On near orthogonality of the Banach frames of the wave packet spaces

Dimitri Bytchenkoff $^{1,2,*}$

$^1$Österreichische Akademie der Wissenschaften, Institut für Schallforschung, Wohllebengasse 12-14, 1040 Wien, Österreich

$^2$Laboratoire d’Energétique et de Mécanique Théorique et Appliquée, Université de Lorraine, 2 avenue de la Forêt de Haye, 54505 Vandoeuvre-lès-Nancy, France

Abstract

In solving scientific, engineering or pure mathematical problems one is often faced with a need to approximate the function of a given class by the linear combination of a preferably small number of functions that are localised one way or another both in the time and frequency domain. Over the last seventy years or so a range of systems of thus localised functions have been developed to allow the decomposition and synthesis of functions of various classes. The most prominent examples of such systems are Gabor functions, wavelets, ridgelets, curvelets, shearlets and wave atoms. We recently introduced a family of quasi-Banach spaces – which we called wave packet spaces – that encompasses all those classes of functions whose elements have sparse expansions in one of the above-mentioned systems, supplied them with Banach frames and provided their atomic decompositions. Herein we prove that the Banach frames and sets of atoms of the wave packet spaces are well localised or, more specifically, that they are near orthogonal.

MSC: 42B35, 42C15, 42C40

Keywords: Decomposition spaces, Banach frames, Atomic decompositions, Pseudo-metric, Near-orthogonality

1. Introduction

In science, engineering and mathematics, functions often need to be decomposed into or synthesised from those localised one way or another both in the time and frequency domain. For this purpose a number of systems of localised real functions of two real variables have been designed and successfully used over the last seventy years or so. The most important of them are Gabor functions $^{[14]}$, wavelets $^{[6]}$, ridgelets $^{[3,11]}$, curvelets $^{[4]}$, shearlets $^{[18,17]}$ and wave atoms $^{[7]}$.

Two essential properties of each of these systems are defined by two parameters known as $\alpha$ and $\beta$. The parameter $\alpha$ defines how the length $(1 + |\xi|)^\alpha$, so to speak, of the Fourier transform of the element of the system along its radial axis of symmetry depends on the absolute value of the frequency $\xi$ at which it is localised. For instance, the length of the Gabor functions, for which $\alpha = 0$, is independent of the frequency, while that of wavelets, for which $\alpha = 1$, is proportional to it. That being so, expanding a given function in a series of Gabor functions will result in identifying higher and lower frequencies present in the function with the same resolution, while in its expansion in a series of wavelets lower frequency will be much better resolved than higher ones. The parameter $\beta$ determines the number $2^{(1-\beta)j}$ of elements of the systems whose Fourier transforms appear at a given frequency scale, defined by the index $j$, and differ from each other only in orientation. This, in its turn, defines directional resolution of different frequencies present in the function that can be achieved by expanding it into a series of a given type of localised functions. For instance, Gabor functions, for which $\beta = 0$, have equally good direction resolution at all frequency scales, while wavelets, for which $\beta = 1$, do not allow one to distinguish different directions in the frequency domain at all.

* Corresponding author. E-mail address: dimitri.bytchenkoff@univ-lorraine.fr (D. Bytchenkoff)
The spaces of functions whose elements are characterised by their sparse decompositions into localised functions that belong to one or another of the systems mentioned above are worth studying in their own right. Not long ago we introduced a family of quasi-Banach spaces – which we called ‘wave packet spaces’ – that encompasses all these spaces, studied their properties, equipped them with Banach frames – which we called wave packet systems – and provided their atomic decompositions [2]. Efficiency of the approximation of functions [12] of various classes in terms of near orthogonality of its elements and prove it for the wave packet systems, in Section 3.

2. Structure of wave packet spaces

The wave packet spaces are quasi-Banach spaces defined by the decomposition method [3] and, as such, are made of three basic building blocks, namely an almost-structured covering \( Q \) of the frequency plane, a regular partition of unity and a \( \alpha \)-\( \beta \) system, in Section 3.

First, let \( C \) be both Banach frames and sets of atoms for the wave packet spaces as long as their parameters \( \alpha \) and \( \beta \) satisfy the condition \( 0 \leq \beta \leq \alpha \leq 1 \), discuss their structure, remind the notion of good localisation of the set of functions in terms of near orthogonality of its elements and prove it for the wave packet systems, in Section 3.

Definition 1. The set \( Q = \{Q_i\}_{i \in I} \) is called an ‘almost structured covering’ of \( \mathbb{R}^2 \) if
1. the number of elements in the sets \( \{i' \in I : Q_{i'} \cap Q_i \neq \emptyset\} \) is uniformly bounded for all \( i \in I \);
2. there is a set \( \{T_i \cdot b_i\}_{i \in I} \) of invertible affine-linear maps and finite sets \( \{Q_i\}_{n=1}^N \) of non-empty open and bounded subsets \( Q'_i \) and \( P_i \) of \( \mathbb{R}^2 \) such that
   (a) \( \overline{P_n} \subset Q'_n \) for \( 1 \leq n \leq N \);
   (b) for each \( i \in I \) there is such an \( n_i \in \{1, ..., N\} \) that \( Q_i = T_i Q'_{n_i} + b_i \);
   (c) there is such a constant \( C > 0 \) that \( \|T^{-1}_iT_{i'}\| \leq C \) for all such \( i \) and \( i' \in I \) that \( Q_i \cap Q_{i'} \neq \emptyset \); and
   (d) \( \bigcup_{i \in I} (T_i P_i + b_i) = \mathbb{R}^2 \).

Definition 2. First, let \( \epsilon \in (0, 1/32) \) and
\[
Q := (-\epsilon, 1 + \epsilon) \times (-1 - \epsilon, 1 + \epsilon) \quad ;
\]
second, let \( 0 \leq \beta \leq \alpha \leq 1 \), \( N := 10 \) and
\[
I_0 := \{0\} \cup \{(jml) \in N \times N_0 \times N_0 : m \leq m_j^{\text{max}} \text{ and } l \leq l_j^{\text{max}}\} \quad ;
\]
where
\[
m_j^{\text{max}} := \lfloor 2(1-\alpha)j-1 \rfloor \quad \text{and} \quad l_j^{\text{max}} := \lfloor N \cdot 2(1-\beta)j \rfloor ;
\]
finally, let \( Q_0 := Q_{000} := B_4(0) \) and
\[
Q_i := Q_{jml} := R_{jl}(A_j Q + B_{jm}) \quad ;
\]
where
\[
A_j = \begin{pmatrix} 2^{\alpha j} & 0 \\ 0 & 2^{\beta j} \end{pmatrix}, \quad B_{jm} = \begin{pmatrix} 2^{j-1} + m \cdot 2^{\alpha j} \\ 0 \end{pmatrix} ,
\]
and
\[
R_{jl} := \begin{pmatrix} \cos \theta_{jl} & -\sin \theta_{jl} \\ \sin \theta_{jl} & \cos \theta_{jl} \end{pmatrix} \quad \text{and} \quad \theta_{jl} := 2\pi \frac{2^{(\beta-1)j}-l}{N} \quad ;
\]
for all \( i = (jml) \in I_0 \setminus \{0\} \). The set \( Q^{(\alpha,\beta)} := \{Q_i\}_{i \in I_0} \) is called ‘wave packet covering’ of \( \mathbb{R}^2 \).
The proof that $Q^{(\alpha,\beta)}$ is indeed a covering of $\mathbb{R}^2$ and that it is almost-structured can be found in Lemma 3.2 and 5.1 of [2] respectively.

**Definition 3.** Let $Q = \{Q_i\}_{i \in I}$ and $\{T_i \bullet + b_i\}_{i \in I}$ be an almost structured covering of $\mathbb{R}^2$ and the set of invertible affine-linear maps associated with it respectively. The set of functions $\Phi = \{\phi_i\}_{i \in I}$ is called 'regular partition of unity subordinate to $Q$', if

1. $\phi_i \in C_c^\infty(\mathbb{R}^2)$ with $\operatorname{supp} \phi_i \subset Q_i$ for all $i \in I$;
2. $\sum_{i \in I} \phi_i \equiv 1$ on $\mathbb{R}^2$; and
3. $\sup_{i \in I} \|\partial^\alpha \phi_i^T\|_{L^\infty} < \infty$ for all $\alpha \in \mathbb{N}_0^2$, where $\phi^T : \mathbb{R}^2 \to \mathbb{C}$, $\xi \mapsto \phi_i(T_i \xi + b_i)$.

According to Theorem 2.4 in [2], there is a regular partition of unity subordinate to the wave packet covering $Q^{\alpha,\beta} = \{Q_i\}_{i \in I}$ as the latter is almost structured.

**Definition 4.** The sequence $w = \{w_i\}_{i \in I}$ of positive numbers is called 'weight'. The weight is called 'Q-moderate' if $Q = \{Q_i\}_{i \in I}$ stands for an almost structured covering of $\mathbb{R}^2$, if there is such a positive number $C$ that $w_i \leq C \cdot w_{i'}$ for all such $i$ and $i' \in I$ that $Q_i \cap Q_{i'} \neq \emptyset$.

**Definition 5.** Let $0 \leq \beta \leq \alpha \leq 1$, $s \in \mathbb{R}$, $I_0$ be as defined by [2], and

$$w_i^s := \begin{cases} 2^{js} & \text{if } i = (jml) \in I_0 \setminus \{(000)\} \\ 1 & \text{if } i = 0 = (000) \end{cases}. \quad (7)$$

Then $w^s = (w_i^s)_{i \in I_0}$ is called 'wave packet weight'.

The proof that $w^s = (w_i^s)_{i \in I_0}$ is $Q^{(\alpha,\beta)}$-moderate can be found in Lemma 6.1 of [2].

**Definition 6.** Let $Q = \{Q_i\}_{i \in I_0}$, $\Phi = \{\phi_i\}_{i \in I_0}$, and $w = (w_i)_{i \in I_0}$ be an almost structured covering of $\mathbb{R}^2$, a regular partition of unity subordinate to $Q$ and a $Q$-moderate weight respectively and let $p, q \in (0, \infty]$. Then

$$D(Q, L^p, \ell^q_w) := \left\{ g \in Z' : \|g\|_{D(Q, L^p, \ell^q_w)} < \infty \right\}$$

where $Z'$ stands for the topological dual of $Z := \mathcal{F}(C_c^\infty(\mathbb{R}^2)) \subset \mathcal{S}(\mathbb{R}^2)$ and $\|g\|_{D(Q, L^p, \ell^q_w)}$ for the quasi-norm

$$\|g\|_{D(Q, L^p, \ell^q_w)} := \left\| \left( w_i \cdot \|\mathcal{F}^{-1}(\varphi_i \cdot \hat{g})\|_{L^p} \right)_{i \in I_0} \right\|_{\ell^q} \in [0, \infty] \quad (8)$$

is called 'decomposition space'.

As explained in [2], thus defined decomposition space is a quasi-Banach space. In this definition, which differs slightly from that introduced originally [3], the functions $\{\mathcal{F}^{-1}(\varphi_i \cdot \hat{g})\}_{i \in I_0}$ play the role of components of the function $g$ localised in the frequency domain. Therefore, for [3] to be finite and so for the function $g$ to belong to $D(Q, L^p, \ell^q_w)$, these components must be $p$-integrable and their contributions $\{\|\mathcal{F}^{-1}(\varphi_i \cdot \hat{g})\|_{L^p}\}_{i \in I_0}$ to the the norm $\| \cdot \|_{D(Q, L^p, \ell^q_w)}$ weighted with $w_i$ must be $q$-summable. With all these definitions we can now introduce the wave packet space.

**Definition 7.** Let $0 \leq \beta \leq \alpha \leq 1$, $s \in \mathbb{R}$, $p$ and $q \in (0, \infty]$, and $Q^{\alpha,\beta}$ and $w_i^s$ be the wave packet covering and weight respectively. The wave packet space with parameters $\alpha, \beta, p, q$ and $s$ is defined as the decomposition space

$$\mathcal{W}^{p,q}_s(\alpha, \beta) := D(Q^{\alpha,\beta}, L^p, \ell^q_w).$$

The parameter $s$ determines what weight $w_i^s$ will be assigned to the contribution of the $i$-th component $\mathcal{F}^{-1}(\varphi_i \cdot \hat{g})$ of $g$ to the norm [3]. Therefore varying the value of $s$ amounts to making $\mathcal{W}^{p,q}_s(\alpha, \beta)$ more or, for that matter, less regular, since $i$ determines, among other things, the absolute value of the frequency at which the $i$-th component is localised. More details on the wave packet spaces, their properties and embeddings in each other or more classical function spaces such as Besov or Sobolev spaces can be found in [2].
3. Banach frames and sets of atoms of wave packet spaces and their good localisation

To be able to decompose or synthesise the elements of $W_{s}^{p,q}(\alpha, \beta)$ we shall need Banach frames and sets of atoms of it. Here are the definitions of these two notions [13].

**Definition 8.** A set $\{\psi_{i}\}_{i \in I}$ in the dual space $X'$ of a quasi-Banach space $X$ is called *Banach frame* for $X$ if there is a well-defined bounded map, called *analysis operator*, $A : X \to x, f \mapsto \{\langle \psi_{i}, f \rangle\}_{i \in I}$ where $x := \{\langle \psi_{i}, f \rangle\}_{i \in I} : f \in X$ is a solid Banach subspace of $C^{I}$ and there is such a bounded linear map $A_{j}^{-1} : x \to X$ that $A_{j}^{-1} \circ A = I_{X}$ where $I_{X}$ stands for an identity operator on $X$.

**Definition 9.** A set $\{\phi_{i}\}_{i \in I}$ in a quasi-Banach space $X$ is called *set of atoms* in $X$ if there is a well-defined bounded map, called *synthesis operator*, $S : x \to X, \{c_{i}\}_{i \in I} \mapsto \sum_{i \in I} c_{i} \phi_{i}$ where the coefficient space $x := \{c_{i}\}_{i \in I}$ associated with $\{\phi_{i}\}_{i \in I}$ is a solid subspace of $C^{I}$ and there is such a bounded linear map $S_{r}^{-1} : x \to X$ that $S \circ S_{r}^{-1} = I_{X}$ where $I_{X}$ stands for an identity operator on $X$. The series expansion $g = \sum_{i \in I} c_{i} \phi_{i}$ of a given function $g \in X$ where $\{\phi_{i}\}_{i \in I}$ is a set of atoms is called atomic decomposition of $g$.

In Definition 9.1 of [2] we first introduced what we called 'wave packet system' and then, in Theorems 9.3 and 9.4 of the same report, established the conditions on the prototype functions of the elements of the system under which it will constitute either a Banach frame or a set of atoms for the wave packet space. By integrating these conditions with the definition, we shall now redefine the wave packet system so that it will automatically form both a Banach frame and a set of atoms of the corresponding wave packet space.

**Definition 10.** First, let $0 \leq \beta \leq \alpha < 1, s_{0} \leq 0, p_{0}$ and $q_{0} \in (0, 1]$. Second, let $t := (t_{1}, t_{2}) \in \mathbb{R}^{2}, \gamma(t)$ and $\psi(t) \in C^{1}(\mathbb{R}^{2}), \partial^{\gamma} \gamma(t)$ and $\partial^{\gamma} \psi(t) \in L^{1}(\mathbb{R}^{2}) \cap L^{\infty}(\mathbb{R}^{2})$ for any $a \in \mathbb{N}_{0} \times \mathbb{N}_{0}$ such that $|a| \leq 1$,

$$
\sup_{t \in \mathbb{R}^{2}} (1 + |t|)^{10+2p_{0}} \cdot |\gamma(t)| < \infty \quad \text{and} \quad \sup_{t \in \mathbb{R}^{2}} (1 + |t|)^{10+2p_{0}} \cdot |\psi(t)| < \infty .
$$

(9)

Third, let $\xi := (\xi_{1}, \xi_{2}) \in \mathbb{R}^{2}, \hat{\gamma}(\xi)$ and $\hat{\psi}(\xi) \in C^{\infty}(\mathbb{R}^{2}), \hat{\gamma}(\xi) \neq 0$ for any $\xi \in \overline{B}_{4}(0), \hat{\psi}(\xi) \neq 0$ for any $\xi \in [-\epsilon, 1 + \epsilon] \times [-1 - \epsilon, 1 + \epsilon]$,

$$
|\partial^{\gamma} \hat{\gamma}(\xi)| \lesssim (1 + |\xi|)^{-\kappa} \cdot (1 + |\xi_{1}|)^{-\kappa_{1}} \cdot (1 + |\xi_{2}|)^{-\kappa_{2}}
$$

(10)

and

$$
|\partial^{\gamma} \hat{\psi}(\xi)| \lesssim (1 + |\xi|)^{-\kappa} \cdot (1 + |\xi_{1}|)^{-\kappa_{1}} \cdot (1 + |\xi_{2}|)^{-\kappa_{2}}
$$

(11)

for any $a \in \mathbb{N}_{0} \times \mathbb{N}_{0}$ such that $|a| \leq 12$, any $b \in \mathbb{N}_{0} \times \mathbb{N}_{0}$ such that $|b| \leq 1, \kappa \geq 10, \kappa_{1} \geq 2$ and $\kappa_{2} \geq 10$. More precise, but rather cumbersome definitions of $a, b, \kappa, \kappa_{1}$ and $\kappa_{2}$, which all depend on $\alpha, \beta, p_{0}$ and $q_{0}$, can be found in the statements of Theorems 9.3 and 9.4 of [2]. Fourth, let

$$
I := \{(000k) : k \in \mathbb{Z}^{2}\} \cup \{(jm\ell k) \in \mathbb{N} \times \mathbb{N}_{0} \times \mathbb{N}_{0} \times \mathbb{Z}^{2} : m \leq m_{j}^{max} \text{ and } l \leq l_{j}^{max}\}
$$

(12)

with $m_{j}^{max}$ and $l_{j}^{max}$ as defined in (3) and

$$
\psi_{i}(t) := \gamma(t - \delta \cdot k) \left| \det A_{j} \right|^{1/2} \exp \left(-2\pi i (R_{j} B_{jm} \ell, t) \right) \cdot \left( \psi \circ A_{j} \circ R_{j}^{-1} \right) (t - \delta \cdot R_{j} A_{j}^{-1} k) \quad \text{if } (jm \ell) = (000)
$$

(13)

with $A_{j}, B_{jm}$ and $R_{j}$ as defined in (5) and (6). Then the set $W(\alpha, \beta) := \{\psi_{i}\}_{i \in I}$ is called 'wave packet system'.

Theorems 9.3 and 9.4 of [2] indicate, as a corollary, that thus defined wave packets are both Banach frames and sets of atoms of $W_{s}^{p,q}(\alpha, \beta)$ with $s \in [-s_{0}, s_{0}], p \in [p_{0}, \infty], q \in [q_{0}, \infty]$ as long as $\delta \in (0, \delta_{0}]$ with $\delta_{0} = \delta_{0}(\alpha, \beta, s_{0}, p_{0}, q_{0}, \gamma, \psi) > 0$ and that the coefficient space that is associated with $W_{s}^{p,q}(\alpha, \beta)$ through $\{\psi_{i}\}_{i \in I}$ is given by the following definition.
Definition 11. Let $0 \leq \beta \leq \alpha < 1$, $s \in \mathbb{R}$, $p$ and $q \in (0, \infty]$, and $I_0$ be as defined in [2]. The set of sequences of complex numbers

$$C^p_s q(\alpha, \beta) := \{c = \{c^{(k)}_i\}_{i \in I_0, k \in \mathbb{Z}^2} \in \mathbb{C}^{I_0 \times \mathbb{Z}^2} : \|c\|_{C^p_s q} < \infty\}$$

where

$$\|c\|_{C^p_s q} := \left\|\left(w_i \cdot \|c^{(i)}_k\|_{l_p}\right)_{i \in I_0, k \in \mathbb{Z}^2}\right\|_q \in [0, \infty]$$

is called 'coefficient space' associated with the wave packet space $\mathcal{W}^p_s q(\alpha, \beta)$.

The Fourier transform of $\psi_i(t)$ can be calculated to be

$$\hat{\psi}_i(\xi) := \begin{cases} e^{-2\pi i(j \cdot k \cdot \xi)} \cdot \hat{\gamma}(\xi) & \text{if } (jml) = (000) \\
\left|\det A_j\right|^{1/2} \cdot e^{-2\pi i(j \cdot R_j A_j^{-1} k \cdot \xi)} \cdot \left(\hat{\psi} \circ A_j^{-1} \circ R_j^{-1}\right)(\xi - R_j A_j B_{jm}) & \text{if } (jml) \neq (000) \end{cases}.$$  (14)

The function $\gamma(t)$ and $\psi(t)$ play the role of prototypes of the wave packets $\psi_i(t)$ with $i \in I$. Specifically all wave packets $\psi_i(t)$ with $i$ such that $(jml) = (000)$ are generated from $\gamma(t)$ only by shifting it, while the wave packets $\psi_i(t)$ with $i$ such that $(jml) \neq (000)$ are generated from $\psi(t)$ by scaling, modulating, rotating and shifting it. The index $j \in \mathbb{N}$ identifies the measurements $2^{-\alpha_j} \times 2^{-\beta_j}$ of the wave packet $\psi_i(t)$ — compared with those of its prototype — and, indeed, the measurements $2^{\alpha_j} \times 2^{\beta_j}$ of its Fourier transform $\hat{\psi}_i(\xi)$ — compared with those of the Fourier transform of its prototype. The indexes $j$, $m$ and $l$ determine the position $\xi_i$ of the maximum of the Fourier transform $\hat{\psi}_i(\xi)$ of the wave packet $\psi_i(t)$ in the frequency plane, namely

$$\xi_i := \xi_{jml} := \begin{cases} 0 & \text{if } (jml) = (000) \\
R_{jl} B_{jm} = R_{jl} \left(2^{j-1} + m \cdot 2^{\alpha_j}\right) & \text{if } (jml) \neq (000) \end{cases}.$$  (15)

The indexes $j$ and $m$ in particular determine the distance $r_i$ of the maximum of the of Fourier transform $\hat{\psi}_i(\xi)$ of the wave packet $\psi_i(t)$ from the origin of the frequency plane, namely

$$r_i := |\xi_i| := |\xi_{jml}| := \begin{cases} 0 & \text{if } (jml) = (000) \\
2^{j-1} + m \cdot 2^{\alpha_j} & \text{if } (jml) \neq (000) \end{cases}.$$  (16)

while the indexes $j$ and $l$ determine the angle $\theta_i$ at which the Fourier transform $\hat{\psi}_i(\xi)$ of the wave packet $\psi_i(t)$ is inclined to the $\xi_1$-axis in the frequency plane, namely

$$\theta_i := \begin{cases} 0 & \text{if } (jml) = (000) \\
\theta_{jl} & \text{if } (jml) \neq (000) \end{cases}.$$  (17)

with $\theta_{jl}$ as defined by [3]. In other words $r_i$ and $\theta_i$ are polar coordinates of the maximum of $\hat{\psi}_i(\xi)$ in the frequency plane. The indexes $k$, $j$ and $l$ determine the position of the maximum of the wave packet $\psi_i(t)$ in the time plane, namely

$$t_i := t_k^{(jl)} := \begin{cases} \delta \cdot k = \delta \cdot \left(\begin{array}{c} k_1 \\ k_2 \end{array}\right) & \text{if } (jml) = (000) \\
\delta \cdot R_{jl} A_j^{-1} k = \delta \cdot R_{jl} \left(\begin{array}{c} k_1 \cdot 2^{-\alpha_j} \\ k_2 \cdot 2^{-\beta_j}\end{array}\right) & \text{if } (jml) \neq (000) \end{cases}.$$  (18)

Whether the set $\{\psi_i\}_{i \in I}$ of the wave packets can be orthogonalised, say, by imposing additional conditions on the prototypes $\gamma$ and $\phi$ is the subject of a separate ongoing project. At the moment we shall show that $\langle \psi_i | \psi_i \rangle$ tends to zero as $i$ and $i'$ become different. To measure the difference between $i$ and $i'$ we shall need the following definition.
Lemma justifies the name Lemma 1. the Cartesian components of the vectors that can be made into packets \( \psi \).

Definition 12. Let \( i = (jmlk) \) and \( i' = (j'm'l'k') \in I \) where \( I \) as defined by [12], \( \Delta r := |r_i - r_{i'}| \in \mathbb{R}_+ \)
where \( r_i \) and \( r_{i'} \) as defined by [16], \( \Delta \theta := |\theta_i - \theta_{i'}| \in [0, \pi] \) where \( \theta_i \) and \( \theta_{i'} \) as defined by [17],

\[
\Delta t_1 := \delta \cdot \left| |k_1| \cdot 2^{-\alpha j} - |k_{1'}| \cdot 2^{-\alpha j'} \right| \quad (19)
\]

and

\[
\Delta t_2 := \delta \cdot \left| |k_2| \cdot 2^{-\beta j} - |k_{2'}| \cdot 2^{-\beta j'} \right| . \quad (20)
\]

Then

\[
\rho(i, i') := \Delta r + \Delta \theta + \Delta t_1 + \Delta t_2 \quad (21)
\]

will be called pseudo-distance between indexes \( i \) and \( i' \) of the wave packets \( \psi_{i}(t) \) and \( \psi_{i'}(t) \).

Thus the pseudo-distance \( \rho(i, i') \) between the indexes \( i \) and \( i' \) and, indeed, between the wave packets \( \psi_{i} \) and \( \psi_{i'} \), indexed by them, is defined as the sum of the absolute values of the differences \( \Delta r \) and \( \Delta \theta \) between the polar coordinates of the maxima of the Fourier transforms \( \hat{\psi}_{i} (\xi) \) and \( \hat{\psi}_{i'} (\xi) \) of the wave packets \( \psi_{i}(t) \) and \( \psi_{i'}(t) \) and the absolute values of the differences \( \Delta t_1 \) and \( \Delta t_2 \) between the lengths of the Cartesian components of the vectors that can be made into \( t_i \) and \( t_{i'} \), pointing at the maxima of \( \psi_{i}(t) \) and \( \psi_{i'}(t) \), by rotating them clockwise through the angles \( \theta_i \) and \( \theta_{i'} \) respectively. The following Lemma justifies the name pseudo-distance of \( \rho(i, i') \).

Lemma 1. Let \( i, i' \in I \), then

\[
\rho(i, i') \geq 0 , \quad (22)
\]

\[
i = i' \implies \rho(i, i') = 0 , \quad (23)
\]

\[
\rho(i, i') = \rho(i', i) \quad (24)
\]

and

\[
\rho(i, i'') \leq \rho(i, i') + \rho(i', i'') . \quad (25)
\]

Proof. First, the non-negativity of \( \rho(i', i) \) expressed by (22) follows from its being the sum of four non-negative numbers. Second, if \( i = (jmlk) = i' = (j'm'l'k') \), then \( \Delta r, \Delta \theta, \Delta t_1 \) and \( \Delta t_2 \) equal zero according to (16), (17), (19) and (20) respectively. This, in its turn, implies that \( \rho(i, i') = 0 \). Third, the symmetry of \( \rho(i, i') \) expressed by (24) results from that of \( \Delta r, \Delta \theta, \Delta t_1 \) and \( \Delta t_2 \). Fourth, the subadditivity of \( \rho(i, i') \) expressed by (25) results from that of \( \Delta r, \Delta \theta, \Delta t_1 \) and \( \Delta t_2 \), which are absolute values of differences between real numbers. \( \square \)

Thus \( \rho(i, i') \) satisfies the four axioms of pseudo-metric. Unfortunately \( \rho(i, i') = 0 \) does not quite imply that \( i = i' \) to qualify as a proper metric. Indeed, if \( \rho(i, i') = 0 \), then \( \Delta r = 0 \) or, according to (16), \( 2^{j-1} + m \cdot 2^{\alpha j} = 2^{j'-1} + m' \cdot 2^{\alpha j'} \). This, in its turn, implies that \( j = j' \) and therefore \( m = m' \) too. Indeed, let us assume that, say, \( j > j' \) or, in other words, that \( j \geq j' + 1 \). Then, on the one hand

\[
r_i = 2^{j-1} + m \cdot 2^{\alpha j} \geq 2^{j-1} \geq 2^{j'} \quad (26)
\]

and, on the other hand,

\[
r_{i'} = 2^{j'-1} + m' \cdot 2^{\alpha j'} \leq 2^{j'-1} + (m'_{\text{max}} - \omega) \cdot 2^{\alpha j'} = 2^{j'} - \omega \cdot 2^{\alpha j'} < 2^{j'} \quad (27)
\]

where \( 0 < \omega \leq 1 \) in general and \( \omega \) equals one only if \( 2^{(1-\alpha)j'-1} \), in the definition of \( m'_{\text{max}} := \left[ 2^{(1-\alpha)j'-1} \right] \), happens to be even. The last two estimates together imply that \( |r_i - r_{i'}| = \Delta r > 0 \), which contradicts \( \rho(i, i') = 0 \). Furthermore \( \rho(i, i') = 0 \) also implies that \( \Delta \theta, \Delta t_1 \) and \( \Delta t_2 \) equal zero or, according to (17), (19) and (20), that either \( l = l' = 0 \) if \( i \) is such that \((jml) = (000)\) or \( 2^{(\beta-1)j} \cdot l = 2^{(\beta-1)j} \cdot l' \) if it is not, and that

\[
|k_1| \cdot 2^{-\alpha j} = |k_{1'}| \cdot 2^{-\alpha j'}
\]

and

\[
|k_2| \cdot 2^{-\beta j} = |k_{2'}| \cdot 2^{-\beta j'} .
\]
This together with \( j = j \)' indicates that \( l = l' \), \( |k_1| = |k_1'| \) and \( |k_2| = |k_2'| \). Thus, in general, \( \rho(i, i') = 0 \) does not necessarily imply that \( k_1 = k_1' \) and \( k_2 = k_2' \).

We now state and prove the main theorem of this section.

**Theorem 1.** Let \( \psi_i(t) \) and \( \psi_{i'}(t) \) be any two elements of the wave packet system \( W(\alpha, \beta) \) of the wave packet space \( \mathcal{W}^{p,q}_{s}(\alpha, \beta) \) and \( \rho(i, i') \) the pseudo-distance between their indexes \( i \) and \( i' \in I \), then

\[
|\langle \psi_i, \psi_{i'} \rangle| \lesssim (1 + \rho(i, i'))^{-6}. \tag{28}
\]

**Proof.** For the sake of clarity we shall distinguish three possible situations depending on the values of \( i \) and \( i' \) and deal with them separately. Specifically, \( i \) and \( i' \) are such that \( (jml) = (000) = (j'm'l') \), \( (jml) \neq (000) \neq (j'm'l') \) and \( (jml) = (000) \neq (j'm'l') \).

Let \( (jml) = (000) = (j'm'l') \), then \( r_i = r_{i'} = 0 \) and \( \theta_i = \theta_{i'} = 0 \) and so \( \Delta r = 0 \) and \( \Delta \theta = 0 \). From \[9\] we infer that

\[
|\gamma(t)| \lesssim \frac{1}{(1 + |t|)^{12}}. \tag{29}
\]

Furthermore from Section K.1 in \[10\] we know that

\[
\int_{\mathbb{R}^n} \frac{1}{(1 + |x - a|)^m} \cdot \frac{1}{(1 + |x - b|)^m} \, dx \lesssim \frac{1}{(1 + |a - b|)^m} \tag{30}
\]

for \( m > n \). Therefore

\[
|\langle \psi_i | \psi_{i'} \rangle| \lesssim \int |\psi_i| \cdot |\psi_{i'}| \, dt \lesssim \int \frac{1}{(1 + |t - t_i|)^{12}} \cdot \frac{1}{(1 + |t - t_{i'}|)^{12}} \, dt \lesssim \frac{1}{(1 + |t_i - t_{i'}|)^{12}}. \tag{31}
\]

where

\[
\frac{1}{1 +\delta^2 \cdot (k_1 - m) (k_2 - 2k_2)} \lesssim \frac{1}{1 + \Delta t_1 + \Delta t_2} \cdot \tag{32}
\]

Indeed, if \( |k_1 - k_1'| \) equals zero or one, then \( (k_1 - k_1')^2 = |k_1 - k_1'| \), while if \( |k_1 - k_1'| > 1 \), then \( (k_1 - k_1')^2 > |k_1 - k_1'| \), and in any case \( |k_1 - k_1'| \geq |k_1| - |k_1'| \). Therefore

\[
\frac{1}{1 +\delta^2 \cdot (k_1 - m) (k_2 - 2k_2)} \leq \frac{1}{1 + \delta \cdot (k_1 - m) + \delta' \cdot (k_2 - 2k_2)} \leq \frac{1}{1 + \delta \cdot (k_1 - m) + \delta' \cdot (k_2 - 2k_2)} \]

for any positive \( \delta \). Combining (31) with (32) results in (28).

Let now \( (jml) \neq (000) \neq (j'm'l') \). From the first three inequalities (31) – which hold in this case too – we infer that \( \langle \psi_i | \psi_{i'} \rangle \) decreases monotonically as \( t_i - t_{i'} \) increases. Moreover we note that

\[
|t_i - t_{i'}| = |\delta \cdot R_{j} A_{j}^{-1} k - \delta \cdot R_{j'} A_{j'}^{-1} k'| \geq |\Delta t| \tag{33}
\]

where

\[
\Delta t := \left( \frac{\Delta t_1}{\Delta t_2} \right) \cdot \tag{34}
\]
where we replaced the vector \( t_i - t_i' \) by the shorter vector \( \Delta t \), assumed that neither \( \Delta t_1 \) nor \( \Delta t_2 \) equals zero and integrated \( \langle \hat{\psi}_1 | \hat{\psi}_1 \rangle \) by parts six times with respect to \( \xi_1 \) and six times with respect to \( \xi_2 \). If either \( \Delta t_1 \) or \( \Delta t_2 \) equals zero, then integrating six times by parts only with respect to \( \xi_2 \) or \( \xi_1 \) results in

\[
| \langle \psi_1 | \psi_1' \rangle | \leq \frac{1}{(2\pi)^6 \cdot \Delta t_1^5 \cdot \Delta t_2^5} \cdot \int_{\mathbb{R}^2} \left| \frac{\partial^{12}}{\partial \xi_2 \partial \xi_1^5} \right| \left[ \hat{\psi}_1 \cdot \hat{\psi}_1' \right] d\xi
\]

or

\[
| \langle \psi_1 | \psi_1' \rangle | \leq \frac{1}{(2\pi)^6 \cdot \Delta t_1^5 \cdot \Delta t_2^5} \cdot \int_{\mathbb{R}^2} \left| \frac{\partial^{12}}{\partial \xi_1 \partial \xi_2^5} \right| \left[ \hat{\psi}_1 \cdot \hat{\psi}_1' \right] d\xi
\]

respectively. Finally, if both \( \Delta t_1 \) and \( \Delta t_2 \) equal zero, then

\[
\langle \psi_1 | \psi_1' \rangle \leq \int_{\mathbb{R}^2} | \hat{\psi}_1 \cdot \hat{\psi}_1' | d\xi
\]

We now note that

\[
\frac{1}{\Delta t_1} \leq \frac{1}{1 + \Delta t_1}
\]

for any positive \( \Delta t_1 \). Indeed, if \( \Delta t_1 \geq 1 \), then

\[
\frac{1}{\Delta t_1} = \frac{2}{1 + \Delta t_1} \leq \frac{2}{1 + \Delta t_1} \times \frac{1}{1 + \Delta t_1} ;
\]

if \( 0 < \Delta t_1 < 1 \), then there exists such an \( 0 < \epsilon < 1 \) that \( \epsilon < \Delta t_1 \) and

\[
\frac{1}{\Delta t_1} = \frac{2/\epsilon}{(2/\epsilon) \cdot \Delta t_1} \leq \frac{2/\epsilon}{1 + (1/\epsilon) \cdot \Delta t_1} \leq \frac{2/\epsilon}{1 + \Delta t_1} \times \frac{1}{1 + \Delta t_1} .
\]
From (39) we infer that
\[
\frac{1}{\Delta t_1 \cdot \Delta t_2} \lesssim \frac{1}{1 + \Delta t_1} \cdot \frac{1}{1 + \Delta t_2} \lesssim \frac{1}{1 + \Delta t_1 + \Delta t_2}
\] (42)
for any positive \(\Delta t_1\) and \(\Delta t_2\),
\[
\frac{1}{\Delta t_1} \lesssim \frac{1}{1 + \Delta t_1 + \Delta t_2}
\] (43)
for \(\Delta t_1 = 0\) and \(\Delta t_2 > 0\) and
\[
\frac{1}{\Delta t_2} \lesssim \frac{1}{1 + \Delta t_1 + \Delta t_2}
\] (44)
for \(\Delta t_1 > 0\) and \(\Delta t_2 = 0\). Moreover
\[
1 = \frac{1}{1 + \Delta t_1 + \Delta t_2}
\] (45)
if \(\Delta t_1 = \Delta t_2 = 0\).

Using (42), (44), (43) or (45) and replacing \(\partial^2 \theta_1 \partial \xi_2 [\hat{\psi}_1 \cdot \hat{\psi}_2], \partial^4 \theta_2 [\hat{\psi}_1 \cdot \hat{\psi}_2], \partial^6 \theta_1 [\hat{\psi}_1 \cdot \hat{\psi}_2]\) or \(|\hat{\psi}_1 \cdot \hat{\psi}_2|\) in (35), (36), (37) or (38), respectively, by their estimate (11) results in
\[
|\langle \psi_1 | \psi_2 \rangle| \lesssim \frac{1}{(1 + \Delta t_1 + \Delta t_2)^6} \cdot \int_{\mathbb{R}^2} \left[(1 + |\xi - \xi_1|) \cdot (1 + |\xi - \xi_2|)\right]^{-\kappa} \cdot [(1 + |\xi_1 - \xi_{i1}|) \cdot (1 + |\xi_1 - \xi_{i'1}|)]^{-\kappa_1} \cdot [(1 + |\xi_2 - \xi_{i2}|) \cdot (1 + |\xi_2 - \xi_{i'2}|)]^{-\kappa_2} \, d\xi.
\] (46)
Moreover, since \(|\xi - \xi_1| \geq |\xi_1 - \xi_{i1}|, |\xi - \xi_2| \geq |\xi_1 - \xi_{i1}|\), \(\kappa \geq 10, \kappa_1 \geq 2\) and \(\kappa \geq 10\), we conclude that
\[
|\langle \psi_1 | \psi_2 \rangle| \lesssim \frac{1}{(1 + \Delta t_1 + \Delta t_2)^6} \cdot \int_{\mathbb{R}^2} \left[(1 + |\xi - \xi_1|) \cdot (1 + |\xi - \xi_2|)\right]^{-\kappa+4} \cdot [(1 + |\xi_1 - \xi_{i1}|) \cdot (1 + |\xi_1 - \xi_{i'1}|)]^{-\kappa_1+4} \cdot [(1 + |\xi_2 - \xi_{i2}|) \cdot (1 + |\xi_2 - \xi_{i'2}|)]^{-\kappa_2} \, d\xi.
\] (47)
We now estimate each of the three factors of the integrand in this expression.
\[
\frac{1}{1 + |\xi - \xi_1|} \cdot \frac{1}{1 + |\xi - \xi_2|} \lesssim \frac{1}{1 + |\xi - \xi_1|} \cdot \frac{1}{1 + |\xi - \xi_2|} = \frac{1}{1 + |r - r_1|} \cdot \frac{1}{1 + |r - r_2|} = \frac{1}{1 + |r - r_1| + |r - r_2|} \cdot \frac{1}{1 + |r - r_1| + |r - r_2|} = \frac{1}{1 + |r - r_1| + |r - r_2|} \cdot \frac{1}{1 + |r - r_1| + |r - r_2|}.
\] (48)
Similarly
\[
\frac{1}{1 + |\xi_2 - \xi_{i2}|} \cdot \frac{1}{1 + |\xi_2 - \xi_{i'2}|} = \frac{1}{1 + |r \cdot \sin \theta - r_1 \cdot \sin \theta_1|} \cdot \frac{1}{1 + |r \cdot \sin \theta - r_2 \cdot \sin \theta_2|} = \frac{1}{1 + |r \cdot \sin \theta - r_1 \cdot \sin \theta_1|} \cdot \frac{1}{1 + |r \cdot \sin \theta - r_2 \cdot \sin \theta_2|}.
\] (49)
At this point we consider two possible situations. If \(|r_{i'} \cdot \sin \theta_{i'} - r_1 \cdot \sin \theta_1| = 0\), then
\[
\frac{1}{1 + |\xi_2 - \xi_{i2}|} \cdot \frac{1}{1 + |\xi_2 - \xi_{i'2}|} = \frac{1}{1 + |\xi_2 - \xi_{i2}|} \cdot \frac{1}{1 + |\xi_2 - \xi_{i'2}|} \lesssim \frac{1}{1 + |\xi_2 - \xi_{i2}|} \cdot \frac{1}{1 + |\xi_2 - \xi_{i'2}|}.
\] (50)
Furthermore, if $|\theta_i - \theta_j| \in [0, \frac{\pi}{2}]$, then, given $|\phi| = |\sin(\phi - \frac{\pi}{2})|$, we have

$$
\frac{1}{1 + |\xi_1 - \xi_i|} \cdot \frac{1}{1 + |\xi_1 - \xi_i|} \leq \frac{1}{1 + |\cos(\theta_i - \theta_j)|} = \frac{1}{1 + |\cos(\theta_i - \theta_j)|} = \frac{1}{1 + \sin(|\theta_i - \theta_j| - \frac{\pi}{2})} \leq \frac{1}{1 + |\theta_i - \theta_j|}.
$$

Combining (54) and (55) results in

$$
\frac{1}{1 + |\xi_1 - \xi_i|} \cdot \frac{1}{1 + |\xi_1 - \xi_i|} \leq \frac{1}{1 + |\theta_i - \theta_j|} = \frac{1}{1 + |\theta_i - \theta_j|}.
$$
for $|\theta_i - \theta_{i'}| \in [\frac{\pi}{2}, \pi]$. Combining (53) and (56) results in

$$\frac{1}{1 + |\xi_1 - \xi_{i1}|} \cdot \frac{1}{1 + |\xi_1 - \xi_{i'1}|} \cdot \frac{1}{1 + |\xi_2 - \xi_{i2}|} \cdot \frac{1}{1 + |\xi_2 - \xi_{i'2}|} \lesssim \frac{1}{1 + |\theta_i - \theta_{i'}|}$$

(57)

for $|\theta_i - \theta_{i'}| \in [0, \pi]$. Moreover from (30) we know that

$$\int_{\mathbb{R}^2} \frac{1}{(1 + |\xi - \xi_i|)^3} \cdot \frac{1}{(1 + |\xi - \xi_{i'}|)^3} \, d\xi \lesssim \frac{1}{(1 + |\xi_i - \xi_{i'}|)^3} \lesssim \frac{1}{(1 + |r_i - r_{i'}|)^3}.$$  

(58)

Finally combining (47), (48), (57) and (58) results in (28). Using (48) along with (58) rather than the latter alone leads to a much smaller implied constant in (28).

Finally let $(jm\ell) = (000) \neq (j'm'n')$, then the arguments similar to those used in the previous example lead again to the expressions identical with (47), (48) and (57) where $\Delta r = r_{i'}, \Delta \theta = \theta_{i'}$,

$$\Delta t_1 := \delta \cdot |k_1| - |k_{i'}| \cdot 2^{-\alpha j'} \quad \text{and} \quad \Delta t_2 := \delta \cdot |k_2| - |k_{i'}| \cdot 2^{-\alpha j'}.$$  

The estimate (30) also remains valid. Therefore putting them all together again results in (28). $\square$

Theorem 1 indicates that the elements of the Gram matrix $G := \langle \langle \psi_i, \psi_{i'} \rangle \rangle_{(i,j)\in I \times I}$ of the wave packet system $\{\psi_i\}_{i\in I}$ decay polynomially as they deviate from the diagonal or, in other words, that the correlation of the wave packets $\psi_i$ and $\psi_{i'}$ fades as their distance $\rho(i, i')$ in the phase space increases. This can be viewed as near orthogonality, so to speak, of the wave packets as their true orthogonality would amount to their Gram matrix being diagonal. This allows us to conclude that the wave packet system $W(\alpha, \beta) = \{\psi_i\}_{i\in I}$ will be localised in the sense of the following definition of this notion.

**Definition 13.** Let $I \subset \mathbb{R}^d$ be a denumerable set equipped with a metric $\rho$ and $n > d$. The set of functions $\{\psi_i\}_{i\in I}$ is said to be ‘self-localised’ if

$$|\langle \psi_i, \psi_{j} \rangle| \lesssim (1 + \rho(i, i'))^{-n}$$

(59)

for all $i$ and $i' \in I$.

Indeed the elements of the Gram matrix $G$ of the wave packet system $W(\alpha, \beta) = \{\psi_i\}_{i\in I}$ satisfy (59) with $6 = n > d = 5$ as $i$ and $i' \in I \subset \mathbb{Z}^5 \subset \mathbb{R}^5$. Moreover, from the work in [16] we know that, under this assumption, the matrix $G$ will belong to a solid involutive Banach algebra closed under inversion and therefore the wave packet system will be localised according to the definition of this notion given in [9] as well.

4. Conclusions

In this report we first reformulated more neatly the notion of the wave packet system $W(\alpha, \beta)$, originally introduced in [2], so that it now forms automatically a Banach frame and a set of atoms of the corresponding wave packet space $W^p(\alpha, \beta)$. We then introduced the notion of the pseudo-distance between the indexes of the elements of the wave packet system and proved that the absolute value of the scalar product of two elements of the system decays polynomially as the pseudo-distance between their indexes increases. This, in its turn, can be interpreted as near orthogonality or, in other words, good localisation of the elements the wave packet system in the phase space. We believe that this property will make the wave packet systems particularly useful for the efficient approximation of functions of a wide range of classes and discretisation of bounded linear operators.

Acknowledgements

I am very grateful to Peter Balazs, Nicki Holighaus, Georg Tauböck and Luis Daniel Abreu for an interesting and helpful discussion we had after I had reported on this work at a seminar at the Acoustics Research Institute of the Austrian Academy of Sciences. I am also grateful to Felix Voigtlaender for drawing my attention to two reports relevant to this work. Finally I thank the Centre National de la Recherche Scientifique and Österreichische Akademie der Wissenschaften for their support of this work.
References

[1] P. Balazs and K. Gröchenig, *A Guide to localized frames and applications to Galerkin-like representations of operators*, in I. Pesenson, Q. Le Gia, A. Mayeli, H. Mhaskar, D.-X. Zhou, *Frames and Other Bases in Abstract and Function Spaces*, Birkhäuser, Cham (2017).

[2] D. Bytchenkoff and F. Voigtlaender, *Design and properties of wave packet smoothness spaces*, J. Math. Pures Appl. 133 (2020) 185-262.

[3] E. J. Candès and D. L. Donoho, *Ridgelets : a key to higher-dimensional intermittency?*, Phil. Trans. R. Soc. Lond. A. 357 (1999) 2495-2509.

[4] E. J. Candès and D. L. Donoho, *New tight frames of curvelets and optimal representations of objects with piecewise $C^2$ singularities*, Comm. Pure Appl. Math. 57 (2002) 219-266.

[5] E. J. Candès and L. Demanet, *The curvelet representation of wave propagators is optimally sparse*, Comm. Pure Appl. Math. 58 (2004) 1472-1528.

[6] I. Daubechies, *Ten lectures on wavelets*, Society For Industrial and Applied Mathematics (1992).

[7] L. Demanet, L. Ying, *Wave atoms and time upscaling of wave equations*, Numer. Math. 113 (2009) 1-71.

[8] H. G. Feichtinger and P. Gröbner, *Banach spaces of distributions defined by decomposition methode*, Nachr. 123 (1985) 97-120.

[9] M. Fornasier and K. Gröchenig *Intrinsic localization of frames*, Constr. Approx. 22 (2005) 395-415.

[10] K. Graﬁkos, *Classical Fourier analysis*, Springer (2008).

[11] P. Grohs, *Ridgelet-type frame decompositions for Sobolev spaces related to linear transport*, J. Fourier Anal. Appl. 18 (2012) 309-325.

[12] P. Grohs and G. Kutyniok, *Parabolic molecules*, Found. Comput. Math. 14 (2014) 299-337.

[13] K. Gröchenig, *Describing functions : atomic decompositions versus frames*, Mh. Math. 112 (1991) 1-41.

[14] K. Gröchenig, *Foundations of time-frequency analysis*, Birkhäuser, Basel (2001).

[15] K. Gröchenig, *Localized frames are finite unions of Riesz sequences*, Adv. Comp. Math. 18 (2003) 149-157.

[16] S. Jaffard *Propriétés des matrices "bien localisées" près de leur diagonale et quelques applications*, Ann. Inst. Henri Poincaré 7 (1990) 461-476.

[17] P. Kittipoom, G. Kutyniok and W.-Q. Lim *Construction of compactly supported shearlet frames*, Constr. Approx. 35 (2012) 21-72.

[18] G. Kutyniok and W.-Q Lim *Compactly supported shearlets are optimally sparse*, J. Approx. Theory 163 (2011) 1564-1589.