b(X)$. We have $f = S(\lambda_i)_{i \in I} \in \text{Co}Y_{\mathcal{A}} \subset \text{Co}Z_{\mathcal{A}}$ with S as in Theorem 2.3. By [20, Theorem 5.8] (see also [21, Theorem 4.5.15]) we conclude $(\lambda_i)_{i \in I} = (V_g f(x_i))_{i \in I} \in Z_{\mathcal{A}}^b(X)$, a contradiction. Thus, $Y_{\mathcal{A}}^b \subset Z_{\mathcal{A}}^b$. \square

We remark that the embeddings in the previous theorems are automatically continuous by Theorem 2.9 below.

Now we are ready to show the double retract property of the coorbit spaces of invariant elements. Recall that some Banach space B_1 is called a retract of the Banach space B_2 if there exist bounded linear operators $S : B_1 \rightarrow B_2$ and $T : B_2 \rightarrow B_1$ such that $T \circ S = \text{Id}_{B_1}$.

Theorem 2.8. *Choose $X = (x_i)_{i \in I}$ as in Theorem 2.2, i.e., such that $\{\tilde{\pi}(x_i)g\}_{i \in I}$ is a Banach frame for $\text{Co}Y_{\mathcal{A}}$. Then $\text{Co}Y_{\mathcal{A}}$ is a retract of $Y_{\mathcal{A}}^b(X)$ and, conversely, $Y_{\mathcal{A}}^b(X)$ is a retract of a finite direct sum of copies of $\text{Co}Y_{\mathcal{A}}$.*

Proof: Denote by $A : \text{Co}Y_{\mathcal{A}} \rightarrow Y_{\mathcal{A}}^b$ the operator $f \mapsto (\langle f, \tilde{\pi}(x_i)g \rangle)_{i \in I}$. Since $\{\tilde{\pi}(x_i)g\}_{i \in I}$ is a Banach frame with bounded reconstruction operator $\Omega : Y_{\mathcal{A}}^b \rightarrow \text{Co}Y_{\mathcal{A}}$ we have $\Omega \circ A = \text{Id}_{\text{Co}Y_{\mathcal{A}}}$, i.e., $\text{Co}Y_{\mathcal{A}}$ is a retract of $Y_{\mathcal{A}}^b$.

For the converse we choose a compact set $K = K^{-1} = \mathcal{A}(K)$ according to Theorem 2.3 and split the index set I into finitely many subfamilies I_r , $r = 1, \dots, s$, such that $(x_i)_{i \in I_r}$ is K - \mathcal{A} -separated for all r , see Lemma 2.1. We denote by $P_r : Y_{\mathcal{A}}^b \rightarrow Y_{\mathcal{A}}^b$ the projection defined by

$$(P_r(\lambda_j)_{j \in I})_i = \begin{cases} \lambda_i, & \text{if } i \in I_r, \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, define $S_r : Y_{\mathcal{A}}^b \rightarrow \text{Co}Y_{\mathcal{A}}$, $r = 1, \dots, s$, to be the operator $S_r(\lambda_i)_{i \in I} = f$ such that $V_g f(x_i) = \lambda_i$ for all $i \in I_r$. The existence and boundedness of this operator follows from Theorem 2.3. Now we may define the following bounded linear operators between $Y_{\mathcal{A}}^b$ and $\bigoplus_{r=1}^s \text{Co}Y_{\mathcal{A}}$:

$$\widehat{A} : Y_{\mathcal{A}}^b \rightarrow \bigoplus_{r=1}^s \text{Co}Y_{\mathcal{A}}, \quad \Lambda \mapsto (S_1 \circ P_1 \Lambda, \dots, S_s \circ P_s \Lambda)$$

and

$$\widehat{B} : \bigoplus_{r=1}^s \text{Co}Y_{\mathcal{A}} \rightarrow Y_{\mathcal{A}}^b, \quad (f_1, \dots, f_s) \mapsto \sum_{r=1}^s P_r(V_g f(x_i)).$$

It immediately follows that $\widehat{B} \circ \widehat{A} = \text{Id}_{Y_{\mathcal{A}}^b}$. Thus, $Y_{\mathcal{A}}^b$ is a retract of $\bigoplus_{r=1}^s \text{Co}Y_{\mathcal{A}}$. \square

This theorem allows to transform many questions about the invariant coorbit spaces to questions about the corresponding sequence spaces. In order to formulate a particular result we recall that a class \mathcal{J} of operators between arbitrary Banach spaces is called an *operator ideal* if the following conditions are satisfied for the components $\mathcal{J}(E, F) := \mathcal{J} \cap \mathcal{B}(E, F)$, where $\mathcal{B}(E, F)$ denotes the space of bounded operators between Banach spaces E, F (see also [18, p.45]).

- The identity operator $\text{Id}_{\mathbb{C}}$ belongs to \mathcal{J} , where \mathbb{C} is identified with the one-dimensional Banach space.
- If $S_1, S_2 \in \mathcal{J}(E, F)$ then $S_1 + S_2 \in \mathcal{J}(E, F)$.
- If $T \in \mathcal{B}(E_0, E)$, $S \in \mathcal{J}(E, F)$ and $R \in \mathcal{B}(F, F_0)$ then $RST \in \mathcal{J}(E_0, F_0)$.

For instance, the compact operators form an operator ideal.

Theorem 2.9. (a) *The inclusion mapping $J_{\mathcal{A}} : \text{Co}Y_{\mathcal{A}} \hookrightarrow \text{Co}Z_{\mathcal{A}}$ between invariant coorbit spaces is automatically continuous. The same holds for the inclusion mapping $J_{\mathcal{A}}^b : Y_{\mathcal{A}}^b \hookrightarrow Z_{\mathcal{A}}^b$.*

(b) *$J_{\mathcal{A}}$ is compact if and only if $J_{\mathcal{A}}^b$ is compact.*

(c) Let \mathcal{J} be an operator ideal. Then $J_{\mathcal{A}} \in \mathcal{J}$ if and only if $J_{\mathcal{A}}^b \in \mathcal{J}$.

Proof: The proof is exactly the same as the one of Theorem 9.4 in [11] and hence omitted. \square

We remark that this theorem may be used to estimate the entropy numbers and approximation numbers of (compact) embedding operators of coorbit spaces. For instance, the compact operators, whose entropy numbers (approximation numbers) are contained in some $\ell^p(\mathbb{N})$ -space form an operator ideal [18]. So according to the above theorem it suffices to compute entropy or approximation numbers of embeddings between sequence spaces, which is much easier than for function spaces, see also the next section.

Let us now investigate compactness of embeddings for the important special case $Y = L_m^p$. According to Lemma 4.3.1 in [21] (see also [20]) the corresponding sequence space $Y_{\mathcal{A}}^b$ coincides with $\ell_{m_p}^p$, where $m_p(i) = m(x_i)|\mathcal{A}(x_i U)|^{1/p}$. The following lemmas give criterions on the compactness of embeddings between certain weighted ℓ^p -spaces.

Lemma 2.10. *Let $1 \leq p \leq q \leq \infty$ and v, w be some positive weight functions on some infinite index set I .*

- (a) $\ell_w^p(I)$ is continuously embedded into $\ell_v^q(I)$ if and only if $\sup_{i \in I} v(i)/w(i) < \infty$.
- (b) The embedding $\ell_w^p(I) \hookrightarrow \ell_v^q(I)$ is compact if and only if the sequence $(v(i)/w(i))_{i \in I}$ is contained in $c_0(I)$, the space of all sequences vanishing at ∞ .

Proof: The proof is straightforward and thus omitted. \square

Lemma 2.11. *Let $\infty \geq p > q \geq 1$ and v, w be some positive weight functions on I . Set $\beta := 1/q - 1/p > 0$. If $\sum_{i \in I} \left(\frac{v(i)}{w(i)}\right)^{1/\beta} < \infty$ then $\ell_w^p(I)$ is compactly embedded into $\ell_v^q(I)$.*

Proof: The assertion follows from a simple application of Hölder's inequality. \square

Let us apply these lemmas to embeddings of coorbit spaces.

Theorem 2.12. *Suppose \mathcal{G} is not compact. Let v, m be two moderate invariant weight functions, and suppose $1 \leq p \leq q \leq \infty$. Further, let $X = (x_i)_{i \in I}$ be some well-spread set with respect to \mathcal{A} and $U = U^{-1} = \mathcal{A}(U)$ be some relatively compact neighborhood of $e \in \mathcal{G}$. We define the sequence*

$$h_i := \frac{v_p(i)}{m_p(i)} = \frac{v(x_i)}{m(x_i)|\mathcal{A}(x_i U)|^{1/p-1/q}}, \quad i \in I. \quad (2.10)$$

- (a) $(\text{Co}L_m^p)_{\mathcal{A}}$ is continuously embedded into $(\text{Co}L_v^q)_{\mathcal{A}}$ if and only if $h \in \ell^\infty(I)$.
- (b) The embedding in (a) is compact if and only if $h \in c_0(I)$.

Proof: The sequence spaces associated to L_m^p and L_v^q are given by $(L_m^p)_{\mathcal{A}}^b = \ell_{m_p}^p$ and $(L_v^q)_{\mathcal{A}} = \ell_{v_q}^q$ with (see [20] or [21, Lemma 4.3.1])

$$m_p(i) = m(x_i)|\mathcal{A}(x_i U)|^{1/p} \quad \text{and} \quad v_q(i) = v(x_i)|\mathcal{A}(x_i U)|^{1/q}. \quad (2.11)$$

By Lemma 2.10 $\ell_{m_p}^p$ is embedded into $\ell_{v_q}^q$ if and only if the sequence h is contained in ℓ^∞ . Moreover, this embedding is compact if and only if $h \in c_0(I)$. Hence, by Theorem 2.7 we

have $(\text{Co}L_m^p)_\mathcal{A} \hookrightarrow (\text{Co}L_v^q)_\mathcal{A}$ if and only if $h \in \ell^\infty$. Moreover, this embedding is compact if and only if $h \in c_0(I)$. \square

It is also possible to get rid of the well-spread set X in the previous theorem as is shown by the next Lemma.

Lemma 2.13. *With the same notation as in Theorem 2.12 we define the function*

$$H : \mathcal{G} \rightarrow \mathbb{R}, \quad x \mapsto \frac{v(x)}{m(x)|\mathcal{A}(xU)|^{1/p-1/q}}. \quad (2.12)$$

Then we have $h \in \ell^\infty(I)$ if and only if $H \in L^\infty(\mathcal{G})$ and $h \in c_0(I)$ if and only if $H \in C_0(\mathcal{G})$.

Proof: Clearly, if the function H in (2.12) is contained in L^∞ , resp. $C_0(\mathcal{G})$ then $h \in \ell^\infty$ resp. $h \in c_0$. This shows the "if"-part of (a) and (b). The converse part follows from the moderateness of the functions v and m . \square

This theorem gives an abstract explanation to the phenomenon that restricting to functions that possess symmetry may enforce compactness of embeddings. This fact was observed for the Besov and Triebel-Lizorkin spaces of radial distributions recently in [17, 22, 23, 24].

Indeed, in typical situations the function $x \mapsto |\mathcal{A}(xU)|$ has some growth. However, for the trivial subgroup $\mathcal{A} = \{e\}$ this function is constant by definition of the Haar measure. So by restricting to elements in $\text{Co}Y$ which are invariant under a suitable \mathcal{A} , we might be able to enforce the function in (2.12) to belong to C_0 , although the function $v(x)/w(x)$ does not. We formulate this observation for a simple case in the next corollary.

Corollary 2.14. *Let \mathcal{G} be non-compact. Suppose \mathcal{A} is such that $x \mapsto |\mathcal{A}(xU)|^{-1}$ belongs to $C_0(\mathcal{G})$ and let $1 \leq p < q \leq \infty$. Further, let m be some moderate invariant weight function. Then $\text{Co}L_m^p$ is embedded in $\text{Co}L_m^q$, but not compactly. Restricting to \mathcal{A} -invariant elements enforces compactness, i.e., the embedding $(\text{Co}L_m^p)_\mathcal{A} \hookrightarrow (\text{Co}L_m^q)_\mathcal{A}$ is compact.*

Let us now consider the case $p > q$.

Theorem 2.15. *Let $\infty \geq p > q \geq 1$ and v, m be two moderate invariant weight functions. Set $\beta := 1/q - 1/p > 0$. If $\int_{\mathcal{G}} \left(\frac{v(x)}{m(x)}\right)^{1/\beta} dx < \infty$ then $(\text{Co}L_m^p)_\mathcal{A}$ is compactly embedded into $(\text{Co}L_v^q)_\mathcal{A}$.*

Proof: The sequence spaces associated to L_m^p and L_v^q are given by $(L_m^p)_\mathcal{A} = \ell_{m_p}^p$ and $(L_v^q)_\mathcal{A} = \ell_{v_q}^q$ with m_p and v_q as in (2.11). According to Lemma 2.11 and Theorem 2.9 the assertion follows from the following computation:

$$\begin{aligned} \sum_{i \in I} \left(\frac{v_q(i)}{m_p(i)}\right)^{1/\beta} &= \sum_{i \in I} \left(\frac{v(x_i)}{m(x_i)} |\mathcal{A}(x_i U)|^\beta\right)^{1/\beta} \asymp \sum_{i \in I} \left(\frac{v(x_i)}{m(x_i)}\right)^{1/\beta} \int_{\mathcal{A}(x_i U)} dx \\ &\asymp \sum_{i \in I} \int_{\mathcal{A}(x_i U)} \left(\frac{v(x)}{m(x)}\right)^{1/\beta} dx \asymp \int_{\mathcal{G}} \left(\frac{v(x)}{m(x)}\right)^{1/\beta} dx < \infty. \end{aligned}$$

Here the moderateness of v, m and the finite overlap property of the well-spread set $(x_i)_{i \in I}$ was used. \square

Since the condition in the previous theorem is independent of the automorphism group \mathcal{A} , restricting to invariant elements does not give stronger results for embeddings – in contrast to the case $p < q$.

2.4 Linear and Nonlinear approximation

Let us now consider linear approximations of the form $\sum_{i \in N} \lambda_i \tilde{\pi}(x_i)g$ of $f = \sum_{i \in I} \lambda_i \tilde{\pi}(x_i)g$, where $N \subset I$ is a finite set. For some subspace V of some Banach space B we introduce the error of approximation by

$$e(f, V, B) := \inf_{g \in V} \|f - g\|_B \quad \text{for } f \in B. \quad (2.13)$$

Furthermore, for some bounded sequence $\lambda = (\lambda_i)_{i \in I}$ we denote by $(s_n(\lambda))_{n \in \mathbb{N}}$ its non-increasing rearrangement, i.e.,

$$s_n(\lambda) := \inf \{ \sigma \geq 0, \#\{i : |\lambda_i| \geq \sigma\} < n \}. \quad (2.14)$$

Clearly, it holds $s_n(\lambda) \geq s_{n+1}(\lambda)$ for all $n \in \mathbb{N}$.

Theorem 2.16. *Let $1 \leq p \leq q \leq \infty$ and v, m be some moderate weight functions on \mathcal{G} such that the condition in Theorem (b) is satisfied, i.e., the embedding $(\text{Co}L_m^p)_{\mathcal{A}} \hookrightarrow (\text{Co}L_v^q)_{\mathcal{A}}$ is compact. Let $X = (x_i)_{i \in I}$ and $g \in \mathbb{B}_w^{\mathcal{A}}$ such that $\{\tilde{\pi}(x_i)g\}_{i \in I}$ is an atomic decomposition of $(\text{Co}L_m^p)_{\mathcal{A}}$ and $(\text{Co}L_v^q)_{\mathcal{A}}$ as in Theorem 2.12. We denote by $(h_i)_{i \in I}$ the sequence defined in (2.10). Let $\tau : \mathbb{N} \rightarrow I$ realize the non-increasing rearrangement of h , i.e., $s_n(h) = h_{\tau(n)}$, and let $V_n = \text{span}\{\tilde{\pi}(x_{\tau(j)})g, j = 1, \dots, n\}$. Then for all $n \in \mathbb{N}$*

$$e(f, V_n, (\text{Co}L_v^q)_{\mathcal{A}}) \leq C s_{n+1}(h) \|f\|_{(\text{Co}L_w^p)_{\mathcal{A}}} \quad \text{for all } f \in (\text{Co}L_m^p)_{\mathcal{A}}.$$

Proof: Let $\lambda_i = \langle f, e_i \rangle$ with $e_i \in (\mathcal{H}_w^1)_{\mathcal{A}}$ as in Theorem 2.2. It holds $(L_m^p)^{\natural} = \ell_{m^{(p)}}^p$ and $(L_v^q)^{\flat} = \ell_{v^{(q)}}^q$ with $m^{(p)}(i) = m(x_i)|\mathcal{A}(x_i U)|^{1/p-1}$ and $v^{(q)} = v(x_i)|\mathcal{A}(x_i U)|^{1/q-1}$ by Lemma 4.3.1 in [21]. Since $\|(\lambda_i)_{i \in I}\|_{\ell_{m^{(q)}}^q}$ is an equivalent norm on $(\text{Co}L_m^p)_{\mathcal{A}}$ by Theorem 2.2 we obtain for all $f \in (\text{Co}L_m^p)_{\mathcal{A}}$

$$\begin{aligned} \|f - \sum_{j=1}^n \lambda_{\tau(j)} \tilde{\pi}(x_{\tau(j)})g\|_{(\text{Co}L_v^q)_{\mathcal{A}}} &= \left\| \sum_{i \in I} \lambda_i \tilde{\pi}(x_i)g - \sum_{j=1}^n \lambda_{\tau(j)} \tilde{\pi}(x_{\tau(j)})g \right\|_{(\text{Co}L_v^q)_{\mathcal{A}}} \\ &= \left\| \sum_{j=n+1}^{\infty} \lambda_{\tau(j)} \tilde{\pi}(x_{\tau(j)})g \right\|_{(\text{Co}L_v^q)_{\mathcal{A}}} \leq C \left(\sum_{j=n+1}^{\infty} (|\lambda_{\tau(j)}| v^{(q)}(\tau(j)))^q \right)^{1/q} \\ &\leq C \sup_{j > n} \frac{v(x_{\tau(j)})|\mathcal{A}(x_{\tau(j)} U)|^{1/q-1}}{w(x_{\tau(j)})|\mathcal{A}(x_{\tau(j)} U)|^{1/p-1}} \left(\sum_{j=n+1}^{\infty} (|\lambda_{\tau(j)}|^p m^{(p)}(\tau(j))^p) \right)^{1/p} \\ &\leq C' h_{\tau(n+1)} \|f\|_{(\text{Co}L_m^p)_{\mathcal{A}}}. \end{aligned}$$

Hereby, we used also Lemma 2.10. This yields the claim. \square

Remark 2.1. This theorem shows that using elements $\tilde{\pi}(x)g$ instead of the elements $\pi(y)g$ for approximating $f \in \text{Co}(L_w^p)_{\mathcal{A}}$ gives an advantage. Indeed, in typical situations $|\mathcal{A}(xU)|$ is a growing function so that the sequence $h = h_{\mathcal{A}}$ in (2.10) decreases faster than $h_{\{e\}}$, the one for the trivial automorphism group $\{e\}$. This means that the error of linear approximation (measured in the $\text{Co}L_v^q$ -norm) with elements $\tilde{\pi}(x)g$ decreases faster than the one of approximation with elements $\pi(y)g$. Moreover, if $v = w$ and the function $x \mapsto |\mathcal{A}(xU)|^{-1}$ is contained in $C_0(\mathcal{G})$ then the statement in Theorem 2.16 is a significant improvement for invariant f 's since for the approximation with $\pi(y)g$ we only know that the error (measured again in $\text{Co}L_v^q$) converges to 0 (with no information about the speed of convergence and provided $q < \infty$) while we have a more concrete error estimate for the approximation with elements $\tilde{\pi}(x)g$.

Let us finally discuss non-linear approximation. Let $(x_i)_{i \in I}$ be some well-spread set with respect to \mathcal{A} and $g \in \mathbb{B}_w^{\mathcal{A}}$ such that $\{\tilde{\pi}(x_i)g\}_{i \in I}$ forms an atomic decomposition. We denote by

$$\sigma_n(f, \text{Co}Y_{\mathcal{A}}) := \inf_{N \subset I, \#N \leq n} \|f - \sum_{i \in N} \lambda_i \tilde{\pi}(x_i)g\|_{\text{Co}Y_{\mathcal{A}}}$$

the error of best n -term approximation. Hereby, the infimum is also taken over all possible choices of coefficients λ_i . Our task is to find a class of elements for which this error has a certain decay when n tends to ∞ .

The following lemma, which is taken from [15], is useful for this task.

Lemma 2.17. *Let $b = (b_k)_{k \in \mathbb{N}}$ be some non-increasing sequence of positive numbers. Set $\sigma_{n,q}(b) = (\sum_{k=n}^{\infty} b_k^q)^{1/q}$ and for $p, q > 0$ set $\alpha = 1/p - 1/q$. Then for $0 < p < q \leq \infty$ we have*

$$2^{-1/p} \|b\|_{\ell^p(\mathbb{N})} \leq \left(\sum_{n=1}^{\infty} (n^\alpha \sigma_{n,q}(b))^p \frac{1}{n} \right)^{1/p} \leq C \|b\|_{\ell^q(\mathbb{N})}.$$

Theorem 2.18. *Let m, v be some w -moderate weight functions on \mathcal{G} , let $1 \leq p < q \leq \infty$ and define $\alpha = 1/p - 1/q$. Let $(x_i)_{i \in I}$ be some well-spread set with respect to \mathcal{A} such that $\{\tilde{\pi}(x_i)g\}_{i \in I}$ is an atomic decomposition of $(\text{Co}L_m^p)_{\mathcal{A}}$. Assume further that the function H defined in (2.12) is contained in L^∞ , or equivalently $(\text{Co}L_m^p)_{\mathcal{A}} \hookrightarrow (\text{Co}L_v^q)_{\mathcal{A}}$. Then*

$$\left(\sum_{n=1}^{\infty} \frac{1}{n} (n^\alpha \sigma_n(f, (\text{Co}L_v^q)_{\mathcal{A}}))^p \right)^{1/p} \leq C \|f\|_{(\text{Co}L_m^p)_{\mathcal{A}}} \quad \text{for all } f \in (\text{Co}L_m^p)_{\mathcal{A}}. \quad (2.15)$$

Proof: Let $f = \sum_{i \in I} \lambda_i \tilde{\pi}(x_i)g$ with $\lambda_i = \langle f, e_i \rangle$ be an expansion of $f \in (\text{Co}L_m^p)_{\mathcal{A}}$ in terms of the atomic decomposition. Further let $v^{(q)}(i) = v(x_i)|\mathcal{A}(x_i U)|^{1/q-1}$ and $b(i) := |\lambda_i|v^{(q)}(i)$. Let $\tau : \mathbb{N} \rightarrow I$ be a bijection such that $b(\tau(s)) \geq b(\tau(s+1))$ for all $s \in \mathbb{N}$. We obtain

$$\begin{aligned} \sigma_n(f, (\text{Co}L_v^q)_{\mathcal{A}}) &\leq \left\| \sum_{s=n+1}^{\infty} \lambda_{\tau(s)} \tilde{\pi}(x_{\tau(s)})g \right\|_{(\text{Co}L_v^q)_{\mathcal{A}}} \leq C \left(\sum_{s=n+1}^{\infty} |\lambda_{\tau(s)} v^{(q)}(\tau(s))|^q \right)^{1/q} \\ &= C \sigma_{n,q}(b). \end{aligned} \quad (2.16)$$

By Lemma 2.17 and Theorem 2.2 we deduce

$$\begin{aligned} &\left(\sum_{n=1}^{\infty} \frac{1}{n} (n^\alpha \sigma_n(f, (\text{Co}L_v^q)_{\mathcal{A}}))^p \right)^{1/p} \leq C \left(\sum_{n=1}^{\infty} \frac{1}{n} (n^\alpha \sigma_{n,q}(b))^p \right)^{1/p} \\ &\leq C' \left(\sum_{i \in I} (|\lambda_i|v^{(q)}(i))^p \right)^{1/p} = C' \left(\sum_{i \in I} |\lambda_i|^p v(x_i)^p |\mathcal{A}(x_i U)|^{(1/q-1)p} \right)^{1/p} \\ &= C' \left(\sum_{i \in I} |\lambda_i|^p (v(x_i)|\mathcal{A}(x_i U)|^{1/q-1/p})^p |\mathcal{A}(x_i U)|^{p(1/p-1)} \right)^{1/p} \\ &\leq C'' \left(\sum_{i \in I} |\lambda_i|^p (m(x_i)|\mathcal{A}(x_i U)|^{1/p-1})^p \right)^{1/p} \leq C''' \|f\|_{(\text{Co}L_m^p)_{\mathcal{A}}}. \end{aligned}$$

□

Note that (2.15) implies

$$\sigma_n(f, (\text{Co}L_v^q)_\mathcal{A}) \leq Cn^{-\alpha} \quad \text{for all } f \in (\text{Co}L_m^p)_\mathcal{A}.$$

We remark that once again this theorem shows that using the atomic decomposition $\{\tilde{\pi}(x_i)g\}_{i \in I}$ instead of the atomic decomposition $\{\pi(y_j)g\}_{j \in J}$ for approximating elements in $(\text{Co}L_m^p)_\mathcal{A}$ is advantageous. Indeed, the class of functions f for which the error of best n -term approximation with $\{\tilde{\pi}(x_i)g\}_{i \in I}$ has some prescribed decay (i.e. for which an estimate as in (2.15) holds) is strictly larger than the one for $\{\pi(y_j)g\}_{j \in J}$ in typical situations. Indeed, if $x \mapsto |A(xU)|$ is an unbounded function then we may have $(\text{Co}L_m^p)_\mathcal{A} \subset (\text{Co}L_v^q)_\mathcal{A}$ although $\text{Co}L_m^p \not\subset \text{Co}L_v^q$.

Let us note further that Theorem 2.18 does not give a characterization of all elements f satisfying the decay condition (2.15), i.e., we cannot provide a converse inequality of (2.15) (a Bernstein inequality). This is due to the fact that we do not work with a basis but with a frame. In fact, it is a difficult (and open) problem to find classes of frames for which Bernstein inequalities hold.

3 Radial Time Frequency Analysis

3.1 Short Time Fourier Transform and Modulation Spaces

We will now apply the abstract results from the previous section to time frequency analysis of radial functions.

Let $\mathbb{H}_d := \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{T}$ denote the (reduced) Heisenberg group with group law

$$(x, \omega, \tau)(x', \omega', \tau') = (x + x', \omega + \omega', \tau\tau' e^{\pi i(x' \cdot \omega - x \cdot \omega')}).$$

The Heisenberg group is unimodular and has Haar measure

$$\int_{\mathbb{H}_d} f(h) dh = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_0^1 f(x, \omega, e^{2\pi i t}) dt d\omega dx.$$

The Schrödinger representation ρ acting on $\mathcal{H} = L^2(\mathbb{R}^d)$ is described as follows. Let

$$T_x f(t) := f(t - x), \quad \text{and} \quad M_\omega f(t) = e^{2\pi i \omega \cdot t} f(t), \quad x, \omega, t \in \mathbb{R}^d,$$

denote the translation and modulation operator on $L^2(\mathbb{R}^d)$. Then ρ is defined by

$$\rho(x, \omega, \tau) := \tau e^{\pi i x \cdot \omega} T_x M_\omega = \tau e^{-\pi i x \cdot \omega} M_\omega T_x.$$

It is well-known that this is an irreducible unitary and square-integrable representation of \mathbb{H}_d . The corresponding voice transform is essentially the short time Fourier transform:

$$\begin{aligned} V_g f(x, \omega, \tau) &= \langle f, \rho(x, \omega, \tau)g \rangle_{L^2(\mathbb{R}^d)} = \bar{\tau} \int_{\mathbb{R}^d} f(t) \overline{e^{-\pi i x \cdot \omega} M_\omega T_x g(t)} dt \\ &= \bar{\tau} e^{\pi i x \cdot \omega} \int_{\mathbb{R}^d} f(t) \overline{g(t - x)} e^{-2\pi i t \cdot \omega} dt = \bar{\tau} e^{\pi i x \cdot \omega} \text{STFT}_g f(x, \omega). \end{aligned} \quad (3.1)$$

The automorphisms of $\mathbb{R}^d \times \mathbb{R}^d$ that extend to automorphisms of \mathbb{H}_d are given by the elements of the symplectic group $Sp(d)$. The latter is defined as the subgroup of $GL(2d, \mathbb{R})$

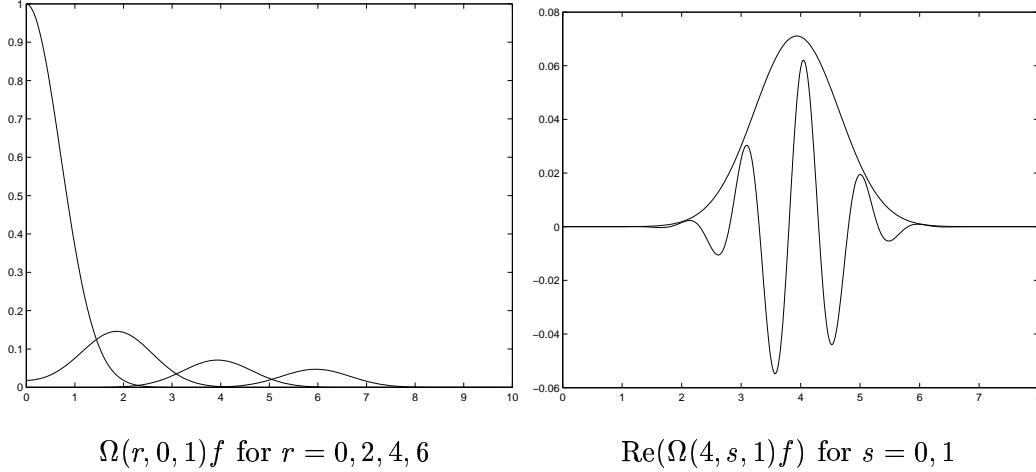


Figure 1: $\Omega(r, s, \cos \alpha)f$ for a Gaussian $f(x) = e^{-|x|^2}$ for different values of r, s, α

leaving invariant the symplectic form $[(x, \omega), (x', \omega')] := x' \cdot \omega - x \cdot \omega'$ (see [14]). A compact subgroup of $Sp(d)$ is given by

$$\mathcal{A} := \left\{ \begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix} \mid A \in SO(d) \right\} \cong SO(d). \quad (3.2)$$

An element $A \in SO(d) \cong \mathcal{A}$ acts on \mathbb{H}_d by $A(x, \omega, \tau) = (Ax, A\omega, \tau)$.

We choose the natural representation σ of $SO(d)$ on $L^2(\mathbb{R}^d)$ given by $\sigma(A)f(t) = f(A^{-1}t)$ for $A \in SO(d), t \in \mathbb{R}^d$. Using the orthogonality of $A \in SO(d)$ we obtain

$$\begin{aligned} \rho(Ax, A\omega, \tau)\sigma(A)f(t) &= \tau e^{-\pi i(Ax \cdot A\omega)} e^{2\pi i A\omega \cdot t} f(A^{-1}(t - Ax)) \\ &= \tau e^{-\pi i(x \cdot \omega)} e^{2\pi i \omega \cdot A^{-1}t} f(A^{-1}t - x) = \sigma(A)\rho(x, \omega, \tau)f(t). \end{aligned}$$

Thus, condition (2.4) is satisfied. Clearly, it holds $\mathcal{H}_{\mathcal{A}} = L^2_{rad}(\mathbb{R}^d)$, the space of radially symmetric L^2 -functions. In the sequel we assume $d \geq 2$.

The operator $\tilde{\rho}$ as defined in (2.5) reads

$$\begin{aligned} \tilde{\rho}(x, \omega, \tau)f(t) &= \int_{SO(d)} \rho(Ax, A\omega, \tau)f(t)dA = \tau e^{\pi i x \cdot \omega} \int_{SO(d)} e^{2\pi i A\omega \cdot t} f(t - Ax)dA \\ &=: \tau e^{\pi i x \cdot \omega} \Omega(x, \omega)f(t), \quad (x, \omega, \tau) \in \mathbb{H}_d. \end{aligned}$$

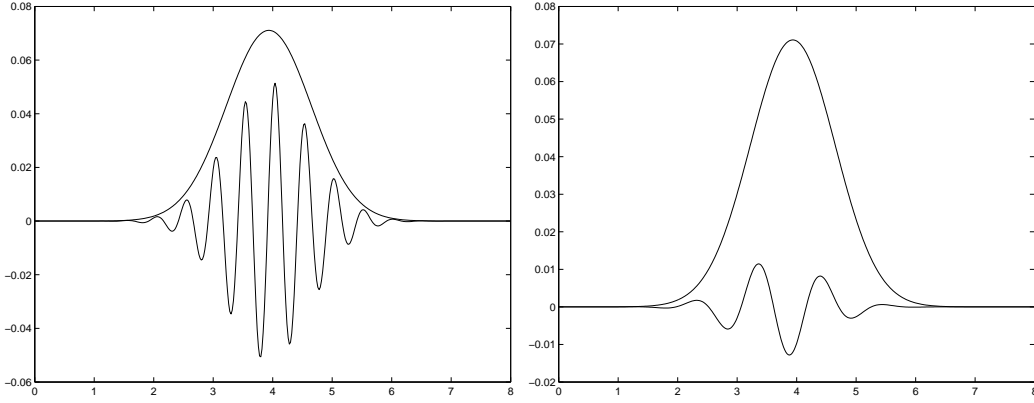
It has been shown in [19, 21] that Ω may be expressed as

$$\begin{aligned} \Omega(x, \omega)f(t) &= \Omega(r, s, \cos \alpha)f_0(\theta) \\ &= \frac{|S^{d-2}|}{|S^{d-1}|} \int_0^\pi f_0(\sqrt{\theta^2 - 2r\theta \cos \phi + r^2}) e^{2\pi i \theta s \cos \alpha \cos \phi} \mathcal{B}_{d-1}(\theta s \sin \alpha \sin \phi) \sin^{d-2} \phi d\phi \end{aligned}$$

where $r = |x|, s = |\omega|, x \cdot \omega = rs \cos \alpha, \theta = |t|$ and $f_0 : [0, \infty) \rightarrow \mathbb{C}$ is such that $f(x) = f_0(|x|)$. Here, \mathcal{B}_{d-1} denotes the spherical Bessel function defined by

$$\mathcal{B}_d(t) = \frac{1}{|S^{d-1}|} \int_{S^{d-1}} e^{2\pi i t \eta \cdot \xi} dS(\xi), \quad \eta \in S^{d-1}. \quad (3.3)$$

(independent of the choice of $\eta \in S^{d-1}$).



Re($\Omega(4, s, 1)f$) for $s = 0, 2$

Re($\Omega(4, s, \cos \alpha)f$) for
 $(s, \alpha) = (0, 0), (1, \pi/6)$

Figure 2: $\Omega(r, s, \cos \alpha)f$ for a Gaussian $f(x) = e^{-|x|^2}$ for different values of r, s, α

Note that the spherical Bessel function (3.3) can be expressed by means of the Bessel function J_α of the first kind

$$\mathcal{B}_d(t) = \Gamma(\alpha + 1)(\pi t)^{-\alpha} J_\alpha(2\pi t), \quad \alpha = \frac{d-2}{2}.$$

In particular, we have

$$\mathcal{B}_1(t) = \cos(2\pi t), \quad \mathcal{B}_2(t) = J_0(2\pi t) \quad \text{and} \quad \mathcal{B}_3(t) = \frac{\sin(2\pi t)}{2\pi t}.$$

Figures 1, 2 illustrate $\Omega(r, s, \cos \alpha)f$ for a Gaussian f and $d = 2$ for certain values of r, s, α . We have always plotted the real part of the functions (as functions of $|x|$, $x \in \mathbb{R}^2$.)

For more details about the short time Fourier transform of radial functions we refer to [19, 21].

Let us now introduce the modulation spaces on \mathbb{R}^d . We consider nonnegative continuous weight functions m on $\mathbb{R}^d \times \mathbb{R}^d$ that satisfy

$$m(x + y, \omega + \xi) \leq C(1 + |x|^2 + |\omega|^2)^{a/2} m(y, \xi), \quad (x, \omega), (y, \xi) \in \mathbb{R}^d \times \mathbb{R}^d.$$

for some constants $C > 0, a \geq 0$. This means that m is a moderate function with respect to $(x, \omega) \mapsto (1 + |x|^2 + |\omega|^2)^{a/2}$, see also [14, Chapter 11.1]. A typical choice is

$$m_s(x, \omega) = (1 + |\omega|)^s, \quad s \in \mathbb{R}. \quad (3.4)$$

Now let g be some non-zero Schwartz function on \mathbb{R}^d , for instance a Gaussian. The short time Fourier transform STFT_g extends to the space $\mathcal{S}'(\mathbb{R}^d)$ of tempered distributions in a natural way. Given $1 \leq p, q \leq \infty$ and m as above the modulation space $M_m^{p,q}$ is defined as the collection of all distributions $f \in \mathcal{S}'(\mathbb{R}^d)$ for which the norm

$$\|f\|_{M_m^{p,q}} = \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |\text{STFT}_g f(x, \omega)|^p m(x, \omega)^p dx \right)^{q/p} d\omega \right)^{1/q} \quad (3.5)$$

is finite. We will sometimes restrict to the case $p = q$ in the sequel and denote $M_m^p = M_m^{p,p}$.

Since $|V_g f(x, \omega, \tau)| = |\text{STFT}_g f(x, \omega)|$ we can easily identify the modulation spaces with coorbit spaces, i.e.

$$M_m^{p,q}(\mathbb{R}^d) = \text{Co}L_m^{p,q}(\mathbb{H}_d),$$

where m is extended to \mathbb{H}_d in a trivial way by $m(x, \omega, \tau) = m(x, \omega)$ and $L_m^{p,q}$ is a weighted mixed norm space whose definition is obvious from (3.5).

Moreover, the elements invariant under \mathcal{A} are clearly the radial distributions, hence

$$(\text{Co}L_m^{p,q})_{SO(d)} = (M_m^{p,q})_{rad}(\mathbb{R}^d) = \{f \in M_m^p, f \text{ is radial}\}.$$

3.2 Radial Gabor Frames

We are interested in atomic decompositions and Banach frames of the form $\{\Omega(x_i, \omega_i)g\}_{i \in I}$ of $M_m^{p,q}$, i.e., we want to apply Theorem 2.2 to this particular situation. To this end we need a covering of $\mathbb{R}^d \times \mathbb{R}^d$ by sets of the form $\mathcal{A}((x_i, \omega_i) + U)$ with $\mathcal{A} \cong SO(d)$ being the automorphism group in (3.2) and $U = U^{-1} = \mathcal{A}(U)$ some relatively compact neighborhood of $0 \in \mathbb{R}^d \times \mathbb{R}^d$.

Given $a, b > 0$ let $U_{a,b} := B(0, a) \times B(0, b) \subset \mathbb{R}^d \times \mathbb{R}^d$, where $B(x, r)$ denotes the closed ball in \mathbb{R}^d of radius r centered at x . Clearly, $U_{a,b}$ is invariant under $SO(d)$. For $(x, \omega) \in \mathbb{R}^d \times \mathbb{R}^d$ we consider the set

$$\begin{aligned} V_{a,b}(x, \omega) &:= SO(d)((x, \omega) + U_{a,b}) \\ &= \{(A(x + y), A(\omega + \xi)), A \in SO(d), y \in B(0, a), \xi \in B(0, b)\}. \end{aligned}$$

By construction this set is again invariant under $SO(d)$. It follows from Lemma 4.1 in [19] that $V_{a,b}(x, \omega)$ depends only on $|x|$, $|\omega|$ and $x \cdot \omega$.

In [21] a relatively separated covering of $\mathbb{R}^d \times \mathbb{R}^d$ of the form $\{V_{a,b}(x_i, \omega_i)\}_{i \in I}$ was constructed. Let

$$N(j, k) := \left\lceil \frac{\pi}{4} \left(\arctan \frac{k \left(3 + \frac{3}{2j} - \left(\frac{3}{4j}\right)^2\right)^{1/2} + j \left(3 + \frac{3}{2k} - \left(\frac{3}{4k}\right)^2\right)^{1/2}}{(jk + \frac{1}{2}(j+k) - \frac{3}{8}(\frac{j}{k} + \frac{k}{j}) + 1)} \right)^{-1} \right\rceil, \quad j, k \in \mathbb{N},$$

$$N(0, k) := N(j, 0) := 0, \quad j, k \in \mathbb{N}_0.$$

We note that asymptotically these numbers behave like

$$N(j, k) \asymp \frac{\pi}{4\sqrt{3}} \frac{jk}{j+k}. \quad (3.6)$$

We further define

$$\begin{aligned} \theta_{j,k}^\ell &:= \cos \alpha_{j,k}^\ell := \sin \frac{\pi \ell}{2N(j,k)}, \quad j, k \in \mathbb{N}, \ell = -N(j,k), \dots, N(j,k), \\ \theta_{0,k}^0 &:= \theta_{j,0}^0 = 1, \quad j, k \in \mathbb{N}_0, \end{aligned}$$

and, finally, with some fixed unit vectors η, ζ with $\eta \cdot \zeta = 0$ we let

$$(x_{j,k,\ell}, \omega_{j,k,\ell}) := (aj\eta, bk(\cos(\alpha_{j,k}^\ell)\eta + \sin(\alpha_{j,k}^\ell)\zeta)), \quad (j, k, \ell) \in I,$$

where the index set is given by

$$I := \{(j, k, \ell), j, k \in \mathbb{N}_0, \ell = -N(j, k), \dots, N(j, k)\}.$$

We further denote by $X_{a,b}$ the collection of all points $\{(x_{j,k,\ell}, \omega_{j,k,\ell}), (j, k, \ell) \in I\}$ emphasizing the dependence on the parameters a, b .

Theorem 3.1. ([21, Theorem 4.7.4]) For all $a, b > 0$ the point set $X_{a,b}$ is a $U_{a,b}$ -dense, well-spread set with respect to $SO(d)$. In particular, it holds

$$\mathbb{R}^d \times \mathbb{R}^d = \bigcup_{(j,k,\ell) \in I} V_{a,b}(x_{j,k,\ell}, \omega_{j,k,\ell}). \quad (3.7)$$

Furthermore, we need some estimation of the Lebesgue-measure of $V_{a,b}(x, \omega)$ in $\mathbb{R}^d \times \mathbb{R}^d$. The following theorem follows from Lemma 4.7.3 in [21], see also [21, p.145].

Theorem 3.2. Let

$$\begin{aligned} \mu_{j,k,\ell} &:= (j+k) \left(jk \cos \frac{\pi\ell}{2N(j,k)} \right)^{d-2}, \quad (j,k,\ell) \in I, \ell \neq \pm N(j,k), j, k \geq 1, \\ \mu_{j,k,\pm N(j,k)} &:= j^{d-1} + k^{d-1} + 1, \quad j, k \in \mathbb{N}_0. \end{aligned}$$

Then there are constants $C_1 = C_1(a, b), C_2 = C_2(a, b)$ such that

$$C_1 \mu_{j,k,\ell} \leq |V_{a,b}(x_{j,k,\ell}, \omega_{j,k,\ell})| \leq C_2 \mu_{j,k,\ell}.$$

With this explicit covering of $\mathbb{R}^d \times \mathbb{R}^d$ at hand we may apply Theorem 2.2 to obtain radial Banach frames for coorbit spaces, see also Theorem 4.7.6 in [21].

Theorem 3.3. Let w be a submultiplicative $SO(d)$ -invariant weight and suppose $g \in (M_w^{1,1})_{rad}$. Then there exist constants $a_0, b_0 > 0$ such that for all positive $a < a_0, b < b_0, 1 \leq p, q \leq \infty$ and all w -moderate $SO(d)$ -invariant weights m the following holds:

(i) The set of functions

$$\{g_{j,k,\ell}\}_{(j,k,\ell) \in I} := \left\{ \Omega \left(aj, bk, \sin \frac{\pi\ell}{2N(j,k)} \right) g, \quad (j,k,\ell) \in I \right\}$$

forms a Banach frame of $(M_m^{p,q})_{rad}$ with corresponding sequence space $(L_m^{p,q})_{SO(d)}^b$.

(ii) The functions $\{g_{j,k,\ell}\}_{(j,k,\ell)}$ form an atomic decomposition of $(M_m^{p,q})_{rad}$ with corresponding sequence space $(L_m^{p,q})_{SO(d)}^h$. In particular, any $f \in (M_m^{p,q})_{rad}$ has an expansion

$$f = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sum_{\ell=-N(j,k)}^{N(j,k)} \lambda_{j,k,\ell}(f) \Omega \left(aj, bk, \sin \frac{\pi\ell}{2N(j,k)} \right) g$$

with norm-convergence whenever $1 \leq p, q < \infty$, and weak-* convergence otherwise. Moreover, we have the norm-equivalence

$$\|(\lambda_{j,k,\ell}(f))_{(j,k,\ell) \in I}\|_{(L_m^{p,q})_{SO(d)}^h} \asymp \|f\|_{(M_m^{p,q})_{rad}}.$$

In the special case $p = q$ we have the following nice description of the sequence spaces. For some moderate weight function m let

$$m_{j,k,\ell} := m(x_{j,k,\ell}, \omega_{j,k,\ell}).$$

Then the following norms are equivalent,

$$\begin{aligned} \|(\lambda_{j,k,\ell})_{(j,k,\ell) \in I}\|_{(L_m^p)_{SO(d)}^b} &\asymp \left(\sum_{(j,k,\ell) \in I} |\lambda_{j,k,\ell}|^p m_{j,k,\ell}^p \mu_{j,k,\ell} \right)^{1/p}, \\ \|(\lambda_{j,k,\ell})_{(j,k,\ell) \in I}\|_{(L_m^p)_{SO(d)}^h} &\asymp \left(\sum_{(j,k,\ell) \in I} |\lambda_{j,k,\ell}|^p m_{j,k,\ell}^p \mu_{j,k,\ell}^{1-p} \right)^{1/p}. \end{aligned}$$

with obvious modification for $p = \infty$.

Let us finally specialize the previous theorem to $\mathcal{H}_{\mathcal{A}} = \text{Co}L_{\mathcal{A}}^2$. In order to have a (Hilbert) frame in the usual sense, we have to renormalize the frame elements, see also [21, Theorem 4.7.7].

Theorem 3.4. *Assume $g \in (S_0)_{rad} = (M_0^{1,1})_{rad}$, $g \neq 0$. Then there exist constants $a_0, b_0 > 0$ such that for all a, b satisfying $0 < a < a_0, 0 < b < b_0$ the functions*

$$\tilde{g}_{j,k,\ell} := \sqrt{\mu_{j,k,\ell}} \Omega \left(aj, bk, \sin \frac{\pi \ell}{2N(j,k)} \right) g, \quad (j, k, \ell) \in I$$

form a (Hilbert-) frame for $L_{rad}^2(\mathbb{R}^d)$.

3.3 Embeddings of radial modulation spaces

Let us now apply the abstract results about embeddings of coorbit spaces to modulation spaces of radial distributions. For simplicity we specialize to the spaces $(M_s^p)_{rad} = (M_{m_s}^p)_{rad}$, $1 \leq p \leq \infty$, $s \in \mathbb{R}$, with the weight function m_s defined in (3.4).

Theorem 3.5. *Let $1 \leq p \leq q \leq \infty$ and $s, t \in \mathbb{R}$. Set $\alpha := 1/p - 1/q > 0$.*

(a) *We have the continuous embedding $(M_s^p)_{rad}(\mathbb{R}^d) \hookrightarrow (M_t^q)_{rad}(\mathbb{R}^d)$ if and only if*

$$t - s \leq \alpha(d - 1).$$

(a) *The embedding $(M_s^p)_{rad}(\mathbb{R}^d) \hookrightarrow (M_t^q)_{rad}(\mathbb{R}^d)$ is compact if and only if*

$$p < q \quad \text{and} \quad t - s < \alpha(d - 1).$$

Proof: According to Theorem 2.12 we need to investigate the sequence

$$h(j, k, \ell) = \frac{m_t(x_{j,k,\ell})}{m_s(x_{j,k,\ell}) |V_{a,b}(x_{j,k,\ell}, \omega_{j,k,\ell})|^{1/p-1/q}} = (1 + bk)^{t-s} |V_{a,b}(x_{j,k,\ell}, \omega_{j,k,\ell})|^{-\alpha}.$$

Hence, by Theorem 3.2 we get the estimation

$$h(j, k, \ell) \leq C(1 + j)^{-(d-1)\alpha} (1 + k)^{t-s-(d-1)\alpha}, \quad (j, k, \ell) \in I$$

Thus, if $t - s - (d - 1)\alpha \leq 0$ then h is contained in ℓ^∞ , and if we have the strict inequality $t - s - (d - 1)\alpha < 0$ and $p < q$ then $h \in c_0$. This shows the "if"-part of (a) and (b). On the other hand we have

$$|V_{a,b}(x_{0,k,N(0,k)}, \omega_{0,k,N(0,k)})| \asymp (1 + k)^{d-1}.$$

Thus, if $t - s - (d - 1)\alpha > 0$ then $h \notin \ell^\infty$ and if $t - s - (d - 1)\alpha \geq 0$ or $p = q$ then $h \notin c_0$. This shows the "only if"-part. \square

We note the interesting special cases

$$\begin{aligned} (M_{-(d-1)/2}^1)_{rad}(\mathbb{R}^d) &\hookrightarrow (M_0^2)_{rad}(\mathbb{R}^d) = L_{rad}^2(\mathbb{R}^d), \\ (S_0)_{rad}(\mathbb{R}^d) &= (M_0^1)_{rad}(\mathbb{R}^d) \hookrightarrow (M_{(d-1)/2}^2)_{rad}(\mathbb{R}^d) = H_{rad}^{(d-1)/2}(\mathbb{R}^d), \end{aligned} \quad (3.8)$$

where $H^s(\mathbb{R}^d)$ denotes the Bessel potential space (Sobolev space) of index s . Moreover, $(S_0)_{rad}(\mathbb{R}^d)$ is compactly embedded into $L_{rad}^2(\mathbb{R}^d)$ for $d \geq 2$.

In particular, the previous theorem shows that for $d \geq 2$, $s \in \mathbb{R}$ and $p < q$ the embedding $(M_s^p)_{rad}(\mathbb{R}^d) \hookrightarrow (M_s^q)_{rad}(\mathbb{R}^d)$ is compact although $M_s^p(\mathbb{R}^d)$ is not compactly embedded into $M_s^q(\mathbb{R}^d)$. Moreover, if t is such that $0 < (t - s) < (d - 1)(1/p - 1/q)$ then $(M_s^p)_{rad}$ is (compactly) embedded into $(M_t^q)_{rad}$ although M_s^p is not even embedded into M_t^q .

So roughly speaking, symmetry enforces compactness of embeddings or even generates embeddings. The first phenomenon was also observed for the Besov and Triebel-Lizorkin spaces [17, 22, 23, 24] while the second phenomenon does not seem to be noticed earlier.

Also Theorem 2.15 can be applied to the modulation spaces. However, since it is not specific to radial functions we omit its application here. We only note that one cannot work with the weight function m_s (because of the integrability condition). One rather has to take the functions $v_s(x, \omega) = (1 + |x| + |\omega|)^s$ punishing also the space variable.

Let us now gain more information about the compact embedding $(M_s^p)_{rad} \hookrightarrow (M_s^q)_{rad}$, $p < q$. Indeed, let us compute the entropy numbers of the embedding operator. For an operator $S \in \mathcal{B}(E, F)$ the n -th entropy number $e_n(S)$ is defined by [18]

$$e_n(S) := \inf\{\epsilon > 0 : \exists y_1, \dots, y_{2^{n-1}} \in F \text{ such that } S(U_E) \subset \cup_{j=1}^{2^{n-1}} y_j + \epsilon U_F\}$$

where U_E, U_F denote the unit balls in E, F . Moreover, we also need Lorentz sequence spaces. For $0 < p \leq \infty$ the Lorentz space $\ell_{p, \infty}$ (sometimes also called weak ℓ^p) consists of those bounded sequences $\lambda = (\lambda_i)$ such that the quasi-norm

$$\|\lambda\|_{p, \infty} := \sup_{n \in \mathbb{N}} n^{1/p} s_n(\lambda)$$

is finite, where $s_n(\lambda)$ denotes the non-increasing rearrangement of λ defined in (2.14). It is well known that the class $\mathcal{L}_{p, \infty}^{(e)}$ of all operators S whose sequence $(e_n(S))_{n \in \mathbb{N}}$ of entropy numbers is contained in $\ell_{p, \infty}$ forms an operator ideal [18, Chapter 14.3], [2, p. 35].

The following theorem is a special case of Proposition 2 in [1].

Theorem 3.6. *Let $1 \leq p \leq q \leq \infty$ and $0 < r \leq \infty$ be given and set $1/s = 1/r + 1/p - 1/q$. Then $(w_i)_{i \in I} \in \ell_{r, \infty}$ implies that the embedding operator $\text{Id} : \ell^p \rightarrow \ell_w^q$ is contained in $\mathcal{L}_{s, \infty}^{(e)}$, i. e., $e_n(\text{Id} : \ell^p \rightarrow \ell_w^q) \leq Cn^{-1/s}$.*

In order to apply this theorem, we first need an estimate of the non-increasing rearrangement of the sequence $(|V_{a,b}(x_{j,k,\ell}, \omega_{j,k,\ell})|^{-1})_{(j,k,\ell) \in I}$.

Lemma 3.7. *Let $b_{j,k,\ell} := |V_{a,b}(x_{j,k,\ell}, \omega_{j,k,\ell})|^{-1}$, $(j, k, \ell) \in I$. Then it holds*

$$s_n(b) \leq Cn^{-\frac{d-1}{3}},$$

i. e., $b \in \ell_{3/(d-1), \infty}$.

Proof: Let us first count the number of indices in $I_n := \{(j, k, \ell) \in I, j + k \leq n\}$ for $n \in \mathbb{N}_0$. By (3.6) we get

$$\begin{aligned} \#I_n &= \sum_{j,k \geq 0, j+k \leq n} 2N(j, k) + 1 \asymp 2n - 1 + \sum_{j,k \geq 1, j+k \leq n} \frac{jk}{j+k} \\ &= 2n - 1 + \sum_{\ell=1}^n \frac{1}{\ell} \sum_{j=1}^{\ell-1} j(\ell - j) \asymp n^3. \end{aligned}$$

Moreover, it follows from Theorem 3.2 that

$$\sup_{(j,k,\ell) \in I \setminus I_n} b(j,k,\ell) \leq Cn^{-(d-1)}$$

and, hence, $s_{\#I_{n+1}}(b) \leq Cn^{-(d-1)}$. Let κ denote the inverse function of $n \mapsto \#I_n$. Since $s_n(b)$ is a non-increasing sequence we obtain

$$s_n(b) \leq C\kappa(n)^{-(d-1)} \leq C'n^{-\frac{d-1}{3}}.$$

This completes the proof. \square

Now we are ready to give an estimate of the entropy numbers of embeddings of radial modulation spaces.

Theorem 3.8. *Let $1 \leq p < q \leq \infty$, $d \geq 2$ and m be some moderate invariant weight function on $\mathbb{R}^d \times \mathbb{R}^d$. Then it holds*

$$e_n(\text{Id} : (M_m^p)_{rad}(\mathbb{R}^d) \rightarrow (M_m^q)_{rad}(\mathbb{R}^d)) \leq Cn^{-\frac{d+2}{3}(1/p-1/q)}, \quad n \in \mathbb{N}.$$

Proof: Since $\mathcal{L}_{s,\infty}^{(e)}$ is an operator ideal, it suffices to estimate the entropy numbers of the embedding J from $(L_m^p)_{SO(d)}^b = \ell_{m_p}^p$ into $(L_m^q)_{SO(d)}^b = \ell_{m_q}^q$ with $m_p(i) = m(z_i)|V_{a,b}(z_i)|^{1/p}$ by Theorem 2.9. Clearly, J can be factorized as

$$J : \ell_{m_p}^p \rightarrow \ell^p \rightarrow \ell_v^q \rightarrow \ell_{m_q}^q, \quad J = D_\sigma^{-1} J' D_\sigma \quad (3.9)$$

where D_σ is the (formal) diagonal operator $(D_\sigma x)_i = \sigma_i x_i$ with $\sigma_i = m_p(i)$. Furthermore,

$$v(i) = \overline{m_p(i)} = |V_{a,b}(z_i)|^{-(1/p-1/q)}$$

and $J' = \text{Id} : \ell^p(I) \rightarrow \ell_v^q(I)$. By Lemma 3.7 we have $v \in \ell_{r,\infty}$ with $1/r = \frac{d-1}{3}(1/p-1/q)$. Theorem 3.6 yields $J' \in \mathcal{L}_{s,\infty}^{(e)}$ with

$$1/s = 1/r + 1/p - 1/q = \left(\frac{d-1}{3} + 1 \right) (1/p - 1/q) = \frac{d+2}{3}(1/p - 1/q).$$

Since $\mathcal{L}_{s,\infty}^{(e)}$ is an operator ideal also $J \in \mathcal{L}_{s,\infty}^{(e)}$ by the factorization (3.9). This concludes the proof. \square

Remark 3.1. Another measure of compactness of operators is provided by the approximation numbers. For an operator $S \in \mathcal{L}(E, F)$ they are defined by

$$a_n(S) = \inf\{\|S - T\|, T \in \mathcal{L}(E, F), \text{rank}(T) \leq n\}.$$

It is easy to see that for weight functions v, w on I and $1 \leq p \leq q \leq \infty$ it holds

$$a_n(\text{Id} : \ell_w^p \rightarrow \ell_v^q) \leq s_{n+1} \left(\left(\frac{v(i)}{w(i)} \right)_{i \in I} \right).$$

Moreover, the class of operators S whose sequence of approximation numbers $(a_n(S))_{n \in \mathbb{N}}$ is contained in $\ell_{s,\infty}$ forms again an operator ideal. So with similar arguments as in the previous proof one can show that

$$a_n(\text{Id} : (M_m^p)_{rad}(\mathbb{R}^d) \rightarrow (M_m^q)_{rad}(\mathbb{R}^d)) \leq Cn^{-\frac{d-1}{3}(1/p-1/q)}$$

for $1 \leq p < q \leq \infty$, $d \geq 2$ and some moderate weight function m .

3.4 Linear and nonlinear approximation

Let us finally apply the abstract results on linear and nonlinear approximation to the radial Gabor-like atomic decompositions of Theorem 3.3. Let us first state the theorem for linear approximation (recall also the definition (2.13) of the error of linear approximation).

Theorem 3.9. *Let $1 \leq p < q \leq \infty$, $d \geq 2$ and m be some moderate weight function. Moreover, let $a, b > 0$ and g such that $(\Omega(aj, bk, \sin \frac{\pi \ell}{2N(j, k)})g)_{(j, k, \ell) \in I} = (g_{j, k, \ell})$ forms an atomic decomposition of $(M_m^p)_{rad}(\mathbb{R}^d)$ and of $(M_m^q)_{rad}$. Further, let $\tau : \mathbb{N} \rightarrow I$ be an ordering that realizes the non-increasing rearrangement of $(h_{j, k, \ell}) = (|V_{a, b}(x_{j, k, \ell}, \omega_{j, k, \ell})|^{-1})$, $(j, k, \ell) \in I$, i.e., $h_{\tau(n)} = s_n(h)$. Denote $V_n := \text{span}\{\Omega(x_{\tau(j)}, \omega_{\tau(j)})g, j = 1, \dots, n\}$. Then*

$$e(f, V_n, (M_m^q)_{rad}) \leq C n^{-\frac{d-1}{3}(1/p-1/q)} \|f\|_{(M_m^p)_{rad}}$$

for all $f \in (M_m^p)_{rad}(\mathbb{R}^d)$.

Proof: The claim follows immediately from Theorem 2.16 together with Lemma 3.7. \square

Of course, an important special case is $q = 2$ and $m = 1$ which corresponds to approximation in $(M^2)_{rad} = L^2_{rad}(\mathbb{R}^d)$. The theorem shows, in particular, that it is advantageous to approximate a radial function with radial Gabor frames rather than with usual Gabor frames, see also Remark 2.1.

Let us now consider non-linear approximation. We assume once more that $a, b > 0$ and g are chosen such that

$$(g_{j, k, \ell})_{(j, k, \ell) \in I} = (\Omega(aj, bk, \sin \frac{\pi \ell}{2N(j, k)})g)_{(j, k, \ell) \in I}$$

forms an atomic decomposition for any of the spaces that we consider. Further we denote

$$\sigma_n(f, (M_t^q)_{rad}) := \inf_{N \subset I, \#N \leq n} \|f - \sum_{(j, k, \ell) \in N} \lambda_{j, k, \ell} g_{j, k, \ell}\|_{(M_t^q)_{rad}}$$

the error of best n -term approximation in M_t^q , $1 \leq q \leq \infty$, $t \in \mathbb{R}$.

Theorem 3.10. *Let $1 \leq p < q \leq \infty$ and $s, t \in \mathbb{R}$ and set $\alpha := 1/p - 1/q > 0$ and assume*

$$s \geq t - (d-1)(1/p - 1/q) = t - \alpha(d-1),$$

i.e., $(M_s^p)_{rad}(\mathbb{R}^d) \hookrightarrow (M_t^q)_{rad}(\mathbb{R}^d)$ by Theorem 3.5. Then

$$\left(\sum_{n=1}^{\infty} (n^\alpha \sigma_n(f, (M_t^q)_{rad}))^p \right)^{1/p} \leq C \|f\|_{(M_s^p)_{rad}} \quad \text{for all } f \in (M_s^p)_{rad}. \quad (3.10)$$

Proof: This follows from Theorem 2.18 in connection with Theorem 3.5. \square

We remark that in the corresponding result for approximating with usual Gabor frames the condition on t, s would be $s \geq t$. However, for $d \geq 2$ we are allowed to choose $s = t - \alpha(d-1) < t$ when approximating with the radial Gabor frames. Thus, the class $(M_s^p)_{rad}$ for which (3.10) is guaranteed to hold for a certain α is larger than the one for approximating an invariant f with usual Gabor frames. Let us illustrate this for the

special case $p = 1, q = 2$ and $t = 0$ corresponding to approximation in $L_{rad}^2(\mathbb{R}^d)$. From the previous theorem it follows that

$$\sigma_n(f, L_{rad}^2(\mathbb{R}^d)) \leq C \|f\|_{(M_{-(d-1)/2}^1)_{rad}} n^{-1/2} \quad \text{for all } f \in (M_{-(d-1)/2}^1)_{rad}(\mathbb{R}^d).$$

Now denote the error of best n -term approximation with usual Gabor frames by

$$\sigma_n^*(f, M_t^q) = \inf_{N \subset \mathbb{Z}^d \times \mathbb{Z}^d, \#N \leq n} \|f - \sum_{(j,k) \in N} \lambda_{j,k} M_{bk} T_{aj} g\|_{M_t^q}$$

where the constants $a, b > 0$ and the function g are chosen according to Theorem 2.2. We conclude from Theorem 2.18 (choosing \mathcal{A} as the trivial automorphism group) that

$$\sigma_n^*(f, L^2(\mathbb{R}^d)) \leq C \|f\|_{M^1} n^{-1/2} \quad \text{for all } f \in M^1(\mathbb{R}^d) = S_0(\mathbb{R}^d),$$

in particular for all $f \in M_{rad}^1(\mathbb{R}^d)$. This is the best result we can get from Theorem 2.18 (see also [15]) and there is no reason why it should extend to the larger space $(M_{-(d-1)/2}^1)_{rad}(\mathbb{R}^d)$.

This shows to some extent that the radial Gabor frames perform better than usual Gabor frames when approximating radial functions.

4 Acknowledgements

The author would like to thank Prof. H.G. Feichtinger for pointing out the papers [17, 22, 23, 24], which motivated to consider also compactness of embeddings of radial modulation spaces. This paper was partly written during a stay at the Mathematical Institute of the University of Wrocław. He would like to thank its members for the warm hospitality. The stay was supported by the European Union's Human Potential Programme under contract HPRN-CT-2001-00273 (HARP). The author also thanks the graduate program "Applied Algorithmic Mathematics" at the Technical University of Munich (funded by the DFG) for its support.

References

- [1] B. Carl, Entropy numbers of diagonal operators with an application to eigenvalue problems, *J. Approx. Theory* **32**, 135–150, 1981.
- [2] B. Carl, I. Stephani, *Entropy, compactness and the approximation of operators*, Cambridge University Press, 1990.
- [3] O. Christensen, *An Introduction to Frames and Riesz Bases*, Birkhäuser Boston, 2003.
- [4] M. Duflo, C.C. Moore, On the regular representation of a nonunimodular locally compact group, *J. Funct. Anal.* **21**, 209–243, 1976.
- [5] H.G. Feichtinger, Banach spaces of distributions of Wiener's type and interpolation, In: *Functional Analysis and Approximation, Proc. Conf., Oberwolfach 1980*, ISNM 60, pages 153–165, 1981.
- [6] H.G. Feichtinger, Banach convolution algebras of Wiener's type. In *Functions, Series, Operators, Proc. Int. Conf., Budapest 1980, Vol. I, Colloq. Math. Soc. Janos Bolyai* 35, pages 509–524. 1983.

- [7] H.G. Feichtinger, Modulation spaces on locally compact groups, Technical report, Vienna, 1983.
- [8] H.G. Feichtinger, P. Gröbner, Banach spaces of distributions defined by decomposition methods I, *Math. Nachr.*, **123**, 97–120, 1985.
- [9] H.G. Feichtinger, K.H. Gröchenig, A unified approach to atomic decompositions via integrable group representations. In: *Proc. Conf. Function Spaces and Applications*, (M. Cwikel et al. eds.) pp. 52–73, *Lect. Notes Math.* 1302, Berlin–Heidelberg–New York: Springer, 1988.
- [10] H.G. Feichtinger, K. Gröchenig, Banach spaces related to integrable group representations and their atomic decompositions I, *J. Funct. Anal.*, **86**, 307–340, 1989.
- [11] H.G. Feichtinger, K. Gröchenig, Banach spaces related to integrable group representations and their atomic decompositions II, *Monatsh. f. Mathematik*, **108**, 129–148, 1989.
- [12] M. Fornasier, H. Rauhut, Continuous frames, function spaces, and the discretization problem, to appear in *J. Fourier Anal. Appl.*, 2005.
- [13] K. Gröchenig, Describing functions: Atomic decomposition versus frames, *Monatsh. f. Mathematik* **112**, 1–41, 1991.
- [14] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhäuser, 2001.
- [15] K. Gröchenig, S. Samarah, Nonlinear approximation with local Fourier bases, *Constr. Approx.*, **16**, 317–331, 2000.
- [16] R.I. Jewett, Spaces with an abstract convolution of measures, *Adv. in Math.* **18**, 1–101, 1975.
- [17] T. Kühn, H. Leopold, W. Sickel, L. Skrzypczak, Entropy numbers of Sobolev embeddings of radial Besov spaces, *J. Approx. Theory*, **121**, 244–268, 2003.
- [18] A. Pietsch, *Operator Ideals*, North-Holland, 1979.
- [19] H. Rauhut, Wavelet transforms associated to group representations and functions invariant under symmetry groups, to appear in *Int. J. Wavelets, Multisc. and Inf. Proc.*, 2005.
- [20] H. Rauhut, Banach frames in coorbit spaces consisting of elements which are invariant under symmetry groups, *Appl. Comput. Harm. Anal.* **18**, 94–122, 2005.
- [21] H. Rauhut, *Time-Frequency and Wavelet Analysis of Functions with Symmetry Properties*, Logos-Verlag, 2005.
- [22] W. Sickel, L. Skrzypczak, Radial subspaces of Besov and Lizorkin-Triebel classes: extended Strauss lemma and compactness of embeddings, *J. Fourier Anal. Appl.*, **6**, 639–662, 2000.
- [23] L. Skrzypczak, Function Spaces in presence of symmetries: Compactness of embeddings, regularity and decay of functions, In: D. Haroske (ed.) et al., *Function Spaces, Differential Operators and Nonlinear Analysis. The Hans Triebel Anniversary Volume*, 453–466, Birkhäuser Basel, 2003.
- [24] L. Skrzypczak, B. Tomasz, Compactness of embeddings of the Trudinger-Strichartz type for rotation invariant functions, *Houston J. Math.*, **27**, 633–647, 2001.