

Gabor representations of evolution operators

*Original*

Gabor representations of evolution operators / Elena, C., Nicola, F., Luigi, R.. - In: TRANSACTIONS OF THE AMERICAN MATHEMATICAL SOCIETY. - ISSN 0002-9947. - STAMPA. - 367:11(2015), pp. 7639-7663. [10.1090/S0002-9947-2015-06302-8]

*Availability:*

This version is available at: 11583/2543363 since: 2020-06-01T11:33:51Z

*Publisher:*

AMERICAN MATHEMATICAL SOCIETY

*Published*

DOI:10.1090/S0002-9947-2015-06302-8

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

# GABOR REPRESENTATIONS OF EVOLUTION OPERATORS

ELENA CORDERO, FABIO NICOLA AND LUIGI RODINO

ABSTRACT. We perform a time-frequency analysis of Fourier multipliers and, more generally, pseudodifferential operators with symbols of Gevrey, analytic and ultra-analytic regularity. As an application we show that Gabor frames, which provide optimally sparse decompositions for Schrödinger-type propagators, reveal to be an even more efficient tool for representing solutions to a wide class of evolution operators with constant coefficients, including weakly hyperbolic and parabolic-type operators. Besides the class of operators, the main novelty of the paper is the proof of *super-exponential* (as opposite to super-polynomial) off-diagonal decay for the Gabor matrix representation.

## 1. INTRODUCTION

A Gabor system is obtained by fixing a function  $g \in L^2(\mathbb{R}^d)$  and considering the time-frequency shifts

$$(1) \quad \pi(\lambda)g = e^{2\pi i n x} g(x - m), \quad \lambda = (m, n) \in \Lambda,$$

for some lattice  $\Lambda \subset \mathbb{R}^{2d}$ . A Gabor system is a Gabor frame if there exist  $A, B > 0$  such that

$$(2) \quad A\|f\|_{L^2}^2 \leq \sum_{\lambda \in \Lambda} |\langle f, \pi(\lambda)g \rangle|^2 \leq B\|f\|_{L^2}^2$$

for every  $f \in L^2(\mathbb{R}^d)$ , see for example [18, 29, 48]. With respect to frames of wavelets, curvelets and shearlets, in (1) dilations are then replaced by modulations. From the analytic point of view, this provides symmetric properties with respect to the Fourier transform and a simple treatment, cf. [22].

Gabor frames turned out to be appropriate tools for many problems in time-frequency analysis, with relevant applications to signal processing and related issues in Numerical Analysis, see for example [7, 43], and references there. More recently, attention has been addressed to the analysis of partial differential equations, especially the Schrödinger, wave and Klein-Gordon equations with constant coefficients [2, 3, 9, 16, 30, 31, 32, 34, 49, 50, 51, 52]; see also the recent survey [40]. The main results and techniques are now also available in the monograph [53]. Here we carry on this investigation.

To give a general setting for our results, let us denote by  $T$  the linear operator providing the solution of a well-posed Cauchy problem for a partial differential equation, in suitable function spaces. Generically, we expect that  $T$  is written in the form

---

2010 *Mathematics Subject Classification.* 35S05, 42C15.

*Key words and phrases.* Pseudodifferential operators, Gelfand-Shilov spaces, short-time Fourier transform, Gabor frames, sparse representations, hyperbolic equations, parabolic equations.

of pseudodifferential operator (PSDO), or Fourier integral operator (FIO). Both for theoretic and numeric purposes often it turns out necessary to decompose the initial datum with respect to a given frame and to reconstruct the solution by superposition, after studying the evolution of each wave packet. In this picture, the operator  $T$  is then regarded as an infinite matrix, which for Gabor frames reads

$$(3) \quad \langle T\pi(\mu)g, \pi(\lambda)g \rangle.$$

The more this matrix is sparse, the more the above representation is tight.

For example, let  $T$  be a pseudodifferential operator with symbol  $\sigma(x, \xi)$  in the class  $S_{0,0}^0$ , namely satisfying

$$(4) \quad |\partial^\alpha \sigma(z)| \leq C_\alpha, \quad z = (x, \xi) \in \mathbb{R}^{2d}.$$

Then, it was proved in [23, 25, 39, 44] that  $T$  has a super-polynomial Gabor decay:

$$(5) \quad |\langle T\pi(\mu)g, \pi(\lambda)g \rangle| \leq C_N(1 + |\lambda - \mu|)^{-N}$$

for every  $N > 0$ , if the window  $g$  is Schwartz.

Instead, when  $T$  is a FIO with phase function of quadratic type, cf. [1], a similar result was shown in [13, 44], with the difference  $|\lambda - \mu|$  replaced by  $|\lambda - \chi(\mu)|$ , where  $\chi$  is the corresponding canonical transformation. Applications were given there to Schrödinger equations with variable coefficients; see also [8, 10, 11, 36, 37].

For strictly hyperbolic equations with variable coefficients it is instead well-known [12, 14] that the corresponding propagators, which are classical FIO with phase function homogeneous of degree 1 in the dual variables, cf. [28, Vol. 4], do not have a sparse Gabor matrix; see also Remark 5.7 below. For these operators the “correct” wave packets are given by curvelet-type waves, tailored to a parabolic scaling, as shown in [4, 17, 20, 27, 42]. All these papers show a super-polynomial decay for the matrix of the propagator.

The present work is devoted to a systematic study in terms of Gabor frames of a general class of constant coefficient evolution operators. Besides the class of operators, the main novelty with respect to the existent literature is the investigation of super-exponential decay of the Gabor matrix (as opposite to super-polynomial).

To be more precise, we consider operators of the form

$$(6) \quad P(\partial_t, D_x) = \partial_t^m + \sum_{k=1}^m a_k(D_x) \partial_t^{m-k}, \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^d$$

where  $a_k(\xi)$ ,  $1 \leq k \leq m$ , are polynomials, whose degree may be arbitrary; as usual,  $D_{x_j} = \frac{1}{2\pi i} \partial_{x_j}$ ,  $j = 1, \dots, d$ . We suppose that the corresponding forward Cauchy problem is well-posed in a very mild sense, namely in  $\mathcal{S}(\mathbb{R}^d)$ . A necessary and sufficient condition for this is given by the forward Hadamard-Petrowsky condition, involving the complex roots  $\tau \in \mathbb{C}$  of the symbol  $P(i\tau, \xi)$ :

$$(7) \quad (\tau, \xi) \in \mathbb{C} \times \mathbb{R}^d, \quad P(i\tau, \xi) = 0 \implies \text{Im } \tau \geq -C,$$

for some constant  $C > 0$ . When the polynomials  $a_k(\xi)$  in (6) have degree  $\leq k$ , this reduces to the definition of (weak) hyperbolicity, but we do not assume this here. There is a wide literature concerning constant coefficients operators, and the study of

basic problems as hypoellipticity, fundamental solutions, etc., has reached the highest level of sophistication, with a combination of techniques from Algebraic Geometry and Mathematical Analysis [28, Vol. II]. Now, we will show that the Gabor matrix of the corresponding propagator enjoys a *super-exponential* decay, and we link the precise decay rate to the algebraic growth of the imaginary part of the roots  $\tau$  of  $P(i\tau, \zeta)$ , as  $\zeta \in \mathbb{C}$ ,  $\text{Im } \zeta \rightarrow +\infty$ .

As a simple and typical model, let us consider the Cauchy problem for the wave equation:

$$(8) \quad \begin{aligned} \partial_t^2 u - \Delta_x u &= 0, & (t, x) \in \mathbb{R} \times \mathbb{R}^d, \\ u(0, x) &= u_0(x), & \partial_t u(0, x) = u_1(x). \end{aligned}$$

We may express the solution in the form

$$(9) \quad u(t, x) = T_t u_1(x) + \partial_t T_t u_0(x),$$

where  $T_t$  is the Fourier multiplier

$$(10) \quad T_t f(x) = \int e^{2\pi i x \xi} \sigma_t(\xi) \widehat{f}(\xi) d\xi,$$

with symbol

$$(11) \quad \sigma_t(\xi) = \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}, \quad \xi \in \mathbb{R}^d.$$

Taking the Gaussian as window function  $g$  one has in fact the Gaussian decay

$$(12) \quad |\langle T_t \pi(\mu)g, \pi(\lambda)g \rangle| \leq C \exp(-\epsilon|\lambda - \mu|^2);$$

(we address to the sequel of the paper for the dependence of  $C$  and  $\epsilon > 0$  on  $t$ ). This is easily obtained from the explicit expression of the forward fundamental solution  $E(t, x) = \mathcal{F}_{\xi \rightarrow x}^{-1} \sigma(t, \xi)$  (see Example 5.5 below), but this approach does not extend to the more general equations above. Instead, we regard here the propagator  $T_t$  as a PSDO where, with respect to the approach via Fourier integral operators, the  $\xi$ -dependent phase functions are absorbed into the symbol (11). Because of its oscillations,  $\sigma_t(\xi)$  then belongs to the non-standard Hörmander's class  $S_{0,0}^0$  defined by (4), and therefore (5) holds for  $T_t$ . Actually, seeking super-exponential decay, is natural to introduce an analytic version of the class  $S_{0,0}^0$ , by assuming for  $s \geq 0$

$$(13) \quad |\partial^\alpha \sigma(z)| \leq C^{|\alpha|+1} (\alpha!)^s, \quad z \in \mathbb{R}^{2d},$$

for a constant  $C > 0$  depending on  $\sigma$ . We say that the symbol  $\sigma$  is Gevrey when  $s > 1$ , analytic when  $s = 1$ , and ultra-analytic if  $s < 1$ .

We will show, in fact, that a pseudodifferential operator  $T$  with symbol satisfying (13) displays a matrix decay

$$(14) \quad |\langle T \pi(\lambda)g, \pi(\mu)g \rangle| \leq C \exp(-\epsilon|\mu - \lambda|^r)$$

for some  $C, \epsilon > 0$ , with  $r = \min\{2, 1/s\}$ , for suitable windows  $g$  (see Theorem 4.4 below). In the case of the Gevrey-analytic symbols, i.e.  $s \geq 1$  in (13) and  $r \leq 1$  in (14), the estimates follow from the results in [25]. Instead, in view of the applications, our attention will be mainly addressed to the ultra-analytic case, i.e.  $0 \leq s < 1$  in (13),  $r > 1$  in (14). Incidentally we note that the class of ultra-analytic functions exhibits

exotic behaviors (e.g. it is not closed by composition), which make it amazing in many respects; we will see an instance at once.

In fact, it turns out that the symbol  $\sigma_t(\xi)$  in (11) satisfies the estimate (13) with  $s = 0$ . This is already a curious remark even in dimension 1, because the function  $\sin(2\pi|\xi|t)$  satisfies those estimates, e.g. for  $|\xi| \geq 1$ , but  $|\xi|^{-1}$  does not, for any  $s < 1$ . It is really surprising that a similar miraculous combination occurs for a wide class of equations of the form (6).

To state the precise result, we refine the forward Hadamard-Petrowsky condition by requiring that the symbol  $P(i\tau, \zeta)$  of the operator in (6) satisfies

$$(\tau, \zeta) \in \mathbb{C} \times \mathbb{C}^d, \quad P(i\tau, \zeta) = 0 \implies \operatorname{Im} \tau \geq -C(1 + |\operatorname{Im} \zeta|)^\nu,$$

for some  $C > 0$ ,  $\nu \geq 1$ . Then we will prove that the propagator  $Tu = E(t, \cdot) * u$ , where  $E(t, \cdot)$  is the forward fundamental solution, has symbol  $\sigma(t, \cdot) = \widehat{E}(t, \cdot)$  satisfying (13) with  $s = 1 - 1/\nu$ , and therefore (14) holds with  $r = \min\{2, \nu/(\nu - 1)\}$ , for suitable windows (Theorem 5.3). We always have  $r > 1$ , i.e. super-exponential decay.

As a special case, we get Gaussian decay ( $r = 2$ ) for *all* hyperbolic equations. As another example, we get  $r = 2k/(2k - 1)$ ,  $k \in \mathbb{N} \setminus \{0\}$ , for the generalized heat equation:

$$(15) \quad \partial_t u + (-\Delta_x)^k u = 0.$$

Accordingly, every column or row of the Gabor matrix, rearranged in decreasing order, displays a similar decay, i.e. we obtain an *exponential-type sparsity*.

As an easy byproduct, we have the continuity of the propagator on modulation spaces ([22, 53] and Section 2.4 below), which for the wave and Klein-Gordon equations was already proved in [2, 3, 9, 52] by other methods. Actually our result for hyperbolic equations is of particular interest when the operator is only *weakly* hyperbolic, where we are not aware of any almost-diagonalization result in the literature (even super-polynomial). In this respect, there is a very small intersection with the wide literature, mentioned above, of curvelet-type representations, which instead deals with variable coefficient *strictly* hyperbolic equations (and super-polynomial matrix decay).

The plan of the paper is the following. In Section 2 we provide the necessary preliminaries, concerning Gelfand-Shilov spaces, time-frequency representations and modulation spaces. Section 3 is devoted to the study of the symbols in (13). The corresponding PSDOs are considered in Section 4, where we prove (14) and give boundedness results in modulation and Gelfand-Shilov spaces. The applications to evolution equations are given in Section 5.

## 2. PRELIMINARIES

**2.1. Notations.** The Schwartz class is denoted by  $\mathcal{S}(\mathbb{R}^d)$ , the space of tempered distributions by  $\mathcal{S}'(\mathbb{R}^d)$ . We use the brackets  $\langle f, g \rangle$  to denote the extension to  $\mathcal{S}'(\mathbb{R}^d) \times \mathcal{S}(\mathbb{R}^d)$  of the inner product  $\langle f, g \rangle = \int f(t) \overline{g(t)} dt$  on  $L^2(\mathbb{R}^d)$ .

Euclidean norm of  $x \in \mathbb{R}^d$  is given by  $|x| = (x_1^2 + \dots + x_d^2)^{1/2}$ , and  $\langle x \rangle = (1 + |x|^2)^{1/2}$ . We write  $xy = x \cdot y$  for the scalar product on  $\mathbb{R}^d$ , for  $x, y \in \mathbb{R}^d$ .

We adopt the usual multi-index notation and recall that if  $\alpha \in \mathbb{N}^d$  we have

$$(16) \quad \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} = 2^{|\alpha|},$$

and

$$(17) \quad |\alpha|! \leq d^{|\alpha|} \alpha!.$$

The Fourier transform is normalized to be  $\hat{f}(\xi) = \mathcal{F}f(\xi) = \int f(t)e^{-2\pi i t \xi} dt$ .

Translation and modulation operators,  $T$  and  $M$  are defined by

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

For  $z = (x, \xi)$  we shall also write

$$\pi(z)f = M_\xi T_x f.$$

The following relations hold

$$M_y T_x = e^{2\pi i x y} T_x M_y, \quad (T_x f)^\wedge = M_{-x} \hat{f}, \quad (M_x f)^\wedge = T_x \hat{f}, \quad x, y \in \mathbb{R}^d, f, g \in L^2(\mathbb{R}^d).$$

For  $0 < p \leq \infty$  and a weight  $m$ , the space  $\ell_m^p(\Lambda)$  is the space of sequences  $a = \{a_\lambda\}_{\lambda \in \Lambda}$  on a lattice  $\Lambda$ , such that

$$\|a\|_{\ell_m^p} := \left( \sum_{\lambda \in \Lambda} |a_\lambda|^p m(\lambda)^p \right)^{1/p} < \infty$$

(with obvious changes when  $p = \infty$ ).

Throughout the paper, we shall use the notation  $A \lesssim B$  to express the inequality  $A \leq cB$  for a suitable constant  $c > 0$ , and  $A \asymp B$  for the equivalence  $c^{-1}B \leq A \leq cB$ .

**2.2. Gelfand-Shilov Spaces.** The Schwartz class  $\mathcal{S}(\mathbb{R}^d)$  does not give quantitative information about how fast a function  $f \in \mathcal{S}(\mathbb{R}^d)$  and its derivatives decay at infinity. This is the main motivation to use subspaces of the Schwartz class, so-called Gelfand-Shilov type spaces, introduced in [21]. Let us recall their definition and main properties, contained in [21, 35].

**Definition 2.1.** *Let there be given  $s, r \geq 0$  and  $A, B > 0$ . The Gelfand-Shilov type space  $S_{r,B}^{s,A}(\mathbb{R}^d)$  is defined by*

$$(18) \quad S_{r,B}^{s,A}(\mathbb{R}^d) = \{f \in \mathcal{S}(\mathbb{R}^d) : |x^\alpha \partial^\beta f(x)| \lesssim A^{|\alpha|} B^{|\beta|} (\alpha!)^r (\beta!)^s, \quad \alpha, \beta \in \mathbb{N}^d\}.$$

*Their projective and inductive limits are denoted by*

$$\Sigma_r^s = \text{proj} \lim_{A>0, B>0} S_{r,B}^{s,A}; \quad S_r^s = \text{ind} \lim_{A>0, B>0} S_{r,B}^{s,A}.$$

The space  $S_r^s(\mathbb{R}^d)$  is nontrivial if and only if  $r + s > 1$ , or  $r + s = 1$  and  $r, s > 0$ . So the smallest nontrivial space with  $r = s$  is provided by  $S_{1/2}^{1/2}(\mathbb{R}^d)$ . Every function of the type  $P(x)e^{-a|x|^2}$ , with  $a > 0$  and  $P(x)$  polynomial on  $\mathbb{R}^d$ , is in the class  $S_{1/2}^{1/2}(\mathbb{R}^d)$ . We observe the trivial inclusions  $S_{r_1}^{s_1}(\mathbb{R}^d) \subset S_{r_2}^{s_2}(\mathbb{R}^d)$  for  $s_1 \leq s_2$  and  $r_1 \leq r_2$ . Moreover, if  $f \in S_r^s(\mathbb{R}^d)$ , also  $x^\delta \partial^\gamma f$  belongs to the same space for every fixed  $\delta, \gamma$ .

The action of the Fourier transform on  $S_r^s(\mathbb{R}^d)$  interchanges the indices  $s$  and  $r$ , as explained in the following theorem.

**Theorem 2.2.** *For  $f \in \mathcal{S}(\mathbb{R}^d)$  we have  $f \in S_r^s(\mathbb{R}^d)$  if and only if  $\hat{f} \in S_s^r(\mathbb{R}^d)$ .*

Therefore for  $s = r$  the spaces  $S_s^s(\mathbb{R}^d)$  are invariant under the action of the Fourier transform.

**Theorem 2.3.** *Assume  $s > 0, r > 0, s+r \geq 1$ . For  $f \in \mathcal{S}(\mathbb{R}^d)$ , the following conditions are equivalent:*

a)  $f \in S_r^s(\mathbb{R}^d)$ .

b) *There exist constants  $A, B > 0$ , such that*

$$\|x^\alpha f\|_{L^\infty} \lesssim A^{|\alpha|}(\alpha!)^r \quad \text{and} \quad \|\xi^\beta \hat{f}\|_{L^\infty} \lesssim B^{|\beta|}(\beta!)^s, \quad \alpha, \beta \in \mathbb{N}^d.$$

c) *There exist constants  $A, B > 0$ , such that*

$$\|x^\alpha f\|_{L^\infty} \lesssim A^{|\alpha|}(\alpha!)^r \quad \text{and} \quad \|\partial^\beta f\|_{L^\infty} \lesssim B^{|\beta|}(\beta!)^s, \quad \alpha, \beta \in \mathbb{N}^d.$$

d) *There exist constants  $h, k > 0$ , such that*

$$\|f e^{h|x|^{1/r}}\|_{L^\infty} < \infty \quad \text{and} \quad \|\hat{f} e^{k|\xi|^{1/s}}\|_{L^\infty} < \infty.$$

A suitable window class for weighted modulation spaces (see the subsequent Definition 2.6) is the Gelfand-Shilov type space  $\Sigma_1^1(\mathbb{R}^d)$ , consisting of functions  $f \in \mathcal{S}(\mathbb{R}^d)$  such that for every constant  $A > 0$  and  $B > 0$

$$(19) \quad |x^\alpha \partial^\beta f(x)| \lesssim A^{|\alpha|} B^{|\beta|} \alpha! \beta!, \quad \alpha, \beta \in \mathbb{N}^d.$$

We have  $S_s^s(\mathbb{R}^d) \subset \Sigma_1^1(\mathbb{R}^d) \subset S_1^1(\mathbb{R}^d)$  for every  $s < 1$ . Observe that the characterization of Theorem 2.3 can be adapted to  $\Sigma_1^1(\mathbb{R}^d)$  by replacing the words “there exist” by “for every” and taking  $r = s = 1$ .

Let us underline the following property, which exhibits two equivalent ways of expressing the decay of a continuous function  $f$  on  $\mathbb{R}^d$ . This follows immediately from [35, Proposition 6.1.5], where the property was formulated for  $f \in \mathcal{S}(\mathbb{R}^d)$ . For the sake of clarity, we shall detail the proof showing the mutual dependence between the constants  $\epsilon$  and  $C$  below.

**Proposition 2.4** ([35, Proposition 6.1.5]). *Consider  $r > 0$  and let  $h$  be a continuous function on  $\mathbb{R}^d$ . Then the following conditions are equivalent:*

(i) *There exists a constant  $\epsilon > 0$  such that*

$$(20) \quad |h(x)| \lesssim \exp(-\epsilon|x|^{1/r}), \quad x \in \mathbb{R}^d,$$

(ii) *There exists a constant  $C > 0$  such that*

$$(21) \quad |x^\alpha h(x)| \lesssim C^{|\alpha|}(\alpha!)^r, \quad x \in \mathbb{R}^d, \alpha \in \mathbb{N}^d.$$

*Proof.* We re-write (20) in the form

$$(22) \quad |h(x)|^{\frac{1}{r}} \lesssim \exp\left(-\frac{\epsilon}{r}|x|^{\frac{1}{r}}\right), \quad x \in \mathbb{R}^d.$$

In turn, (22) can be re-written as

$$(23) \quad \sup_{x \in \mathbb{R}^d} \sum_{n=0}^{\infty} \left(\frac{\epsilon}{r}\right)^n (n!)^{-1} |x|^{\frac{n}{r}} |h(x)|^{\frac{1}{r}} < \infty.$$

Hence the sequence of the terms of the series is uniformly bounded, as well as the sequence of the  $r$ -th powers:

$$\frac{\epsilon^{rn}}{r^{rn}} (n!)^{-r} |x|^n |h(x)|, \quad n \in \mathbb{N},$$

and we obtain

$$|x|^n |h(x)| \lesssim \frac{r^{rn}}{\epsilon^{rn}} (n!)^r, \quad x \in \mathbb{R}^d, \quad n = 0, 1, \dots$$

Hence, writing  $|\alpha| = n$  and applying (17):

$$|x^\alpha h(x)| \lesssim (rd/\epsilon)^{r|\alpha|} (\alpha!)^r = C^{|\alpha|} (\alpha!)^r, \quad x \in \mathbb{R}^d, \quad \alpha \in \mathbb{N}^d,$$

where  $C = (rd/\epsilon)^r$ . Therefore (21) is proved.

In the opposite direction, let (21) be satisfied. Using the following inequalities

$$|x|^n \leq \sum_{|\alpha|=n} \frac{n!}{\alpha!} |x^\alpha|, \quad \sum_{|\alpha|=n} \frac{n!}{\alpha!} = d^n, \quad \text{and} \quad \alpha! \leq |\alpha|!,$$

and the assumption (21) we obtain

$$|x|^n |h(x)| \leq \sum_{|\alpha|=n} \frac{n!}{\alpha!} |x^\alpha h(x)| \lesssim \sum_{|\alpha|=n} \frac{n!}{\alpha!} C^{|\alpha|} (\alpha!)^r \leq C^n (n!)^r d^n = (dC)^n (n!)^r,$$

for every  $x \in \mathbb{R}^d$ ,  $n \in \mathbb{N}$ . Therefore the sequence

$$(dC)^{-n} (n!)^{-r} |x|^n |h(x)|, \quad n \in \mathbb{N}$$

is uniformly bounded for  $x \in \mathbb{R}^d$ , as well as the sequence

$$(dC)^{-\frac{n}{r}} (n!)^{-1} |x|^{\frac{n}{r}} |h(x)|^{\frac{1}{r}}, \quad n \in \mathbb{N}.$$

If we choose  $\epsilon = q(dC)^{-\frac{1}{r}}$ , for a fixed  $q \in (0, 1)$ , we conclude

$$e^{\epsilon |x|^{\frac{1}{r}}} |h(x)|^{\frac{1}{r}} = \sum_{n=0}^{\infty} q^n (dC)^{-\frac{n}{r}} (n!)^{-1} |x|^{\frac{n}{r}} |h(x)|^{\frac{1}{r}} \lesssim \sum_{n=0}^{\infty} q^n.$$

This is (22); hence we get (20) and the proof is complete.  $\square$

It follows from this proof the precise relation between the constants  $\epsilon$  and  $C$ . Indeed, assuming (20), then (21) is satisfied with  $C = (rd/\epsilon)^r$ . Viceversa, (21) implies (20) for any  $\epsilon < r(dC)^{-1/r}$ . The bound is sharp for  $d = 1$ . Also, it follows from the proof that the constant implicit in the notation  $\lesssim$  in (20) depends only on the corresponding one in (21) and viceversa.

The strong dual spaces of  $S_r^s(\mathbb{R}^d)$  and  $\Sigma_1^1(\mathbb{R}^d)$  are called spaces of tempered ultra-distributions and denoted by  $(S_r^s)'(\mathbb{R}^d)$  and  $(\Sigma_1^1)'(\mathbb{R}^d)$ , respectively. Notice that they contain the space of tempered distribution  $\mathcal{S}'(\mathbb{R}^d)$ .

The spaces  $S_r^s(\mathbb{R}^d)$  are nuclear spaces [33], and this property yields a corresponding kernel theorem; cf. [47].

**Theorem 2.5.** *There exists an isomorphism between the space of linear continuous maps  $T$  from  $S_r^s(\mathbb{R}^d)$  to  $(S_r^s)'(\mathbb{R}^d)$  and  $(S_r^s)'(\mathbb{R}^{2d})$ , which associates to every  $T$  a kernel  $K_T \in (S_r^s)'(\mathbb{R}^{2d})$  such that*

$$\langle Tu, v \rangle = \langle K_T, v \otimes \bar{u} \rangle, \quad \forall u, v \in S_r^s(\mathbb{R}^d).$$

$K_T$  is called the kernel of  $T$ .

**2.3. Time-frequency representations.** We recall the basic concepts of time-frequency analysis and refer the reader to [22] for the full details. Consider a distribution  $f \in \mathcal{S}'(\mathbb{R}^d)$  and a Schwartz function  $g \in \mathcal{S}(\mathbb{R}^d) \setminus \{0\}$  (the so-called *window*). The short-time Fourier transform (STFT) of  $f$  with respect to  $g$  is  $V_g f(z) = \langle f, \pi(z)g \rangle$ ,  $z = (x, \xi) \in \mathbb{R}^d \times \mathbb{R}^d$ . The short-time Fourier transform is well-defined whenever the bracket  $\langle \cdot, \cdot \rangle$  makes sense for dual pairs of function or (ultra-)distribution spaces, in particular for  $f \in \mathcal{S}'(\mathbb{R}^d)$  and  $g \in \mathcal{S}(\mathbb{R}^d)$ ,  $f, g \in L^2(\mathbb{R}^d)$ ,  $f \in (\Sigma_1^1)'(\mathbb{R}^d)$  and  $g \in \Sigma_1^1(\mathbb{R}^d)$  or  $f \in (S_r^s)'(\mathbb{R}^d)$  and  $g \in S_r^s(\mathbb{R}^d)$ . Let us recall the covariance formula for the short-time Fourier transform that will be used in the sequel

$$(24) \quad V_g(M_\eta T_y f)(x, \xi) = e^{-2\pi i(\xi - \eta)y} (V_g f)(x - y, \xi - \eta), \quad x, y, \xi, \eta \in \mathbb{R}^d.$$

Another time-frequency representation we shall use is the (cross-)Wigner distribution of  $f, g \in L^2(\mathbb{R}^d)$ , defined as

$$(25) \quad W(f, g)(x, \xi) = \int_{\mathbb{R}^d} f\left(x + \frac{t}{2}\right) \overline{g\left(x - \frac{t}{2}\right)} e^{-2\pi i t \xi} dt.$$

If we set  $\check{g}(t) = g(-t)$ , then the relation between cross-Wigner distribution and short-time Fourier transform is provided by

$$(26) \quad W(f, g)(x, \xi) = 2^d e^{4\pi i x \xi} V_{\check{g}} f(2x, 2\xi).$$

For the discrete description of function spaces and operators we use Gabor frames. Let  $\Lambda = A\mathbb{Z}^{2d}$  with  $A \in GL(2d, \mathbb{R})$  be a lattice of the time-frequency plane. The set of time-frequency shifts  $\mathcal{G}(g, \Lambda) = \{\pi(\lambda)g : \lambda \in \Lambda\}$  for a non-zero  $g \in L^2(\mathbb{R}^d)$  is called a Gabor system. The set  $\mathcal{G}(g, \Lambda)$  is a Gabor frame, if there exist constants  $A, B > 0$  such that

$$(27) \quad A\|f\|_2^2 \leq \sum_{\lambda \in \Lambda} |\langle f, \pi(\lambda)g \rangle|^2 \leq B\|f\|_2^2, \quad \forall f \in L^2(\mathbb{R}^d).$$

If (27) is satisfied, then there exists a dual window  $\gamma \in L^2(\mathbb{R}^d)$ , such that  $\mathcal{G}(\gamma, \Lambda)$  is a frame, and every  $f \in L^2(\mathbb{R}^d)$  possesses the frame expansions

$$f = \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)g \rangle \pi(\lambda)\gamma = \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)\gamma \rangle \pi(\lambda)g$$

with unconditional convergence in  $L^2(\mathbb{R}^d)$ . In conclusion, we list some results about time-frequency analysis of Gelfand-Shilov functions, cf. [6, 15, 26, 45]:

$$(28) \quad f, g \in S_s^s(\mathbb{R}^d), \quad s \geq 1/2 \Rightarrow V_g f \in S_s^s(\mathbb{R}^{2d}),$$

$$(29) \quad f, g \in S_s^s(\mathbb{R}^d), \quad s \geq 1/2 \Rightarrow W(f, g) \in S_s^s(\mathbb{R}^{2d}),$$

$$(30) \quad f, g \in \Sigma_1^1(\mathbb{R}^d) \Rightarrow W(f, g) \in \Sigma_1^1(\mathbb{R}^{2d}).$$

Finally, if  $g \in S_s^s(\mathbb{R}^d)$ ,  $s \geq 1/2$ , then

$$(31) \quad f \in S_s^s(\mathbb{R}^d) \iff |V_g(f)(z)| \lesssim \exp(-\epsilon|z|^{1/s}) \text{ for some } \epsilon > 0.$$

**2.4. Modulation Spaces.** Weighted modulation spaces measure the decay of the STFT on the time-frequency (phase space) plane and were introduced by Feichtinger in the 80's [19] for weight of sub-exponential growth at infinity. The study of weights of exponential growth was developed in [15, 45].

*Weight Functions.* In the sequel  $v$  will always be a continuous, positive, even, submultiplicative function (submultiplicative weight), i.e.,  $v(0) = 1$ ,  $v(z) = v(-z)$ , and  $v(z_1 + z_2) \leq v(z_1)v(z_2)$ , for all  $z, z_1, z_2 \in \mathbb{R}^d$ . Submultiplicativity implies that  $v(z)$  is *dominated* by an exponential function, i.e.

$$(32) \quad \exists C, k > 0 \text{ such that } 1 \leq v(z) \leq Ce^{k|z|}, \quad z \in \mathbb{R}^d.$$

For example, every weight of the form

$$(33) \quad v(z) = e^{s|z|^b} (1 + |z|)^a \log^r(e + |z|)$$

for parameters  $a, r, s \geq 0$ ,  $0 \leq b \leq 1$  satisfies the above conditions.

We denote by  $\mathcal{M}_v(\mathbb{R}^d)$  the space of  $v$ -moderate weights on  $\mathbb{R}^d$ ; these are measurable positive functions  $m$  satisfying  $m(z + \zeta) \leq Cv(z)m(\zeta)$  for every  $z, \zeta \in \mathbb{R}^d$ .

**Definition 2.6.** Given  $g \in \Sigma_1^1(\mathbb{R}^d)$ , a weight function  $m \in \mathcal{M}_v(\mathbb{R}^{2d})$ , and  $1 \leq p, q \leq \infty$ , the modulation space  $M_m^{p,q}(\mathbb{R}^d)$  consists of all tempered ultra-distributions  $f \in (\Sigma_1^1)'(\mathbb{R}^d)$  such that  $V_g f \in L_m^{p,q}(\mathbb{R}^{2d})$  (weighted mixed-norm spaces). The norm on  $M_m^{p,q}(\mathbb{R}^d)$  is

$$(34) \quad \|f\|_{M_m^{p,q}} = \|V_g f\|_{L_m^{p,q}} = \left( \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |V_g f(x, \xi)|^p m(x, \xi)^p dx \right)^{q/p} d\xi \right)^{1/q}$$

(obvious changes if  $p = \infty$  or  $q = \infty$ ).

We observe that for  $f, g \in \Sigma_1^1(\mathbb{R}^d)$  the above integral is convergent and thus  $\Sigma_1^1(\mathbb{R}^d) \subset M_m^{p,q}(\mathbb{R}^d)$ ,  $1 \leq p, q \leq \infty$ , cf. [15], with dense inclusion when  $p, q < \infty$ , cf. [6]. When  $p = q$ , we simply write  $M_m^p(\mathbb{R}^d)$  instead of  $M_m^{p,p}(\mathbb{R}^d)$ . The spaces  $M_m^{p,q}(\mathbb{R}^d)$  are Banach spaces and every nonzero  $g \in M_v^1(\mathbb{R}^d)$  yields an equivalent norm in (34) and so  $M_m^{p,q}(\mathbb{R}^d)$  is independent on the choice of  $g \in M_v^1(\mathbb{R}^d)$ .

We also recall the inversion formula for the STFT (see [22, Proposition 11.3.2] and [15, Proposition 2.6] for exponential weights): assume  $g \in M_v^1(\mathbb{R}^d) \setminus \{0\}$ ,  $f \in M_m^{p,q}(\mathbb{R}^d)$ , then

$$(35) \quad f = \frac{1}{\|g\|_2^2} \int_{\mathbb{R}^{2d}} V_g f(x, \xi) M_\xi T_x g dx d\xi,$$

and the equality holds in  $M_m^{p,q}(\mathbb{R}^d)$  (observe that  $M_v^1(\mathbb{R}^d) \subset M^2(\mathbb{R}^d) = L^2(\mathbb{R}^d)$ , so  $\|g\|_2 < \infty$ ).

## 3. GEVREY-ANALYTIC AND ULTRA-ANALYTIC SYMBOLS

In this section we characterize the smoothness and the growth of a function  $f$  on  $\mathbb{R}^d$  in terms of the decay of its STFT  $V_g f$ , for a suitable window  $g$ . In the proofs we shall detail the relations among the constants  $C_f, C_g, C_{f,g}$  and  $\epsilon$  which come into play, since it could be useful for numerical purposes.

**Theorem 3.1.** *Consider  $s > 0$ ,  $m \in \mathcal{M}_v(\mathbb{R}^d)$ ,  $g \in M_{v \otimes 1}^1(\mathbb{R}^d) \setminus \{0\}$  such that for some  $C_g > 0$ ,*

$$(36) \quad \|\partial^\alpha g\|_{L_v^1(\mathbb{R}^d)} \lesssim C_g^{|\alpha|} (\alpha!)^s, \quad \alpha \in \mathbb{N}^d.$$

For  $f \in C^\infty(\mathbb{R}^d)$  the following conditions are equivalent:

(i) *There exists a constant  $C_f > 0$  such that*

$$(37) \quad |\partial^\alpha f(x)| \lesssim m(x) C_f^{|\alpha|} (\alpha!)^s, \quad x \in \mathbb{R}^d, \alpha \in \mathbb{N}^d.$$

(ii) *There exists a constant  $C_{f,g} > 0$  such that*

$$(38) \quad |\xi^\alpha V_g f(x, \xi)| \lesssim m(x) C_{f,g}^{|\alpha|} (\alpha!)^s, \quad (x, \xi) \in \mathbb{R}^{2d}, \alpha \in \mathbb{N}^d.$$

(iii) *There exists a constant  $\epsilon > 0$  such that*

$$(39) \quad |V_g f(x, \xi)| \lesssim m(x) \exp(-\epsilon |\xi|^{1/s}), \quad (x, \xi) \in \mathbb{R}^{2d}, \alpha \in \mathbb{N}^d.$$

If the equivalent conditions (37), (38), (39) are satisfied, we say that  $f$  is a Gevrey symbol when  $s > 1$ , an analytic symbol if  $s = 1$  and an ultra-analytic symbol when  $s < 1$ .

*Proof.* (i)  $\Rightarrow$  (ii). We can use Leibniz' formula and write

$$\partial^\alpha (f T_x g)(t) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \partial^{\alpha-\beta} f(t) T_x(\partial^\beta g)(t), \quad t \in \mathbb{R}^d, \alpha \in \mathbb{N}^d.$$

Let us estimate the  $L^1(\mathbb{R}^d)$ -norm of the products  $\partial^{\alpha-\beta} f T_x(\partial^\beta g)$ . Using the positivity and  $v$ -moderateness of the weight  $m$ , Hölder inequality and the assumptions (37) and (36),

$$\begin{aligned} \|\partial^{\alpha-\beta} f T_x(\partial^\beta g)\|_{L^1} &\leq \|\partial^{\alpha-\beta} f\|_{L_{1/m}^\infty} \|m T_x(\partial^\beta g)\|_{L^1} \\ &\lesssim C_f^{|\alpha-\beta|} ((\alpha-\beta)!)^s \|v(\cdot-x) \partial^\beta g(\cdot-x)\|_{L^1} m(x) \\ &\lesssim C_f^{|\alpha-\beta|} ((\alpha-\beta)!)^s C_g^{|\beta|} (\beta!)^s m(x) \\ &\leq C^{|\alpha|} (\alpha!)^s m(x) \end{aligned}$$

where  $C := \max\{C_f, C_g\}$  and we used  $(\alpha-\beta)! \beta! \leq \alpha!$ . These estimates tell us, in particular, that  $V_g f(x, \xi) = \mathcal{F}(f T_x g)(\xi)$  is well-defined. We can exchange partial derivatives and Fourier transform as follows

$$\xi^\alpha V_g f(x, \xi) = \frac{1}{(2\pi i)^{|\alpha|}} \mathcal{F}(\partial^\alpha (f T_x g))(\xi) = \frac{1}{(2\pi i)^{|\alpha|}} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} \mathcal{F}(\partial^{\alpha-\beta} f T_x(\partial^\beta g))(\xi).$$

Using

$$|\mathcal{F}(\partial^{\alpha-\beta} f T_x(\partial^\beta g))(\xi)| \leq \|\mathcal{F}(\partial^{\alpha-\beta} f T_x(\partial^\beta g))\|_{L^\infty} \leq \|\partial^{\alpha-\beta} f T_x(\partial^\beta g)\|_{L^1},$$

the majorizations above and (16),

$$|\xi^\alpha V_g f(x, \xi)| \lesssim \frac{m(x)}{(2\pi)^{|\alpha|}} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} C^{|\alpha|} (\alpha!)^s = \frac{m(x)}{\pi^{|\alpha|}} C^{|\alpha|} (\alpha!)^s.$$

This proves (38), with constant  $C_{f,g} := \frac{\max\{C_f, C_g\}}{\pi}$ .

(ii)  $\Rightarrow$  (i). We use the inversion formula (35), observing that (38) implies  $f \in M_{m \otimes 1}^{\infty, 1}(\mathbb{R}^d)$ , so that the equality in (35) holds a.e. and we can assume that it holds everywhere since  $f$  is smooth. Let us consider the partial derivatives of  $f$  and exchange them with the integrals, the estimates below will provide the justification of this operation. So, formally, we can write

$$\partial^\alpha f(t) = \frac{1}{\|g\|_2^2} \int_{\mathbb{R}^{2d}} V_g f(x, \xi) \partial^\alpha (M_\xi T_x g)(t) dx d\xi, \quad t \in \mathbb{R}^d.$$

Using Leibniz' formula  $\partial^\alpha (M_\xi T_x g) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (2\pi i \xi)^\beta M_\xi T_x (\partial^{\alpha-\beta} g)$  we estimate

$$|\partial^\alpha f(t)| \leq \frac{1}{\|g\|_2^2} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (2\pi)^{|\beta|} \int_{\mathbb{R}^{2d}} |V_g f(x, \xi) \xi^\beta| \cdot |T_x \partial^{\alpha-\beta} g(t)| dx d\xi, \quad t \in \mathbb{R}^d.$$

We set

$$I_{\alpha, \beta}(t) := \int_{\mathbb{R}^{2d}} |V_g f(x, \xi) \xi^\beta| \cdot |T_x \partial^{\alpha-\beta} g(t)| dx d\xi$$

and prove that, for every fixed  $t \in \mathbb{R}^d$ ,  $I_{\alpha, \beta}(t)$  are absolutely convergent integrals. Using  $1 + |\xi|^{d+1} \leq c_d \sum_{|\gamma| \leq d+1} |\xi^\gamma|$ , where  $c_d$  depends only on the dimension  $d$ , and assumption (38),

$$\begin{aligned} I_{\alpha, \beta}(t) &= \int_{\mathbb{R}^{2d}} |V_g f(x, \xi) \xi^\beta| \frac{1 + |\xi|^{d+1}}{1 + |\xi|^{d+1}} |T_x \partial^{\alpha-\beta} g(t)| dx d\xi \\ &\lesssim \int_{\mathbb{R}^{2d}} \sum_{|\gamma| \leq d+1} |V_g f(x, \xi) \xi^{\beta+\gamma}| \frac{1}{1 + |\xi|^{d+1}} |T_x \partial^{\alpha-\beta} g(t)| dx d\xi \\ &\lesssim \sum_{|\gamma| \leq d+1} C_{f,g}^{|\beta+\gamma|} ((\beta + \gamma)!)^s \left( \int_{\mathbb{R}^d} \frac{1}{1 + |\xi|^{d+1}} d\xi \right) \left( \int_{\mathbb{R}^d} m(x) |T_x \partial^{\alpha-\beta} g(t)| dx \right). \end{aligned}$$

Now,  $(\beta + \gamma)! \leq 2^{|\beta|+|\gamma|} (\beta!) (\gamma!) \leq 2^{d+1} (d+1)! 2^{|\beta|} \beta!$  and  $C_{f,g}^{|\beta+\gamma|} = C_{f,g}^{d+1} C_{f,g}^{|\beta|}$  so, forgetting about the constants depending only on  $d$ ,  $f$  and  $g$ ,

$$\begin{aligned} I_{\alpha, \beta}(t) &\lesssim C_{f,g}^{|\beta|} 2^{s\beta} (\beta!)^s \left( \int_{\mathbb{R}^d} m(x) |T_x \partial^{\alpha-\beta} g(t)| dx \right) \\ &\lesssim C_{f,g}^{|\beta|} 2^{s\beta} (\beta!)^s m(t) \left( \int_{\mathbb{R}^d} v(x-t) |\partial^{\alpha-\beta} g(t-x)| dx \right) \\ &= C_{f,g}^{|\beta|} 2^{s\beta} (\beta!)^s m(t) \|\partial^{\alpha-\beta} g\|_{L_v^1}. \end{aligned}$$

The estimate above and the assumption (36) on  $g$  allow the following majorization:

$$\begin{aligned} |\partial^\alpha f(t)| &\lesssim \frac{m(t)}{\|g\|_2^2} \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} (2\pi)^{|\beta|} (2^s C_{f,g})^{|\beta|} (\beta!)^s C_g^{|\alpha-\beta|} ((\alpha-\beta)!)^s \\ &\leq \frac{m(t)}{\|g\|_2^2} (\alpha!)^s \max\{C_{f,g}, C_g\}^{|\alpha|} 2^{s|\alpha|} (2\pi)^{|\alpha|} 2^{|\alpha|} \\ &\lesssim m(t) C_f^\alpha (\alpha!)^s, \end{aligned}$$

where  $C_f := 2^{s+2}\pi \max\{C_{f,g}, C_g\}$ , and we used  $(\beta)!(\alpha-\beta)! \leq \alpha!$  and (16).

(ii)  $\Leftrightarrow$  (iii). The equivalence is an immediate consequence of Proposition 2.4 and the subsequent remarks on the constants, where, for every fixed  $x \in \mathbb{R}^d$ , we choose  $h(\xi) := V_g f(x, \xi)/m(x)$  and  $r = s$ .  $\square$

A natural question is whether we may find window functions satisfying (36). To this end, we recall the following characterization of Gelfand-Shilov spaces.

**Proposition 3.2.** *Let  $g \in \mathcal{S}(\mathbb{R}^d)$ . We have  $g \in S_r^s(\mathbb{R}^d)$ , with  $s, r > 0$ ,  $r + s \geq 1$ , if and only if there exist constants  $A > 0$ ,  $\epsilon > 0$  such that*

$$|\partial^\alpha g(x)| \lesssim A^{|\alpha|} (\alpha!)^s \exp(-\epsilon|x|^{1/r}), \quad x \in \mathbb{R}^d, \alpha \in \mathbb{N}^d.$$

We have  $g \in \Sigma_1^1(\mathbb{R}^d)$  if and only if, for every  $A > 0$ ,  $\epsilon > 0$ ,

$$|\partial^\alpha g(x)| \lesssim A^{|\alpha|} \alpha! \exp(-\epsilon|x|), \quad x \in \mathbb{R}^d, \alpha \in \mathbb{N}^d.$$

*Proof.* The first part of the statement is in [35, Proposition 6.1.7]. For the second part, the assumption  $g \in \Sigma_1^1(\mathbb{R}^d)$  means that for every  $A > 0, B > 0$ ,

$$|x^\beta \partial^\alpha g(x)| \lesssim A^{|\alpha|} B^{|\beta|} \alpha! \beta!, \quad x \in \mathbb{R}^d, \alpha, \beta \in \mathbb{N}^d.$$

Therefore the function  $h = \frac{\partial^\alpha g}{A^{|\alpha|} \alpha!}$  satisfies (21) in Proposition 2.4 for every  $C > 0$ , and the estimate (20) is then satisfied for every  $\epsilon > 0$  (see the observations after the proof of Proposition 2.4 for the uniformity of the constants). This gives the claim.  $\square$

Hence every  $g \in S_r^s(\mathbb{R}^d)$  with  $s > 0$ ,  $0 < r < 1$ ,  $s + r \geq 1$ , satisfies (36) for every submultiplicative weight  $v$  (see (32)). The same holds true if  $g \in \Sigma_1^1(\mathbb{R}^d)$  and  $s \geq 1$ .

#### 4. ALMOST DIAGONALIZATION FOR PSEUDODIFFERENTIAL OPERATORS

In this section we extend some results of [25] to the case of pseudodifferential operators having ultra-analytic symbols. Our result covers also the Gevrey-analytic case which was already discussed in [25].

The Weyl form  $\sigma^w$  of a pseudodifferential operator (the so-called Weyl operator or Weyl transform) with symbol  $\sigma(x, \xi)$  on  $\mathbb{R}^{2d}$  is formally defined by

$$(40) \quad \sigma^w f(x) = \int_{\mathbb{R}^d} \sigma\left(\frac{x+y}{2}, \xi\right) e^{2\pi i(x-y)\xi} f(y) dy d\xi.$$

Using the Wigner distribution (25), we can recast the definition as follows

$$(41) \quad \langle \sigma^w f, g \rangle = \langle \sigma, W(f, g) \rangle.$$

Since for  $f, g \in \mathcal{S}(\mathbb{R}^d)$  we have  $W(f, g) \in \mathcal{S}(\mathbb{R}^{2d})$ , (41) defines for  $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$  a continuous map  $\sigma^w : \mathcal{S}(\mathbb{R}^d) \rightarrow \mathcal{S}'(\mathbb{R}^d)$ . The Schwartz kernel  $K$  of  $\sigma^w$  is given by

$$(42) \quad K(x, y) = \mathcal{F}_{\xi \rightarrow x-y}^{-1} \sigma \left( \frac{x+y}{2}, \xi \right) \in \mathcal{S}'(\mathbb{R}^{2d}).$$

On the other hand, in view of the kernel theorem in  $\mathcal{S} - \mathcal{S}'$ , every linear continuous map from  $\mathcal{S}(\mathbb{R}^d)$  to  $\mathcal{S}'(\mathbb{R}^{2d})$  can be represented by means of a kernel  $K \in \mathcal{S}'(\mathbb{R}^{2d})$ , hence as in (40), (41) with symbol

$$(43) \quad \sigma(x, \xi) = \mathcal{F}_{y \rightarrow \xi} K \left( x + \frac{y}{2}, x - \frac{y}{2} \right).$$

Same arguments are valid in Gelfand-Shilov classes with  $s = r \geq 1/2$  (see [33, 47]). Namely, considering  $\sigma \in (S_s^s)'(\mathbb{R}^{2d})$ ,  $f, g \in S_s^s(\mathbb{R}^d)$  in (40), (41), we have  $\sigma^w : S_s^s(\mathbb{R}^d) \rightarrow (S_s^s)'(\mathbb{R}^d)$  continuously, in view of (29). The kernel  $K$  of  $\sigma^w$ , given by (42), belongs to  $(S_s^s)'(\mathbb{R}^{2d})$ . In the opposite direction, in view of Theorem 2.5, every linear continuous map from  $S_s^s(\mathbb{R}^d)$  to  $(S_s^s)'(\mathbb{R}^d)$  can be represented in the form (40), (41), with  $\sigma \in (S_s^s)'(\mathbb{R}^{2d})$  given by (43). The same holds for the couple of spaces  $\Sigma_1^1, (\Sigma_1^1)'$ .

The crucial relation between the action of the Weyl operator  $\sigma^w$  on time-frequency shifts and the short-time Fourier transform of its symbol, contained in [23, Lemma 3.1] can now be extended to Gelfand-Shilov spaces and their dual spaces as follows.

**Lemma 4.1.** *Consider  $s \geq 1/2$ ,  $g \in S_s^s(\mathbb{R}^d)$ ,  $\Phi = W(g, g)$ . Then, for  $\sigma \in (S_s^s)'(\mathbb{R}^{2d})$ ,*

$$(44) \quad |\langle \sigma^w \pi(z)g, \pi(w)g \rangle| = \left| V_{\Phi} \sigma \left( \frac{z+w}{2}, j(w-z) \right) \right| = |V_{\Phi} \sigma(u, v)|$$

and

$$(45) \quad |V_{\Phi} \sigma(u, v)| = \left| \langle \sigma^w \pi \left( u - \frac{1}{2}j^{-1}(v) \right) g, \pi \left( u + \frac{1}{2}j^{-1}(v) \right) g \rangle \right|,$$

where  $j(z_1, z_2) = (z_2, -z_1)$ . Moreover, the same results hold true if we replace the space  $S_s^s(\mathbb{R}^d)$  with the space  $\Sigma_1^1(\mathbb{R}^d)$ .

*Proof.* Since  $\Phi = W(g, g) \in S_s^s(\mathbb{R}^{2d})$  for  $g \in S_s^s(\mathbb{R}^d)$  the duality  $\langle \sigma, \pi(u, v)\Phi \rangle_{(S_s^s)' \times S_s^s}$  is well-defined so that the short-time Fourier transform  $V_{\Phi} \sigma(u, v)$  makes sense. The same pattern applies to the case  $g \in \Sigma_1^1(\mathbb{R}^d)$ . The rest of the proof is analogous to [23, Lemma 3.1].  $\square$

We now exhibit a characterization of the operator  $\sigma^w$  in terms of its continuous Gabor matrix. The symbol  $\sigma$  belongs to the classes defined in Section 3, the dimension being now  $2d$ .

**Theorem 4.2.** *Let  $s \geq 1/2$ , and  $m \in \mathcal{M}_v(\mathbb{R}^{2d})$ . If  $1/2 \leq s < 1$  consider a window function  $g \in S_s^s(\mathbb{R}^d)$ ; otherwise, if  $s \geq 1$ , assume either  $g \in \Sigma_1^1(\mathbb{R}^d)$ , or  $g \in S_s^s(\mathbb{R}^d)$  and the following growth condition on the weight  $v$ :*

$$v(z) \lesssim \exp(\epsilon |z|^{1/s}), \quad z \in \mathbb{R}^{2d},$$

for every  $\epsilon > 0$ . Then the following properties are equivalent for  $\sigma \in \mathcal{C}^\infty(\mathbb{R}^{2d})$ :

(i) *The symbol  $\sigma$  satisfies*

$$(46) \quad |\partial^\alpha \sigma(z)| \lesssim m(z) C^{|\alpha|} (\alpha!)^s, \quad \forall z \in \mathbb{R}^{2d}, \forall \alpha \in \mathbb{N}^{2d}.$$

(ii) *There exists  $\epsilon > 0$  such that*

$$(47) \quad |\langle \sigma^w \pi(z)g, \pi(w)g \rangle| \lesssim m \left( \frac{w+z}{2} \right) \exp(-\epsilon |w-z|^{1/s}), \quad \forall z, w \in \mathbb{R}^{2d}.$$

*Proof.* (i)  $\Rightarrow$  (ii). Proposition 3.2 applied to the window  $\Phi = W(g, g)$  in Lemma 4.1, which lives in the space  $S_s^s(\mathbb{R}^{2d})$  since  $g \in S_s^s(\mathbb{R}^d)$  by (29), and the assumptions on  $v$  (recall also (32)) imply that  $\Phi$  satisfies the assumptions of Theorem 3.1. Hence, using the equivalence (37)  $\Leftrightarrow$  (39), the assumption (46) is equivalent to the following decay estimate of the corresponding short-time Fourier transform

$$|V_{\Phi} \sigma(u, v)| \lesssim m(u) e^{-\epsilon |v|^{1/s}}, \quad u, v \in \mathbb{R}^{2d},$$

for a suitable  $\epsilon > 0$ , which combined with (44) yields

$$|V_{\Phi} \sigma \left( \frac{z+w}{2}, j(w-z) \right)| \lesssim m \left( \frac{w+z}{2} \right) e^{-\epsilon |j(w-z)|^{1/s}} = m \left( \frac{w+z}{2} \right) e^{-\epsilon |w-z|^{1/s}}$$

that is (ii).

(ii)  $\Rightarrow$  (i). Relation (45) and the decay assumption (47) give

$$\begin{aligned} |V_{\Phi} \sigma(u, v)| &= \left| \langle \sigma^w \pi \left( u - \frac{1}{2} j^{-1}(v) \right) g, \pi \left( u + \frac{1}{2} j^{-1}(v) \right) g \rangle \right| \\ &\lesssim m(u) e^{-\epsilon |j^{-1}(v)|^{1/s}} = m(u) e^{-\epsilon |v|^{1/s}} \end{aligned}$$

and using the equivalence (37)  $\Leftrightarrow$  (39) we obtain the claim.  $\square$

Of course, from (47) we deduce the discrete Gabor matrix decay in (50) below. The viceversa requires more effort. Indeed, we appeal to a recent result obtained by Gröchenig and Lyubarskii in [24]. They find sufficient conditions on the lattice  $\Lambda = A\mathbb{Z}^2$ ,  $A \in GL(2, \mathbb{R})$ , such that  $g = \sum_{k=0}^n c_k H_k$ , with  $H_k$  Hermite function, forms a so-called Gabor (super)frame  $\mathcal{G}(g, \Lambda)$ . Besides they prove the existence of dual windows  $\gamma$  that belong to the space  $S_{1/2}^{1/2}(\mathbb{R})$  (cf. [24, Lemma 4.4]). This theory transfers to the  $d$ -dimensional case by taking a tensor product  $g = g_1 \otimes \cdots \otimes g_d \in S_{1/2}^{1/2}(\mathbb{R}^d)$  of windows as above, which defines a Gabor frame on the lattice  $\Lambda_1 \times \cdots \times \Lambda_d$  and possesses a dual window  $\gamma = \gamma_1 \otimes \cdots \otimes \gamma_d$  in the same space  $\in S_{1/2}^{1/2}(\mathbb{R}^d)$ . Let us simply call *Gabor super-frame*  $\mathcal{G}(g, \Lambda)$  for  $L^2(\mathbb{R}^d)$  a Gabor frame with the above properties. The Gabor super-frames are the key for the discretization of the kernel in (47). First, we need the preliminary result below, which reflects the algebra property of Gelfand-Shilov spaces. For  $s \geq 1/2$ ,  $\epsilon > 0$ , we define the following weight functions:

$$(48) \quad w_{s,\epsilon}(z) := \exp(-\epsilon |z|^{1/s}), \quad z \in \mathbb{R}^{2d}.$$

**Lemma 4.3.** *Let  $\Lambda \subset \mathbb{R}^{2d}$  be a lattice of  $\mathbb{R}^{2d}$ . Then the sampling  $\{w_{s,\epsilon}(\lambda)\}_{\lambda \in \Lambda}$  of (48) (defined on  $\mathbb{R}^{2d}$ ) satisfies*

$$(49) \quad (w_{s,\epsilon} * w_{s,\epsilon})(\lambda) := \sum_{\nu \in \Lambda} w_{s,\epsilon}(\lambda - \nu) w_{s,\epsilon}(\nu) \lesssim w_{s,\epsilon 2^{-1/s}}(\lambda).$$

*Proof.* We use the arguments of [22, Lemma 11.1.1(c)]. For  $\lambda \in \Lambda$ , we divide the lattice  $\Lambda$  into the subsets  $N_\lambda = \{\nu \in \Lambda : |\lambda - \nu| \leq |\lambda|/2\}$  and  $N_\lambda^c = \{\nu \in \Lambda : |\lambda - \nu| > |\lambda|/2\}$ . For  $\nu \in N_\lambda$ ,  $|\nu| \geq |\lambda|/2$  and  $|\nu|^{1/s} \geq (|\lambda|/2)^{1/s}$ , so

$$(w_{s,\epsilon} * w_{s,\epsilon})(\lambda) \leq e^{-(\epsilon 2^{-\frac{1}{s}})|\lambda|^{\frac{1}{s}}} \left( \sum_{\nu \in N_\lambda} e^{-\epsilon|\lambda-\nu|^{\frac{1}{s}}} + \sum_{\nu \in N_\lambda^c} e^{-\epsilon|\nu|^{\frac{1}{s}}} \right) \lesssim e^{-(\epsilon 2^{-\frac{1}{s}})|\lambda|^{\frac{1}{s}}}.$$

This concludes the proof.  $\square$

**Theorem 4.4.** *Let  $\mathcal{G}(g, \Lambda)$  a Gabor super-frame for  $L^2(\mathbb{R}^d)$ . Consider  $m \in \mathcal{M}_v(\mathbb{R}^{2d})$ ,  $s \geq 1/2$ , and a symbol  $\sigma \in C^\infty(\mathbb{R}^{2d})$ . Then the following properties are equivalent:*

- (i) *There exists  $\epsilon > 0$  such that the estimate (47) holds.*
- (ii) *There exists  $\epsilon > 0$  such that*

$$(50) \quad |\langle \sigma^w \pi(\mu)g, \pi(\lambda)g \rangle| \lesssim m \left( \frac{\lambda + \mu}{2} \right) \exp(-\epsilon|\lambda - \mu|^{1/s}), \quad \forall \lambda, \mu \in \Lambda.$$

*Proof.* It remains to show that (ii)  $\Rightarrow$  (i). The pattern of [23, Theorem 3.2] can be adapted to this proof by using a Gabor super-frame  $\mathcal{G}(g, \Lambda)$ , with a dual window  $\gamma \in S_{1/2}^{1/2}(\mathbb{R}^d)$ .

Let  $\mathcal{Q}$  be a symmetric relatively compact fundamental domain of the lattice  $\Lambda \subset \mathbb{R}^{2d}$ . Given  $w, z \in \mathbb{R}^{2d}$ , we can write them uniquely as  $w = \lambda + u$ ,  $z = \mu + u'$ , for  $\lambda, \mu \in \Lambda$  and  $u, u' \in \mathcal{Q}$ . Using the Gabor reproducing formula for the time-frequency shift  $\pi(u)g \in S_{1/2}^{1/2}(\mathbb{R}^d)$  we can write

$$\pi(u)g = \sum_{\nu \in \Lambda} \langle \pi(u)g, \pi(\nu)\gamma \rangle \pi(\nu)g.$$

Inserting the prior expansions in the assumption (50),

$$(51) \quad \begin{aligned} & |\langle \sigma^w \pi(\mu + u')g, \pi(\lambda + u)g \rangle| \\ & \leq \sum_{\nu, \nu' \in \Lambda} |\langle \sigma^w \pi(\mu + \nu')g, \pi(\lambda + \nu)g \rangle| |\langle \pi(u')g, \pi(\nu')\gamma \rangle| |\langle \pi(u)g, \pi(\nu)\gamma \rangle| \\ & \lesssim \sum_{\nu, \nu' \in \Lambda} m \left( \frac{\lambda + \mu + \nu + \nu'}{2} \right) e^{-\epsilon|\lambda + \nu - \mu - \nu'|^{1/s}} |V_\gamma g(\nu' - u')| |V_\gamma g(\nu - u)|. \end{aligned}$$

Since the window functions  $g, \gamma$  are both in  $S_{1/2}^{1/2}(\mathbb{R}^d)$ , the STFT  $V_\gamma g$  is in  $S_{1/2}^{1/2}(\mathbb{R}^{2d})$  in view of (28). Thus there exists  $h > 0$  such that  $|V_\gamma g(z)| \lesssim e^{-h|z|^2}$ , for every  $z \in \mathbb{R}^{2d}$ . In particular, being  $\mathcal{Q}$  relatively compact and  $u \in \mathcal{Q}$ ,

$$|V_\gamma g(\nu - u)| \lesssim e^{-h|\nu - u|^2} \leq \sup_{u \in \mathcal{Q}} e^{-h|\nu - u|^2} \lesssim e^{-h|\nu|^2}.$$

The assumption  $m \in \mathcal{M}_v(\mathbb{R}^{2d})$  yields

$$m \left( \frac{\lambda + \mu + \nu + \nu'}{2} \right) \lesssim m \left( \frac{\lambda + \mu}{2} \right) v \left( \frac{\nu}{2} \right) v \left( \frac{\nu'}{2} \right)$$

and, for every  $0 < \tilde{h} < h$ ,

$$v\left(\frac{\nu}{2}\right) e^{-h|\nu|^2} \lesssim e^{-\tilde{h}|\nu|^2}, \quad \forall \nu \in \Lambda.$$

Inserting these estimates in (51),

$$\begin{aligned} & |\langle \sigma^w \pi(\mu + u')g, \pi(\lambda + u)g \rangle| \\ & \lesssim m\left(\frac{\lambda + \mu}{2}\right) \sum_{\nu, \nu' \in \Lambda} e^{-\epsilon|\lambda + \nu - \mu - \nu'|^{\frac{1}{s}}} v\left(\frac{\nu}{2}\right) e^{-h|\nu|^2} v\left(\frac{\nu'}{2}\right) e^{-h|\nu'|^2} \\ & \lesssim m\left(\frac{\lambda + \mu}{2}\right) \sum_{\nu, \nu' \in \Lambda} e^{-\epsilon|\lambda + \nu - \mu - \nu'|^{\frac{1}{s}}} e^{-\tilde{h}|\nu|^2} e^{-\tilde{h}|\nu'|^2} \\ (52) \quad & \leq m\left(\frac{\lambda + \mu}{2}\right) \sum_{\nu, \nu' \in \Lambda} e^{-b(|\lambda + \nu - \mu - \nu'|^{\frac{1}{s}} + |\nu|^{\frac{1}{s}} + |\nu'|^{\frac{1}{s}})} \end{aligned}$$

with  $b = \epsilon$  for  $s > 1/2$  whereas  $b = \min\{\epsilon, \tilde{h}\}$  for  $s = 1/2$  (there may be a loss of decay). We observe that the row (52) can be rewritten as

$$m\left(\frac{\lambda + \mu}{2}\right) (w_{s,b} * w_{s,b} * w_{s,b})(\lambda - \mu),$$

where  $w_{s,b}(\lambda) = e^{-b|\lambda|^{1/s}}$ . Now we apply Lemma 4.3 twice and we obtain

$$(53) \quad (52) \lesssim m\left(\frac{\lambda + \mu}{2}\right) e^{-\tilde{\epsilon}|\lambda - \mu|^{\frac{1}{s}}}$$

with  $\tilde{\epsilon} = b2^{-2/s}$ .

If  $w, z \in \mathbb{R}^{2d}$  and  $w = \lambda + u$ ,  $z = \mu + u'$ ,  $\lambda, \mu \in \Lambda$ ,  $u, u' \in \mathcal{Q}$ , then  $\lambda - \mu = w - z + u' - u$  and  $u' - u \in \mathcal{Q} - \mathcal{Q}$ , which is a relatively compact set, thus

$$(54) \quad e^{-\tilde{\epsilon}|\lambda - \mu|^{\frac{1}{s}}} \lesssim \sup_{u \in \mathcal{Q} - \mathcal{Q}} e^{-\tilde{\epsilon}|w - z + u|^{\frac{1}{s}}} \lesssim e^{-\tilde{\epsilon}|w - z|^{\frac{1}{s}}}.$$

Finally, the  $v$ -moderateness of the weight  $m \in \mathcal{M}_v(\mathbb{R}^{2d})$ , together with the fact that  $v$  is continuous and the set  $\mathcal{Q} + \mathcal{Q}$  is relatively compact, let us write

$$(55) \quad \begin{aligned} m\left(\frac{\lambda + \nu}{2}\right) &= m\left(\frac{w + z}{2} - \frac{u + u'}{2}\right) \lesssim m\left(\frac{w + z}{2}\right) v\left(-\frac{u + u'}{2}\right) \\ &\lesssim m\left(\frac{w + z}{2}\right) \sup_{u \in \mathcal{Q} + \mathcal{Q}} v\left(\frac{u}{2}\right) \lesssim m\left(\frac{w + z}{2}\right) \end{aligned}$$

Combining the estimates (54) and (55) with (53) we obtain (47), with the parameter  $\epsilon = \tilde{\epsilon}$  which appears in (53).  $\square$

**4.1. Sparsity of the Gabor matrix.** The operators  $\sigma^w$  which satisfy Theorem 4.2, say with  $m = v = 1$ , enjoy a fundamental sparsity property. Indeed, let  $\mathcal{G}(g, \Lambda)$  be a Gabor frame for  $L^2(\mathbb{R}^d)$ , with  $g \in S_s^s(\mathbb{R}^d)$ ,  $s \geq 1/2$ . Then, as we saw,

$$(56) \quad |\langle \sigma^w \pi(\mu)g, \pi(\lambda)g \rangle| \leq C \exp(-\epsilon|\lambda - \mu|^{1/s}), \quad \forall \lambda, \mu \in \Lambda,$$

with suitable constants  $C > 0$ ,  $\epsilon > 0$ . This gives at once an exponential-type sparsity, in the sense precised by the following proposition (cf. [4, 27] for the more standard notion of super-polynomial sparsity).

**Proposition 4.5.** *Let the Gabor matrix  $\langle \sigma^w \pi(\mu)g, \pi(\lambda)g \rangle$  satisfy (56). Then it is sparse in the following sense. Let  $a$  be any column or row of the matrix, and let  $|a|_n$  be the  $n$ -largest entry of the sequence  $a$ . Then,  $|a|_n$  satisfies*

$$|a|_n \leq C \exp(-\epsilon n^{1/(2ds)}), \quad n \in \mathbb{N}$$

for some constants  $C > 0$ ,  $\epsilon > 0$ .

*Proof.* By a discrete analog of Proposition 2.4 it suffices to prove that

$$n^\alpha |a|_n \leq C^{\alpha+1} (\alpha!)^{2ds}, \quad \alpha \in \mathbb{N}.$$

On the other hand we have

$$n^{\frac{1}{p}} \cdot |a|_n \leq \|a\|_{\ell^p},$$

for every  $0 < p \leq \infty$ . Hence by (56) and setting  $p = 1/\alpha$  we obtain

$$n^\alpha |a|_n \leq \left( \sum_{\lambda \in \Lambda} e^{-\epsilon p |\lambda - \mu|^{\frac{1}{s}}} \right)^{\frac{1}{p}} = \left( \sum_{\lambda \in \Lambda} e^{-\epsilon p |\lambda|^{\frac{1}{s}}} \right)^{\frac{1}{p}}.$$

Let  $\mathcal{Q}$  be a fundamental domain of the lattice  $\Lambda$ . Then if  $x \in \lambda + \mathcal{Q}$ ,  $\lambda \in \Lambda$ , we have  $|x| \leq |\lambda| + C_0$ , therefore  $|x|^{1/s} \leq C_1(|\lambda|^{1/s} + 1)$ . Hence

$$\begin{aligned} \sum_{\lambda \in \Lambda} e^{-\epsilon p |\lambda|^{\frac{1}{s}}} &\leq C_2 \int_{\mathbb{R}^{2d}} e^{-\epsilon p |x|^{\frac{1}{s}}} dx = \int_{\mathbb{S}^{2d-1}} d\sigma \int_0^{+\infty} e^{-\epsilon p \rho^{\frac{1}{s}}} \rho^{2d-1} d\rho \\ &= \frac{C_3 s}{(\epsilon p)^{2ds}} \int_0^{+\infty} e^{-t} t^{2ds-1} dt = \frac{C_3 s \Gamma(2ds)}{(\epsilon p)^{2ds}} = \frac{C_4}{p^{2ds}} \end{aligned}$$

Finally, by Stirling's formula,

$$n^\alpha |a|_n \leq \frac{C_4^{1/p}}{p^{\frac{2ds}{p}}} \leq C_5^{\alpha+1} (\alpha!)^{2ds}.$$

□

**4.2. Boundedness of pseudodifferential operators on modulation and Gelfand-Shilov spaces.** We list here some continuity results on modulation and Gelfand-Shilov spaces which follow easily from our basic Theorem 4.2.

**Proposition 4.6.** *Let  $s \geq 1/2$  and consider a symbol  $\sigma \in C^\infty(\mathbb{R}^{2d})$  satisfying the estimates*

$$(57) \quad |\partial^\alpha \sigma(z)| \lesssim C^{|\alpha|} (\alpha!)^s,$$

for some  $C > 0$ . Let  $m \in \mathcal{M}_v(\mathbb{R}^{2d})$  and, if  $s \geq 1$ , assume the weight  $v$  satisfies

$$v(z) \lesssim \exp(\epsilon |z|^{1/s}), \quad z \in \mathbb{R}^{2d}$$

for every  $\epsilon > 0$ . Then the Weyl operator  $\sigma^w$  extends to a bounded operator on  $\mathcal{M}_m^{p,q}(\mathbb{R}^d)$ .

*Proof.* Let  $g \in \Sigma_1^1(\mathbb{R}^d)$  with  $\|g\|_{L^2} = 1$ . From the inversion formula (35),

$$V_g(\sigma^w f)(u) = \int_{\mathbb{R}^{2d}} \langle \sigma^w \pi(z)g, \pi(u)g \rangle V_g f(z) dz.$$

The desired result thus follows if we can prove that the map  $M(\sigma)$  defined by

$$M(\sigma)G(u) = \int_{\mathbb{R}^{2d}} \langle \sigma^w \pi(z)g, \pi(u)g \rangle G(z) dz$$

is continuous from  $L_m^{p,q}(\mathbb{R}^{2d})$  into  $L_m^{p,q}(\mathbb{R}^{2d})$ . The characterization of Theorem 4.2 assures the existence of an  $\epsilon > 0$ , such that

$$|V_g(\sigma^w f)| = |M(\sigma)V_g(f)| \lesssim (w_{s,\epsilon} * |V_g f|)(u), \quad u \in \mathbb{R}^{2d}$$

where the weight function  $w_{s,\epsilon}$  on  $\mathbb{R}^{2d}$  is defined in (48). The desired conclusion then follows from the relation  $L_m^{p,q} * L_v^1 \hookrightarrow L_m^{p,q}$ , for  $w_{s,\epsilon} \in L_v^1$  by the growth assumption on  $v$  (if  $s < 1$ , (32) suffices).  $\square$

**Proposition 4.7.** *Let  $s \geq 1/2$ , and consider a symbol  $\sigma \in \mathcal{C}^\infty(\mathbb{R}^{2d})$  that satisfies (57). Then the Weyl operator  $\sigma^w$  is bounded on  $S_s^s(\mathbb{R}^d)$ .*

*Proof.* Fix a window function  $g \in S_{1/2}^{1/2}(\mathbb{R}^d)$ . For  $f \in S_s^s(\mathbb{R}^d)$  we have  $V_g f \in S_s^s(\mathbb{R}^{2d})$  in view of (28). Hence there exists a constant  $h > 0$  such that  $|V_g f|(z) \lesssim e^{-h|z|^{1/s}}$ , for every  $z \in \mathbb{R}^{2d}$ . Thus, taking  $\tilde{\epsilon} = \min\{\epsilon, h\}$  and using a continuous version of Lemma 4.3, for every  $u \in \mathbb{R}^{2d}$ ,

$$|V_g(\sigma^w f)(u)| \lesssim (w_{s,\epsilon} * w_{s,h})(u) \lesssim (w_{s,\tilde{\epsilon}} * w_{s,\tilde{\epsilon}})(u) \lesssim w_{s,\tilde{\epsilon}2^{-1/s}}(u).$$

Using (31) we obtain the claim.  $\square$

Boundedness results for pseudodifferential operators on Gelfand-Shilov spaces are also contained in [46].

## 5. APPLICATIONS TO EVOLUTION EQUATIONS

In this section we apply the above almost diagonalization result to the propagators for certain constant coefficient evolution equations.

We consider an operator

$$(58) \quad P(\partial_t, D_x) = \partial_t^m + \sum_{k=1}^m a_k(D_x) \partial_t^{m-k}, \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^d,$$

where  $a_k(\xi)$ ,  $1 \leq k \leq m$ , are polynomials. They may be non-homogeneous, and their degree may be arbitrary. Here we set  $D_{x_j} = \frac{1}{2\pi i} \partial_{x_j}$ ,  $j = 1, \dots, d$ .

We are interested in the forward Cauchy problem

$$\begin{cases} P(\partial_t, D_x)u = 0, & (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d \\ \partial_t^k u(0, x) = u_k(x), & 0 \leq k \leq m-1, \end{cases}$$

where  $u_k \in \mathcal{S}(\mathbb{R}^d)$ ,  $0 \leq k \leq m-1$ . Hence we suppose that the *forward Hadamard-Petrowsky condition* is satisfied: *There exists a constant  $C > 0$  such that*

$$(59) \quad (\tau, \xi) \in \mathbb{C} \times \mathbb{R}^d, \quad P(i\tau, \xi) = 0 \implies \text{Im } \tau \geq -C.$$

This is a sufficient and necessary condition for the above Cauchy problem with Schwartz data to be well posed [38, Section 3.10]. By taking the Fourier transform with respect to  $x$  and using classical results for the fundamental solution to ordinary differential operators [41, pp. 126-127], one sees that the solution is then given by

$$u(t, x) = \sum_{k=0}^{m-1} \partial_t^k E(t, \cdot) * \left( u_{m-1-k} + \sum_{j=1}^{m-k-1} a_j(D_x) u_{m-k-1-j} \right).$$

Here  $E(t, x) = \mathcal{F}_{\xi \rightarrow x}^{-1} \sigma(t, \xi)$ , where  $\sigma(t, \xi)$  is the unique solution to

$$\left( \partial_t^m + \sum_{k=1}^m a_k(\xi) \partial_t^{m-k} \right) \sigma(t, \xi) = \delta(t)$$

supported in  $[0, +\infty) \times \mathbb{R}^d$ . The distribution  $E(t, x)$  is therefore the fundamental solution of  $P$  supported in  $[0, +\infty) \times \mathbb{R}^d$ .

The study of the Cauchy problem is therefore reduced to that of the Fourier multiplier

$$(60) \quad \sigma^w(t, D_x) = \sigma(t, D_x) f = \mathcal{F}^{-1} \sigma(t, \cdot) \mathcal{F} f = E(t, \cdot) * f.$$

(For Fourier multipliers the Weyl and Kohn-Nirenberg quantizations give the same operator).

**Example 5.1.** Here are some classical operators and the symbols  $\sigma(t, \xi)$  of the corresponding propagators for  $t \geq 0$  ( $\sigma(t, \xi) = 0$  for  $t < 0$ ):

$$\text{Wave operator } \partial_t^2 - \Delta; \quad \sigma(t, \xi) = \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}.$$

$$\text{Klein-Gordon operator } \partial_t^2 - \Delta + m^2, \text{ with } m > 0; \quad \sigma(t, \xi) = \frac{\sin(t\sqrt{4\pi^2|\xi|^2+m^2})}{\sqrt{4\pi^2|\xi|^2+m^2}}.$$

$$\text{Heat operator } \partial_t - \Delta; \quad \sigma(t, \xi) = e^{-4\pi^2|\xi|^2 t}.$$

We now introduce an assumption under which the multiplier  $\sigma(t, D_x)$  falls in the class of pseudodifferential operators considered above.

*Assume that there are constants  $C > 0$ ,  $\nu \geq 1$  such that*

$$(61) \quad (\tau, \zeta) \in \mathbb{C} \times \mathbb{C}^d, \quad P(i\tau, \zeta) = 0 \implies \text{Im } \tau \geq -C(1 + |\text{Im } \zeta|)^\nu.$$

It is clear that this condition is stronger than the forward Hadamard-Petrowsky condition.

**Theorem 5.2.** *Assume  $P$  satisfies (61) for some  $C > 0$ ,  $\nu \geq 1$ . Then the symbol  $\sigma(t, \xi)$  of the corresponding propagator  $\sigma(t, D_x)$  in (60) satisfies the following estimates:*

$$(62) \quad |\partial_\xi^\alpha \sigma(t, \xi)| \leq C^{(t+1)|\alpha|+t} (\alpha!)^s, \quad \xi \in \mathbb{R}^d, \quad t \geq 0, \quad \alpha \in \mathbb{N}^d,$$

with  $s = 1 - 1/\nu$ , for a new constant  $C > 0$ .

*Proof.* It is well known (see e.g. [5, Proposition 1.3.2, Lemma 1.3.3], [38, Section 3.10]) that  $\sigma(t, \xi)$  extends to an entire analytic function in the second variable, and that the estimates (61) imply the bound

$$(63) \quad |\sigma(t, \zeta)| \leq e^{Ct(1+|\operatorname{Im} \zeta|)^\nu}, \quad \zeta \in \mathbb{C}^d.$$

Now, given  $\xi \in \mathbb{R}^d$ , we consider the polydisk  $B(\xi, R) = \prod_{j=1}^d B_j(\xi_j, R) = \{\zeta \in \mathbb{C}^d : |\zeta_j - \xi_j| \leq R, 1 \leq j \leq d\}$ , with  $R = (1 + |\alpha|)^{1/\nu}$ ,  $\alpha \in \mathbb{N}^d$ . Observe that (63) implies

$$(64) \quad \sup_{\zeta \in B(\xi, R)} |\sigma(t, \zeta)| \leq e^{Ct(1+\sqrt{d}R)^\nu}.$$

The Cauchy's Generalized Integral Formula

$$\partial_\xi^\alpha \sigma(t, \xi) = \frac{\alpha!}{(2\pi i)^d} \int \cdots \int_{\partial B_1(\xi_1, R) \times \cdots \times \partial B_d(\xi_d, R)} \frac{\sigma(t, \zeta_1, \dots, \zeta_d)}{(\zeta_1 - \xi_1)^{\alpha_1+1} \cdots (\zeta_d - \xi_d)^{\alpha_d+1}} d\zeta_1 \cdots d\zeta_d$$

and the estimate (64) yield

$$(65) \quad |\partial_\xi^\alpha \sigma(t, \xi)| \leq \frac{\alpha! e^{Ct(1+\sqrt{d}R)^\nu}}{R^{|\alpha|}} \leq C_1^{t(|\alpha|+1)} \frac{\alpha!}{(1 + |\alpha|)^{|\alpha|/\nu}}.$$

Using Stirling formula and  $1/\nu = 1 - s$ , we have

$$\frac{1}{(1 + |\alpha|)^{|\alpha|/\nu}} \leq \frac{C_2^{|\alpha|}}{(\alpha!)^{1-s}},$$

which combined with (65) provides the desired majorization (62).  $\square$

Observe that the assumption  $\nu \geq 1$  in the above theorem implies  $0 \leq s < 1$ .

Combining Theorem 5.2 and Theorem 4.2 we obtain at once our main application.

**Theorem 5.3.** *Assume  $P$  satisfies (61) for some  $C > 0$ ,  $\nu \geq 1$ , and set  $r = \min\{2, \nu/(\nu - 1)\}$ . If  $g \in S_{1/r}^{1/r}(\mathbb{R}^d)$  then  $\sigma(t, D_x)$  in (60) satisfies*

$$(66) \quad |\langle \sigma(t, D_x) \pi(z)g, \pi(w)g \rangle| \leq C \exp(-\epsilon|w - z|^r), \quad \forall z, w \in \mathbb{R}^{2d},$$

for some  $\epsilon > 0$  and for a new constant  $C > 0$ . The constants  $\epsilon$  and  $C$  are uniform when  $t$  lies in bounded subsets of  $[0, +\infty)$ .

Notice that in (66) we have  $r > 1$ , so that we always obtain super-exponential decay.

We now show that, if  $P(\partial_t, D_x)$  is any hyperbolic operator, then (61) is satisfied with  $\nu = 1$ , and hence the above theorem applies with Gaussian decay ( $r = 2$  in (66)), for windows  $g \in S_{1/2}^{1/2}(\mathbb{R}^d)$ .

We recall that the operator  $P(\partial_t, D_x)$  is called hyperbolic with respect to  $t$  if the direction  $N = (1, 0, \dots, 0) \in \mathbb{R} \times \mathbb{R}^d$  is non characteristic for  $P$  (i.e. its principal symbol – the higher order homogeneous part in the symbol – does not vanish at  $N$ ), and  $P$  satisfies the forward Hadamard-Petrowsky condition (59). It follows then that the operators  $a_k(D_x)$  in (58) must have degree  $\leq k$  and  $P$  has order  $m$ .

The wave and Klein-Gordon operators are of course the most important examples of hyperbolic operators. We emphasize, however, that  $P$  is *not* required to be strictly hyperbolic, namely the roots of the principal symbol are allowed to coincide. For

example, the operator  $P = \partial_t^2 - \sum_{j,k=1}^d a_{j,k} \partial_{x_j} \partial_{x_k}$ , is hyperbolic if the matrix  $a_{j,k}$  is real, symmetric and positive *semi*-definite.

**Proposition 5.4.** *Assume  $P(\partial_t, D_x)$  is hyperbolic with respect to  $t$ . Then the condition (61) is satisfied with  $\nu = 1$  for some  $C > 0$ , and hence*

$$(67) \quad |\langle \sigma(t, D_x) \pi(z)g, \pi(w)g \rangle| \leq C \exp(-\epsilon|w - z|^2), \quad \forall z, w \in \mathbb{R}^{2d},$$

if  $g \in S_{1/2}^{1/2}(\mathbb{R}^d)$ , for some  $\epsilon > 0$  and for a new constant  $C > 0$ .

*Proof.* Denote by  $P_m$  the principal symbol of  $P$  and  $\Gamma(P, N)$  the component of  $N$  in  $\{\theta \in \mathbb{R} \times \mathbb{R}^d : P_m(\theta) \neq 0\}$ . It follows from the hyperbolicity assumption (see e.g. [28, Proposition 12.4.4]) that the symbol  $Q(\tau, \xi) = P(i\tau, \xi)$  of  $P(\partial_t, D_x)$  satisfies

$$Q(\Xi + \tau N + \sigma\theta) \neq 0 \quad \text{if } \Xi \in \mathbb{R}^{d+1}, \text{ Im } \tau < -C, \text{ Im } \sigma \leq 0, \theta \in \Gamma(P, N),$$

for the same constant  $C$  which appears in (59). We then deduce that

$$(\tau, \zeta) \in \mathbb{C} \times \mathbb{C}^d, \quad Q(\tau, \zeta) = P(i\tau, \zeta) = 0 \implies (\text{Im } \tau, \text{Im } \zeta) \notin -CN - \Gamma(P, N),$$

Indeed, if  $(\tau, \zeta) = \Xi + i(-CN - \theta)$ , with  $\theta \in \Gamma(P, N)$ ,  $\Xi \in \mathbb{R}^{d+1}$ , then  $\theta - \epsilon N \in \Gamma(P, N)$  is  $\epsilon$  is small enough because  $\Gamma(P, N)$  is open, and then

$$P(\tau, \zeta) = P(\Xi - i(C + \epsilon)N - i(\theta - \epsilon N)) \neq 0.$$

Hence  $-(C + \text{Im } \tau, \text{Im } \zeta) \notin \Gamma(P, N)$ , and since the cone  $\Gamma(P, N)$  is open, this implies

$$\frac{-(C + \text{Im } \tau, \text{Im } \zeta) \cdot N}{|(C + \text{Im } \tau, \text{Im } \zeta)|} = \frac{-(C + \text{Im } \tau)}{|(C + \text{Im } \tau, \text{Im } \zeta)|} \leq C_1,$$

for some constant  $C_1 < 1$ , which gives

$$\text{Im } \tau \geq -C - \frac{C_1}{1 - C_1} |\text{Im } \zeta|.$$

□

**Example 5.5.** Consider the wave operator  $P = \partial_t^2 - \Delta$  in  $\mathbb{R} \times \mathbb{R}^d$ , hence  $\sigma(t, \xi) = \frac{\sin(2\pi|\xi|t)}{2\pi|\xi|}$ . Using the expression for its forward fundamental solution we can estimate directly the matrix decay. Consider, for simplicity, the case of dimension  $d \leq 3$ . We take  $g(x) = 2^{d/4} e^{-\pi|x|^2}$  as window function, which is allowed because  $g \in S_{1/2}^{1/2}(\mathbb{R}^d)$ , and moreover  $\|g\|_{L^2} = 1$  (Gaussian functions minimize the Heisenberg uncertainty so that they are, generally speaking, a natural choice for wave-packet decompositions). We claim that

$$|\langle \sigma(t, D_x) M_\xi T_x g, M_{\xi'} T_{x'} g \rangle| \leq t e^{-\frac{\pi}{2} [|\xi' - \xi|^2 + (|x' - x| - t)_+^2]}, \quad x, x', \xi, \xi' \in \mathbb{R}^d, \quad d \leq 3,$$

where  $(\cdot)_+$  denote positive part. This agrees with Proposition 5.4, with the constants made explicit.

In fact, in dimension  $d \leq 3$  we know that

$$\sigma(t, D_x) f(x) = \int f(x - y) d\mu_t(y)$$

where  $\mu_t$  is a positive Borel measure, supported in the ball  $B(0, t)$ , with total mass =  $t$  (see e.g. [38, Chapter 4]). To be precise, we have  $d\mu_t(y) = (1/2)\chi_{[-t, t]} dy$  in dimension

1,  $d\mu_t(y) = (2\pi)^{-1}(t^2 - |y|^2)_+^{-1/2}dy$  in dimension 2, and  $d\mu_t = (4\pi t)^{-1}d\sigma_{\partial B(0,t)}$  in dimension 3 (surface measure).

Hence, using  $T_y M_\xi = e^{-2\pi i y \xi} M_\xi T_y$ ,

$$\langle \sigma(t, D_x) M_\xi T_x g, M_{\xi'} T_{x'} g \rangle = \int e^{-2\pi i y \xi} \langle M_\xi T_{x+y} g, M_{\xi'} T_{x'} g \rangle d\mu_t(y).$$

An explicit computation shows that

$$|\langle \pi(z)g, \pi(w)g \rangle| = e^{-\frac{\pi}{2}|w-z|^2}, \quad z, w \in \mathbb{R}^{2d},$$

so that

$$\begin{aligned} |\langle \sigma(t, D_x) M_\xi T_x g, M_{\xi'} T_{x'} g \rangle| &\leq \int e^{-\frac{\pi}{2}[|\xi' - \xi|^2 + |x' - x - y|^2]} d\mu_t(y) \\ &\leq t e^{-\frac{\pi}{2}[|\xi' - \xi|^2 + (|x' - x| - t)_+^2]}, \end{aligned}$$

which gives the claim.

Consider now the Gabor frame  $\mathcal{G}(g, \Lambda)$ , with  $g(x) = 2^{d/4}e^{-\pi|x|^2}$ ,  $\Lambda = \mathbb{Z}^d \times (1/2)\mathbb{Z}^d$  ([22, Theorem 7.5.3]), and the Gabor matrix

$$T_{m', n', m, n} = \langle \sigma(t, D_x) M_n T_m g, M_{n'} T_{m'} g \rangle, \quad (m, n), (m', n') \in \Lambda.$$

We therefore have

$$|T_{m', n', m, n}| \leq \tilde{T}_{m', n', m, n} := t e^{-\frac{\pi}{2}[|n' - n|^2 + (|m' - m| - t)_+^2]}, \quad (m, n), (m', n') \in \Lambda, \quad d \leq 3.$$

We now present a class of examples of operators which satisfy Theorem 5.3 and are not hyperbolic.

**Example 5.6.** Consider the operator

$$(68) \quad P(\partial_t, D_x) = \partial_t + (-\Delta)^k,$$

with  $k \geq 1$  integer. When  $k = 1$  we have the heat operator. Its symbol is the polynomial

$$P(i\tau, \zeta) = i\tau + (4\pi^2 \zeta^2)^k,$$

We claim that it satisfies (61) with  $\nu = 2k$ .

If  $\tau \in \mathbb{C}$ ,  $\xi, \eta \in \mathbb{R}^d$  and  $P(i\tau, \xi + i\eta) = 0$  then

$$\operatorname{Im} \tau = (2\pi)^{2k} \operatorname{Re}(|\xi|^2 + 2i\xi \cdot \eta - |\eta|^2)^k = (2\pi)^{2k} |\xi|^{2k} + Q(\xi, \eta),$$

where  $Q(\xi, \eta)$  is a homogeneous polynomial of degree  $2k$ , with  $Q(\xi, 0) = 0$ . Hence, for every  $\epsilon > 0$  and some constant  $C_\epsilon > 0$ ,

$$|Q(\xi, \eta)| \lesssim \sum_{j+l=2k, j \geq 0, l \geq 1} |\xi|^j |\eta|^l \leq \epsilon |\xi|^{2k} + C_\epsilon |\eta|^{2k},$$

where we applied the inequality  $ab \leq \frac{(\epsilon a)^p}{p} + \frac{(b/\epsilon)^q}{q}$  (with  $a = |\xi|^j$ ,  $b = |\eta|^l$ ,  $p = 2k/j$  and  $q = 2k/l$ ) to the terms of the sum with  $j \geq 1$ . Taking  $\epsilon = (2\pi)^{2k}/2$  we get

$$\operatorname{Im} \tau \geq \frac{(2\pi)^{2k}}{2} |\xi|^{2k} - C |\eta|^{2k} \geq -C |\eta|^{2k},$$

which proves the claim.

Hence, for the operator  $P$  in (68), Theorem 5.3 applies with  $\nu = 2k$  and  $r = 2k/(2k - 1)$ .

**Remark 5.7.** Let us add a few words about the Gabor analysis of hyperbolic problems with variable coefficients. In this regards, the example of the transport equation with variable coefficients, standard test for numerical schemes, is somewhat discouraging. In fact, for a fixed time  $t > 0$ , the operator  $T$  solving the Cauchy problem turns out to be a change of variables, and Gabor frames are not suited to follow the corresponding non-linear propagation of singularities. Heuristically this can be seen as follows. Decomposition of a function in terms of Gabor atoms  $\pi(\lambda)g$  corresponds, geometrically, to a uniform partition of the time-frequency space (or phase space, in PDE's terminology) into boxes, each atom occupying a box, loosely speaking ([22, pag 211]). The correspondence principle in Quantum Mechanics suggests one to follow the PDE's evolution in terms of its classical analog, namely the Hamiltonian flow in phase space. The propagator has therefore a sparse Gabor matrix if its Hamiltonian flow moves the above mentioned boxes, but introduces *only a controlled number of overlaps*. However, this is not the case for changes of variables, as one sees easily by direct inspection. More rigorously, it follows from the results in [12, 14] that Hörmander-type Fourier integral operators, in particular the changes of variables, do not have a sparse Gabor representation.

However, Theorems 4.2 and 4.4 consent also applications to certain equations with variable coefficients. As elementary example consider the transport equation

$$\begin{cases} \partial_t u - i \left( \sum_{j=1}^d a_j D_{x_j} + \sum_{j=1}^d b_j x_j \right) u = 0 \\ u(0, x) = u_0(x), \end{cases}$$

with  $a_j, b_j \in \mathbb{R}$ . The fundamental solution of the problem has symbol

$$\sigma(t, x, \xi) = \exp \left[ it \sum_{j=1}^d (a_j \xi_j + b_j x_j) \right]$$

satisfying (46) with  $m = 1$ ,  $s = 0$ , and the preceding arguments apply. More generally, it seems possible to treat similarly the pseudodifferential problems of the form

$$\partial_t u - ia^w(x, D)u = 0, \quad u(0, x) = u_0(x),$$

where  $a(x, \xi)$  is real-valued and  $\partial_{x, \xi} a(x, \xi)$ , as well as higher order derivatives, are bounded in  $\mathbb{R}^{2d}$ , satisfying suitable estimates.

#### ACKNOWLEDGEMENTS

The authors would like to thank Prof. Karlheinz Gröchenig for his inspiring comments about Gabor frames of Hermite functions.

#### REFERENCES

- [1] K. Asada and D. Fujiwara. On some oscillatory integral transformations in  $L^2(\mathbf{R}^n)$ . *Japan. J. Math. (N.S.)*, 4(2):299–361, 1978.
- [2] A. Bényi and K.A. Okoudjou. Local well-posedness of nonlinear dispersive equations on modulation spaces. *Bull. Lond. Math. Soc.*, 41(3): 549–558, 2009.

- [3] A. Bényi, K. Gröchenig, K.A. Okoudjou and L.G. Rogers. Unimodular Fourier multipliers for modulation spaces. *J. Funct. Anal.*, 246(2):366–384, 2007.
- [4] E. J. Candès and L. Demanet. The curvelet representation of wave propagators is optimally sparse. *Comm. Pure Appl. Math.*, 58:1472–1528, 2004.
- [5] L. Cattabriga, *Alcuni problemi per equazioni differenziali lineari con coefficienti costanti*, Quaderni dell’Unione Matematica Italiana, 24, Pitagora, Bologna, 1983.
- [6] E. Cordero. Gelfand-Shilov Window Classes for Weighted Modulation Spaces. *Int. Tran. Spec. Funct.*, 18(11):809–817, 2007.
- [7] E. Cordero, H. Feichtinger and F. Luef. Banach Gelfand Triples for Gabor Analysis, in *Pseudo-differential operators*, Lecture Notes in Math., Springer Berlin, 1949:1–33, 2008.
- [8] E. Cordero and K. Gröchenig and F. Nicola. Approximation of Fourier integral operators by Gabor multipliers. *J. Fourier Anal. Appl.*, 18(4):661–684, 2012.
- [9] E. Cordero and F. Nicola. Remarks on Fourier multipliers and applications to the wave equation. *J. Math. Anal. Appl.*, 353(2):583–591, 2009.
- [10] E. Cordero and F. Nicola. Boundedness of Schrödinger type propagators on modulation spaces. *J. Fourier Anal. Appl.*, 16(3):311–339, 2010.
- [11] E. Cordero, F. Nicola and L. Rodino, Sparsity of Gabor representation of Schrödinger propagators. *Appl. Comput. Harmon. Anal.*, 26(3):357–370, 2009.
- [12] E. Cordero, F. Nicola and L. Rodino, Boundedness of Fourier integral operators in  $\mathcal{FL}^p$  spaces. *Trans. Amer. Math. Soc.*, 361(11):6049–6071, 2009.
- [13] E. Cordero, F. Nicola and L. Rodino. Time-frequency analysis of Fourier integral operators. *Commun. Pure Appl. Anal.*, 9(1):1–21, 2010.
- [14] E. Cordero, F. Nicola and L. Rodino, On the global boundedness of Fourier integral operators, *Annals of Global Analysis and Geometry*, 38(4):373–398, 2010.
- [15] E. Cordero, S. Pilipović, L. Rodino and N. Teofanov. Localization operators and exponential weights for modulation spaces. *Mediterranean J. Math.*, 2(4):381–394, 2005.
- [16] E. Cordero and D. Zucco. The Cauchy Problem for the Vibrating Plate Equation in Modulation Spaces. *J. Pseudo-Differ. Op. and Appl.*, 2:343–354, 2011.
- [17] A. Córdoba and C. Fefferman. Wave packets and Fourier integral operators. *Comm. Partial Differential Equations*, 3(11):979–1005, 1978.
- [18] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inform. Theory*, 36(5):961–1005, 1990.
- [19] H. G. Feichtinger, Modulation spaces on locally compact abelian groups, *Technical Report, University Vienna, 1983*, and also in *Wavelets and Their Applications*, M. Krishna, R. Radha, S. Thangavelu, editors, Allied Publishers, 99–140, 2003.
- [20] D. A. Geba and D. Tataru. A Phase Space Transform Adapted to the Wave Equation. *Communications in Partial Differential Equations*, 32:1065–1101, 2007.
- [21] I. M. Gelfand, G. E. Shilov. *Generalized Functions II, III*. Academic Press, 1967.
- [22] K. Gröchenig. *Foundations of time-frequency analysis*. Birkhäuser Boston, Inc., Boston, MA, 2001.
- [23] K. Gröchenig. Time-Frequency Analysis of Sjöstrand’s Class. *Rev. Mat. Iberoamericana*, 22(2):703–724, 2006.
- [24] K. Gröchenig and Y. Lyubarskii. Gabor (super)frames with Hermite functions. *Math. Ann.*, 345:267–286, 2009.
- [25] K. Gröchenig and Z. Rzesotnik. Banach algebras of pseudodifferential operators and their almost diagonalization. *Ann. Inst. Fourier*, 58(7):2279–2314, 2008.
- [26] K. Gröchenig and G. Zimmermann. Spaces of test functions via the STFT. *Journal of Function Spaces and Applications*, 2(1): 25–53, 2004.
- [27] K. Guo and D. Labate. Sparse shearlet representation of Fourier integral operators. *Electron. Res. Announc. Math. Sci.*, 14:7–19, 2007.
- [28] L. Hörmander. *The analysis of linear partial differential operators*, Vol. 2, Vol. 4, Springer, 1983, 1985.
- [29] A. J. E. M. Janssen. Duality and biorthogonality for Weyl-Heisenberg frames. *J. Fourier Anal. Appl.*, 1(4):403–436, 1995.

- [30] K. Kato, M. Kobayashi and S. Ito. Representation of Schrödinger operator of a free particle via short time Fourier transform and its applications. *Tohoku Math. J.*, 64:223–231, 2012.
- [31] K. Kato, M. Kobayashi and S. Ito. Remark on wave front sets of solutions to Schrödinger equation of a free particle and a harmonic oscillator. *SUT J.Math.*, 47:175–183, 2011.
- [32] K. Kato, M. Kobayashi and S. Ito. Remarks on Wiener Amalgam space type estimates for Schrödinger equation. 41–48, RIMS Kôkyûroku Bessatsu, B33, Res. Inst. Math. Sci. (RIMS), Kyoto, 2012.
- [33] B. S. Mitjagin. Nuclearity and other properties of spaces of type  $S$ . *Amer. Math. Soc. Transl.*, 93(2):45–59, 1970.
- [34] A. Miyachi, F. Nicola, S. Rivetti, A. Tabacco and N. Tomita. Estimates for unimodular Fourier multipliers on modulation spaces, *Proc. Amer. Math. Soc.*, 137:3869–3883, 2009.
- [35] F. Nicola and L. Rodino. *Global Pseudo-Differential Calculus on Euclidean Spaces*. Pseudo-Differential Operators. Theory and Applications, 4, Birkhäuser Verlag, Basel, 2010.
- [36] J. Qian and L. Ying. Fast Multiscale Gaussian Wavepacket Transforms and Multiscale Gaussian Beams for the Wave Equation. *Multiscale Model. Simul.*, 8(5): 1803–1837, 2010.
- [37] J. Qian and L. Ying. Fast Gaussian wavepacket transforms and Gaussian beams for the Schrödinger equation. *J. Computational Physics*, 229:7848–7873, 2010.
- [38] J. Rauch, *Partial differential equations*, Springer, 1991.
- [39] R. Rochberg and K. Tachizawa. Pseudodifferential operators, Gabor frames, and local trigonometric bases. In *Gabor analysis and algorithms*, pages 171–192. Birkhäuser Boston, Boston, MA, 1998.
- [40] M. Ruzhansky, M. Sugimoto, B. Wang. Modulation spaces and nonlinear evolution equations. In *Evolution Equations of Hyperbolic and Schrödinger Type*, 267–283, Progress in Mathematics, Vol. 301, Birkhäuser, 2012.
- [41] L. Schwartz, *Mathematics for the physical sciences*, Dover, 2008.
- [42] H. F. Smith. A parametrix construction for wave equations with  $C^{1,1}$  coefficients. *Annales de l’institut Fourier*, 48(3):797–835, 1998.
- [43] T. Strohmer. Pseudodifferential operators and Banach algebras in mobile communications. *Appl. Comput. Harmon. Anal.*, 20(2):237–249, 2006.
- [44] D. Tataru. Phase space transforms and microlocal analysis. in “*Phase space analysis of partial differential equations*. Vol. II, 505–524, Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa, 2004.
- [45] N. Teofanov, Ultradistributions and time-frequency analysis, in *Pseudo-differential Operators and Related Topics, Operator Theory: Advances and Applications*, P. Boggiatto, L. Rodino, J. Toft, M.W. Wong, editors, Birkhäuser, 164:173–191, 2006.
- [46] J. Toft. The Bargmann transform on modulation and Gelfand-Shilov spaces with applications to Toeplitz and pseudo-differential operators. *J. Pseudo-Differ. Oper. Appl.*, 3:145–227, 2012.
- [47] F. Trèves. *Topological Vector Spaces, Distributions and Kernels*. Academic Press, New York, 1967.
- [48] D. F. Walnut. Lattice size estimates for Gabor decompositions. *Monatsh. Math.*, 115(3):245–256, 1993.
- [49] B. Wang. Sharp global well-posedness for non-elliptic derivative Schrödinger equations with small rough data. arXiv:1012.0370.
- [50] B. Wang, Z. Lifeng and G. Boling. Isometric decomposition operators, function spaces  $E_{p,q}^\lambda$  and applications to nonlinear evolution equations. *J. Funct. Anal.*, 233(1):1–39, 2006.
- [51] B. Wang and C. Huang. Frequency-uniform decomposition method for the generalized BO, KdV and NLS equations. *J. Differential Equations*, 239(1):213–250, 2007.
- [52] B. Wang and H. Hudzik. The global Cauchy problem for the NLS and NLKG with small rough data. *J. Differential Equations*, 231:36–73, 2007.
- [53] B. Wang, Z. Huo, C. Hao and Z. Guo. *Harmonic analysis method for nonlinear evolution equations*. I. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2011.

DIPARTIMENTO DI SCIENZE MATEMATICHE, POLITECNICO DI TORINO, CORSO DUCA DEGLI ABRUZZI  
24, 10129 TORINO, ITALY

DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI TORINO, VIA CARLO ALBERTO 10, 10123 TORINO,  
ITALY

*E-mail address:* `elena.cordero@unito.it`

*E-mail address:* `fabio.nicola@polito.it`

*E-mail address:* `luigi.rodino@unito.it`