View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by OTHES

universität wien

DISSERTATION

Bilinear Time-Frequency Distributions and Pseudodifferential Operators

Verfasser

Dipl.Math. Dominik Bayer

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien, im August 2010

Studienkennzahl lt. Studienblatt: A 091 405
Studienrichtung lt. Studienblatt: Mathematik
Betreuer: Prof. Karlheinz Gröchenig

Contents

Acknowledgements

No man is an island, entire of itself. Since this is particularly true for me, I want to thank many people and institutions that provided support and help in various ways and at various times.

I am grateful to the following institutions and people that helped in the making of this work:

- The Department of Mathematics at the Technical University (TU), Munich, especially Prof. Gero Friesecke and the people at research unit M7;
- The Institute for Biomathematics and Biometry (IBB) at the Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg (formally called GSF National Research Center for Environment and Health); in particular I would like to thank Prof. Rupert Lasser, the head of the institute, for his kind support, as well as Prof. Hartmut Führ, now at RWTH Aachen, for many good times and discussions not only, but also, about mathematics;
- The Mathematics department of the University of Vienna, in particular the Numerical Harmonic Analysis Group (NuHAG); thanks go not only to Prof. Hans Georg Feichtinger but all the staff there (which are by far too many to list explicitly);
- The Acoustics Research Institute at the Austrian Academy of Sciences, Vienna, especially to Dr. Peter Balazs.

This work was partially funded by

the European Centre of Time-Frequency Analysis (Marie Curie Excellence Grant MEXT-CT-2004-517154)
and
the WWTF project MULAC ('Frame Multipliers: Theory and Application in Acoustics', MA07-025)

Particularly warm thanks go to the following persons:

- Prof. Domenico Castrigiano at TU Munich, who brought me on the way;
- Prof. Karlheinz Gröchenig, my advisor, who guided me all along the (long, long indeed) way to a happy end, and whose friendship and patience I severely tested;
- Prof. Gerhard Winkler at IBB, who knows why.

Finally, this work is dedicated to Marie (of course), who is the single most important and best thing in my life.

Introduction

Time-frequency analysis (or Gabor analysis) is a modern branch of harmonic analysis with applications, among others, in signal processing, audio engineering, psychoacoustics as well as in theoretical physics and pure mathematics. As such, it is concerned with its own version of the basic general idea of harmonic analysis, namely the decomposition of mathematical objects (like functions) into parts that are in some sense more elementary or simpler than the original object. In classical Fourier analysis, functions (or signals) are decomposed and represented as superpositions of pure frequencies. This is accomplished by using the Fourier transform, that can be interpreted as giving the frequency distribution of the original function. In the Gabor case, functions (or signals) are decomposed into pieces that are time-frequency shifts of some given analyzing window function. The corresponding transform is the short-time Fourier transform, which can be interpreted as giving a time-frequency distribution of the transformed function. It is, however, not the only time-frequency distribution that is in use. In quantum mechanics, the Wigner distribution has been around for quite some time, its first appearance in [44]. It was later derived independently in [41] and [29]. In [43], it is used in connection with questions of quantization. The Weyl calculus associated with the Wigner distribution is used to analyze pseudodifferential operators in [15], [21] and [24]. Further time-frequency methods are applied to pseudodifferential operaors in [34]. The Rihacek distribution plays a prominent part in connection with the Kohn-Nirenberg correspondence, [27]. All these time-frequency distributions are very similar structurally, have similar properties and lead to similar theories. It is therefore quite natural to try to find a general framework into which the different distributions can be integrated as particular cases. To provide such a framework is one of the aims of this work. The class of bilinear time-frequency distributions introduced in the first chapter is sufficiently general to contain the short-time Fourier transform, the Wigner and the Rihacek distributions as special cases. A bilinear time-frequency distribution is a bilinear mapping

$$\begin{aligned} f, g \quad \mapsto \quad \mathrm{TF}_A(f,g)(x,\omega) &= \mathcal{F}^2\mathcal{T}_A(f \otimes \overline{g})(x,\omega) \\ &= \int_{\mathbb{R}^d} (f \otimes \overline{g})(A \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)) e^{-2\pi i \omega \cdot y} \, dy, \end{aligned}$$

that is a coordinate transform with transformation matrix A followed by a partial Fourier transform, applied to the tensor product of the functions f and g.

Which properties carry over to the general distributions? Can one set up a pseudodifferential calculus associated to a general time-frequency distribution that has similar properties as the Weyl calculus associated to the Wigner distribution? Is it possible to identify necessary or sufficient conditions on the distributions to guarantee nice behavior of the associated calculus? The first part (chapters 1 and 2) tries to answer these questions. It turns out that many time-frequency distributions allow a rich theory analogous to the existing calculi, although some pathological cases must be discarded. A key property seems to be right resp. left regularity of the transformation matrix A, as defined in the first chapter.

The second part (chapter 2) is concerned with another type of pseudodifferential operator, namely time-frequency localization operators, first introduced in [9] and [10]. These are multipliers for the short-time Fourier transform, acting on a function f by

$$f \quad \mapsto \quad V_{\varphi_1}^*(a \cdot V_{\varphi_2} f) = \iint_{\mathbb{R}^{2d}} a(x,\omega) \cdot V_{\varphi_1} f(x,\omega) \, M_\omega T_x \varphi_2 \, dx d\omega.$$

The window functions φ_1, φ_2 are taken from some function space, that determines the mapping and boundedness properties of the operator. The function a is called the symbol of the localization operator. The main result of this part is to show a connection with the so-called Berezin transform, an object that is of some importance in complex analysis. This connection gives a powerful tool to examine the question how large the class of localization operators (with symbols from some prescribed class, e.g. from $L^2(\mathbb{R}^{2d})$) is compared to larger classes of operators (e.g. the Hilbert-Schmidt class). Can arbitrary operators be approximated by localization operators, and in what topology? It turns out that in many cases the set of localization operators is dense in a larger class of operators, with respect to either the norm topology or the weak-$*$ topology. There occurs an interesting phenomenon, however. For symbols from, say, Lebesgue spaces $L^p(\mathbb{R}^{2d})$, the exponents $1 \leq p \leq 2$ give stronger results than the exponents $2 \leq p \leq \infty$. For symbols from modulation spaces, the situation is similar. In this regard, further results may be possible in the future.

This work is structured as follows.
The *first chapter* introduces the notion of (generalized) bilinear time-frequency distribution. This is a direct generalization of some well-known time-frequency distributions that are in wide use in mathematics and physics, such as the short-time Fourier transform (also called the Gabor transform) or the Wigner distribution. Some basic properties of bilinear time-frequency distributions are presented. These usually generalize the according properties of the short-time Fourier transform. The time-frequency distributions considered can be parametrized by real $2d \times 2d$ coefficient matrices. Conditions on these matrices yielding nicely behaved distributions, namely the notions of right resp. left regularity, are identified and defined. Two important technical tools for future use are provided, namely the covariance property and the so-called 'magic' formula. Again, these generalize well-known formulas for the short-time Fourier transform to the more general case of bilinear time-frequency distributions. The former are contained in the latter as special cases.
In the *second chapter*, we present the pseudodifferential operator calculus associated with bilinear time-frequency distributions. This is motivated by the Kohn-Nirenberg correspondence (associated to the Rihacek distribution) and the Weyl calculus (associated to the Wigner distribution) and generalizes these for general bilinear time-frequency distributions. For a large class of distributions, most of the desirable properties of the aforementioned classic pseudodifferential calculi carry over to the more general situation. Well-known mapping and Schatten class properties for the short-time Fourier transform are proved in a more general setting. Particular importance is given to various boundedness theorems on modulation spaces for pseudodifferential operators associated to well-behaved bilinear time-frequency distributions.
The *third chapter* is devoted to the study of time-frequency localization operators. Well-known mapping properties of localization operators are summarized in a unified framework. The connection to the Berezin transform is shown and used to prove some new results on the density of the set of local-

ization operators within larger spaces of operators with respect to different topologies, e.g. the space of all bounded operators equipped with either the norm topology or the weak-$*$ topology.

Chapter 1

Bilinear Time-Frequency Distributions

In this chapter, we define a particular class of general bilinear time-frequency distributions that is the basic object of all our subsequent considerations. We show that some well-known time-frequency distributions, like the short-time Fourier transform or the Wigner distribution, can be subsumed under our general framework. Moreover, many elementary properties enjoyed by the aforementioned distributions have analogues in the general setting.

1.1 Definition

Definition 1.1.1 (Bilinear Time-Frequency Distribution). *Let $f, g \in L^2(\mathbb{R}^d)$ and $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d \times 2d}$ with $\det A \neq 0$. Define the* **bilinear time-frequency distribution** *of f and g to be*

$$\mathrm{TF}_A(f, g) := \mathcal{F}_2 \mathcal{T}_A(f \otimes \overline{g}).$$

In this definition, $\mathcal{F}_2$ denotes the partial Fourier transform of a function of two variables with respect to the second variable, and $\mathcal{T}_A$ is the coordinate transformation with matrix A.

If $\mathcal{T}_A(f \otimes \overline{g})$ is an integrable function with respect to the second variable over

$\mathbb{R}^d$, we can represent the partial Fourier transform as an integral:

$$\begin{aligned}\mathrm{TF}_A(f,g)(x,\omega) &= \int_{\mathbb{R}^d} (f\otimes\overline{g})(A\left(\begin{smallmatrix}x\\ y\end{smallmatrix}\right))e^{-2\pi i\omega\cdot y}\,dy\\ &= \int_{\mathbb{R}^d} f(A_{11}x+A_{12}y)\overline{g(A_{21}x+A_{22}y)}e^{-2\pi i\omega\cdot y}\,dy\end{aligned}$$

for $(x,\omega)\in\mathbb{R}^{2d}$. This holds for example when f and g are Schwartz functions.

Important and well-known **examples** are

- **Short-time Fourier transform:**

  $$V(f,g)(x,\omega) = \int_{\mathbb{R}^d} f(y)\overline{g(y-x)}e^{-2\pi i\omega\cdot y}\,dy = \mathrm{TF}_A(f,g)(x,\omega)$$

  with matrix $A = \left(\begin{smallmatrix}0 & I\\ -I & I\end{smallmatrix}\right)$. This is the basic time-frequency distribution used in time-frequency analysis. For a short overview over the short-time Fourier transform and its most basic properties see the appendix.

- **Wigner distribution:**

  $$W(f,g)(x,\omega) = \int_{\mathbb{R}^d} f(x+\frac{y}{2})\overline{g(x-\frac{y}{2})}e^{-2\pi i\omega\cdot y}\,dy = \mathrm{TF}_A(f,g)(x,\omega)$$

  with matrix $A = \begin{pmatrix}I & \frac{1}{2}I\\ I & -\frac{1}{2}I\end{pmatrix}$. This is undoubtedly the most popular time-frequency distribution in signal analysis. It was, however, first introduced in 1932 by E. Wigner in a paper on quantum mechanics ([44]).

- **α-Wigner distribution:**

  $$\begin{aligned}W_\alpha(f,g)(x,\omega) &= \int_{\mathbb{R}^d} f(x+(1-\alpha)y)\overline{g(x-\alpha y)}e^{-2\pi i\omega\cdot y}\,dy\\ &= \mathrm{TF}_A(f,g)(x,\omega)\end{aligned}$$

  with matrix $A = \left(\begin{smallmatrix}I & (1-\alpha)I\\ I & -\alpha I\end{smallmatrix}\right)$ and $\alpha\in(0,1)$. This is a (less symmetric) variant of the Wigner distribution. The ordinary Wigner distribution corresponds to the value $\alpha=\frac{1}{2}$.

- **Rihacek distribution:**

$$
\begin{aligned}
R(f,g)(x,\omega) &= f(x)\overline{\hat{g}(\omega)}e^{-2\pi i x\cdot\omega} \\
&= \int_{\mathbb{R}^d} f(x)\overline{g(x-y)}e^{-2\pi i y\cdot\omega}\,dy = \mathrm{TF}_A(f,g)(x,\omega)
\end{aligned}
$$

with matrix $A = \begin{pmatrix} I & 0 \\ I & -I \end{pmatrix}$.

- **(Radar) ambiguity function:**

$$
A(f,g)(x,\omega) = \int_{\mathbb{R}^d} f(y+\frac{x}{2})\overline{g(y-\frac{x}{2})}e^{-2\pi i y\cdot\omega}\,dy = \mathrm{TF}_A(f,g)(x,\omega)
$$

with matrix $A = \begin{pmatrix} \frac{1}{2}I & I \\ -\frac{1}{2}I & I \end{pmatrix}$.

1.2 Elementary Properties

After having defined the general bilinear time-frequency distributions in the previous section, we are now going to examine some of their most basic properties.

Theorem 1.2.1. *If* $f, g \in L^2(\mathbb{R}^d)$ *and* $A \in \mathbb{R}^{2d\times 2d}$ *with* $\det A \neq 0$, *then* $\mathrm{TF}_A(f,g) \in L^2(\mathbb{R}^{2d})$.

Proof. This is obvious, since the tensor product $f \otimes \overline{g} \in L^2(\mathbb{R}^{2d})$ and $\mathcal{T}_A$ and $\mathcal{F}_2$ are bounded linear operators from $L^2(\mathbb{R}^{2d})$ into itself. □

The set $\{\mathrm{TF}_A(f,g)|\ f,g \in L^2(\mathbb{R}^d)\}$ is in fact a complete subset of $L^2(\mathbb{R}^{2d})$, i.e. its linear span is dense in $L^2(\mathbb{R}^{2d})$:

Theorem 1.2.2. *Denote the set* $\{\mathrm{TF}_A(f,g)|\ f,g \in L^2(\mathbb{R}^d)\} \subseteq L^2(\mathbb{R}^{2d})$ *by* S. *If* $\det A \neq 0$, *then*

$$\overline{\mathit{span}(S)} = L^2(\mathbb{R}^{2d}),$$

i.e. S *is a complete subset of* $L^2(\mathbb{R}^{2d})$.

Proof. Assume $F \in L^2(\mathbb{R}^{2d})$ such that

$$\langle F, \mathrm{TF}_A(f,g)\rangle = 0$$

for all $f, g \in L^2(\mathbb{R}^d)$. We have to show that this implies $F = 0$ in $L^2(\mathbb{R}^{2d})$. Now

$$\begin{aligned}\langle F, \mathrm{TF}_A(f,g)\rangle &= \langle F, \mathcal{F}_2\mathcal{T}_A(f\otimes\overline{g})\rangle\\ &= \langle \mathcal{T}_A^*\mathcal{F}_2^* F, f\otimes\overline{g}\rangle\\ &= 0\end{aligned}$$

for all $f, g \in L^2(\mathbb{R}^d)$ if and only if

$$\mathcal{T}_A^*\mathcal{F}_2^* F = 0 \text{ in } L^2(\mathbb{R}^{2d}),$$

since $\{f\otimes\overline{g}|\ f,g \in L^2(\mathbb{R}^d)\} \subseteq L^2(\mathbb{R}^{2d})$ is a complete subset of $L^2(\mathbb{R}^{2d})$. But this is equivalent to

$$F = 0 \in L^2(\mathbb{R}^{2d}),$$

since $\mathcal{T}_A$ and $\mathcal{F}_2$ are bounded invertible operators from $L^2(\mathbb{R}^{2d})$ onto $L^2(\mathbb{R}^{2d})$. □

We can also show (with identical proof), that $\{\mathrm{TF}_A(e_n, e_m) |\, n, m \in \mathbb{N}\}$ is a complete subset of $L^2(\mathbb{R}^{2d})$, if $(e_n)_{n\in\mathbb{N}}$ is an arbitrary orthonormal basis of $L^2(\mathbb{R}^d)$ (because then the family $(e_n \otimes \overline{e_m})_{n,m\in\mathbb{N}}$ is an orthonormal basis of $L^2(\mathbb{R}^{2d})$). In Section 1.3 we will see that this set is actually an orthonormal basis of $L^2(\mathbb{R}^{2d})$ itself.

Theorem 1.2.3. *Let* $\det A \neq 0$. *Then the bilinear mapping* $\mathrm{TF}_A : L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^{2d})$, $(f, g) \mapsto \mathrm{TF}_A(f, g)$ *is (jointly) continuous.*

Proof. Let $f, g, f_n, g_n \in L^2(\mathbb{R}^d)$, $n = 1, 2, \ldots$, with $\|f_n - f\| \to 0$ and $\|g_n - g\| \to 0$ for $n \to \infty$. Then

$$\begin{aligned}
\|\mathrm{TF}_A(f, g) - \mathrm{TF}_A(f_n, g_n)\| &= \|\mathcal{F}_2\mathcal{T}_A(f \otimes \overline{g}) - \mathcal{F}_2\mathcal{T}_A(f_n \otimes \overline{g_n})\| \\
&= \|\mathcal{T}_A(f \otimes \overline{g}) - \mathcal{T}_A(f_n \otimes \overline{g_n})\| \qquad \text{(since } \mathcal{F}_2 \text{ is unitary)} \\
&= \frac{1}{|\det A|}\|f \otimes \overline{g} - f_n \otimes \overline{g_n}\| \qquad \text{(by Lemma A.3.2)}
\end{aligned}$$

and

$$\begin{aligned}
\|f \otimes \overline{g} - f_n \otimes \overline{g_n}\| &\leq \|f \otimes \overline{g} - f \otimes \overline{g_n}\| + \|f \otimes \overline{g_n} - f_n \otimes \overline{g_n}\| \\
&= \|f \otimes \overline{(g - g_n)}\| + \|(f - f_n) \otimes \overline{g_n}\| \\
&= \|f\| \cdot \underbrace{\|g - g_n\|}_{\to 0} + \underbrace{\|f - f_n\|}_{\to 0} \cdot \underbrace{\|g_n\|}_{\to \|g\|} \\
&\longrightarrow 0
\end{aligned}$$

for $n \to \infty$. □

Theorem 1.2.4. *If* $f, g \in \mathcal{S}(\mathbb{R}^d)$, *then* $\mathrm{TF}_A(f, g) \in \mathcal{S}(\mathbb{R}^{2d})$.

Proof. If $f, g \in \mathcal{S}(\mathbb{R}^d)$, then the tensor product $f \otimes \overline{g} \in \mathcal{S}(\mathbb{R}^{2d})$, hence $\mathrm{TF}_A(f, g) = \mathcal{F}_2\mathcal{T}_A(f \otimes \overline{g}) \in \mathcal{S}(\mathbb{R}^{2d})$ by Lemma A.3.5 and Lemma A.4.4, respectively. □

In the next theorem, certain bilinear time-frequency distributions are shown to have a representation associated to the short-time Fourier transform, thus having very similar (pleasant) properties.

Theorem 1.2.5 (Representation)**.** *Let* $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d\times 2d}$ *with* $\det A \neq 0$ *and let both* A_{12} *and* A_{22} *be invertible matrices in* $\mathbb{R}^{d\times d}$*. Then the following representation holds for all* $f, g \in L^2(\mathbb{R}^d)$*,* $x, \omega \in \mathbb{R}^d$*:*

$$\mathrm{TF}_A(f,g)(x,\omega) = \frac{e^{2\pi i (A_{12}^*)^{-1}\omega\cdot A_{11}x}}{|\det A_{12}|} \langle f, M_d T_c \tilde{g}\rangle ,$$

with $\tilde{g}(z) := g(A_{22}A_{12}^{-1}z)$ *and*

$$d = d(\omega) = (A_{12}^*)^{-1}\omega, \qquad c = c(x) = (A_{11} - A_{12}A_{22}^{-1}A_{21})x.$$

Proof. If A_{12} and A_{22} are invertible, the functions $f'(y) = f(A_{12}y)$ and $g'(y) = g(A_{22}y)$ are well-defined and in $L^2(\mathbb{R}^d)$ for f and g in $L^2(\mathbb{R}^d)$. Then $f(A_{11}x + A_{12}y) = T_{-A_{12}^{-1}A_{11}x} f'(y)$ and $g(A_{21}x + A_{22}y) = T_{-A_{22}^{-1}A_{21}x} g'(y)$ are in $L^2(\mathbb{R}^d)$ as well, so the integral

$$\mathrm{TF}_A(f,g)(x,\omega) = \int_{\mathbb{R}^d} f(A_{11}x + A_{12}y)\overline{g(A_{21}x + A_{22}y)}e^{-2\pi i\omega\cdot y}\,dy$$

makes sense pointwisely for all $x, \omega \in \mathbb{R}^d$. The stated representation now follows from the change of variables $z = A_{11}x + A_{12}y$. □

Corollary 1.2.6. *Under the assumptions of Theorem 1.2.5,* $\mathrm{TF}_A(f,g)(x,\omega)$ *is a continuous function on* $\mathbb{R}^{2d}$*.*

Proof. Let $x_n \to x$ and $\omega_n \to \omega$ in $\mathbb{R}^d$. Put $d(\omega) = (A_{12}^*)^{-1}\omega$ and $c(x) = (A_{11} - A_{12}A_{22}^{-1}A_{21})x$ as in Theorem 1.2.5, then $d(\omega_n) \to d(\omega)$ and $c(x_n) \to c(x)$ for $n \to \infty$, since the functions d and c are continuous. It follows that $M_{d(\omega_n)}T_{c(x_n)}\tilde{g}$ converges to $M_{d(\omega)}T_{c(x)}\tilde{g}$ in $L^2(\mathbb{R}^d)$. Finally, with Theorem 1.2.5, we conclude $\mathrm{TF}_A(f,g)(x_n,\omega_n) \to \mathrm{TF}_A(f,g)(x,\omega)$. □

Corollary 1.2.7. *Under the assumptions of Theorem 1.2.5, the function* $\mathrm{TF}_A(f,g)(x,\omega)$ *is bounded:*

$$|\mathrm{TF}_A(f,g)(x,\omega)| \leq \frac{\|f\|\cdot\|g\|}{|\det A_{12}|^{1/2}\cdot|\det A_{22}|^{1/2}}$$

for all $(x,\omega) \in \mathbb{R}^{2d}$*.*

Proof. We have

$$\|\tilde{g}\| = \|\mathcal{T}_{A_{22}A_{12}^{-1}} g\| = \frac{|\det A_{12}|^{1/2}}{|\det A_{22}|^{1/2}} \|g\|,$$

hence

$$\begin{aligned} |\mathrm{TF}_A(f,g)(x,\omega)| &= \left| \frac{e^{2\pi i (A_{12}^*)^{-1}\omega \cdot A_{11}x}}{|\det A_{12}|} \langle f, M_d T_c \tilde{g} \rangle \right| \\ &\leq \frac{1}{|\det A_{12}|} \|f\| \cdot \|\tilde{g}\| \\ &= \frac{\|f\| \cdot \|g\|}{|\det A_{12}|^{1/2} \cdot |\det A_{22}|^{1/2}}. \end{aligned}$$

□

Corollary 1.2.8 (Riemann-Lebesgue)**.** *Under the assumptions of Theorem 1.2.5,* $\mathrm{TF}_A(f,g)$ *vanishes at infinity, i.e.*

$$\lim_{|(x,\omega)| \to \infty} |\mathrm{TF}_A(f,g)(x,\omega)| = 0.$$

Proof. We have to show that

$$\lim_{|(x,\omega)| \to \infty} |\left\langle f, M_{d(\omega)} T_{c(x)} \tilde{g} \right\rangle| = 0.$$

Let $\Omega \subset \mathbb{R}^d$ be an arbitrary subset. Then

$$\begin{aligned} |\langle f, M_d T_c \tilde{g} \rangle| &= |\int_{\mathbb{R}^d} f(t) \overline{\tilde{g}(t-c)} e^{-2\pi i d \cdot t}\, dt| \\ &\leq \int_{\mathbb{R}^d} |f(t)| \cdot |\tilde{g}(t-c)|\, dt \\ &= \int_{\Omega} |f(t)| \cdot |\tilde{g}(t-c)|\, dt + \int_{\mathbb{R}^d \setminus \Omega} |f(t)| \cdot |\tilde{g}(t-c)|\, dt. \end{aligned}$$

By the Cauchy-Schwarz Inequality, we get

$$|\langle f, M_d T_c \tilde{g} \rangle| \leq \left(\int_{\Omega} |f|^2 \right)^{\frac{1}{2}} \left(\int_{\Omega} |T_c \tilde{g}|^2 \right)^{\frac{1}{2}} + \left(\int_{\mathbb{R}^d \setminus \Omega} |f|^2 \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d \setminus \Omega} |T_c \tilde{g}|^2 \right)^{\frac{1}{2}}.$$

Now let $\varepsilon > 0$ be given. Choose $\Omega \subset \mathbb{R}^d$ in such a way that Ω is compact and both $\left(\int_\Omega |f|^2\right)^{1/2} \geq (1-\varepsilon)\|f\|$ and $\left(\int_\Omega |\tilde{g}|^2\right)^{1/2} \geq (1-\varepsilon)\|\tilde{g}\|$. (One could use e.g. $\Omega = B_R = \{t \in \mathbb{R}^d : |t| \leq r\}$ for sufficiently large r, the closed ball with radius r.) Since Ω is compact, there exists a constant $K > 0$ depending on f and $\tilde{g}$ such that $\Omega \cap c + \Omega = \emptyset$ for all $c \in \mathbb{R}^d$ with $|c| > K$. This yields

$$\left(\int_\Omega |f|^2\right)^{1/2} \left(\int_\Omega |T_c\tilde{g}|^2\right)^{1/2} = \left(\int_\Omega |f|^2\right)^{1/2} \left(\int_{c+\Omega} |\tilde{g}|^2\right)^{1/2} \leq \varepsilon \cdot \|f\| \cdot \|\tilde{g}\|$$

and

$$\left(\int_{\mathbb{R}^d\setminus\Omega} |f|^2\right)^{1/2} \left(\int_{\mathbb{R}^d\setminus\Omega} |T_c\tilde{g}|^2\right)^{1/2} \leq \varepsilon \cdot \|f\| \cdot \|\tilde{g}\|,$$

so

$$|\left\langle f, M_d T_c \tilde{g}\right\rangle| \leq 2\varepsilon \cdot \|f\| \cdot \|\tilde{g}\|$$

for any d and all $|c| > K$ in $\mathbb{R}^d$.

Next observe that by Plancherel's Formula and the canonical commutation relation

$$|\left\langle f, M_d T_c \tilde{g}\right\rangle| = |\left\langle \hat{f}, \widehat{M_d T_c \tilde{g}}\right\rangle| = |\left\langle \hat{f}, T_d M_{-c} \hat{\tilde{g}}\right\rangle| = |\left\langle \hat{f}, M_{-c} T_d \hat{\tilde{g}}\right\rangle|.$$

By the same argument as above we conclude that there exists a constant $K' > 0$ depending on $\hat{f}$ and $\hat{\tilde{g}}$ such that

$$|\left\langle f, M_d T_c \tilde{g}\right\rangle| \leq 2\varepsilon \cdot \|\hat{f}\| \cdot \|\hat{\tilde{g}}\| = 2\varepsilon \cdot \|f\| \cdot \|\tilde{g}\|$$

for any c and all $|d| > K'$ in $\mathbb{R}^d$.

So we find that

$$|\left\langle f, M_d T_c \tilde{g}\right\rangle| \leq 2\varepsilon \cdot \|f\| \cdot \|\tilde{g}\|$$

for all $(c, d) \in \mathbb{R}^{2d}$ with $|c| > K$ or $|d| > K'$, that means outside the compact set $B_K \times B_{K'} \subset \mathbb{R}^{2d}$. So

$$\lim_{|(c,d)|\to\infty} |\left\langle f, M_d T_c \tilde{g}\right\rangle| = 0.$$

It remains to prove that $|(x, \omega)| \to \infty$ implies $|(c(x), d(\omega))| \to \infty$. We have $|d(\omega)| = |(A_{12}^*)^{-1}\omega| \to \infty$ for $|\omega| \to \infty$ if and only if $(A_{12}^*)^{-1}$ is invertible, and $|c(x)| = |(A_{11} - A_{12}A_{22}^{-1}A_{21})x| \to \infty$ for $|x| \to \infty$ if and only if $A_{11} -$

$A_{12}A_{22}^{-1}A_{21}$ is invertible. The former is true by assumption, the latter follows from

$$\begin{pmatrix} A_{11}-A_{12}A_{22}^{-1}A_{21} & 0 \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

and

$$\begin{aligned} 0 \neq \det A = \det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} &= \det \begin{pmatrix} A_{11}-A_{12}A_{22}^{-1}A_{21} & 0 \\ A_{21} & A_{22} \end{pmatrix} \\ &= \det(A_{11} - A_{12}A_{22}^{-1}A_{21}) \cdot \det A_{22}, \end{aligned}$$

therefore $\det(A_{11} - A_{12}A_{22}^{-1}A_{21}) \neq 0$. □

Theorem 1.2.9. *Suppose* $\det A_{22} \neq 0$*, but* $\det A_{12} = 0$*, i.e. the matrix* A_{12} *is not invertible. Then* $\mathrm{TF}_A(f,g)$ *is not necessarily a continuous function; more precisely, there always exist functions* $f, g \in L^2(\mathbb{R}^d)$ *such that* $\mathrm{TF}_A(f,g)$ *is not a continuous function on* $\mathbb{R}^{2d}$*.*

Proof. Choose an orthonormal basis $v_1, \ldots, v_l$ of $\ker A_{12}$ and extend it to an orthonormal basis $v_1, \ldots, v_l, v_{l+1}, \ldots, v_d$ of $\mathbb{R}^d$. The matrix V consisting of the vectors $v_1, \ldots, v_d$ as columns is an orthogonal matrix. The vectors $w_{l+1} := A_{12}v_{l+1}, \ldots, w_d := A_{12}v_d$ are linearly independent; extend this set of vectors to a basis $w_1, \ldots, w_l, w_{l+1}, \ldots, w_d$ of $\mathbb{R}^d$. The matrix formed with the vectors $w_1, \ldots, w_d$ as columns is denoted by W.
Now consider the orthogonal coordinate transformation $y = Vz$, $y, z \in \mathbb{R}^d$. This yields

$$\begin{aligned} A_{12}y = A_{12}Vz = A_{12}\sum_{j=1}^{d} z_j v_j = \sum_{j=1}^{d} z_j A_{12} v_j &= \sum_{j=l+1}^{d} z_j w_j \\ = W \cdot \begin{pmatrix} 0 \\ \vdots \\ 0 \\ z_{l+1} \\ \vdots \\ z_d \end{pmatrix}. \end{aligned}$$

We can express the vector $A_{11}x$ in terms of the basis $w_1, \ldots, w_d$:

$$A_{11}x = W\xi = \sum_{j=1}^{d} \xi_j w_j.$$

Now let

$$\tilde{f} := (r_1 \otimes r_2 \otimes \ldots \otimes r_d)(t) = r_1(t_1) \cdot \ldots \cdot r_d(t_d)$$

be a tensor product of arbitrary functions $r_1, \ldots, r_d \in L^2(\mathbb{R})$, set

$$f := \tilde{f} \circ W^{-1}$$

and let $g \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ be an arbitrary function. Note that under this assumptions we always have $f, g \in L^2(\mathbb{R}^d)$. Denoting

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ z_{l+1} \\ \vdots \\ z_d \end{pmatrix} = z',$$

we calculate

$$A_{11}x + A_{12}y = W\xi + Wz' = W(\xi + z'),$$

therefore

$$\begin{aligned} f(A_{11}x + A_{12}y) &= f(W(\xi + z')) = \tilde{f}(\xi + z') \\ &= \prod_{j=1}^{l} r_j(\xi_j) \cdot \prod_{j=l+1}^{d} r_j(\xi_j + z_j). \end{aligned}$$

Using the substitution $y = Vz$, we find

$$\begin{aligned} \mathrm{TF}_A(f,g)(x,\omega) &= \int_{\mathbb{R}^d} f(A_{11}x + A_{12}y)\overline{g(A_{21}x + A_{22}y)}e^{-2\pi i\omega\cdot y}\, dy \\ &= \int_{\mathbb{R}^d} f(A_{11}x + A_{12}Vz)\overline{g(A_{21}x + A_{22}Vz)}e^{-2\pi i\omega\cdot Vz}\, dz \\ &= \int_{\mathbb{R}^d} \prod_{j=1}^{l} r_j(\xi_j) \cdot \prod_{j=l+1}^{d} r_j(\xi_j + z_j)\overline{g(A_{21}x + A_{22}Vz)}e^{-2\pi i\omega\cdot Vz}\, dz \\ &= \prod_{j=1}^{l} r_j(\xi_j) \int_{\mathbb{R}^d} \prod_{j=l+1}^{d} r_j(\xi_j + z_j)\overline{g(A_{21}x + A_{22}Vz)}e^{-2\pi i\omega\cdot Vz}\, dz. \end{aligned}$$

Resubstituting x for ξ by the formula $\xi = W^{-1}A_{11}x$ gives

$$\begin{aligned}\mathrm{TF}_A(f,g)(x,\omega) &= H(x)\int_{\mathbb{R}^d} G(x,z)e^{-2\pi i V^*\omega\cdot z}\,dz\\ &= H(x)\cdot F(x,\omega)\end{aligned}$$

with

$$H(x) = \prod_{j=1}^{l} r_j((W^{-1}A_{11}x)_j),$$

$$G(x,z) = \left(\prod_{j=l+1}^{d} r_j((W^{-1}A_{11}x)_j + z_j)\right)\cdot\overline{g(A_{21}x + A_{22}Vz)}$$

and

$$F(x,\omega) = \int_{\mathbb{R}^d} G(x,z)e^{-2\pi i V^*\omega\cdot z}\,dz.$$

Observe that $F(x,\omega)$ is well-defined since both A_{22} and V are invertible and hence $G(x,z)$ is integrable with respect to z for all $x \in \mathbb{R}^d$. If we choose $g \in \mathcal{S}(\mathbb{R}^d)$ and $r_{l+1},\ldots,r_d \in \mathcal{S}(\mathbb{R})$, it is not hard to see that $F(x,\omega)$ is a continuous function on $\mathbb{R}^{2d}$. An appropriate choice ensures that F is not identically zero. Choose a point $(x_0,\omega_0) \in \mathbb{R}^{2d}$ with $F(x_0,\omega_0) \neq 0$. Since F is continuous, $F(x,\omega) \neq 0$ in a neighborhood of (x_0,ω_0). Then, by a suitable choice of functions $r_1,\ldots,r_l \in L^2(\mathbb{R})$, we clearly can achieve $H(x)$ to be uncontinuous at x_0. But then $\mathrm{TF}_A(f,g)$ cannot be continuous at (x_0,ω_0), because otherwise

$$H(x) = \frac{\mathrm{TF}_A(f,g)(x,\omega_0)}{F(x,\omega_0)}$$

(in some neighborhood of x_0) would be continuous at x_0, a contradiction. □

The above theorems and corollaries show the relevance of various assumptions on the invertibility or noninvertibility of the submatrices A_{11}, A_{12}, A_{21} and A_{22} of $A = \left(\begin{smallmatrix} A_{11} & A_{12}\\ A_{21} & A_{22}\end{smallmatrix}\right)$. We will come across this phenomenon several times in this work. In order to simplify terminology, we make a general definition.

Definition 1.2.10 (Left- and Right-Regularity)**.** *A matrix* $A = \left(\begin{smallmatrix} A_{11} & A_{12}\\ A_{21} & A_{22}\end{smallmatrix}\right) \in \mathbb{R}^{2d\times 2d}$ *is called*

- **left-regular**, *if the submatrices* $A_{11}, A_{21} \in \mathbb{R}^{d\times d}$ *are invertible;*

- **right-regular**, *if the submatrices* $A_{12}, A_{22} \in \mathbb{R}^{d\times d}$ *are invertible.*

The next theorem expresses a connection between right-regularity and left-regularity that will be useful later.

Theorem 1.2.11. *Let* $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d\times 2d}$ *be invertible. Let* $B = \left(\begin{smallmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{smallmatrix}\right) = (A^{-1})^*$. *Then* A *is right-regular (respectively left-regular) if and only if* B *is left-regular (respectively right-regular).*

Proof. We show that A is right-regular if and only if B is left-regular. The analogous statement where right-regular and left-regular are interchanged follows from this by changing the roles of A and B (observe that if $B = (A^{-1})^*$, then also $A = (B^{-1})^*$).
In order to simplify the notation, we write

$$A = \left(\begin{smallmatrix} X & U \\ Y & V \end{smallmatrix}\right) \quad \text{and} \quad A^{-1} = C = \left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right).$$

Since $B = C^* = \left(\begin{smallmatrix} P^* & R^* \\ Q^* & S^* \end{smallmatrix}\right)$, B is left-regular if and only P and Q are invertible. Thus we have to show that U and V are invertible if and only if P and Q are invertible.
Assume first that U and V are invertible. Then, since

$$CA = \left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right)\left(\begin{smallmatrix} X & U \\ Y & V \end{smallmatrix}\right) = \left(\begin{smallmatrix} I & 0 \\ 0 & I \end{smallmatrix}\right),$$

we have

$$PU + QV = 0$$

and

$$PX + QY = I.$$

The former is equivalent to $QV = -PU$ and $Q = -PUV^{-1}$; inserting this in the latter yields $PX - PUV^{-1}Y = I$, thus

$$P(X - UV^{-1}Y) = I. \tag{$*$}$$

On the other hand, we also have

$$AC = \left(\begin{smallmatrix} X & U \\ Y & V \end{smallmatrix}\right)\left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right) = \left(\begin{smallmatrix} I & 0 \\ 0 & I \end{smallmatrix}\right),$$

hence

$$XP + UR = I$$

and

$$YP + VR = 0.$$

Here the latter yields $VR = -YP$ and $R = -V^{-1}YP$, which upon inserting into the former gives $XP - UV^{-1}YP = I$, thus

$$(X - UV^{-1}Y)P = I. \qquad (**)$$

The equalities $(*)$ and $(**)$ together show that P is invertible with inverse $P^{-1} = X - UV^{-1}Y$, hence also $Q = -PUV^{-1}$ is invertible.
For the opposite implication, assume P and Q are invertible. Then it is not hard to see that from

$$AC = \left(\begin{smallmatrix} X & U \\ Y & V \end{smallmatrix}\right)\left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right) = \left(\begin{smallmatrix} I & 0 \\ 0 & I \end{smallmatrix}\right) = \left(\begin{smallmatrix} P & Q \\ R & S \end{smallmatrix}\right)\left(\begin{smallmatrix} X & U \\ Y & V \end{smallmatrix}\right) = CA$$

it follows that

$$\left(\begin{smallmatrix} R^* & P^* \\ S^* & Q^* \end{smallmatrix}\right)\left(\begin{smallmatrix} U^* & V^* \\ X^* & Y^* \end{smallmatrix}\right) = \left(\begin{smallmatrix} I & 0 \\ 0 & I \end{smallmatrix}\right) = \left(\begin{smallmatrix} U^* & V^* \\ X^* & Y^* \end{smallmatrix}\right)\left(\begin{smallmatrix} R^* & P^* \\ S^* & Q^* \end{smallmatrix}\right).$$

By what we have already shown, the invertibility of P^* and Q^* implies the invertibility of U^* and V^*. Hence, P and Q invertible implies U and V invertible. □

Theorem 1.2.12 (Interchanging f and g)**.** *For $f, g \in L^2(\mathbb{R}^d)$ and $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d\times 2d}$,*

$$\mathrm{TF}_A(g,f)(x,\omega) = \overline{\mathrm{TF}_B(f,g)(x,\omega)}$$

with $B = \left(\begin{smallmatrix} A_{21} & -A_{22} \\ A_{11} & -A_{12} \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right) \cdot A \cdot \left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right)$.

Proof. Observe that

$$\begin{aligned} \mathcal{T}_A(g \otimes \overline{f})(x,y) &= g(A_{11}x + A_{12}y)\overline{f(A_{21}x + A_{22}y)} \\ &= \overline{f(A_{21}x + A_{22}y)\overline{g(A_{11}x + A_{12}y)}} = \overline{\mathcal{T}_{B'}(f \otimes \overline{g})(x,y)} \end{aligned}$$

with $B' = \left(\begin{smallmatrix} A_{21} & A_{22} \\ A_{11} & A_{12} \end{smallmatrix}\right)$. It follows

$$\begin{aligned} \mathrm{TF}_A(g,f)(x,\omega) &= \mathcal{F}_2\mathcal{T}_A(g \otimes \overline{f})(x,\omega) = \mathcal{F}_2\overline{\mathcal{T}_{B'}(f \otimes \overline{g})}(x,\omega) \\ &= \overline{\mathcal{F}_2\mathcal{T}_{B'}(f \otimes \overline{g})(x,-\omega)} = \overline{\mathcal{F}_2\mathcal{T}_B(f \otimes \overline{g})(x,\omega)} \end{aligned}$$

with $B = \left(\begin{smallmatrix} A_{21} & -A_{22} \\ A_{11} & -A_{12} \end{smallmatrix}\right)$. □

One may ask whether it is possible to consider also bilinear time-frequency distributions defined by using the partial Fourier transform $\mathcal{F}_1$ in the *first* argument as opposed to our definition using $\mathcal{F}_2$. The following shows that both formulations are essentially equivalent and that both definitions lead to completely analogous theories.

In order to simplify the notation, we introduce:

Definition 1.2.13 (Flip)**.** *Let $F(x,y)$ be a function on $\mathbb{R}^{2d}$. The* **flip** *operator is defined by*

$$\tilde{F}(x,y) := F(y,x),$$

i.e. interchanges the arguments x and y.

The flip operator is a specific coordinate transformation:

$$\tilde{F}(x,y) = F(y,x) = \mathcal{T}_{\tilde{I}} F(x,y)$$

with matrix $\tilde{I} = \left(\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right)$.

Lemma 1.2.14. *Let $F = F(x,y) \in L^2(\mathbb{R}^{2d})$. Then*

$$\mathcal{F}_1 F(\xi, y) = \mathcal{F}_2 \tilde{F}(y, \xi).$$

Proof. Let $F \in \mathcal{S}(\mathbb{R}^{2d})$. Then

$$\begin{aligned} \mathcal{F}_1 F(\xi, y) &= \int_{\mathbb{R}^d} F(x,y) e^{-2\pi i x \cdot \xi} \, dx \\ &= \int_{\mathbb{R}^d} \tilde{F}(y,x) e^{-2\pi i x \cdot \xi} \, dx = \mathcal{F}_2 \tilde{F}(y, \xi). \end{aligned}$$

Since $\mathcal{S}(\mathbb{R}^{2d}) \subset L^2(\mathbb{R}^{2d})$ is a dense subspace, the assertion holds for all $f \in L^2(\mathbb{R}^{2d})$ by the standard density argument. $\square$

Lemma 1.2.15. *Let $f, g \in L^2(\mathbb{R}^d)$ and $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d \times 2d}$ with $\det A \neq 0$. Then*

$$\mathcal{F}_1 \mathcal{T}_A (f \otimes \overline{g})(\xi, y) = \mathcal{F}_2 \mathcal{T}_B (f \otimes \overline{g})(y, \xi) = \mathcal{F}_2 \widetilde{\mathcal{T}_B (f \otimes \overline{g})}(\xi, y)$$

with $B = \left(\begin{smallmatrix} A_{12} & A_{11} \\ A_{22} & A_{21} \end{smallmatrix}\right) = A \cdot \left(\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right)$.

Proof. By the previous Lemma 1.2.14,

$$\begin{aligned}\mathcal{F}_1\mathcal{T}_A(f\otimes\overline{g})(\xi,y) &= \mathcal{F}_2\left(\mathcal{T}_A\widetilde{(f\otimes\overline{g})}\right)(y,\xi)\\ &= \mathcal{F}_2\mathcal{T}_{\tilde{I}}\mathcal{T}_A(f\otimes\overline{g})(y,\xi),\end{aligned}$$

where $\tilde{I}$ denotes the permutation matrix $\left(\begin{smallmatrix}0 & I\\ I & 0\end{smallmatrix}\right)$. But

$$\mathcal{T}_{\tilde{I}}\mathcal{T}_A = \mathcal{T}_{A\cdot\tilde{I}} = \mathcal{T}_B,$$

which concludes the proof. □

Theorem 1.2.16 (Fourier Transform of a Time-Frequency Distribution)**.** *For $f,g\in L^2(\mathbb{R}^d)$ and $A=\left(\begin{smallmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{smallmatrix}\right)\in\mathbb{R}^{2d\times 2d}$,*

$$\widehat{\mathrm{TF}_A(f,g)}(\xi,\eta) = \mathrm{TF}_B(f,g)(\eta,\xi) = \widetilde{\mathrm{TF}_B(f,g)}(\xi,\eta).$$

with $B=\left(\begin{smallmatrix}-A_{12} & A_{11}\\ -A_{22} & A_{21}\end{smallmatrix}\right)=A\cdot\left(\begin{smallmatrix}0 & I\\ -I & 0\end{smallmatrix}\right)$.

Proof. Assume again that $f,g\in\mathcal{S}(\mathbb{R}^d)$. Then $\mathrm{TF}_A(f,g)\in\mathcal{S}(\mathbb{R}^{2d})$ as well, by Theorem 1.2.4. Using Fubini's Theorem and the Fourier Inversion Formula, we compute

$$\begin{aligned}\widehat{\mathrm{TF}_A(f,g)}(\xi,\eta) &= \iint_{\mathbb{R}^{2d}} \mathcal{F}_2\mathcal{T}_A(f\otimes\overline{g})(x,\omega)e^{-2\pi i(x\cdot\xi+\omega\cdot\eta)}\,dxd\omega\\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathcal{F}_2\mathcal{T}_A(f\otimes\overline{g})(x,\omega)e^{-2\pi i\omega\cdot\eta}\,d\omega\cdot e^{-2\pi ix\cdot\xi}\,dx\\ &= \int_{\mathbb{R}^d}\mathcal{T}_A(f\otimes\overline{g})(x,-\eta)e^{-2\pi ix\cdot\xi}\,dx\\ &= \mathcal{F}_1\mathcal{T}_A(f\otimes\overline{g})(\xi,-\eta)\\ &= \mathcal{F}_2\mathcal{T}_{\tilde{B}}(f\otimes\overline{g})(-\eta,\xi)\end{aligned}$$

with $\tilde{B}=A\cdot\left(\begin{smallmatrix}0 & I\\ I & 0\end{smallmatrix}\right)$ as in Lemma 1.2.15.
The last expression can also be written as

$$\mathcal{F}_2\mathcal{T}_{\tilde{B}}(f\otimes\overline{g})(-\eta,\xi) = \mathcal{F}_2\mathcal{T}_B(f\otimes\overline{g})(\eta,\xi)$$

where B denotes the matrix

$$\tilde{B}\cdot\left(\begin{smallmatrix}-I & 0\\ 0 & I\end{smallmatrix}\right) = A\cdot\left(\begin{smallmatrix}0 & I\\ I & 0\end{smallmatrix}\right)\cdot\left(\begin{smallmatrix}-I & 0\\ 0 & I\end{smallmatrix}\right) = A\cdot\left(\begin{smallmatrix}0 & I\\ -I & 0\end{smallmatrix}\right).$$

□

1.3 Orthogonality Relation and Inversion Formula

Bilinear time-frequency distributions take two functions as their arguments: they map a pair of functions of d variables to a single function of $2d$ variables. In many situations, however, it is more convenient to take up another viewpoint. One may consider one of the two functions as a window function, a fixed parameter of the mapping and not a variable. That means instead of looking at

$$\mathrm{TF}_A : L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^{2d})$$

we now consider a mapping

$$\mathrm{TF}_{A,g} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^{2d}), \quad f \mapsto \mathrm{TF}_{A,g}(f) = \mathrm{TF}_A(f, g).$$

In this sense the function $\mathrm{TF}_{A,g}(f)$ can be thought of as a transformation of f very much similar to other well-known integral transforms like e.g. the Fourier transform.

In this section we will prove two important properties of such time-frequency transformations, that are analogues of well-known properties enjoyed by the Fourier transform. The first is the orthogonality relation for time-frequency distributions, a counterpart of Parseval's formula. In the presence of such a formula there can usually be derived an inversion theorem that allows to reconstruct the original function f from its time-frequency transformation $\mathrm{TF}_{A,g}(f)$.

Theorem 1.3.1 (Orthogonality Relation). *Let* $f_1, f_2, g_1, g_2 \in L^2(\mathbb{R}^d)$ *and* $A \in \mathbb{R}^{2d\times 2d}$ *with* $\det A \neq 0$. *Then*

$$\langle \mathrm{TF}_A(f_1, g_1), \mathrm{TF}_A(f_2, g_2)\rangle = \frac{1}{|\det A|} \langle f_1, f_2\rangle \, \overline{\langle g_1, g_2\rangle}.$$

Proof. Since $\mathcal{F}_2$ is a unitary operator on $L^2(\mathbb{R}^{2d})$ and $\mathcal{T}_A$ is unitary up to a constant factor,

$$\begin{aligned}\langle \mathrm{TF}_A(f_1,g_1),\mathrm{TF}_A(f_2,g_2)\rangle &= \langle \mathcal{F}_2\mathcal{T}_A(f_1\otimes\overline{g_1}),\mathcal{F}_2\mathcal{T}_A(f_2\otimes\overline{g_2})\rangle\\ &= \langle \mathcal{T}_A(f_1\otimes\overline{g_1}),\mathcal{T}_A(f_2\otimes\overline{g_2})\rangle\\ &= \frac{1}{|\det A|}\,\langle f_1\otimes\overline{g_1},f_2\otimes\overline{g_2}\rangle\\ &= \frac{1}{|\det A|}\,\langle f_1,f_2\rangle\,\overline{\langle g_1,g_2\rangle}.\end{aligned}$$

□

In particular we have

$$\|\mathrm{TF}_{A,g}f\| = \frac{1}{|\det A|^{1/2}}\cdot\|f\|\cdot\|g\|;$$

this shows that $\mathrm{TF}_{A,g}$ is a multiple of an isometry on $L^2(\mathbb{R}^d)$.

If $(e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of $L^2(\mathbb{R}^d)$, then

$$\langle \mathrm{TF}_A(e_n,e_m),\mathrm{TF}_A(e_k,e_l)\rangle = \frac{1}{|\det A|}\,\langle e_n,e_k\rangle\,\langle e_m,e_l\rangle = \frac{1}{|\det A|}\delta_{n,k}\delta_{m,l}.$$

Thus $\{\mathrm{TF}_A(e_n,e_m)|\ n,m\in\mathbb{N}\}$ is an orthogonal family in $L^2(\mathbb{R}^{2d})$. By Theorem 1.2.2, it is also a complete subset, hence (up to the constant factor $\frac{1}{|\det A|}$) an orthogonal basis for $L^2(\mathbb{R}^{2d})$.

Next we prove an explicit formula for the adjoint of $\mathrm{TF}_{A,g}$.

Theorem 1.3.2. *The adjoint of the operator* $\mathrm{TF}_{A,g}$ *is given by*

$$\mathrm{TF}^*_{A,g}: L^2(\mathbb{R}^{2d})\to L^2(\mathbb{R}^d),\quad \mathrm{TF}^*_{A,g}H(x) = \int_{\mathbb{R}^d}\mathcal{T}_C\mathcal{F}_2H(x,y)\cdot g(y)\,dy,$$

where C *denotes the matrix* $C = A^{-1}\cdot\left(\begin{smallmatrix} I & 0\\ 0 & -I\end{smallmatrix}\right)$.

Proof. For convenience, denote

$$\int_{\mathbb{R}^d}\mathcal{T}_C\mathcal{F}_2H(x,y)\cdot g(y)\,dy = h(x);$$

note that if $H \in L^2(\mathbb{R}^{2d})$, then by Fubini's Theorem $h(x)$ is defined for almost all $x \in \mathbb{R}^d$ and h is a well-defined function in $L^2(\mathbb{R}^d)$.
Now let $f \in L^2(\mathbb{R}^d)$ and $H \in L^2(\mathbb{R}^{2d})$. Then

$$\langle \mathrm{TF}_{A,g} f, H \rangle = \langle \mathcal{F}_2 \mathcal{T}_A (f \otimes \overline{g}), H \rangle = \langle \mathcal{T}_A (f \otimes \overline{g}), \mathcal{F}_2^* H \rangle .$$

But $\mathcal{F}_2^* = \mathcal{T}_B \mathcal{F}_2$ with $B = \left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right)$ by Lemma A.4.2. This yields

$$\begin{aligned} \langle \mathcal{T}_A (f \otimes \overline{g}), \mathcal{F}_2^* H \rangle &= \langle \mathcal{T}_A (f \otimes \overline{g}), \mathcal{T}_B \mathcal{F}_2 H \rangle = \langle f \otimes \overline{g}, \mathcal{T}_{A^{-1}} \mathcal{T}_B \mathcal{F}_2 H \rangle \\ &= \langle f \otimes \overline{g}, \mathcal{T}_C \mathcal{F}_2 H \rangle \end{aligned}$$

with $C = A^{-1} B = A^{-1} \cdot \left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right)$.
Finally, by Fubini's Theorem, we get

$$\begin{aligned} \langle f \otimes \overline{g}, \mathcal{T}_C \mathcal{F}_2 H \rangle &= \iint_{\mathbb{R}^{2d}} f(x) \overline{g(y)} \cdot \overline{\mathcal{T}_C \mathcal{F}_2 H(x,y)} \, dx dy \\ &= \int_{rd} f(x) \overline{\int_{\mathbb{R}^d} \mathcal{T}_C \mathcal{F}_2 H(x,y) \cdot g(y) \, dy} \, dx \\ &= \langle f, h \rangle . \end{aligned}$$

□

Now assume that $A : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is a bounded operator that is a constant multiple of an isometry, i.e. $\|Af\| = c \cdot \|f\|$ for all $f \in L^2(\mathbb{R}^d)$. There is the following canonical method for inverting A:

$$f = \frac{1}{c^2} A^* A f \qquad \text{for all } f \in L^2(\mathbb{R}^d).$$

The proof is very easy: we have

$$\langle A^* A f, h \rangle = \langle Af, Ah \rangle$$

for all $f, h \in L^2(\mathbb{R}^d)$. But by polarization,

$$\begin{aligned} \langle Af, Ah \rangle &= \frac{1}{4} \sum_{\zeta^4 = 1} \zeta \cdot \|Af + \zeta Ah\|^2 \qquad \text{(sum over the fourth roots of unity)} \\ &= \frac{1}{4} \sum_{\zeta^4 = 1} \zeta \cdot c^2 \cdot \|f + \zeta h\|^2 \\ &= c^2 \langle f, h \rangle . \end{aligned}$$

This yields

$$\langle A^*Af, h\rangle = c^2 \langle f, h\rangle$$

for all $f, h \in L^2(\mathbb{R}^d)$, and therefore

$$A^*Af = c^2 f$$

for all $f \in L^2(\mathbb{R}^d)$.

Slightly more general, we compute for bilinear time-frequency distributions:

$$\left\langle \mathrm{TF}^*_{A,\gamma}\mathrm{TF}_{A,g}f, h\right\rangle = \langle \mathrm{TF}_{A,g}f, \mathrm{TF}_{A,\gamma}h\rangle = \langle \gamma, g\rangle \langle f, h\rangle$$

for arbitrary $f, h, g, \gamma \in L^2(\mathbb{R}^d)$. This implies the following inversion formula:

$$f = \frac{1}{\langle \gamma, g\rangle}\mathrm{TF}^*_{A,\gamma}\mathrm{TF}_{A,g}f$$

for all $f \in L^2(\mathbb{R}^d)$ (where we assume $\langle \gamma, g\rangle \neq 0$).

A more explicit version of this is presented in the next theorem.

Theorem 1.3.3 (Inversion Formula)**.** *Let $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{smallmatrix}\right) \in \mathbb{R}^{2d\times 2d}$ be invertible and right-regular. Let $g, \gamma \in L^2(\mathbb{R}^d)$ be such that $\langle g, \gamma\rangle \neq 0$. Then the following inversion formula holds for all $f \in L^2(\mathbb{R}^d)$:*

$$f = \frac{1}{\langle g, \gamma\rangle}\iint_{\mathbb{R}^{2d}} \mathrm{TF}_{A,\gamma}f(x,\omega)\frac{-e^{2\pi i (A_{12}^*)^{-1}\omega\cdot A_{11}x}}{|\det A_{12}|}M_{d(\omega)}T_{c(x)}\tilde{g}\, dx d\omega,$$

with $\tilde{g}(z) := g(A_{22}A_{12}^{-1}z)$ and

$$d(\omega) = (A_{12}^*)^{-1}\omega, \qquad c(x) = (A_{11} - A_{12}A_{22}^{-1}A_{21})x,$$

as in Theorem 1.2.5. The integral is to be understood as a vector valued integral in $L^2(\mathbb{R}^d)$, defined in a weak sense.

Proof. The statement is an immediate consequence of the representation formula Theorem 1.2.5. Denote the function in $L^2(\mathbb{R}^d)$ defined by the vector valued integral on the righthand side by $\tilde{f}$ for the moment. Then by the weak definition of vector valued integration, we have for all $h \in L^2(\mathbb{R}^d)$

$$\begin{aligned}\left\langle \tilde{f}, h \right\rangle &= \frac{1}{\langle g, \gamma \rangle} \iint_{\mathbb{R}^{2d}} \mathrm{TF}_{A,\gamma} f(x,\omega) \frac{-e^{2\pi i (A_{12}^*)^{-1}\omega \cdot A_{11} x}}{|\det A_{12}|} \left\langle M_{d(\omega)} T_{c(x)} \tilde{g}, h \right\rangle dx d\omega \\ &= \frac{1}{\langle g, \gamma \rangle} \iint_{\mathbb{R}^{2d}} \mathrm{TF}_{A,\gamma} f(x,\omega) \overline{\frac{e^{2\pi i (A_{12}^*)^{-1}\omega \cdot A_{11} x}}{|\det A_{12}|} \left\langle h, M_{d(\omega)} T_{c(x)} \tilde{g} \right\rangle} dx d\omega \\ &= \frac{1}{\langle g, \gamma \rangle} \iint_{\mathbb{R}^{2d}} \mathrm{TF}_{A,\gamma} f(x,\omega) \overline{\mathrm{TF}_{A,g} h(x,\omega)} \, dx d\omega.\end{aligned}$$

The orthogonality relation yields

$$\begin{aligned}\frac{1}{\langle g, \gamma \rangle} &\iint_{\mathbb{R}^{2d}} \mathrm{TF}_{A,\gamma} f(x,\omega) \overline{\mathrm{TF}_{A,g} h(x,\omega)} \, dx d\omega \\ &= \frac{1}{\langle g, \gamma \rangle} \langle \mathrm{TF}_{A,\gamma} f, \mathrm{TF}_{A,g} h \rangle \\ &= \langle f, h \rangle ;\end{aligned}$$

since this is true for arbitrary $h \in L^2(\mathbb{R}^d)$, we conclude $f = \tilde{f}$. □

Finally, we show a little result about a reproducing kernel property of the image subspaces of the transforms $\mathrm{TF}_{A,g}(L^2(\mathbb{R}^d))$.

Theorem 1.3.4 (Reproducing Kernel)**.** *Let $g \in L^2(\mathbb{R}^d) \setminus \{0\}$ and $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d \times 2d}$ be invertible and right-regular. The image $\mathrm{TF}_{A,g}(L^2(\mathbb{R}^d)) \subseteq L^2(\mathbb{R}^{2d})$ of the transformation $\mathrm{TF}_{A,g} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^{2d})$ is a closed subspace of $L^2(\mathbb{R}^{2d})$ consisting of continuous functions. It is a reproducing kernel Hilbert space. The kernel is given by*

$$K_{x_0,\omega_0}(x,\omega) = \frac{|\det A| \cdot e^{2\pi i (A_{12}^*)^{-1}\omega \cdot A_{11} x}}{|\det A_{12}| \cdot \|g\|^2} \mathrm{TF}_{A,g}(M_{d(\omega)} T_{c(x)} \tilde{g})(x,\omega),$$

where $\tilde{g}(z) := g(A_{22} A_{12}^{-1} z)$ and

$$d = d(\omega) = (A_{12}^*)^{-1}\omega, \qquad c = c(x) = (A_{11} - A_{12} A_{22}^{-1} A_{21}) x.$$

Proof. Since $\|\mathrm{TF}_{A,g} f\| = \|f\| \cdot \|g\|$, it is clear that the image of $\mathrm{TF}_{A,g}$ is a closed subspace of $L^2(\mathbb{R}^{2d})$. By Corollary 1.2.6, $\mathrm{TF}_{A,g} f$ is continuous for every $f \in L^2(\mathbb{R}^d)$. In order to prove that $\mathrm{TF}_{A,g}(L^2(\mathbb{R}^d))$ has a reproducing kernel, we have to show that point evaluations are bounded linear functionals on this space. Let $(x_0, \omega_0) \in \mathbb{R}^{2d}$ be arbitrary, then by the Representation Theorem 1.2.5

$$\begin{aligned}
|\mathrm{TF}_{A,g} f(x_0, \omega_0)| &= \frac{1}{|\det A_{12}|} \left| \left\langle f, M_{d(\omega_0)} T_{c(x_0)} \tilde{g} \right\rangle \right| \\
&\leq \frac{1}{|\det A_{12}|} \|f\| \cdot \|M_{d(\omega_0)} T_{c(x_0)} \tilde{g}\| \\
&\leq \frac{1}{|\det A_{12}|} \|f\| \cdot \|\tilde{g}\|.
\end{aligned}$$

Now

$$\|\tilde{g}(\cdot)\| = \|g(A_{22} A_{12}^{-1}) \cdot)\| = \frac{|\det A_{12}|}{|\det A_{22}|} \|g\|,$$

so

$$|\mathrm{TF}_{A,g} f(x_0, \omega_0)| \leq \frac{1}{|\det A_{22}|} \|f\| \cdot \|g\|.$$

Using again the Representation Theorem 1.2.5 and the orthogonality relation, it is easy to verify the reproducing property of the stated kernel

$$\langle \mathrm{TF}_{A,g} f, K_{x_0,\omega_0} \rangle = \mathrm{TF}_{A,g} f(x_0, \omega_0)$$

for all $(x_0, \omega_0) \in \mathbb{R}^{2d}$, $f \in L^2(\mathbb{R}^d)$, by direct calculation. □

1.4 The Uncertainty Principle

The prototype of a qualitative uncertainty principle for the Fourier transform is the following well-known theorem ([2], [1]).

Theorem 1.4.1 (Benedicks' Theorem)**.** *Let $f \in L^1(\mathbb{R}^d)$ be such that both the sets*

$$\{x : f(x) \neq 0\}$$

and

$$\{\xi : \hat{f}(\xi) \neq 0\}$$

have finite Lebesgue measure. Then $f = 0$. □

This statement was subsequently extended to some of the classical time-frequency distributions, in particular the short-time Fourier transform and the Wigner distribution. This was done independently by several authors, see [25], [26], [45]. A good survey is given in [18].

Theorem 1.4.2 ([25], [26], [45], [18])**.** *Let $f, g \in L^2(\mathbb{R}^d)$. Then the following are equivalent:*

1. *The support of the short-time Fourier transform supp$(V(f,g))$ has finite Lebesgue measure.*
2. *The support of the Wigner distribution supp$(W(f,g))$ has finite Lebesgue measure.*
3. *Either $f = 0$ or $g = 0$.* □

Since well-behaved generalized time-frequency distributions can be transformed to short-time Fourier transforms by the Representation Theorem 1.2.5, we immediately have the analogous statement for such time-frequency distributions:

Theorem 1.4.3 (Uncertainty Principle)**.** *Let $f, g \in L^2(\mathbb{R}^d)$ and $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d \times 2d}$ with $\det A \neq 0$ be right-regular. If the support of $\mathrm{TF}_A(f,g)$ has finite Lebesgue measure, then necessarily $f = 0$ or $g = 0$.*

Proof. By Theorem 1.2.5, we have

$$\begin{aligned}|\mathrm{TF}_A(f,g)(x,\omega)| &= \frac{1}{|\det A_{12}|}\cdot\left|\left\langle f, M_{d(\omega)}T_{c(x)}\tilde{g}\right\rangle\right| \\ &= \frac{1}{|\det A_{12}|}\cdot|V(f,\tilde{g})(c(x),d(\omega))|\end{aligned}$$

with $\tilde{g}(z) = g(A_{22}A_{12}^{-1}z)$ and $c(x) = Cx$, $d(\omega) = D\omega$ for some invertible matrices $C, D \in \mathbb{R}^{d\times d}$. In particular, $\mathrm{TF}_A(f,g)(x,\omega) \neq 0$ if and only if $V(f,\tilde{g})(Cx, D\omega) \neq 0$. But this implies that the support of $\mathrm{TF}_A(f,g)$ has finite measure if and only if the support of $V(f,\tilde{g})$ has finite measure. By the preceding theorem, this is equivalent to either f or $\tilde{g}$ being zero. Since $\tilde{g} = 0$ if and only if $g = 0$, the theorem follows. □

1.5 Covariance Property

The covariance property is an important technical tool for the following. It clarifies how general time-frequency distributions $\mathrm{TF}_A(f,g)$ behave under time-frequency shifts of the functions f and g.

Theorem 1.5.1 (Covariance Property)**.** *Let $f,g \in L^2(\mathbb{R}^d)$, $u,v,\eta,\gamma \in \mathbb{R}^d$ and $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d\times 2d}$. Then the following covariance formula holds:*

$$\begin{aligned}\mathrm{TF}_A(M_\eta T_u f, M_\gamma T_v g)(x,\omega) &= e^{2\pi i\sigma\cdot s} M_{\left(\begin{smallmatrix} \rho \\ -s \end{smallmatrix}\right)} T_{\left(\begin{smallmatrix} r \\ \sigma \end{smallmatrix}\right)} \mathrm{TF}_A(f,g)(x,\omega) \\ &= e^{2\pi i\sigma\cdot s} e^{2\pi i(x\cdot\rho - \omega\cdot s)} \mathrm{TF}_A(f,g)(x-r,\omega-\sigma)\end{aligned}$$

with

$$\left(\begin{smallmatrix} r \\ s \end{smallmatrix}\right) = A^{-1}\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)$$

and

$$\left(\begin{smallmatrix} \rho \\ \sigma \end{smallmatrix}\right) = A^*\left(\begin{smallmatrix} \eta \\ -\gamma \end{smallmatrix}\right) = \begin{pmatrix} A_{11}^*\eta - A_{21}^*\gamma \\ A_{12}^*\eta - A_{22}^*\gamma \end{pmatrix}.$$

Proof. Observe that

$$\mathrm{TF}_A(M_\eta T_u f, M_\gamma T_v g) = \mathcal{F}_2\mathcal{T}_A(M_{\left(\begin{smallmatrix} \eta \\ -\gamma \end{smallmatrix}\right)} T_{\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)} f\otimes\overline{g}).$$

Using Lemma A.3.3 and Lemma A.4.3 yields

$$\begin{aligned}\mathcal{F}_2\mathcal{T}_A(M_{\left(\begin{smallmatrix} \eta \\ -\gamma \end{smallmatrix}\right)} T_{\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)} f\otimes\overline{g}) &= \mathcal{F}_2 M_{\left(\begin{smallmatrix} \rho \\ \sigma \end{smallmatrix}\right)} \mathcal{T}_A(T_{\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)} f\otimes\overline{g}), \qquad \text{with } \left(\begin{smallmatrix} \rho \\ \sigma \end{smallmatrix}\right) = A^*\left(\begin{smallmatrix} \eta \\ -\gamma \end{smallmatrix}\right) \\ &= M_{\left(\begin{smallmatrix} \rho \\ 0 \end{smallmatrix}\right)} T_{\left(\begin{smallmatrix} 0 \\ \sigma \end{smallmatrix}\right)} \mathcal{F}_2\mathcal{T}_A(T_{\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)} f\otimes\overline{g}) \\ &= M_{\left(\begin{smallmatrix} \rho \\ 0 \end{smallmatrix}\right)} T_{\left(\begin{smallmatrix} 0 \\ \sigma \end{smallmatrix}\right)} \mathcal{F}_2 T_{\left(\begin{smallmatrix} r \\ s \end{smallmatrix}\right)} \mathcal{T}_A(f\otimes\overline{g}), \qquad \text{with } \left(\begin{smallmatrix} r \\ s \end{smallmatrix}\right) = A^{-1}\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right) \\ &= M_{\left(\begin{smallmatrix} \rho \\ 0 \end{smallmatrix}\right)} T_{\left(\begin{smallmatrix} 0 \\ \sigma \end{smallmatrix}\right)} M_{\left(\begin{smallmatrix} 0 \\ -s \end{smallmatrix}\right)} T_{\left(\begin{smallmatrix} r \\ 0 \end{smallmatrix}\right)} \mathcal{F}_2\mathcal{T}_A(f\otimes\overline{g});\end{aligned}$$

with the canonical commutation relation, this simplifies to

$$\mathrm{TF}_A(M_\eta T_u f, M_\gamma T_v g) = e^{2\pi i\sigma\cdot s} M_{\left(\begin{smallmatrix} \rho \\ -s \end{smallmatrix}\right)} T_{\left(\begin{smallmatrix} r \\ \sigma \end{smallmatrix}\right)} \mathcal{F}_2\mathcal{T}_A(f\otimes\overline{g}).$$

□

1.6 Marginal Densities and Cohen's Class

One possible way of thinking about joint time-frequency distributions is in terms of quantum mechanics as a joint probability density of position and momenentum for a particle whose state is described by the quantum mechanical wave function $f \in L^2(\mathbb{R}^3)$. In fact the quadratic Wigner distribution $W(f, f)$ was introduced with this idea in mind. This interpretation is, however, rather convenient heuristics than true in a strict mathematical sense. The (quadratic) generalized time-frequency distributions that we consider are sometimes lacking the most important features of a probability density function, e.g.

1. correct marginal densities;
2. positivity.

The question arises whether for certain special choices of the matrix A some or all of these requirements can be met.

Lemma 1.6.1 (Marginal Densities)**.** *Let* $f \in \mathcal{S}(\mathbb{R}^d)$*. Then*

$$\int_{\mathbb{R}^d} \mathrm{TF}_A(f,f)(x,\omega)\, d\omega = f(A_{11}x)\overline{f(A_{21}x)}$$

and

$$\int_{\mathbb{R}^d} \mathrm{TF}_A(f,f)(x,\omega)\, dx = \frac{1}{|\det A|} \widehat{f}(B_{12}\omega)\overline{\widehat{f}(-B_{22}\omega)}$$

with $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = (A^{-1})^*$.

Proof. If $f \in \mathcal{S}(\mathbb{R}^d)$ then $\mathrm{TF}_A(f,f) \in \mathcal{S}(\mathbb{R}^{2d})$. Using the Fourier Inversion Formula, we find

$$\begin{aligned}\int_{\mathbb{R}^d} \mathrm{TF}_A(f,f)(x,\omega)\, d\omega &= \int_{\mathbb{R}^d} \mathcal{F}_2\mathcal{T}_A(f\otimes\overline{f})(x,\omega)e^{2\pi i 0\cdot\omega}\, d\omega \\ &= \mathcal{F}_2^{-1}\mathcal{F}_2\mathcal{T}_A(f\otimes\overline{f})(x,0) \\ &= \mathcal{T}_A(f\otimes\overline{f})(x,0) \\ &= f(A_{11}x)\overline{f(A_{21}x)}.\end{aligned}$$

For the second part we proceed in a similar way:

$$\begin{aligned}\int_{\mathbb{R}^d} \mathrm{TF}_A(f,f)(x,\omega)\,dx &= \int_{\mathbb{R}^d} \mathcal{F}_2\mathcal{T}_A(f\otimes\overline{f})(x,\omega)e^{-2\pi i 0\cdot x}\,dx \\ &= \mathcal{F}_1\mathcal{F}_2\mathcal{T}_A(f\otimes\overline{f})(0,\omega) \\ &= \widehat{\mathcal{T}_A(f\otimes\overline{f})}(0,\omega).\end{aligned}$$

Lemma A.2.4 shows

$$\begin{aligned}\widehat{\mathcal{T}_A(f\otimes\overline{f})}(0,\omega) &= \frac{1}{|\det A|}\mathcal{T}_{(A^{-1})^*}(\widehat{f\otimes\overline{f}})(0,\omega) \\ &= \frac{1}{|\det A|}\mathcal{T}_{(A^{-1})^*}(\widehat{f}\otimes\widehat{\overline{f}})(0,\omega) \\ &= \frac{1}{|\det A|}\mathcal{T}_{(A^{-1})^*}(\widehat{f}\otimes\widehat{\tilde{f}})(0,\omega),\end{aligned}$$

where $\tilde{h}(z) = \overline{h(-z)}$ denotes the usual Fourier involution.
Putting $(A^{-1})^* = B = \left(\begin{smallmatrix} B_{11} & B_{12} \\ B_{21} & B_{22}\end{smallmatrix}\right)$, the last expression equals

$$\frac{1}{|\det A|}\mathcal{T}_{(A^{-1})^*}(\widehat{f}\otimes\widehat{\tilde{f}})(0,\omega) = \frac{1}{|\det A|}\widehat{f}(B_{12}\omega)\overline{\widehat{f}(-B_{22}\omega)}.$$

□

Theorem 1.6.2. *Let $A \in \mathbb{R}^{2d\times 2d}$ be of the form $A = \left(\begin{smallmatrix} I & I+V \\ I & V\end{smallmatrix}\right)$ for some arbitrary matrix $V \in \mathbb{R}^{d\times d}$. Then we have*

$$\int_{\mathbb{R}^d} \mathrm{TF}_A(f,f)(x,\omega)\,d\omega = |f(x)|^2$$

and

$$\int_{\mathbb{R}^d} \mathrm{TF}_A(f,f)(x,\omega)\,dx = |\widehat{f}(\omega)|^2$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$, $x,\omega \in \mathbb{R}^d$.

Proof. If $A = \left(\begin{smallmatrix} I & I+V \\ I & V\end{smallmatrix}\right)$ for some $V \in \mathbb{R}^{d\times d}$, then $|\det A| = 1$ since

$$\det A = \det\left(\begin{smallmatrix} I & I+V \\ I & V\end{smallmatrix}\right) = \det\left(\begin{smallmatrix} I & I+V \\ 0 & -I\end{smallmatrix}\right) = 1^d\cdot(-1)^d.$$

One easily verifies that

$$\begin{pmatrix} I & I+V \\ I & V \end{pmatrix} \cdot \begin{pmatrix} I+V & -V \\ I & -I \end{pmatrix} = \begin{pmatrix} I+V & -V \\ I & -I \end{pmatrix} \cdot \begin{pmatrix} I & I+V \\ I & V \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

so

$$A^{-1} = \begin{pmatrix} I+V & -V \\ I & -I \end{pmatrix}.$$

When we denote

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = (A^{-1})^* = \begin{pmatrix} (I+V)^* & I^* \\ -V^* & -I^* \end{pmatrix},$$

then obviously

$$A_{11} = A_{21} = I, \qquad B_{12} = -B_{22} = I;$$

the statement follows now from Lemma 1.6.1. □

There is a class of bilinear time-frequency distributions with particularly nice properties that is closely related to the Wigner distribution and inherits many of its desirable features. Cohen's class consists of appropriately smoothed versions of the quadratic Wigner distribution.

Definition 1.6.3 (Cohen's Class)**.** *A time-frequency disribution $Q(f,g)$ belongs to* **Cohen's class** *if it satisfies a relation of the form*

$$Q(f,f) = W(f,f) * \sigma$$

for some distribution $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$, where $W(f,f)$ denotes the quadratic Wigner distribution.

Without proof, we cite the following result, cf [17]:

Theorem 1.6.4. *Suppose that a bilinear time-frequency distribution $Q(f,g) \in L^2(\mathbb{R}^{2d})$, $f,g \in L^2(\mathbb{R}^d)$, is covariant, i.e.*

$$Q(M_\omega T_x f, M_\omega T_x f) = T_{(x,\omega)} Q(f,f),$$

and satisfies the weak continuity condition

$$|Q(f,g)(0,0)| \leq c \cdot \|f\| \cdot \|g\|$$

for some $c \geq 0$ and all $f,g \in L^2(\mathbb{R}^d)$. Then $Q(f,g)$ belongs to Cohen's class. □

With this sufficient condition, we can prove the following:

Theorem 1.6.5. *The time-frequency representation* $\mathrm{TF}_A(f,g)$ *belongs to Cohen's class if the matrix* A *is of the form*

$$A = \begin{pmatrix} I & B \\ I & B-I \end{pmatrix},$$

where B *denotes any invertible matrix in* $\mathbb{R}^{d\times d}$ *such that also* $B-I$ *is invertible.*

Proof. In order to use the preceding theorem, we have to show, that the stated condition on the matrix A implies covariance and weak continuity of $\mathrm{TF}(f,g)$. The latter is clear from Corollary 1.2.7, since

$$|\mathrm{TF}(f,g)(0,0)| \leq \frac{\|f\|\cdot\|g\|}{|\det B|^{1/2}\cdot|\det(I-B)|^{1/2}}.$$

The former follows from the covariance formula of Theorem 1.5.1. We have

$$\mathrm{TF}_A(M_\eta T_u f, M_\eta T_u g)(x,\omega) = e^{2\pi i\sigma\cdot s} M_{\left(\begin{smallmatrix}\rho \\ -s\end{smallmatrix}\right)} T_{\left(\begin{smallmatrix}r \\ \sigma\end{smallmatrix}\right)} \mathrm{TF}_A(f,g)(x,\omega)$$

with

$$\begin{pmatrix}\rho \\ \sigma\end{pmatrix} = A^* \begin{pmatrix}\eta \\ -\eta\end{pmatrix} = \begin{pmatrix} I^*\eta - I^*\eta \\ B^*\eta-(B-I)^*\eta\end{pmatrix} = \begin{pmatrix}0 \\ \eta\end{pmatrix}$$

and

$$\begin{pmatrix}r \\ s\end{pmatrix} = A^{-1}\begin{pmatrix}u \\ u\end{pmatrix}.$$

The inverse matrix A^{-1} is given explicitely as

$$A^{-1} = \begin{pmatrix} I-B & B \\ I & -I\end{pmatrix},$$

which can be verified by direct computation. Therefore

$$\begin{pmatrix}r \\ s\end{pmatrix} = \begin{pmatrix} I-B & B \\ I & -I\end{pmatrix}\begin{pmatrix}u \\ u\end{pmatrix} = \begin{pmatrix}(I-B)u+Bu \\ u-u\end{pmatrix} = \begin{pmatrix}u \\ 0\end{pmatrix}.$$

Plugging this into the covariance formula above yields

$$\begin{aligned}\mathrm{TF}_A(M_\eta T_u f, M_\eta T_u g)(x,\omega) &= M_{\left(\begin{smallmatrix}0 \\ 0\end{smallmatrix}\right)} T_{\left(\begin{smallmatrix}u \\ \eta\end{smallmatrix}\right)} \mathrm{TF}_A(f,g)(x,\omega) \\ &= T_{\left(\begin{smallmatrix}u \\ \eta\end{smallmatrix}\right)} \mathrm{TF}_A(f,g)(x,\omega),\end{aligned}$$

that is the property of covariance. □

1.7 The "Magic Formula"

In this section, we give a result that will prove to be an extremely useful tool in the following. Several versions of this identity have appeared in the literature, see e.g. [6] for the case of Rihacek distributions. The formula is called "magic" since it constitutes the universal technical backbone of many proofs in time-frequency analysis.

Theorem 1.7.1 (Magic Formula). *Let* $f, g, \phi, \psi \in L^2(\mathbb{R}^d)$, $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \in \mathbb{R}^{2d\times 2d}$ *with* $|\det A| \neq 0$, *and* $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$, $\zeta = \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}$.
Then

$$V\left(TF_A(f,g), TF_A(\phi,\psi)\right)(z,\zeta) \; = \; e^{-2\pi i z_2\zeta_2} V(f,\phi)(u,\eta)\overline{V(g,\psi)(v,\gamma)},$$

where

$$\begin{pmatrix} u \\ v \end{pmatrix} = A \cdot \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ \zeta_2 \end{pmatrix}$$

and

$$\begin{pmatrix} \eta \\ \gamma \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \cdot (A^{-1})^* \cdot \begin{pmatrix} \zeta_1 \\ z_2 \end{pmatrix}.$$

Proof. We start on the right-hand side:

$$\begin{aligned} V(f,g)(u,\eta)\overline{V(\phi,\psi)(v,\gamma)} &= \langle f, M_\eta T_u \phi\rangle \overline{\langle g, M_\gamma T_v \psi\rangle} \\ &= \langle \mathrm{TF}_A(f,g), \mathrm{TF}_A(M_\eta T_u\phi, M_\gamma T_v\psi)\rangle \end{aligned}$$

by the orthogonality relation Theorem 1.3.1.
The covariance formula Theorem 1.5.1 gives for the second term in the inner product

$$\mathrm{TF}_A(M_\eta T_u\phi, M_\gamma T_v\psi) = e^{2\pi i\sigma\cdot s} M_{\begin{pmatrix} \rho \\ -s \end{pmatrix}} T_{\begin{pmatrix} r \\ \sigma \end{pmatrix}} \mathrm{TF}_A(\phi,\psi),$$

with

$$\begin{pmatrix} r \\ s \end{pmatrix} = A^{-1} \begin{pmatrix} u \\ v \end{pmatrix}$$

and

$$\begin{pmatrix} \rho \\ \sigma \end{pmatrix} = A^* \begin{pmatrix} \eta \\ -\gamma \end{pmatrix} = A^* \cdot \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \cdot \begin{pmatrix} \eta \\ \gamma \end{pmatrix}.$$

Inserting this in the inner product, we find

$$\begin{aligned}&\langle \mathrm{TF}_A(f,g), \mathrm{TF}_A(M_\eta T_u\phi, M_\gamma T_v\psi)\rangle\\ &\quad= \left\langle \mathrm{TF}_A(f,g), e^{2\pi i\sigma\cdot s} M_{\left(\begin{smallmatrix}\rho\\ -s\end{smallmatrix}\right)} T_{\left(\begin{smallmatrix}r\\ \sigma\end{smallmatrix}\right)} \mathrm{TF}_A(\phi,\psi)\right\rangle\\ &\quad= e^{-2\pi i\sigma\cdot s}\left\langle \mathrm{TF}_A(f,g), M_{\left(\begin{smallmatrix}\rho\\ -s\end{smallmatrix}\right)} T_{\left(\begin{smallmatrix}r\\ \sigma\end{smallmatrix}\right)} \mathrm{TF}_A(\phi,\psi)\right\rangle.\end{aligned}$$

But this is just the same as

$$\begin{aligned}&e^{-2\pi i\sigma\cdot s}\left\langle \mathrm{TF}_A(f,g), M_{\left(\begin{smallmatrix}\rho\\ -s\end{smallmatrix}\right)} T_{\left(\begin{smallmatrix}r\\ \sigma\end{smallmatrix}\right)} \mathrm{TF}_A(\phi,\psi)\right\rangle\\ &\qquad= e^{-2\pi i\sigma\cdot s} V(\mathrm{TF}_A(f,g), \mathrm{TF}_A(\phi,\psi))\left(\left(\begin{smallmatrix}r\\ \sigma\end{smallmatrix}\right), \left(\begin{smallmatrix}\rho\\ -s\end{smallmatrix}\right)\right).\end{aligned}$$

The stated result now follows easily by identifying

$$z = \left(\begin{smallmatrix}z_1\\ z_2\end{smallmatrix}\right) = \left(\begin{smallmatrix}r\\ \sigma\end{smallmatrix}\right)$$

and

$$\zeta = \left(\begin{smallmatrix}\zeta_1\\ \zeta_2\end{smallmatrix}\right) = \left(\begin{smallmatrix}\rho\\ -s\end{smallmatrix}\right).$$

□

As an immediate application, we give two regularity results that will be needed later.

Theorem 1.7.2. *Let* $A \in \mathbb{R}^{2d\times 2d}$ *be invertible and* $f, g \in M^1(\mathbb{R}^d)$*. Then* $\mathrm{TF}_A(f,g) \in M^1(\mathbb{R}^{2d})$ *and*

$$\|\mathrm{TF}_A(f,g)\|_{M^1} \le C \cdot \|f\|_{M^1} \cdot \|g\|_{M^1}$$

for some constant $C > 0$.

Proof. Let $\Phi \in \mathcal{S}(\mathbb{R}^{2d})$. Then

$$\|\mathrm{TF}_A(f,g)\|_{M^1} \le C \cdot \|V_\Phi(\mathrm{TF}_A(f,g))\|_{L^1}.$$

Choose $\Phi = \mathrm{TF}_A(\varphi,\varphi) \in \mathcal{S}(\mathbb{R}^{2d})$ for some $\varphi \in \mathcal{S}(\mathbb{R}^d)$. Then the Magic Formula 1.7.1 yields

$$\begin{aligned}V_\Phi(\mathrm{TF}_A(f,g))(z,\zeta) &= V(\mathrm{TF}_A(f,g), \mathrm{TF}_A(\varphi,\varphi))(z,\zeta)\\ &= e^{-2\pi i z_2\zeta_2} V(f,\phi)(u,\eta)\overline{V(g,\phi)(v,\gamma)}\end{aligned}$$

with

$$\begin{pmatrix} u \\ v \\ \eta \\ \gamma \end{pmatrix} = \begin{pmatrix} A \cdot \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \cdot (A^{-1})^* \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ \zeta_2 \\ \zeta_1 \\ z_2 \end{pmatrix} = B \cdot \begin{pmatrix} z_1 \\ \zeta_2 \\ \zeta_1 \\ z_2 \end{pmatrix}.$$

The matrix $B \in \mathbb{R}^{4d \times 4d}$ is invertible, we have

$$\begin{aligned} |\det B| &= \left|\det\left(A \cdot \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\right)\right| \cdot \left|\det\left(\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \cdot (A^{-1})^*\right)\right| \\ &= |\det A| \cdot |\det (A^{-1})^*| \\ &= 1. \end{aligned}$$

Thus we may apply a linear coordinate transform to obtain

$$\begin{aligned} \|V_\Phi(\mathrm{TF}_A(f,g))\|_{L^1} &= \iint_{\mathbb{R}^{4d}} |V_\Phi(\mathrm{TF}_A(f,g))(z,\zeta)|\, dz d\zeta \\ &= \iint_{\mathbb{R}^{4d}} |V(f,\varphi)(u,\eta)| \cdot |V(g,\varphi)(v,\gamma)|\, dz d\zeta \\ &= \underbrace{\frac{1}{|\det B|}}_{=1} \iint_{\mathbb{R}^{4d}} |V(f,\varphi)(u,\eta)| \cdot |V(g,\varphi)(v,\gamma)|\, du d\eta dv d\gamma \\ &= \int_{\mathbb{R}^{2d}} |V(f,\varphi)(u,\eta)|\, du d\eta \cdot \int_{\mathbb{R}^{2d}} |V(g,\varphi)(v,\gamma)|\, dv d\gamma \\ &= \|V(f,\varphi)\|_{L^1} \cdot \|V(g,\varphi)\|_{L^1} \\ &\leq C \cdot \|f\|_{M^1} \cdot \|g\|_{M^1} \end{aligned}$$

by Fubini's Theorem. This yields the desired conclusion. □

The last result in this section is concerned with a local and global regularity propery of bilinear time-frequency distributions that can be formulated appropriately in terms of Wiener amalgam spaces, cf. appendix.

Theorem 1.7.3 (Local-Global Regularity Property)**.** *Let $A \in \mathbb{R}^{2d \times 2d}$ be invertible and right-regular. Denote $B = (A^{-1})^*$. Let $f \in M^p(\mathbb{R}^d)$ and $g \in M^1(\mathbb{R}^d)$. Then $\mathrm{TF}_A(f,g)$ belongs to the Wiener amalgam space $W(\mathcal{F}L^1, L^p)$ and*

$$\|\mathrm{TF}_A(f,g)\|_{W(\mathcal{F}L^1,L^p)} \leq C \cdot \|f\|_{M^p} \cdot \|g\|_{M^1}$$

for all $f \in M^p(\mathbb{R}^d)$ and $g \in M^1(\mathbb{R}^d)$ (with some generic constant $C > 0$).

Proof. As in the proof before, choose $\Phi = \mathrm{TF}_A(\varphi,\varphi) \in \mathcal{S}(\mathbb{R}^{2d})$ for some $\varphi \in \mathcal{S}(\mathbb{R}^d)$, φ a test function with compact support that generates a partition of unity. The Wiener amalgam norm of $\mathrm{TF}_A(f,g)$ is given by

$$\begin{aligned}
\|\mathrm{TF}_A(f,g)\|_{W(\mathcal{F}L^1,L^p)} &= \left(\int_{\mathbb{R}^{2d}} \|\mathrm{TF}_A(f,g)\cdot T_z\overline{\Phi}\|^p_{\mathcal{F}L^1}\,dz\right)^{1/p}\\
&= \left(\int_{\mathbb{R}^{2d}} \|(\mathrm{TF}_A\widehat{(f,g)\cdot T_z}\overline{\Phi})\|^p_{L^1}\,dz\right)^{1/p}\\
&= \left(\int_{\mathbb{R}^{2d}}\left(\int_{\mathbb{R}^{2d}} |(\mathrm{TF}_A\widehat{(f,g)\cdot T_z}\overline{\Phi})(\zeta)|\,d\zeta\right)^p dz\right)^{1/p}\\
&= \left(\int_{\mathbb{R}^{2d}}\left(\int_{\mathbb{R}^{2d}} |V(\mathrm{TF}_A(f,g),\Phi)(z,\zeta)|\,d\zeta\right)^p dz\right)^{1/p}\\
&= \left(\int_{\mathbb{R}^{2d}}\left(\int_{\mathbb{R}^{2d}} |V(\mathrm{TF}_A(f,g),\mathrm{TF}_A(\varphi,\varphi))(z,\zeta)|\,d\zeta\right)^p dz\right)^{1/p}.
\end{aligned}$$

Using the Magic Formula 1.7.1, we find for the inner integral

$$\begin{aligned}
&\int_{\mathbb{R}^{2d}} |V(\mathrm{TF}_A(f,g),\mathrm{TF}_A(\varphi,\varphi))(z,\zeta)|\,d\zeta\\
&\qquad = \iint_{\mathbb{R}^{2d}} |V(f,\varphi)(u,\eta)|\cdot|V(g,\varphi)(v,\gamma)|\,d\zeta_1 d\zeta_2
\end{aligned}$$

with

$$\begin{pmatrix} u\\ \eta\end{pmatrix} = \begin{pmatrix} -A_{12}\zeta_2 + A_{11}z_1\\ B_{11}\zeta_1 + B_{12}z_2\end{pmatrix}$$

and

$$\begin{pmatrix} v\\ \gamma\end{pmatrix} = \begin{pmatrix} -A_{22}\zeta_2 + A_{21}z_1\\ -B_{21}\zeta_1 - B_{22}z_2\end{pmatrix}.$$

The coordinate transform

$$\begin{pmatrix} s_1\\ s_2\end{pmatrix} = \begin{pmatrix} u\\ \eta\end{pmatrix} = \begin{pmatrix} -A_{12}\zeta_2 + A_{11}z_1\\ B_{11}\zeta_1 + B_{12}z_2\end{pmatrix} = \begin{pmatrix} 0 & -A_{12}\\ B_{11} & 0\end{pmatrix}\begin{pmatrix} \zeta_1\\ \zeta_2\end{pmatrix} + \begin{pmatrix} A_{11}z_1\\ B_{12}z_2\end{pmatrix}$$

resp.

$$\begin{pmatrix} \zeta_1\\ \zeta_2\end{pmatrix} = \begin{pmatrix} B_{11}^{-1}s_2 - B_{11}^{-1}B_{12}z_2\\ -A_{12}^{-1}s_1 + A_{12}^{-1}A_{11}z_1\end{pmatrix} = \begin{pmatrix} 0 & B_{11}^{-1}\\ -A_{12}^{-1} & 0\end{pmatrix}\begin{pmatrix} s_1\\ s_2\end{pmatrix} + \begin{pmatrix} -B_{11}^{-1}B_{12}z_2\\ A_{12}^{-1}A_{11}z_1\end{pmatrix}$$

(whose tedious details are left to the reader) yields

$$\begin{aligned}
&\iint_{\mathbb{R}^{2d}} |V(f,\varphi)\left(\begin{smallmatrix} u \\ \eta \end{smallmatrix}\right)| \cdot |V(g,\varphi)\left(\begin{smallmatrix} v \\ \gamma \end{smallmatrix}\right)|\, d\zeta_1 d\zeta_2 \\
&= \frac{1}{|\det A_{12}| \cdot |\det B_{11}|} \quad \cdot \\
&\quad \cdot \quad \iint_{\mathbb{R}^{2d}} |V(f,\varphi)\left(\begin{smallmatrix} s_1 \\ s_2 \end{smallmatrix}\right)| \cdot |V(g,\varphi)\left(\begin{smallmatrix} -A_{22}A_{12}^{-1}(w_1-s_1) \\ B_{21}B_{11}^{-1}(w_2-s_2) \end{smallmatrix}\right)|\, ds_1 ds_2
\end{aligned}$$

with

$$w_1 = (A_{11} - A_{12}A_{22}^{-1}A_{21})z_1$$

and

$$w_2 = (B_{12} - B_{11}B_{21}^{-1}B_{22})z_2.$$

Note that if A is right-regular, then B is left-regular by Theorem 1.2.11, so the coordinate transform is allowed.
If we denote

$$F_1(s_1, s_2) = |V(f,\varphi)(s_1, s_2)|$$

and

$$F_2(s_1, s_2) = |V(g,\varphi)(-A_{22}A_{12}^{-1}s_1, B_{21}B_{11}^{-1}s_2)| = |(\mathcal{T}_R V(g,\varphi))(s_1, s_2)|$$

with invertible matrix

$$R = \left(\begin{smallmatrix} -A_{22}A_{12}^{-1} & 0 \\ 0 & B_{21}B_{11}^{-1} \end{smallmatrix}\right) \in \mathbb{R}^{2d\times 2d},$$

then the inner integral can be written as a convolution

$$\begin{aligned}
&\int_{\mathbb{R}^{2d}} |V(\mathrm{TF}_A(f,g), \mathrm{TF}_A(\varphi,\varphi))(z,\zeta)|\, d\zeta \\
&\quad = \frac{1}{|\det A_{12}| \cdot |\det B_{11}|} \iint_{\mathbb{R}^{2d}} F_1(s_1,s_2) \cdot F_2(w_1 - s_1, w_2 - s_2)\, ds_1 ds_2 \\
&\quad = \frac{1}{|\det A_{12}| \cdot |\det B_{11}|} \,(F_1 * F_2)\,(w_1, w_2).
\end{aligned}$$

Hence

$$
\begin{aligned}
&\|\mathrm{TF}_A(f,g)\|_{W(\mathcal{F}L^1,L^p)} \\
&= \frac{1}{|\det A_{12}|\cdot|\det B_{11}|}\left(\iint_{\mathbb{R}^{2d}} |\left(F_1 * F_2\right)(w_1,w_2)|^p\, dz_1 dz_2\right)^{1/p} \\
&= \frac{1}{|\det A_{12}|\cdot|\det B_{11}|} \quad \cdot \\
&\quad \cdot \quad \left(\iint_{\mathbb{R}^{2d}} |\left(F_1 * F_2\right)\left(\begin{smallmatrix}(A_{11}-A_{12}A_{22}^{-1}A_{21})z_1 \\ (B_{12}-B_{11}B_{21}^{-1}B_{22})z_2\end{smallmatrix}\right)|^p\, dz_1 dz_2\right)^{1/p}.
\end{aligned}
$$

A similar argument as at the end of the proof of Corollary 1.2.8 shows that both the matrices $A_{11}-A_{12}A_{22}^{-1}A_{21}$ and $B_{12}-B_{11}B_{21}^{-1}B_{22}$ are invertible with

$$
\det(A_{11}-A_{12}A_{22}^{-1}A_{21}) = \frac{\det A}{\det A_{22}}
$$

and

$$
\det(B_{12}-B_{11}B_{21}^{-1}B_{22}) = -\frac{\det B}{\det B_{21}}.
$$

Another coordinate transform thus leads to

$$
\begin{aligned}
&\|\mathrm{TF}_A(f,g)\|_{W(\mathcal{F}L^1,L^p)} \\
&= \frac{|\det A_{22}|\cdot|\det B_{21}|}{|\det A_{12}|\cdot|\det B_{11}|\cdot|\det A|\cdot|\det B|}\left(\iint_{\mathbb{R}^{2d}} |\left(F_1 * F_2\right)(w_1,w_2)|^p\, dw_1 dw_2\right)^{1/p} \\
&= \frac{|\det A_{22}|\cdot|\det B_{21}|}{|\det A_{12}|\cdot|\det B_{11}|}\|F_1 * F_2\|_{L^p}
\end{aligned}
$$

(note that $\det B = \det\left(A^{-1}\right)^* = \frac{1}{\det A}$). But now everything finally follows from Young's Inequality: obviously, we have $F_1 \in L^p(\mathbb{R}^{2d})$ (since, by assumtion, $f \in M^p(\mathbb{R}^d)$) and $F_2 \in L^1(\mathbb{R}^{2d})$ (since F_2 is a coordinate transform of $V(g,\varphi)$, which in turn is contained in $L^1(\mathbb{R}^{2d})$, because $g \in M^1(\mathbb{R}^d)$ by assumtion). Thus

$$
F_1 * F_2 \in L^p(\mathbb{R}^{2d}) * L^1(\mathbb{R}^{2d}) \subseteq L^p(\mathbb{R}^{2d})
$$

and

$$\begin{aligned}\|F_1 * F_2\|_{L^p} &\le \|F_1\|_{L^p} \cdot \|F_2\|_{L^1} \\ &= \|V(f,\varphi)\|_{L^p} \cdot \|\mathcal{T}_R V(g,\varphi)\|_{L^1} \\ &= \|V(f,\varphi)\|_{L^p} \cdot \frac{1}{|\det R|}\|V(g,\varphi)\|_{L^1} \\ &= \|V(f,\varphi)\|_{L^p} \cdot \frac{|\det A_{12}| \cdot |\det B_{11}|}{|\det A_{22}| \cdot |\det B_{21}|}\|V(g,\varphi)\|_{L^1} \\ &\le C\frac{|\det A_{12}| \cdot |\det B_{11}|}{|\det A_{22}| \cdot |\det B_{21}|}\|f\|_{M^p} \cdot \|g\|_{M^1}.\end{aligned}$$

Putting it all together, all the determinants cancel:

$$\|\mathrm{TF}_A(f,g)\|_{W(\mathcal{F}L^1,L^p)} \le C \cdot \|f\|_{M^p} \cdot \|g\|_{M^1},$$

which finishes the proof. □

Chapter 2

Pseudodifferential Operators

2.1 Motivation: Kohn-Nirenberg Correspondence and Weyl Calculus

The calculus of pseudodifferential operators originated in 1965 with the work of Kohn and Nirenberg, [27]. The fundamental idea is to generalize linear partial differential operators in the following way. Assume

$$Af(x) = \sum_{|\alpha| \le n} a_\alpha(x) \partial^\alpha f(x)$$

is a linear partial differential operator of order n, acting on $f \in \mathcal{S}(\mathbb{R}^d)$. Here α denotes a multiindex $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, and the coefficient functions a_α are usually assumed to be C^∞. Using the Fourier inversion formula, we have

$$\partial^\alpha f(x) = \int_{\mathbb{R}^d} \widehat{\partial^\alpha f}(\omega) e^{2\pi i \omega \cdot x}\, d\omega = \int_{\mathbb{R}^d} (2\pi i \omega)^\alpha \widehat{f}(\omega) e^{2\pi i \omega \cdot x}\, d\omega,$$

hence

$$
\begin{aligned}
Af(x) &= \sum_{|\alpha|\le n} a_\alpha(x)\partial^\alpha f(x) \\
&= \sum_{|\alpha|\le n} a_\alpha(x)\int_{\mathbb{R}^d}(2\pi i\omega)^\alpha \widehat{f}(\omega)e^{2\pi i\omega\cdot x}\,d\omega \\
&= \int_{\mathbb{R}^d}\Big(\sum_{|\alpha|\le n} a_\alpha(x)(2\pi i\omega)^\alpha\Big)\widehat{f}(\omega)e^{2\pi i\omega\cdot x}\,d\omega \\
&= \int_{\mathbb{R}^d}\sigma(x,\omega)\widehat{f}(\omega)e^{2\pi i\omega\cdot x}\,d\omega
\end{aligned}
$$

with

$$\sigma(x,\omega) = \sum_{|\alpha|\le n} a_\alpha(x)(2\pi i\omega)^\alpha$$

the so-called *symbol* of the operator, a polynomial of degree n in ω with the coefficients being smooth functions in x. We extend this approach by allowing as symbols much more general functions or even tempered distributions. This leads to the following definition.

Definition 2.1.1 (Kohn-Nirenberg Correspondence)**.** *Let $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$. The **Kohn-Nirenberg correspondence** maps σ to the operator $\sigma^{KN} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ defined by*

$$\sigma^{KN} f(x) := \int_{\mathbb{R}^d}\sigma(x,\xi)\hat{f}(\xi)e^{2\pi i\xi\cdot x}\,d\xi$$

for $f \in \mathcal{S}(\mathbb{R}^d)$.
The distribution σ is called the (Kohn-Nirenberg) symbol *of the operator σ^{KN}.*

We want to bring the methods of time-frequency analysis into play. As it turns out, the Kohn-Nirenberg correspondence is closely connected with a well-known bilinear time-frequency distribution, the Rihacek distribution.

Proposition 2.1.2. *Let $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$ and $f, g \in \mathcal{S}(\mathbb{R}^d)$. Then*

$$\left\langle \sigma^{KN} f, g\right\rangle = \langle \sigma, R(g,f)\rangle$$

with $R(g,f)(x,\omega)$ the Rihacek distribution.

Proof. If $\sigma \in \mathcal{S}(\mathbb{R}^{2d})$, we can evaluate the integrals explicitly:

$$\begin{aligned} \left\langle \sigma^{KN} f, g \right\rangle &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \sigma(x,\xi) \widehat{f}(\xi) e^{2\pi i \xi \cdot x}\, d\xi \right) \cdot \overline{g(x)}\, dx \\ &= \iint_{\mathbb{R}^{2d}} \sigma(x,\xi) \cdot \overline{\overline{\widehat{f}(\xi)} g(x) e^{-2\pi i \xi \cdot x}}\, d\xi dx \\ &= \langle \sigma, R(g,f) \rangle . \end{aligned}$$

The general case of a distributional symbol follows from this by the usual density argument. □

Observe that by the preceding proposition the mapping $f \mapsto \sigma^{KN} f$ from $\mathcal{S}(\mathbb{R}^d)$ into $\mathcal{S}'(\mathbb{R}^d)$ is continuous (with respect to the weak∗-topology on $\mathcal{S}'(\mathbb{R}^d)$). Indeed, if $f_n \to f$ in $\mathcal{S}(\mathbb{R}^d)$, then for arbitrary $g \in \mathcal{S}(\mathbb{R}^d)$ we have $R(g, f_n) \to R(g, f)$ in $\mathcal{S}(\mathbb{R}^{2d})$, thus $\left\langle \sigma^{KN} f_n, g \right\rangle \to \left\langle \sigma^{KN} f, g \right\rangle$.

In the following, any linear mapping from $\mathcal{S}(\mathbb{R}^d)$ into $\mathcal{S}'(\mathbb{R}^d)$ that is continuous with respect to the weak∗-topology will be called a **pseudodifferential operator**. This obviously includes linear continuous mappings from $\mathcal{S}(\mathbb{R}^d)$ into $\mathcal{S}(\mathbb{R}^d)$.

There are several other ways for associating pseudodifferential operators with distributional symbols. Probably the most prominent among these is the Weyl calculus. This was originally devised as a quantization rule in mathematical physics in the 1930s, cf. [43].

Definition 2.1.3 (Weyl Transform)**.** *Let $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$. The **Weyl transform** maps σ to the pseudodifferential operator $\sigma^W : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ defined by*

$$\begin{aligned} \sigma^W f(x) &= \iint_{\mathbb{R}^{2d}} \widehat{\sigma}(\xi, u) M_{\xi/2} T_{-u} M_{\xi/2} f(x)\, du d\xi \\ &= \iint_{\mathbb{R}^{2d}} \widehat{\sigma}(\xi, u) e^{-\pi i \xi \cdot u} T_{-u} M_{\xi} f(x)\, du d\xi, \qquad f \in \mathcal{S}(\mathbb{R}^d). \end{aligned}$$

The distribution σ is called the (Weyl) symbol *of the operator σ^W.*

The Weyl transform $\sigma^W f$ of a Schwartz function f is a tempered distribution acting on a Schwartz function g in the sense of

$$\left\langle \sigma^W f, g \right\rangle = \iint_{\mathbb{R}^{2d}} \widehat{\sigma}(\xi, u) e^{-\pi i \xi \cdot u} \left\langle T_{-u} M_{\xi} f, g \right\rangle\, du d\xi.$$

Observe that

$$\begin{aligned} e^{-\pi i\xi\cdot u}\langle T_{-u}M_\xi f, g\rangle &= e^{-\pi i\xi\cdot u}\langle f, M_{-\xi}T_u g\rangle \\ &= e^{-\pi i\xi\cdot u}V(f,g)(u,-\xi) \end{aligned}$$

is a Schwartz function on $\mathbb{R}^{2d}$ for $f, g \in \mathcal{S}(\mathbb{R}^d)$. Thus the above is well defined for tempered distributions $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$. In fact the following holds:

Proposition 2.1.4. *Let $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$ and $f, g \in \mathcal{S}(\mathbb{R}^d)$. The action of the pseudodifferential operator $\sigma^W f$ on the Schwartz function g is given by*

$$\left\langle \sigma^W f, g\right\rangle = \left\langle \widehat{\sigma}, \widetilde{\mathrm{TF}_B(g,f)}\right\rangle,$$

with matrix $B = \begin{pmatrix} -\frac{1}{2}I & I \\ \frac{1}{2}I & I \end{pmatrix}$.

Proof. We have

$$\left\langle \sigma^W f, g\right\rangle = \left\langle \widehat{\sigma}(\xi,u), \overline{e^{-\pi i\xi\cdot u}V(f,g)(u,-\xi)}\right\rangle.$$

A short calculation for the right hand expression yields

$$\begin{aligned} \overline{e^{-\pi i\xi\cdot u}V(f,g)(u,-\xi)} &= e^{\pi i\xi\cdot u}\int_{\mathbb{R}^d} g(t-u)\overline{f(t)}e^{-2\pi i\xi\cdot t}\,dt \\ &= \int_{\mathbb{R}^d} g(t-u)\overline{f(t)}e^{-2\pi i\xi\cdot(t-\frac{u}{2})}\,dt \\ &= \int_{\mathbb{R}^d} g(s-\frac{u}{2})\overline{f(s+\frac{u}{2})}e^{-2\pi i\xi\cdot s}\,ds, \end{aligned}$$

by the substitution $s = t - \frac{u}{2}$. This last expression equals

$$\int_{\mathbb{R}^d} g(s-\frac{u}{2})\overline{f(s+\frac{u}{2})}e^{-2\pi i\xi\cdot s}\,ds = \mathrm{TF}_B(g,f)(u,\xi) = \widetilde{\mathrm{TF}_B(g,f)}(\xi,u),$$

where B denotes the matrix $\begin{pmatrix} -\frac{1}{2}I & I \\ \frac{1}{2}I & I \end{pmatrix}$, as claimed. ∎

The preceding proposition gives again a close connection to time-frequency analysis. This connection is even more striking, as seen in the following proposition.

Proposition 2.1.5. *Let $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$ and $f, g \in \mathcal{S}(\mathbb{R}^d)$. Then*

$$\left\langle \sigma^W f, g \right\rangle = \langle \sigma, W(g,f) \rangle$$

with $W(g,f)(x,\omega)$ the Wigner distribution.

Proof. We have

$$\left\langle \sigma^W f, g \right\rangle = \left\langle \widehat{\sigma}, \widetilde{\mathrm{TF}_B(g,f)} \right\rangle$$

with matrix $B = \begin{pmatrix} -\frac{1}{2}I & I \\ \frac{1}{2}I & I \end{pmatrix}$. By Theorem 1.2.16,

$$\widetilde{\mathrm{TF}_B(g,f)} = \widehat{\mathrm{TF}_A(g,f)}$$

with $B = A \cdot \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$. Hence

$$\begin{aligned} A &= B \cdot \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} -\frac{1}{2}I & I \\ \frac{1}{2}I & I \end{pmatrix} \cdot \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \\ &= \begin{pmatrix} I & \frac{1}{2}I \\ I & -\frac{1}{2}I \end{pmatrix}. \end{aligned}$$

This coordinate transformation gives precisely the Wigner distribution: $\mathrm{TF}_A(g,f) = W(g,f)$, thus

$$\left\langle \sigma^W f, g \right\rangle = \left\langle \widehat{\sigma}, \widehat{W(g,f)} \right\rangle = \langle \sigma, W(g,f) \rangle$$

by Plancherel's Theorem. □

This representation proves in particular that $\sigma^W : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ is in fact a pseudodifferential operator in the sense defined above, i.e. continuous.

Thus we find that both presented classical pseudodifferential calculi can be represented in a completely analogous fashion with bilinear time-frequency distributions, the Kohn-Nirenberg correspondence with the Rihacek distribution as

$$\left\langle \sigma^{KN} f, g \right\rangle = \langle \sigma, R(g,f) \rangle, \qquad f, g \in \mathcal{S}(\mathbb{R}^d),$$

the Weyl calculus with the Wigner distribution as

$$\left\langle \sigma^W f, g \right\rangle = \langle \sigma, W(g,f) \rangle, \qquad f, g \in \mathcal{S}(\mathbb{R}^d).$$

The idea to generalize this to arbitrary bilinear time-frequency distributions is very close at hand. This forms the content of the present chapter and will be taken up in the next section.

We present just one more useful representation for the Weyl calculus. This expression gives a rather explicit form of the associated operator.

Proposition 2.1.6.

$$\sigma^W f(t) = \iint_{\mathbb{R}^{2d}} \sigma(\frac{t+\xi}{2}, \omega) f(\xi) e^{2\pi i (t-\xi)\cdot\omega} \, d\xi d\omega$$

for $f \in \mathcal{S}(\mathbb{R}^d)$, $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$.

Proof. Consider the Wigner distribution

$$W(g,f)(x,\omega) = \int_{\mathbb{R}^d} g(x+\frac{y}{2}) \overline{f(x-\frac{y}{2})} e^{-2\pi i y\cdot\omega} \, dy.$$

The substitution $t = x + \frac{y}{2}$ yields

$$W(g,f)(x,\omega) = 2^d \int_{\mathbb{R}^d} g(t) \overline{f(2x-t)} e^{-4\pi i (t-x)\cdot\omega} \, dt.$$

Now let $g \in \mathcal{S}(\mathbb{R}^d)$. Then

$$\begin{aligned}
\left\langle \sigma^W f, g \right\rangle &= \langle \sigma, W(g,f) \rangle \\
&= \iint_{\mathbb{R}^{2d}} \sigma(x,\omega) \cdot \left(2^d \int_{\mathbb{R}^d} \overline{g(t)} f(2x-t) e^{4\pi i (t-x)\cdot\omega} \, dt \right) dx d\omega \\
&= \int_{\mathbb{R}^d} \left(2^d \iint_{\mathbb{R}^{2d}} \sigma(x,\omega) f(2x-t) e^{4\pi i (t-x)\cdot\omega} \, dx d\omega \right) \overline{g(t)} \, dt
\end{aligned}$$

by Fubini's Theorem. We apply the substitution $\xi = 2x - t$, i.e $x = \frac{t+\xi}{2}$, to the inner integral to get

$$\begin{aligned}
2^d \iint_{\mathbb{R}^{2d}} &\sigma(x,\omega) f(2x-t) e^{4\pi i (t-x)\cdot\omega} \, dx d\omega \\
&= \iint_{\mathbb{R}^{2d}} \sigma(\frac{t+\xi}{2}, \omega) f(\xi) e^{4\pi i (t - \frac{t+\xi}{2})\cdot\omega} \, d\xi d\omega \\
&= \iint_{\mathbb{R}^{2d}} \sigma(\frac{t+\xi}{2}, \omega) f(\xi) e^{2\pi i (t-\xi)\cdot\omega} \, d\xi d\omega.
\end{aligned}$$

□

An excellent account of the Weyl calculus and the theory of pseudodifferential operators in textbook form can be found in Folland's book [15]. Other good sources for informations on pseudodifferential operators in the classic "hard analysis" style are the books by Hörmander [23] and Shubin [36].

2.2 Pseudodifferential Operators Associated with Bilinear Time-Frequency Distributions

We begin our study of pseudodifferential operators with time-frequency methods. This approach has become more and more popular in recent years. On the one hand, time-frequency distributions are naturally connected to the classical pseudodifferential calculi, as we have seen in the preceding section. Thus it seems equally natural to use time-frequency tools to analyse the properties of pseudodifferential operators. On the other hand, the time-frequency viewpoint seems to allow for avoiding some of the more technical "hard analysis" parts of the classical machinery.

Our presentation relies heavily on some excellent expositions of time-frequency analysis and its application to pseudodifferential operators. First and foremost, there has to be mentioned the book by Gröchenig [17] and his article [19], that served as a blueprint for almost everything that is to come. Other inspiration was provided by [28].

Proposition 2.2.1. *Let* $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d\times 2d}$ *be invertible and* $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$ *be a tempered distribution.*
The mapping $f \mapsto \sigma^A f$ *given by*

$$\left\langle \sigma^A f, g \right\rangle := \left\langle \sigma, \mathrm{TF}_A(g,f) \right\rangle, \qquad f,g \in \mathcal{S}(\mathbb{R}^d),$$

is well-defined, linear and continuous from $\mathcal{S}(\mathbb{R}^d)$ *to* $\mathcal{S}'(\mathbb{R}^d)$ *(with the weak∗-topology).*

Proof. Let $f,g \in \mathcal{S}(\mathbb{R}^d)$. Then $\mathrm{TF}_A(g,f) \in \mathcal{S}(\mathbb{R}^{2d})$ by Proposition 1.2.4, hence the expression $\left\langle \sigma^A f, g \right\rangle = \left\langle \sigma, \mathrm{TF}_A(g,f) \right\rangle$ is well-defined. It is obviously linear in f and conjugate linear in g. Thus for a fixed $f \in \mathcal{S}(\mathbb{R}^d)$ the mapping $g \mapsto \left\langle \sigma^A f, g \right\rangle$ from $\mathcal{S}(\mathbb{R}^d)$ to $\mathbb{C}$ is conjugate linear. It is continuous, since if $g_n \to g$ in $\mathcal{S}(\mathbb{R}^d)$, then $\mathrm{TF}_A(g_n,f) \to \mathrm{TF}_A(g,f)$ in $\mathcal{S}(\mathbb{R}^{2d})$, hence

$$\left\langle \sigma^A f, g_n \right\rangle = \left\langle \sigma, \mathrm{TF}_A(g_n,f) \right\rangle \to \left\langle \sigma, \mathrm{TF}_A(g,f) \right\rangle = \left\langle \sigma^A f, g \right\rangle.$$

Therefore $\sigma^A f \in \mathcal{S}'(\mathbb{R}^d)$.
Now if $f_n \to f$ in $\mathcal{S}(\mathbb{R}^d)$, then we have for arbitrary fixed $g \in \mathcal{S}(\mathbb{R}^d)$ that

$\mathrm{TF}_A(g, f_n) \to \mathrm{TF}_A(g, f)$ in $\mathcal{S}(\mathbb{R}^{2d})$, which yields

$$\left\langle \sigma^A f_n, g \right\rangle = \langle \sigma, \mathrm{TF}_A(g, f_n) \rangle \to \langle \sigma, \mathrm{TF}_A(g, f) \rangle = \left\langle \sigma^A f, g \right\rangle,$$

hence $\sigma^A f_n \to \sigma^A f$ in the weak∗-topology on $\mathcal{S}'(\mathbb{R}^d)$. So the mapping $f \mapsto \sigma^A f$ is indeed continuous from $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}'(\mathbb{R}^d)$, equipped with the weak∗-topology. □

Definition 2.2.2 (ΨDO Associated with a Bilinear Time-Frequency Distribution)**.** *The pseudodifferential operator defined in the preceding proposition*

$$\sigma^A : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d), \quad f \mapsto \sigma^A f$$

given by

$$\left\langle \sigma^A f, g \right\rangle = \langle \sigma, \mathrm{TF}_A(g, f) \rangle, \qquad f, g \in \mathcal{S}(\}rd), \sigma \in \mathcal{S}'(\mathbb{R}^{2d}),$$

is called the ***pseudodifferential operator with symbol*** σ ***associated with the bilinear time-frequency distribution*** $\mathbf{TF}_A$ *or* ***pseudodifferential operator with symbol*** σ ***associated with*** A*, for short.*

We give some elementary properties.

Proposition 2.2.3 (Adjoint operator)**.** *Let* $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$ *and* $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d \times 2d}$ *be invertible.*
Then

$$(\sigma^A)^* = \tau^B$$

with

$$\tau = \overline{\sigma}$$

and

$$B = \left(\begin{smallmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{smallmatrix}\right) = \left(\begin{smallmatrix} A_{21} & -A_{22} \\ A_{11} & -A_{12} \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right) \cdot A \cdot \left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right).$$

More explicitly:

$$\left\langle \sigma^A f, g \right\rangle = \left\langle f, \tau^B g \right\rangle = \overline{\langle \tau^B g, f \rangle}$$

for all $f, g \in \mathcal{S}(\mathbb{R}^d)$.

Proof. Let $f, g \in \mathcal{S}(\mathbb{R}^d)$ and B be given as above. We compute

$$\begin{aligned}
\mathrm{TF}_B(f,g)(x,\omega) &= \int_{\mathbb{R}^d} f(B_{11}x + B_{12}y)\overline{g(B_{21}x + B_{22}y)}e^{-2\pi i\omega\cdot y}\,dy \\
&= \int_{\mathbb{R}^d} f(A_{21}x - A_{22}y)\overline{g(A_{11}x - A_{12}y)}e^{-2\pi i\omega\cdot y}\,dy \\
&= \int_{\mathbb{R}^d} \overline{g(A_{11}x + A_{12}y)}f(A_{21}x + A_{22}y)e^{2\pi i\omega\cdot y}\,dy \\
&= \overline{\int_{\mathbb{R}^d} g(A_{11}x + A_{12}y)\overline{f(A_{21}x + A_{22}y)}e^{-2\pi i\omega\cdot y}\,dy} \\
&= \overline{\mathrm{TF}_A(g,f)(x,\omega)}.
\end{aligned}$$

Thus for $\tau = \overline{\sigma}$ we get

$$\begin{aligned}
\overline{\langle \tau^B g, f\rangle} &= \overline{\langle \tau, \mathrm{TF}_B(f,g)\rangle} \\
&= \overline{\iint_{\mathbb{R}^{2d}} \tau(x,\omega)\cdot \overline{\mathrm{TF}_B(f,g)(x,\omega)}\,dxd\omega} \\
&= \iint_{\mathbb{R}^{2d}} \overline{\tau(x,\omega)}\cdot \mathrm{TF}_B(f,g)(x,\omega)\,dxd\omega \\
&= \iint_{\mathbb{R}^{2d}} \sigma(x,\omega)\cdot \overline{\mathrm{TF}_A(g,f)(x,\omega)}\,dxd\omega \\
&= \langle \sigma, \mathrm{TF}_A(g,f)\rangle \\
&= \left\langle \sigma^A f, g\right\rangle.
\end{aligned}$$

□

Corollary 2.2.4. *Suppose A is of the form $A = \left(\begin{smallmatrix} U & V \\ U & -V \end{smallmatrix}\right)$ with matrices $U, V \in \mathbb{R}^{d\times d}$.*
Then σ^A is self-adjoint for real symbols $\sigma = \overline{\sigma} \in \mathcal{S}'(\mathbb{R}^{2d})$, i.e if $\sigma = \overline{\sigma}$, then $\sigma^A = (\sigma^A)^*$.

Proof. Obviously, the matrix $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right)$ is of the form $\left(\begin{smallmatrix} U & V \\ U & -V \end{smallmatrix}\right)$ if and only if $A_{11} = A_{21}$ and $A_{22} = -A_{12}$. But this is equivalent to

$$\left(\begin{smallmatrix} 0 & I \\ I & 0 \end{smallmatrix}\right)\cdot A \cdot \left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right) = A.$$

Thus if $\sigma = \overline{\sigma} \in \mathcal{S}'(\mathbb{R}^{2d})$, then $\tau = \sigma$ and $B = A$ in the above theorem and so $(\sigma^A)^* = \tau^B = \sigma^A$ is self-adjoint. □

We make a short **Remark:**
If the matrix $A \in \mathbb{R}^{2d\times 2d}$ is invertible and of the form $\left(\begin{smallmatrix} U & V \\ U & -V \end{smallmatrix}\right)$ mentioned above, then the matrices $U, V \in \mathbb{R}^{d\times d}$ are necessarily invertible themselves. This follows from

$$\det \left(\begin{smallmatrix} U & V \\ U & -V \end{smallmatrix}\right) = \det \left(\begin{smallmatrix} 2U & 0 \\ U & -V \end{smallmatrix}\right) = 2^d \cdot (-1)^d \cdot \det U \cdot \det V \neq 0.$$

As an **example**, the Weyl calculus gives rise to pseudodifferential operators that are self-adjoint for real symbols, since the matrix $\begin{pmatrix} I & \frac{1}{2}I \\ I & -\frac{1}{2}I \end{pmatrix}$ of the Wigner distribution satisfies the assumptions of Corollary 2.2.4. This nice property is not shared by the Kohn-Nirenberg correspondence, since the matrix $\left(\begin{smallmatrix} I & 0 \\ I & -I \end{smallmatrix}\right)$ of the Rihacek distribution does not have the required form.

In a certain sense that is made precise by the following theorem, all pseudodifferential calculi associated with arbitrary invertible matrices A are equivalent.

The first of the following representations is the well-known classical *Schwartz Kernel Theorem* (cf. [22] or [16]), but stated for tempered distributions rather than general ones. We will not give a proof of this special case but rather refer the interested reader to [32], which goes back to [37], or the lecture notes of Christoph Thiele, [38]. It can also be derived from the classical version mentioned above.

Theorem 2.2.5 (Kernel theorem)**.** *Let T be a continuous linear operator mapping $\mathcal{S}(\mathbb{R}^d)$ into $\mathcal{S}'(\mathbb{R}^d)$ (with the weak∗-topology). Let $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix}\right) \in \mathbb{R}^{2d\times 2d}$ be invertible.*
Then there exist tempered distributions $k, \sigma, F \in \mathcal{S}'(\mathbb{R}^{2d})$ such that the following representations hold:

(a) as a (generalized) integral operator:

$$\langle Tf, g\rangle = \left\langle k, g \otimes \overline{f}\right\rangle, \qquad f, g \in \mathcal{S}(\mathbb{R}^d);$$

(b) as a pseudodifferential operator associated with the time-frequency distribution TF_A with matrix A:

$$T = \sigma^A;$$

(c) as a superposition of time-frequency shifts:

$$T = \iint_{\mathbb{R}^{2d}} F(x,\omega) \cdot M_\omega T_x \, dx d\omega.$$

Proof. We skip (a) and proceed to (b).
Set $\sigma = \mathcal{F}_2 \mathcal{T}_A k \in \mathcal{S}'(\mathbb{R}^{2d})$ with the tempered distribution k from (a). Then

$$\begin{aligned} \langle Tf, g \rangle &= \left\langle k, g \otimes \overline{f} \right\rangle \\ &= \left\langle \mathcal{F}_2 \mathcal{T}_A k, \mathcal{F}_2 \mathcal{T}_A (g \otimes \overline{f}) \right\rangle \\ &= \langle \sigma, \mathrm{TF}_A(g,f) \rangle \\ &= \left\langle \sigma^A f, g \right\rangle \end{aligned}$$

for all $f, g \in \mathcal{S}(\mathbb{R}^d)$. So $T = \sigma^A$.
Finally, (c) is a special case of (b).
The vector-valued integral

$$T = \iint_{\mathbb{R}^{2d}} F(x,\omega) \cdot M_\omega T_x \, dx d\omega$$

is interpreted in the weak sense to mean

$$\begin{aligned} \langle Tf, g \rangle &= \iint_{\mathbb{R}^{2d}} F(x,\omega) \cdot \langle M_\omega T_x f, g \rangle \, dx d\omega \\ &= \iint_{\mathbb{R}^{2d}} F(x,\omega) \cdot \overline{V(g,f)(x,\omega)} \, dx d\omega \\ &= \langle F, V(g,f) \rangle \end{aligned}$$

for all $f, g \in \mathcal{S}(\mathbb{R}^d)$. But $V(g,f) = \mathrm{TF}_A(g,f)$ for the matrix $A = \left(\begin{smallmatrix} 0 & I \\ -I & I \end{smallmatrix} \right)$. So simply choose $F = \sigma^A$ for this matrix A as in (b). □

Next we consider some mapping properties of the correspondence $\sigma \mapsto \sigma^A$ from the symbol to the operator. We are particularly interested in the question under what conditions σ^A gives a bounded linear operator on $L^2(\mathbb{R}^d)$. In the following theorems, we always assume that $A = \left(\begin{smallmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{smallmatrix} \right) \in \mathbb{R}^{2d}$ is an invertible matrix.

We begin with symbols in L^p-spaces.

Theorem 2.2.6. *Let $\sigma \in L^1(\mathbb{R}^{2d})$. Let A be right-regular. Then $\sigma^A \in B(L^2(\mathbb{R}^d))$ is a bounded operator on $L^2(\mathbb{R}^d)$ and*

$$\|\sigma^A\|_{B(L^2)} \leq \frac{\|\sigma\|_{L^1}}{|\det(A_{12})|^{1/2} \cdot |\det(A_{22})|^{1/2}}.$$

Proof. Let $f, g \in L^2(\mathbb{R}^d)$. Then

$$\begin{aligned} |\left\langle \sigma^A f, g \right\rangle| &= |\iint_{\mathbb{R}^{2d}} \sigma(x,\omega) \cdot \overline{\mathrm{TF}_A(g,f)(x,\omega)}\, dx d\omega| \\ &\leq \iint_{\mathbb{R}^{2d}} |\sigma(x,\omega)| \cdot |\mathrm{TF}_A(g,f)(x,\omega)|\, dx d\omega \\ &\leq \|\sigma\|_{L^1} \cdot \|\mathrm{TF}_A(g,f)\|_{L^\infty}. \end{aligned}$$

By assumption, A is right-regular, hence by Corollary 1.2.7,

$$\|\mathrm{TF}_A(g,f)\|_{L^\infty} \leq \frac{\|f\| \cdot \|g\|}{|\det(A_{12})|^{1/2} \cdot |\det(A_{22})|^{1/2}}.$$

The estimate

$$|\left\langle \sigma^A f, g \right\rangle| \leq \frac{\|\sigma\|_{L^1}}{|\det(A_{12})|^{1/2} \cdot |\det(A_{22})|^{1/2}} \cdot \|f\| \cdot \|g\|$$

proves that $\sigma^A f \in L^2(\mathbb{R}^d)$ and $\sigma^A : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is bounded, and also yields the norm estimate. □

Theorem 2.2.7. *Let $\sigma \in L^2(\mathbb{R}^{2d})$. Then $\sigma^A \in B(L^2(\mathbb{R}^d))$ is a bounded operator on $L^2(\mathbb{R}^d)$ and*

$$\|\sigma^A\|_{B(L^2)} \leq \frac{\|\sigma\|_{L^2}}{|\det A|^{1/2}}.$$

Proof. We have

$$\begin{aligned} |\left\langle \sigma^A f, g \right\rangle| &= |\langle \sigma, \mathrm{TF}_A(g,f)\rangle| \\ &\leq \|\sigma\|_{L^2} \cdot \|\mathrm{TF}_A(g,f)\|_{L^2} \end{aligned}$$

by the Cauchy-Schwarz Inequality. Using Theorem 1.3.1 (or the remark immediately following it), we continue

$$\|\sigma\|_{L^2} \cdot \|\mathrm{TF}_A(g,f)\|_{L^2} = \|\sigma\|_{L^2} \frac{1}{|\det A|^{1/2}} \cdot \|f\| \cdot \|g\|.$$

This gives the desired conclusion. □

For symbols in modulation spaces, we can even prove some Schatten-class properties. The following line of argument imitates the very elegant proof found in [19] for the special case of the Kohn-Nirenberg correspondence.

Theorem 2.2.8 (Trace class for symbol in M^1)**.** *Let $\sigma \in M^{1,1}(\mathbb{R}^{2d})$. Then $\sigma^A \in B(L^2(\mathbb{R}^d))$ is a bounded operator on $L^2(\mathbb{R}^d)$ and belongs to the trace class $\mathcal{S}^1(L^2(\mathbb{R}^d))$ with*

$$\|\sigma^A\|_{\mathcal{S}^1} \leq C \cdot \|\sigma\|_{M^1}$$

with some constant $C > 0$ independent of σ and of A.

Proof. Let $\sigma \in M^{1,1}(\mathbb{R}^{2d}) = M^1(\mathbb{R}^{2d}) \subseteq L^2(\mathbb{R}^{2d})$. The inversion formula for the short-time Fourier transform implies that σ can be written as a vector-valued integral

$$\sigma = \iint_{\mathbb{R}^{4d}} V_\Phi \sigma(z,\zeta) \cdot M_\zeta T_z \Phi \, dz d\zeta$$

with any $\Phi \in L^2(\mathbb{R}^{2d})$.
Now observe that for all $f, g \in L^2(\mathbb{R}^d)$

$$\begin{aligned} \left\langle \sigma^A f, g \right\rangle &= \langle \sigma, \mathrm{TF}_A(g,f) \rangle \\ &= \iint_{\mathbb{R}^{4d}} V_\Phi \sigma(z,\zeta) \cdot \langle M_\zeta T_z \Phi, \mathrm{TF}_A(g,f) \rangle \, dz d\zeta \end{aligned}$$

by the weak interpretation of the above integral.
The term $\langle M_\zeta T_z \Phi, \mathrm{TF}_A(g,f) \rangle$ can be interpreted as $\left\langle (M_\zeta T_z \Phi)^A f, g \right\rangle$, with $(M_\zeta T_z \Phi)^A$ the pseudodifferential operator associated to the symbol $M_\zeta T_z \Phi \in L^2(\mathbb{R}^{2d})$. Hence

$$\left\langle \sigma^A f, g \right\rangle = \iint_{\mathbb{R}^{4d}} V_\Phi \sigma(z,\zeta) \cdot \left\langle (M_\zeta T_z \Phi)^A f, g \right\rangle \, dz d\zeta.$$

Thus

$$\sigma^A f = \iint_{\mathbb{R}^{4d}} V_\Phi\sigma(z,\zeta)\cdot (M_\zeta T_z\Phi)^A f\,dzd\zeta$$

as vector-valued integral in $L^2(\mathbb{R}^d)$. But this means that σ^A can be written as an operator-valued integral

$$\sigma^A = \iint_{\mathbb{R}^{4d}} V_\Phi\sigma(z,\zeta)\cdot (M_\zeta T_z\Phi)^A\,dzd\zeta,$$

that is as a continuous weighted superposition of elementary operators of the form $(M_\zeta T_z\Phi)^A$.

Let us look closer at these elementary building blocks of the integral. We have

$$\begin{aligned}\left\langle (M_\zeta T_z\Phi)^A f, g\right\rangle &= \langle M_\zeta T_z\Phi, \mathrm{TF}_A(g,f)\rangle\\ &= \overline{V(\mathrm{TF}_A(g,f),\Phi)(z,\zeta)}.\end{aligned}$$

Choose $\Phi = \mathrm{TF}_A(\varphi,\varphi)$ with $\phi \in L^2(\mathbb{R}^d)$. Then the Magic Formula 1.7.1 yields

$$\begin{aligned}\overline{V(\mathrm{TF}_A(g,f),\Phi)(z,\zeta)} &= \overline{V(\mathrm{TF}_A(g,f),\mathrm{TF}_A(\varphi,\varphi))(z,\zeta)}\\ &= e^{2\pi i z_2\zeta_2}\overline{V(g,\varphi)(u,\eta)}V(f,\varphi)(v,\gamma)\\ &= e^{2\pi i z_2\zeta_2}\langle f, M_\gamma T_v\varphi\rangle\langle M_\eta T_u\varphi, g\rangle\end{aligned}$$

with some $u,v,\gamma,\eta\in\mathbb{R}^d$ depending continuously on z and ζ. This implies

$$(M_\zeta T_z\Phi)^A f = e^{2\pi i z_2\zeta_2}\langle f, M_\gamma T_v\varphi\rangle M_\eta T_u\varphi,$$

so the operator

$$(M_\zeta T_z\Phi)^A : L^2(\mathbb{R}^d)\to L^2(\mathbb{R}^d),\quad f\mapsto e^{2\pi i z_2\zeta_2}\langle f, M_\gamma T_v\varphi\rangle M_\eta T_u\varphi$$

is in fact a rank one operator, in particular a trace class operator. The trace class norm is given by

$$\begin{aligned}\|(M_\zeta T_z\Phi)^A\|_{\mathcal{S}^1} &= \|e^{2\pi i z_2\zeta_2}\langle \bullet, M_\gamma T_v\varphi\rangle M_\eta T_u\varphi\|_{\mathcal{S}^1}\\ &= \|\langle \bullet, M_\gamma T_v\varphi\rangle M_\eta T_u\varphi\|_{\mathcal{S}^1}\\ &= \|M_\eta T_u\varphi\otimes\overline{M_\gamma T_v\varphi}\|_{\mathcal{S}^1}\\ &= \|M_\gamma T_v\varphi\|\cdot\|M_\eta T_u\varphi\|\\ &= \|\varphi\|^2\end{aligned}$$

independently of (z,ζ). Thus

$$\sigma^A = \iint_{\mathbb{R}^{4d}} V_\Phi\sigma(z,\zeta)\cdot(M_\zeta T_z\Phi)^A\,dzd\zeta$$

is an operator-valued integral in the space $\mathcal{S}^1(L^2(\mathbb{R}^d))$ of trace class operators. For $\sigma\in M^1(\mathbb{R}^{2d})$, we have $V_\Phi\sigma\in L^1(\mathbb{R}^{4d})$, thus the integrand is absolutely integrable with respect to the $\mathcal{S}^1$-norm:

$$\begin{aligned}
&\iint_{\mathbb{R}^{4d}} \|V_\Phi\sigma(z,\zeta)\cdot(M_\zeta T_z\Phi)^A\|_{\mathcal{S}^1}\,dzd\zeta \\
&\qquad = \iint_{\mathbb{R}^{4d}} |V_\Phi\sigma(z,\zeta)|\cdot\|(M_\zeta T_z\Phi)^A\|_{\mathcal{S}^1}\,dzd\zeta \\
&\qquad = \iint_{\mathbb{R}^{4d}} |V_\Phi\sigma(z,\zeta)|\cdot\|\varphi\|^2\,dzd\zeta \\
&\qquad = \|V_\Phi\sigma\|_{L^1}\cdot\|\varphi\|^2 \\
&\qquad \le C'\cdot\|\varphi\|^2\cdot\|\sigma\|_{M^1} \\
&\qquad = C\cdot\|\sigma\|_{M^1}
\end{aligned}$$

with the constant C depending only on Φ resp. φ.
So finally $\sigma^A\in\mathcal{S}^1(L^2(\mathbb{R}^d))$ and

$$\|\sigma^A\|_{\mathcal{S}^1} \le \iint_{\mathbb{R}^{4d}} \|V_\Phi\sigma(z,\zeta)\cdot(M_\zeta T_z\Phi)^A\|_{\mathcal{S}^1}\,dzd\zeta = C\cdot\|\sigma\|_{M^1}.$$

□

We also have the following generalization of a theorem that was first proved for the Weyl calculus in [30].

Theorem 2.2.9 (Hilbert-Schmidt operator for symbol in M^2). *If $\sigma\in M^{2,2}(\mathbb{R}^{2d}) = L^2(\mathbb{R}^{2d})$, then we have $\sigma^A\in\mathcal{S}^2$, i.e. σ^A is a Hilbert-Schmidt operator. Furthermore,*

$$\|\sigma^A\|_{\mathcal{S}^2} = |\det A|^{1/2}\cdot\|\sigma\|_{L^2} \le C\cdot|\det A|^{1/2}\cdot\|\sigma\|_{M^2}$$

with some constant $C>0$ independent of σ and of A.

Proof. Let $\sigma \in L^2(\mathbb{R}^{2d})$. As a continuous pseudodifferential operator from $\mathcal{S}(\mathbb{R}^d)$ to $\mathcal{S}'(\mathbb{R}^d)$, σ^A can be represented by the Schwartz Kernel Theorem 2.2.5 as a (generalized) integral operator: there exists a tempered distribution $k \in \mathcal{S}'(\mathbb{R}^{2d})$ such that

$$\left\langle \sigma^A f, g \right\rangle = \langle \sigma, \mathrm{TF}_A(g, f) \rangle = \left\langle k, g \otimes \overline{f} \right\rangle.$$

The distributional symbols σ and k are related by the formula

$$\sigma = \mathcal{F}_2 \mathcal{T}_A k, \qquad \text{or} \qquad k = \mathcal{T}_{A^{-1}} \mathcal{F}_2^{-1} \sigma.$$

Since $\mathcal{F}_2$ (and $\mathcal{F}_2^{-1}$) and $\mathcal{T}_A$ (and $\mathcal{T}_A^{-1} = \mathcal{T}_{A^{-1}}$) are operators from $L^2(\mathbb{R}^{2d})$ onto $L^2(\mathbb{R}^{2d})$, we conclude that $k \in L^2(\mathbb{R}^{2d})$. But this implies that σ^A is in fact a true integral operator:

$$\begin{aligned}
\left\langle \sigma^A f, g \right\rangle &= \left\langle k, g \otimes \overline{f} \right\rangle \\
&= \iint_{\mathbb{R}^{2d}} k(x, \omega) \overline{g(x)} f(\omega)\, dx d\omega \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} k(x, \omega) f(\omega)\, d\omega \right) \overline{g(x)}\, dx \\
&= \langle h, g \rangle
\end{aligned}$$

with

$$h(x) = \int_{\mathbb{R}^d} k(x, \omega) f(\omega)\, d\omega.$$

Thus $\sigma^A f(x) = h(x)$ and $\sigma^A : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is an integral operator with kernel $k \in L^2(\mathbb{R}^{2d})$, hence a Hilbert-Schmidt operator with norm

$$\begin{aligned}
\|\sigma^A\|_{\mathcal{S}^2} &= \|k\|_{L^2} \\
&= \|\mathcal{T}_{A^{-1}} \mathcal{F}_2^{-1} \sigma\|_{L^2} \\
&= \frac{1}{|\det A^{-1}|^{1/2}} \cdot \|\mathcal{F}_2^{-1} \sigma\|_{L^2} \\
&= |\det A|^{1/2} \cdot \|\sigma\|_{L^2}
\end{aligned}$$

by Lemma A.3.2 and the fact that $\mathcal{F}_2$ is unitary, Lemma A.4.2. □

2.3 Boundedness of Pseudodifferential Operators

We have seen in the preceeding section that pseudodifferential operators associated with bilinear time-frequency distributions are especially well behaved if the symbol is from a modulation space. This supports once more the folkloristic saying that modulation spaces are most perfectly suited for practicing time-frequency analysis. Some more evidence for this view will be given by the following theorem, which is the main result in this chapter.

Theorem 2.3.1 (Boundedness of ΨDOs)**.** *Let* $\sigma \in M^{\infty,1}(\mathbb{R}^{2d})$ *and* $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \in \mathbb{R}^{2d\times 2d}$ *be invertible and left-regular.*
Then the associated pseudodifferential operator σ^A *is bounded on all modulation spaces* $M^{p,q}(\mathbb{R}^d)$, $1 \le p, q \le \infty$. *In particular,* σ^A *is bounded on* $L^2(\mathbb{R}^d)$. *Furthermore, we have*

$$\begin{aligned} &\|\sigma^A\|_{B(M^{p,q})} \\ &\le C \cdot \|\sigma\|_{M^{\infty,1}} \cdot \frac{1}{|\det A_{21}|^{1/p} |\det B_{22}|^{1/q}} \cdot \frac{1}{|\det A_{11}|^{1/p'} |\det B_{12}|^{1/q'}}. \end{aligned}$$

Proof. Set $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = (A^{-1})^*$. By Proposition 1.2.11, A is left-regular if and only if A_{11} and A_{21} are invertible if and only if B is right-regular if and only if B_{12} and B_{22} are invertible.
Let $f, g \in \mathcal{S}(\mathbb{R}^d)$. Then

$$\begin{aligned} |\left\langle \sigma^A f, g \right\rangle| &= |\langle \sigma, \mathrm{TF}_A(g,f)\rangle| \\ &= |\langle V_\Phi \sigma, V_\Phi \mathrm{TF}_A(g,f)\rangle| \end{aligned}$$

for any $\Phi \in \mathcal{S}(\mathbb{R}^{2d})$.
If $\sigma \in M^{\infty,1}(\mathbb{R}^{2d})$, then $V_\Phi\sigma \in L^{\infty,1}(\mathbb{R}^{2d})$, and Hölder's Inequality for mixed norm spaces implies

$$\begin{aligned} |\langle V_\Phi \sigma, V_\Phi \mathrm{TF}_A(g,f)\rangle| &\le \|V_\Phi\sigma\|_{L^{\infty,1}} \cdot \|V_\Phi \mathrm{TF}_A(g,f)\|_{L^{1,\infty}} \\ &\le C \cdot \|\sigma\|_{M^{\infty,1}} \cdot \|V_\Phi \mathrm{TF}_A(g,f)\|_{L^{1,\infty}} \end{aligned}$$

with a constant C depending on Φ.
Choose Φ to be equal to $\mathrm{TF}_A(\varphi,\varphi)$ for some $\varphi \in \mathcal{S}(\mathbb{R}^d)$, for instance φ a

Gaussian. Then $\Phi \in \mathcal{S}(\mathbb{R}^{2d})$.
With the help of the Magic Formula 1.7.1, we find

$$\begin{aligned} V_\Phi \mathrm{TF}_A(g,f)(z,\zeta) &= V(\mathrm{TF}_A(g,f), \mathrm{TF}_A(\varphi,\varphi))(z,\zeta) \\ &= e^{-2\pi i z_2\cdot\zeta_2} \cdot V(g,\varphi)(u,\eta) \cdot \overline{V(f,\varphi)(v,\gamma)} \end{aligned}$$

with

$$\left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right) = A \cdot \left(\begin{smallmatrix} z_1 \\ -\zeta_2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} A_{11}z_1 - A_{12}\zeta_2 \\ A_{21}z_1 - A_{22}\zeta_2 \end{smallmatrix}\right)$$

and

$$\left(\begin{smallmatrix} \eta \\ \gamma \end{smallmatrix}\right) = \left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right) \cdot B \cdot \left(\begin{smallmatrix} \zeta_1 \\ z_2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} B_{12}z_2 + B_{11}\zeta_1 \\ -B_{22}z_2 - B_{21}\zeta_1 \end{smallmatrix}\right).$$

Hence we continue

$$\begin{aligned} \|V_\Phi \mathrm{TF}_A(g,f)\|_{L^{1,\infty}} &= \|V(\mathrm{TF}_A(g,f), \mathrm{TF}_A(\varphi,\varphi))\|_{L^{1,\infty}} \\ &= \sup_{\zeta\in\mathbb{R}^{2d}} \left(\int_{\mathbb{R}^{2d}} |V(\mathrm{TF}_A(g,f), \mathrm{TF}_A(\varphi,\varphi))(z,\zeta)|\, dz \right) \\ &= \sup_{\zeta\in\mathbb{R}^{2d}} \iint_{\mathbb{R}^{2d}} |V(g,\varphi)(A_{11}z_1 - A_{12}\zeta_2, B_{12}z_2 + B_{11}\zeta_1)| \cdot \\ &\qquad \cdot |V(f,\varphi)(A_{21}z_1 - A_{22}\zeta_2, -B_{22}z_2 - B_{21}\zeta_1)|\, dz_1 dz_2. \end{aligned}$$

Now denote by p', q' the conjugate exponents to p, q, respectively (i.e. such that $\frac{1}{p} + \frac{1}{p'} = 1$ and $\frac{1}{q} + \frac{1}{q'} = 1$). Applying once more Hölder's Inequality for mixed norm spaces, we estimate

$$\begin{aligned} &\iint_{\mathbb{R}^{2d}} |V(g,\varphi)(u,\eta)| \cdot |V(f,\varphi)(v,\gamma)|\, dz_1 dz_2 \\ &\quad \le \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V(g,\varphi)(u,\eta)|^{p'}\, dz_1 \right)^{q'/p'} dz_2 \right)^{1/q'} \cdot \\ &\qquad \cdot \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V(f,\varphi)(v,\gamma)|^{p}\, dz_1 \right)^{q/p} dz_2 \right)^{1/q} \\ &\quad = \|V(g,\varphi)(u,\eta)\|_{L^{p',q'}} \cdot \|V(f,\varphi)(v,\gamma)\|_{L^{p,q}} \end{aligned}$$

(with a slight abuse of notation). The two terms in the last line can be estimated further:

$$\begin{aligned}
&\|V(f,\varphi)(v,\gamma)\|_{L^{p,q}}\\
&= \left(\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|V(f,\varphi)(v,\gamma)|^p\,dz_1\right)^{q/p}\,dz_2\right)^{1/q}\\
&= \left(\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|V(f,\varphi)(A_{21}z_1-A_{22}\zeta_2,-B_{22}z_2-B_{21}\zeta_1)|^p\,dz_1\right)^{q/p}\,dz_2\right)^{1/q}\\
&\underset{(*)}{=} \left(\int_{\mathbb{R}^d}\left(\frac{1}{|\det A_{21}|}\int_{\mathbb{R}^d}|V(f,\varphi)(s_1,-B_{22}z_2-B_{21}\zeta_1)|^p\,ds_1\right)^{q/p}\,dz_2\right)^{1/q}\\
&= \frac{1}{|\det A_{21}|^{1/p}}\cdot\left(\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|V(f,\varphi)(s_1,-B_{22}z_2-B_{21}\zeta_1)|^p\,ds_1\right)^{q/p}\,dz_2\right)^{1/q}\\
&\underset{(*)}{=} \frac{1}{|\det A_{21}|^{1/p}}\cdot\left(\frac{1}{|\det B_{22}|}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|V(f,\varphi)(s_1,s_2)|^p\,ds_1\right)^{q/p}\,ds_2\right)^{1/q}\\
&= \frac{1}{|\det A_{21}|^{1/p}\cdot|\det B_{22}|^{1/q}}\cdot\|V(f,\varphi)\|_{L^{p,q}}\\
&\le \frac{1}{|\det A_{21}|^{1/p}\cdot|\det B_{22}|^{1/q}}\cdot C'\cdot\|f\|_{M^{p,q}}
\end{aligned}$$

with the constant C' depending only on φ and p,q. Note that in the lines marked with $(*)$ the coordinate changes are permitted since the matrices A_{21} and B_{22} are invertible by assumption.
In a completely analogous fashion, we find

$$\|V(g,\varphi)(u,\eta)\|_{L^{p',q'}} \le \frac{1}{|\det A_{11}|^{1/p'}\cdot|\det B_{12}|^{1/q'}}\cdot C''\cdot\|g\|_{M^{p',q'}}$$

with the constant C'' depending only on φ and p',q'.
Thus putting it all together, we have

$$\begin{aligned}
&|\left\langle\sigma^A f,g\right\rangle|\\
&\le C\cdot\|\sigma\|_{M^{\infty,1}}\cdot\frac{\|f\|_{M^{p,q}}}{|\det A_{21}|^{1/p}|\det B_{22}|^{1/q}}\cdot\frac{\|g\|_{M^{p',q'}}}{|\det A_{11}|^{1/p'}|\det B_{12}|^{1/q'}}
\end{aligned}$$

with some generic constant C that does not depend on f or g. By duality, this proves that σ^A extends to a bounded operator from $M^{p,q}$ to $M^{p,q}$. □

With this theorem, we are finally able to strengthen considerably the Schatten class results of the last section.

Theorem 2.3.2 (Schatten class for symbol in $M^{p,1}$). *Let $\sigma \in M^{p,1}(\mathbb{R}^{2d})$, $1 \leq p \leq \infty$, and $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \in \mathbb{R}^{2d \times 2d}$ be invertible and left-regular. Then the associated pseudodifferential operator σ^A belongs to the Schatten p-class on $L^2(\mathbb{R}^d)$, i.e. $\sigma^A \in \mathcal{S}^p(L^2(\mathbb{R}^d))$, and*

$$\|\sigma^A\|_{\mathcal{S}^p} \leq C \cdot \|\sigma\|_{M^{p,1}},$$

where the constant $C > 0$ depends only on A.

Proof. The result (together with the norm estimate) follows immediately from Theorem 2.2.8 and Theorem 2.3.1 by using complex interpolation on the spaces $M^{p,1} = [M^{1,1}, M^{\infty,1}]_\theta$ and $\mathcal{S}^p = [\mathcal{S}^1, B(L^2)]_\theta$. □

The following table summarizes this section's results on the mapping properties of the correspondence $\sigma \mapsto \sigma^A$ from symbols to associated pseudodifferential operators with different matrices A.

Symbol	**Matrix** A	**Operator** σ^A
$\mathcal{S}'(\mathbb{R}^{2d})$	invtbl.	ΨDO $\mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$
$L^1(\mathbb{R}^{2d})$	invtbl., right-reg.	$B(L^2(\mathbb{R}^d))$
$L^2(\mathbb{R}^{2d}) = M^2(\mathbb{R}^{2d})$	invtbl.	$\mathcal{S}^2(L^2(\mathbb{R}^d))$
$M^1(\mathbb{R}^{2d})$	invtbl.	$\mathcal{S}^1(L^2(\mathbb{R}^d))$
$M^{\infty,1}(\mathbb{R}^{2d})$	invtbl., left-reg.	$B(M^{p,q})$, $1 \leq p < \infty$
$M^{p,1}(rdd)$, $1 \leq p < \infty$	invtbl., left-reg.	$\mathcal{S}^p(L^2(\mathbb{R}^d))$

Table 2.1: Pseudodifferential operators associated to A with different symbols

Chapter 3

Time-Frequency Localization Operators and the Berezin Transform

3.1 Time-Frequency Localization Operators

Time-frequency localization operators in the form considered here were first introduced and studied by Daubechies, [9], and Ramanathan and Topiwala, [31]. They were used as a mathematical tool to extract specific features from the time-frequency representation of a signal on phase space. In physics, such operators had been around for quite a long time in connection with questions of quantization, under the name "anti-Wick operators" in the work of Berezin, [3]. They had also appeared earlier in the theory of pseudodifferential operators, cf. [8]. See also the book [36] by Shubin.

The fundamental idea behind the concept of localization operator is that of a multiplier for the short-time Fourier transform. Let f be a given function on $\mathbb{R}^d$, a "signal" (whatever that may be). We perform an analysis by a short-time Fourier transform with window φ_2 to obtain a time-frequency representation of f on phase space $\mathbb{R}^d \times \mathbb{R}^d$. In order to extract interesting parts of this representation, we may apply a time-frequency filtering procedure by multiplying with a suitable "mask", the so-called symbol, a function

a on $\mathbb{R}^d \times \mathbb{R}^d$. Finally, we do a synthesis by means of an adjoint short-time Fourier transform (possibly with some other window φ_1) to get again a representation of the filtered signal in the time domain. The whole process can be summarized as follows:

$$f \mapsto V_{\varphi_1}^*(a \cdot V_{\varphi_2} f) = \iint_{\mathbb{R}^{2d}} a(x,\omega) \cdot V_{\varphi_1} f(x,\omega)\, M_\omega T_x \varphi_2 \, dx d\omega.$$

We will see shortly that this definition yields a class of pseudodifferential operators with many interesting and nice properties, especially concerning the boundedness on Hilbert space $L^2(\mathbb{R}^d)$ and Schatten class properties.

We start by making the above definition rigorous. Time-frequency localization operators may be defined for many different classes of symbols and windows. Most results in this section are well-known and may be found e.g. in [4], [6] and [7].

We begin with the general case of tempered distributions as symbols. The windows are then required to be Schwartz functions.

Definition 3.1.1 (Localization Operators as Pseudodifferential Operators)**.** *Let $a \in \mathcal{S}'(\mathbb{R}^{2d})$ and $\varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R}^d)$. The* **localization operator with symbol** a **and windows** φ_1, φ_2 *is the pseudodifferential operator $A_a^{\varphi_1,\varphi_2} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ defined by*

$$\langle A_a^{\varphi_1,\varphi_2} f, g\rangle := \left\langle a, \overline{V_{\varphi_2} f} V_{\varphi_1} g \right\rangle$$

for $f, g \in \mathcal{S}(\mathbb{R}^d)$.

The term on the r.h.s. is reasonably defined, since for $f, g, \varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R}^d)$, we have $V_{\varphi_2} f, V_{\varphi_1} g \in \mathcal{S}(\mathbb{R}^{2d})$ and in particular $\overline{V_{\varphi_2} f} V_{\varphi_1} g \in \mathcal{S}(\mathbb{R}^{2d})$. Further, if f_n converges to f in the topology of $\mathcal{S}(\mathbb{R}^d)$, then $V_{\varphi_2} f_n$ converges to $V_{\varphi_2} f$ in $\mathcal{S}(\mathbb{R}^{2d})$, and $\overline{V_{\varphi_2} f_n} V_{\varphi_1} g$ converges to $\overline{V_{\varphi_2} f} V_{\varphi_1} g$. Therefore, $\langle A_a^{\varphi_1,\varphi_2} f_n, g\rangle \to \langle A_a^{\varphi_1,\varphi_2} f, g\rangle$, so the mapping $A_a^{\varphi_1,\varphi_2} : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$ is continuous (with respect to the usual metric topology in $\mathcal{S}(\mathbb{R}^d)$ and the weak$*$-topology in $\mathcal{S}'(\mathbb{R}^d)$).

By the Kernel Theorem 2.2.5, we know that $A_a^{\varphi_1,\varphi_2}$ can be written as a Weyl operator σ^W for an appropriate Weyl symbol $\sigma \in \mathcal{S}'(\mathbb{R}^{2d})$. We will derive an explicit formula for σ. As a preparation, we will need two lemmata.

Lemma 3.1.2 (Fourier Transform of a Wigner Distribution)**.** *Let $f, g \in L^2(\mathbb{R}^d)$. Then the Fourier transform of the cross Wigner distribution of f and g can be expressed in terms of the short-time Fourier transform as*

$$\widehat{W(f,g)}(y,\eta) = e^{-\pi i y\cdot\eta} \cdot V(f,g)(-\eta, y), \qquad \textit{for all } y,\eta \in \mathbb{R}^d.$$

Proof. The Wigner distribution $W(f,g)(x,\omega)$ is just the generalized bilinear time-frequency distribution $\mathrm{TF}_A(f,g)(x,\omega)$ with matrix $A = \left(\begin{smallmatrix} I & \frac{1}{2}I \\ I & -\frac{1}{2}I \end{smallmatrix}\right)$. By Theorem 1.2.16, the Fourier transform of this bilinear distribution is given by

$$\widehat{\mathrm{TF}_A(f,g)}(y,\eta) = \mathrm{TF}_B(f,g)(\eta, y)$$

with the matrix $B = A \cdot \left(\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -\frac{1}{2}I & I \\ \frac{1}{2}I & I \end{smallmatrix}\right)$. Hence

$$\begin{aligned}\widehat{W(f,g)}(y,\eta) &= \mathrm{TF}_B(f,g)(\eta,y) \\ &= \int_{\mathbb{R}^d} f(-\frac{\eta}{2} + t) \cdot \overline{g(\frac{\eta}{2} + t)} \cdot e^{-2\pi i y\cdot t}\, dt.\end{aligned}$$

The substitution $s = t - \frac{\eta}{2}$ yields

$$\begin{aligned}\widehat{W(f,g)}(y,\eta) &= \int_{\mathbb{R}^d} f(s) \cdot \overline{g(s+\eta)} \cdot e^{-2\pi i y\cdot(s+\frac{\eta}{2})}\, ds \\ &= e^{-\pi i y\cdot\eta} \int_{\mathbb{R}^d} f(s) \cdot \overline{g(s+\eta)} \cdot e^{-2\pi i y\cdot s}\, ds \\ &= e^{-\pi i y\cdot\eta} V(f,g)(-\eta, y).\end{aligned}$$

□

Lemma 3.1.3 (Fourier Transform of a Product of Short-Time Fourier Transforms)**.** *Let $f_1, f_2, g_1, g_2 \in L^2(\mathbb{R}^d)$. The the following holds:*

$$\widehat{\left(V(f_1,g_1) \cdot \overline{V(f_2,g_2)}\right)}(y,\eta) = V(f_1,f_2)(-\eta,y) \cdot \overline{V(g_1,g_2)(-\eta,y)}$$

for all $y, \eta \in \mathbb{R}^d$.

Proof. The Fourier transform yields

$$\begin{aligned}
&\Big(V(f_1,g_1)\cdot\widehat{\overline{V(f_2,g_2)}}\Big)(y,\eta)\\
&\quad=\iint_{\mathbb{R}^{2d}} V(f_1,g_1)(x,\omega)\cdot\overline{V(f_2,g_2)(x,\omega)}\cdot e^{-2\pi i(x\cdot y+\omega\cdot\eta)}\,dxd\omega\\
&\quad=\left\langle V(f_1,g_1),V(f_2,g_2)\cdot e^{2\pi i(\bullet\cdot y+\bullet\cdot\eta)}\right\rangle.
\end{aligned}$$

Now

$$\begin{aligned}
&V(f_2,g_2)(x,\omega)\cdot e^{2\pi i(x\cdot y+\omega\cdot\eta)}\\
&\quad=\int_{\mathbb{R}^d} f_2(t)\overline{g_2(t-x)}e^{-2\pi i\omega\cdot t}\,dt\cdot e^{2\pi i(x\cdot y+\omega\cdot\eta)}\\
&\quad=\int_{\mathbb{R}^d} f_2(s+\eta)\overline{g_2(s+\eta-x)}e^{-2\pi i\omega\cdot(s+\eta)}\,ds\cdot e^{2\pi i(x\cdot y+\omega\cdot\eta)}\\
&\quad=\int_{\mathbb{R}^d} f_2(s+\eta)\overline{g_2(s+\eta-x)}e^{-2\pi i\omega\cdot s}\,ds\cdot e^{2\pi ix\cdot y}\\
&\quad=\int_{\mathbb{R}^d} e^{-2\pi is\cdot y}(M_yT_{-\eta}f_2)(s)e^{2\pi i(s-x)\cdot y}\overline{(M_yT_{-\eta}g_2)(s-x)}e^{-2\pi i\omega\cdot s}\,ds\cdot e^{2\pi ix\cdot y}\\
&\quad=\int_{\mathbb{R}^d}(M_yT_{-\eta}f_2)(s)\overline{(M_yT_{-\eta}g_2)(s-x)}e^{-2\pi i\omega\cdot s}\,ds\\
&\quad=V(M_yT_{-\eta}f_2,M_yT_{-\eta}g_2)(x,\omega)
\end{aligned}$$

with the substitution $t=s+\eta$. Therefore

$$\begin{aligned}
&\Big(V(f_1,g_1)\cdot\widehat{\overline{V(f_2,g_2)}}\Big)(y,\eta)\\
&\quad=\langle V(f_1,g_1),V(M_yT_{-\eta}f_2,M_yT_{-\eta}g_2)\rangle\\
&\quad=\langle f_1,M_yT_{-\eta}f_2\rangle\,\overline{\langle g_1,M_yT_{-\eta}g_2\rangle}\\
&\quad=V(f_1,f_2)(-\eta,y)\cdot\overline{V(g_1,g_2)(-\eta,y)}
\end{aligned}$$

by the orthogonality relation for the short-time Fourier transform. □

Now for the Weyl symbol of a localization operator.

Theorem 3.1.4 (Connection with the Weyl Calculus)**.** *Let $a\in\mathcal{S}'(\mathbb{R}^{2d})$ and $\varphi_1,\varphi_2\in\mathcal{S}(\mathbb{R}^d)$. Then the localization operator $A_a^{\varphi_1,\varphi_2}$ possesses the Weyl* *symbol $a * W(\varphi_1, \varphi_2)$, where $W(\varphi_1, \varphi_2)$ denotes the Wigner distribution of φ_1, φ_2; in other words,*

$$A_a^{\varphi_1,\varphi_2} = \sigma^W$$

with

$$\sigma = a * W(\varphi_1, \varphi_2) \in \mathcal{S}'(\mathbb{R}^{2d}).$$

Proof. Let $\sigma = a * W(\varphi_1, \varphi_2) \in \mathcal{S}'(\mathbb{R}^{2d})$. Let $f, g \in \mathcal{S}(\mathbb{R}^d)$. Then

$$\begin{aligned}
\left\langle \sigma^W f, g \right\rangle &= \langle \sigma, W(g,f) \rangle \\
&= \langle a * W(\varphi_1, \varphi_2), W(g,f) \rangle \\
&\underset{\text{Parseval}}{=} \left\langle \widehat{a} \cdot \widehat{W(\varphi_1, \varphi_2)}, \widehat{W(g,f)} \right\rangle \\
&= \left\langle \widehat{a}, \overline{\widehat{W(\varphi_1, \varphi_2)}} \cdot \widehat{W(g,f)} \right\rangle \\
&\underset{\text{Parseval}}{=} \left\langle a, \mathcal{F}^{-1}\Big(\overline{\widehat{W(\varphi_1, \varphi_2)}} \cdot \widehat{W(g,f)} \Big) \right\rangle .
\end{aligned}$$

Now by Lemma 3.1.2,

$$\widehat{W(g,f)}(y,\eta) = e^{-\pi i y \cdot \eta} V(g,f)(-\eta, y)$$

and

$$\overline{\widehat{W(\varphi_1, \varphi_2)}(y,\eta)} = e^{\pi i y \cdot \eta} \overline{V(\varphi_1, \varphi_2)(-\eta, y)}.$$

Hence

$$\begin{aligned}
\Big(\overline{\widehat{W(\varphi_1, \varphi_2)}} \cdot \widehat{W(g,f)} \Big)(y,\eta) &= \Big(\overline{V(\varphi_1, \varphi_2)} \cdot V(g,f) \Big)(-\eta, y) \\
&= \mathcal{F}\Big(\overline{V(f, \varphi_2)} \cdot V(g, \varphi_1) \Big)(y,\eta)
\end{aligned}$$

and

$$\mathcal{F}^{-1}\Big(\overline{\widehat{W(\varphi_1, \varphi_2)}} \cdot \widehat{W(g,f)} \Big) = \Big(\overline{V(f, \varphi_2)} \cdot V(g, \varphi_1) \Big)(x,\omega)$$

by Lemma 3.1.3. Thus

$$\begin{aligned}
\left\langle \sigma^W f, g \right\rangle &= \left\langle a, \overline{V(f, \varphi_2)} \cdot V(g, \varphi_1) \right\rangle \\
&= \left\langle A_a^{\varphi_1,\varphi_2} f, g \right\rangle .
\end{aligned}$$

□

As a general principle, if we restrict the class of symbols to some subset of the tempered distributions, we may allow windows from a larger set than the Schwartz functions. We consider first symbols in L^p-spaces. In this case, $L^2(\mathbb{R}^d)$ is a suitable class of windows.

Definition 3.1.5 (Localization Operators with Symbols in L^p-Spaces)**.** *Let $a \in L^p(\mathbb{R}^{2d})$ and $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^d)$. The* **localization operator with symbol** *a* **and windows** *φ_1, φ_2 on $L^2(\mathbb{R}^d)$ is the bounded linear operator on $L^2(\mathbb{R}^d)$ defined by*

$$A_a^{\varphi_1,\varphi_2} f := V_{\varphi_1}^*(a \cdot V_{\varphi_2} f), \qquad f \in L^2(\mathbb{R}^d).$$

In order to make sense, this definition needs to be explained and interpreted in an appropriate way.

Assume first that $a \in L^\infty(\mathbb{R}^{2d})$.

This yields indeed a bounded operator on $L^2(\mathbb{R}^d)$: V_{φ_2} and $V_{\varphi_1}^*$ are bounded operators from $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^{2d})$ and from $L^2(\mathbb{R}^{2d})$ to $L^2(\mathbb{R}^d)$, respectively, and multiplication with a function in $L^\infty(\mathbb{R}^{2d})$ is bounded $L^2(\mathbb{R}^{2d}) \to L^2(\mathbb{R}^{2d})$.

We calculate

$$\begin{aligned} ||A_a^{\varphi_1,\varphi_2} f|| &= ||V_{\varphi_1}^*(a \cdot V_{\varphi_2} f)|| \\ &\leq ||V_{\varphi_1}^*||_{L^2\to L^2} \cdot ||a||_\infty \cdot ||V_{\varphi_2}||_{L^2\to L^2} \cdot ||f|| \\ &= ||\varphi_1|| \cdot ||\varphi_2|| \cdot ||a||_\infty \cdot ||f||, \end{aligned}$$

so $A_a^{\varphi_1,\varphi_2} \in B(L^2(\mathbb{R}^d))$ and $||A_a^{\varphi_1,\varphi_2}||_{B(L^2)} \leq ||\varphi_1|| \cdot ||\varphi_2|| \cdot ||a||_\infty$.

The mapping $\mathcal{A} : L^\infty(\mathbb{R}^{2d}) \to B(L^2(\mathbb{R}^d))$, $a \mapsto \mathcal{A}a := A_a^{\varphi_1,\varphi_2}$ is a bounded linear operator with $||\mathcal{A}||_{L^\infty \to B(L^2)} \leq ||\varphi_1|| \cdot ||\varphi_2||$.

Next suppose that $a \in L^1(\mathbb{R}^{2d})$.

If $\varphi_2 \in L^2(\mathbb{R}^d)$ and $f \in L^2(\mathbb{R}^d)$, then $V_{\varphi_2} f \in L^\infty(\mathbb{R}^{2d})$ and $||V_{\varphi_2} f||_\infty \leq ||f|| \cdot ||\varphi_2||$, so V_{φ_2} is a bounded linear operator from $L^2(\mathbb{R}^d)$ to $L^\infty(\mathbb{R}^{2d})$ with operator norm $||V_{\varphi_2}||_{L^2\to L^\infty} \leq ||\varphi_2||$. Multiplication with a function in $L^1(\mathbb{R}^{2d})$ yields a function in $L^1(\mathbb{R}^{2d})$, whereas $V_{\varphi_1}^*$ is the adjoint of $V_{\varphi_1} : L^2(\mathbb{R}^d) \to L^\infty(\mathbb{R}^{2d})$ and is therefore bounded from $L^1 \to L^2$ with operator norm $||V_{\varphi_1}^*||_{L^1\to L^2} = ||V_{\varphi_1}||_{L^2\to L^\infty} \leq ||\varphi_1||$.

We estimate the operator norm of $A_a^{\varphi_1,\varphi_2}$:

$$\begin{aligned}||A_a^{\varphi_1,\varphi_2} f|| &= ||V_{\varphi_1}^*(a \cdot V_{\varphi_2} f)|| \\ &\leq ||V_{\varphi_1}^*||_{L^1 \to L^2} \cdot ||a \cdot V_{\varphi_2} f||_1 \\ &\leq ||V_{\varphi_1}^*||_{L^1 \to L^2} \cdot ||a||_1 \cdot ||V_{\varphi_2} f||_\infty \\ &\leq ||\varphi_1|| \cdot ||\varphi_2|| \cdot ||a||_1 \cdot ||f||,\end{aligned}$$

hence again $A_a^{\varphi_1,\varphi_2} \in B(L^2(\mathbb{R}^d))$ and $||A_a^{\varphi_1,\varphi_2}||_{B(L^2)} \leq ||\varphi_1|| \cdot ||\varphi_2|| \cdot ||a||_1$.

We denote the mapping from symbol to operator again by $\mathcal{A} : L^1(\mathbb{R}^{2d}) \to B(L^2(\mathbb{R}^d))$; the preceding estimate shows that also in this case $\mathcal{A}$ is a bounded linear operator with $||\mathcal{A}||_{L^1 \to B(L^2)} \leq ||\varphi_1|| \cdot ||\varphi_2||$.

Finally, assume $a \in L^p(\mathbb{R}^{2d})$, $1 < p < \infty$.

In this case, we use interpolation. The mapping $\mathcal{A}$ is well-defined as an operator from $L^1(\mathbb{R}^{2d})$ as well as from $L^\infty(\mathbb{R}^{2d})$ into the space $B(L^2(\mathbb{R}^d))$. On the subset $L^1(\mathbb{R}^{2d}) \cap L^\infty(\mathbb{R}^{2d})$ the two definitions coincide. Using the complex interpolation method on $L^p = [L^1, L^\infty]_\theta$, the mapping $\mathcal{A}$ may then be extended to a bounded operator $\mathcal{A} : L^p(\mathbb{R}^{2d}) \to B(L^2(\mathbb{R}^d))$ on each of the intermediate spaces $L^p(\mathbb{R}^{2d})$, $1 < p < \infty$. The theorem yields the following estimate for the norm:

$$\begin{aligned}||\mathcal{A}||_{L^p \to B(L^2)} &\leq (||\varphi_1|| \cdot ||\varphi_2||)^{1-\theta} \cdot (||\varphi_1|| \cdot ||\varphi_2||)^\theta \cdot ||a||_p \\ &\leq ||\varphi_1|| \cdot ||\varphi_2|| \cdot ||a||_p,\end{aligned}$$

where the constant $\theta \in [0, 1]$ depending on p cancels in the end. In this way we can explain the localization operator $A_a^{\varphi_1,\varphi_2}$ also for symbols $a \in L^p(\mathbb{R}^{2d})$ as

$$A_a^{\varphi_1,\varphi_2} f = (\mathcal{A}a)(f), \qquad f \in L^2(\mathbb{R}^d).$$

The preceding considerations give rise to the following

Corollary 3.1.6. *Let $a \in L^p(\mathbb{R}^{2d})$, $1 \leq p \leq \infty$, and $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^d)$. The localization operator $A_a^{\varphi_1,\varphi_2}$ satisfies the norm estimate*

$$||A_a^{\varphi_1,\varphi_2} f|| \leq ||\varphi_1|| \cdot ||\varphi_2|| \cdot ||a||_p \cdot ||f||$$

for all $f \in L^2(\mathbb{R}^d)$. *Equivalently,*

$$||A_a^{\varphi_1,\varphi_2}||_{B(L^2)} \leq ||\varphi_1|| \cdot ||\varphi_2|| \cdot ||a||_p$$

and

$$||\mathcal{A}||_{L^p \to B(L^2)} \leq ||\varphi_1|| \cdot ||\varphi_2||.$$

□

In case of L^p-symbols, we can prove even more than mere boundedness. These operators satisfy certain compactness and Schatten p-class properties.

Lemma 3.1.7 (Compactness for Symbols with Compact Support)**.** *Let* $a \in L^\infty(\mathbb{R}^{2d})$ *have compact support, i.e., there is a compact set* $K \subseteq \mathbb{R}^{2d}$ *such that* $a(x,\omega) = 0$ *for* $(x,\omega) \in \mathbb{R}^{2d} \setminus K$ *almost everywhere. Let* $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^d)$. *Then the localization operator* $A_a^{\varphi_1,\varphi_2} \in B(L^2(\mathbb{R}^d))$ *is a compact operator on* $L^2(\mathbb{R}^d)$.

Proof. Denote by

$$M_a : L^2(\mathbb{R}^{2d}) \to L^2(\mathbb{R}^{2d}), \quad F(x,\omega) \mapsto (M_a F)(x,\omega) := a(x,\omega) \cdot F(x,\omega)$$

the multiplication operator with the function a. Since a is bounded by assumption, M_a is a bounded operator. We have

$$A_a^{\varphi_1,\varphi_2} = V_{\varphi_1}^* \circ M_a \circ V_{\varphi_2},$$

thus it suffices to show that $M_a \circ V_{\varphi_2}$ is compact.
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $L^2(\mathbb{R}^d)$ that converges weakly to zero, $f_n \xrightarrow{w} 0$ for $n \to \infty$. We will show that $A_a^{\varphi_1,\varphi_2} f_n \to 0$ in the norm for $n \to \infty$.
To this end, consider $a \cdot V_{\varphi_2} f_n$. Since $a \in L^\infty(\mathbb{R}^{2d})$ and $V_{\varphi_2} f_n \in L^2(\mathbb{R}^{2d})$, $a \cdot V_{\varphi_2} f_n \in L^2(\mathbb{R}^{2d})$. The norm is given by

$$\begin{aligned} ||a \cdot V_{\varphi_2} f_n||_{L^2}^2 &= \iint_{\mathbb{R}^{2d}} |a(x,\omega)|^2 \cdot |V_{\varphi_2} f_n(x,\omega)|^2 \, dx d\omega \\ &= \iint_K |a(x,\omega)|^2 \cdot |V_{\varphi_2} f_n(x,\omega)|^2 \, dx d\omega. \end{aligned}$$

We have for every $(x,\omega) \in K$

$$\begin{aligned} |a(x,\omega)|^2 \cdot |V_{\varphi_2} f_n(x,\omega)|^2 \\ &\leq ||a||^2_{L^\infty} \cdot \underbrace{|\langle f_n, M_\omega T_x \varphi_2\rangle|^2}_{\underset{n\to\infty}{\longrightarrow} 0} \\ &\underset{n\to\infty}{\longrightarrow} 0, \end{aligned}$$

since $f_n \xrightarrow{w} 0$, i.e. $\langle f_n, g\rangle \to 0$ for every $g \in L^2(\mathbb{R}^d)$.
Thus the integrand converges to zero pointwisely on K.
Furthermore, weakly convergent sequences are norm bounded, i.e., there is a constant $C > 0$ such that $||f_n|| \leq C$ for every $n \in \mathbb{N}$. Thus

$$\begin{aligned} |a(x,\omega)|^2 \cdot |V_{\varphi_2} f_n(x,\omega)|^2 \\ &\leq ||a||^2_{L^\infty} \cdot ||\varphi_2||^2 \cdot ||f_n||^2 \\ &\leq C^2 \cdot ||a||^2_{L^\infty} \cdot ||\varphi_2||^2 \\ &\in L^1(K) \end{aligned}$$

is an integrable majorant over K independent of n.
Hence

$$||a \cdot V_{\varphi_2} f_n||^2 \to 0, \qquad n \to \infty,$$

by the Dominated Convergence Theorem, that means $a \cdot V_{\varphi_2} f_n \to 0$ in $L^2(\mathbb{R}^{2d})$. But then also $A_a^{\varphi_1,\varphi_2} f_n = V^*_{\varphi_1}(a \cdot V_{\varphi_2} f_n) \to 0$ in $L^2(\mathbb{R}^d)$. □

Theorem 3.1.8 (Compactness for L^p-Symbols)**.** *Let $a \in L^p(\mathbb{R}^{2d})$, $1 \leq p < \infty$, and $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^d)$.*
Then the localization operator $A_a^{\varphi_1,\varphi_2} \in B(L^2(\mathbb{R}^d))$ is a compact operator on $L^2(\mathbb{R}^d)$.

Proof. For every $N \in \mathbb{N}$, set

$$A_N := \{(x,\omega) \in \mathbb{R}^{2d} |\, |a(x,\omega)| \leq N \text{ and } ||(x,\omega)|| \leq N\}.$$

We have $A_N \subseteq A_{N+1}$ for all $N \in \mathbb{N}$ and $\cup_{N\in\mathbb{N}} A_N = \mathbb{R}^{2d}$.
Now set

$$a_N(x,\omega) := a(x,\omega) \cdot \chi_{A_N}, \qquad N \in \mathbb{N},$$

where χ_{A_N} denotes the characteristic function of the set A_N. Then $a_N \in L^p(\mathbb{R}^{2d}) \cap L^\infty(\mathbb{R}^{2d})$ and has compact support. Thus the localization operator

$A_{a_N}^{\varphi_1,\varphi_2}$ is a compact operator by Lemma 3.1.7. Furthermore, $a_N \to a$ in L^p, i.e. $\lim_{N\to\infty} ||a_N - a||_{L^p} = 0$.
Now obviously

$$A_a^{\varphi_1,\varphi_2} f - A_{a_N}^{\varphi_1,\varphi_2} f = A_{a-a_N}^{\varphi_1,\varphi_2} f$$

for all $f \in L^2(\mathbb{R}^d)$, thus

$$||A_a^{\varphi_1,\varphi_2} - A_{a_N}^{\varphi_1,\varphi_2}||_{B(L^2)} \leq ||a - a_N||_{L^p} \cdot ||\varphi_1|| \cdot ||\varphi_2|| \to 0$$

for $N \to \infty$, by Corollary 3.1.6. So $A_a^{\varphi_1,\varphi_2}$ is a uniform limit of compact operators and therefore compact. □

Theorem 3.1.9 (Schatten Class for L^p-Symbols)**.** *If $a \in L^p(\mathbb{R}^{2d})$, $1 \leq p < \infty$, then $A_a^{\varphi_1,\varphi_2} \in \mathcal{S}^p(L^2(\mathbb{R}^d))$, the Schatten p-class, and*

$$||A_a^{\varphi_1,\varphi_2}||_{\mathcal{S}^p} \leq ||a||_{L^p} \cdot ||\varphi_1|| \cdot ||\varphi_2||.$$

Proof. From the discussion above resp. Corollary 3.1.6, we know that if $a \in L^\infty(\mathbb{R}^{2d})$, then $A_a^{\varphi_1,\varphi_2} \in B(L^2(\mathbb{R}^d)) = \mathcal{S}^\infty(L^2(\mathbb{R}^d))$ and

$$||A_a^{\varphi_1,\varphi_2}||_{\mathcal{S}^\infty} = ||A_a^{\varphi_1,\varphi_2}||_{B(L^2)} \leq ||a||_{L^\infty} \cdot ||\varphi_1|| \cdot ||\varphi_2||.$$

That means that the mapping $\mathcal{A} : L^\infty \to \mathcal{S}^\infty$, $a \mapsto \mathcal{A}a = A_a^{\varphi_1,\varphi_2}$ is bounded linear with norm

$$||\mathcal{A}||_{L^\infty \to \mathcal{S}^\infty} \leq ||\varphi_1|| \cdot ||\varphi_2||.$$

Next consider $a \in L^1(\mathbb{R}^{2d})$. To show that $A_a^{\varphi_1,\varphi_2}$ is a trace class operator we use the criterion in Theorem A.7.9.
By Theorem 3.1.8, $A_a^{\varphi_1,\varphi_2}$ is a compact operator.
Let $(e_k)_{k\in\mathbb{N}}$ be an arbitrary orthonormal basis of $L^2(\mathbb{R}^d)$. Then

$$\begin{aligned}
&\sum_{k\in\mathbb{N}} |\left\langle A_a^{\varphi_1,\varphi_2} e_k, e_k \right\rangle| \\
&\quad = \sum_{k\in\mathbb{N}} |\left\langle V_{\varphi_1}^*(a \cdot V_{\varphi_2} e_k), e_k \right\rangle| \\
&\quad = \sum_{k\in\mathbb{N}} |\left\langle a \cdot V_{\varphi_2} e_k, V_{\varphi_1} e_k \right\rangle| \\
&\quad = \sum_{k\in\mathbb{N}} |\int_{\mathbb{R}^{2d}} a(x,\omega) \cdot V(e_k,\varphi_2)(x,\omega) \cdot \overline{V(e_k,\varphi_1)(x,\omega)}\, dxd\omega| \\
&\quad \leq \sum_{k\in\mathbb{N}} \int_{\mathbb{R}^{2d}} |a(x,\omega)| \cdot |V(e_k,\varphi_2)(x,\omega)| \cdot |V(e_k,\varphi_1)(x,\omega)|\, dxd\omega.
\end{aligned}$$

Since all the integrands involved are nonnegative, we may change the order of summation and integration by Fubini's resp. Tonelli's Theorem to continue

$$\sum_{k\in\mathbb{N}} \int_{\mathbb{R}^{2d}} |a(x,\omega)| \cdot |V(e_k,\varphi_2)(x,\omega)| \cdot |V(e_k,\varphi_1)(x,\omega)| \, dx d\omega$$
$$= \int_{\mathbb{R}^{2d}} |a(x,\omega)| \cdot \sum_{k\in\mathbb{N}} |V(e_k,\varphi_2)(x,\omega)| \cdot |V(e_k,\varphi_1)(x,\omega)| \, dx d\omega.$$

Now, using the Cauchy-Schwarz Inequality,

$$\begin{aligned}
&\sum_{k\in\mathbb{N}} |V(e_k,\varphi_2)(x,\omega)| \cdot |V(e_k,\varphi_1)(x,\omega)| \\
&\quad \le \left(\sum_{k\in\mathbb{N}} |V(e_k,\varphi_2)(x,\omega)|^2 \right)^{1/2} \cdot \left(\sum_{k\in\mathbb{N}} |V(e_k,\varphi_1)(x,\omega)|^2 \right)^{1/2} \\
&\quad = \left(\sum_{k\in\mathbb{N}} | \langle e_k, M_\omega T_x \varphi_2 \rangle |^2 \right)^{1/2} \cdot \left(\sum_{k\in\mathbb{N}} | \langle e_k, M_\omega T_x \varphi_1 \rangle |^2 \right)^{1/2} \\
&\quad = \|M_\omega T_x \varphi_2\| \cdot \|M_\omega T_x \varphi_1\| \\
&\quad = \|\varphi_2\| \cdot \|\varphi_1\|
\end{aligned}$$

since $(e_k)_{k\in\mathbb{N}}$ is an orthonormal basis.
Hence

$$\begin{aligned}
&\sum_{k\in\mathbb{N}} | \langle A_a^{\varphi_1,\varphi_2} e_k, e_k \rangle | \\
&\quad \le \int_{\mathbb{R}^{2d}} |a(x,\omega)| \, dx d\omega \cdot \|\varphi_2\| \cdot \|\varphi_1\| \\
&\quad = \|a\|_{L^1} \cdot \|\varphi_2\| \cdot \|\varphi_1\|.
\end{aligned}$$

By Theorem A.7.9, we conclude

$$A_a^{\varphi_1,\varphi_2} \in \mathcal{S}^1(L^2(\mathbb{R}^d))$$

and

$$\|A_a^{\varphi_1,\varphi_2}\|_{\mathcal{S}^1} = \sup_{(e_k)_{k\in\mathbb{N}} \text{ ONB}} \sum_{k\in\mathbb{N}} | \langle A_a^{\varphi_1,\varphi_2} e_k, e_k \rangle | \le \|a\|_{L^1} \cdot \|\varphi_2\| \cdot \|\varphi_1\|.$$

That is to say that the mapping $\mathcal{A}$ also satisfies $\mathcal{A} : L^1(\mathbb{R}^{2d}) \to \mathcal{S}^1(L^2(\mathbb{R}^d))$ with

$$\|\mathcal{A}\|_{L^1\to\mathcal{S}^1} \leq \|\varphi_1\| \cdot \|\varphi_2\|.$$

Finally, for $a \in L^p(\mathbb{R}^{2d})$ we use again complex interpolation between Lebesgue spaces L^p on the one hand and Schatten p-classes $\mathcal{S}^p$ on the other hand. We thus find that $\mathcal{A}a = A_a^{\varphi_1,\varphi_2} \in \mathcal{S}^p(L^2(\mathbb{R}^d))$ for $a \in L^p(\mathbb{R}^{2d})$ with norm estimate

$$\begin{aligned}\|\mathcal{A}\|_{L^p\to\mathcal{S}^p} &\leq \|\mathcal{A}\|_{L^1\to\mathcal{S}^1}^{\theta} \cdot \|\mathcal{A}\|_{L^\infty\to\mathcal{S}^\infty}^{1-\theta} \\ &= (\|\varphi_1\| \cdot \|\varphi_2\|)^\theta \cdot (\|\varphi_1\| \cdot | < \varphi_2\|)^{1-\theta} \\ &= \|\varphi_1\| \cdot \|\varphi_2\|\end{aligned}$$

with $[L^1, L^\infty]_\theta = L^p$, $[\mathcal{S}^1, \mathcal{S}^\infty]_\theta = \mathcal{S}^p$.
So

$$\|A_a^{\varphi_1,\varphi_2}\|_{\mathcal{S}^p} = \|\mathcal{A}a\|_{\mathcal{S}^p} \leq \|\mathcal{A}\|_{L^p\to\mathcal{S}^p} \cdot \|a\|_{L^p} = \|\varphi_1\| \cdot \|\varphi_2\| \cdot \|a\|_{L^p}.$$

□

Finally, we want to look at symbols taken from modulation spaces. We expect once more that modulation spaces behave nicely as symbols for localization operators, in view of the very definition of these operators in terms of time-frequency analysis. The next two theorems show that modulation spaces indeed meet our expectation. The modulation space $M^1(\mathbb{R}^d)$, Feichtinger's Algebra, turns out to be the appropriate window class.

Theorem 3.1.10 (Localization Operators with Symbols in Modulation Spaces)**.** *If $a \in M^\infty(\mathbb{R}^{2d})$ and $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$, then $A_a^{\varphi_1,\varphi_2}$ is a bounded operator on all the modulation spaces $M^{p,q}(\mathbb{R}^d)$, $1 \leq p, q \leq \infty$, in particular on $L^2(\mathbb{R}^d) = M^{2,2}(\mathbb{R}^d)$. Furthermore,*

$$\|A_a^{\varphi_1,\varphi_2}\|_{B(M^{p,q})} \leq C \cdot \|a\|_{M^{\infty,1}} \cdot \|\varphi_1\|_{M^1} \cdot \|\varphi_2\|_{M^1}.$$

Proof. We know that $A_a^{\varphi_1,\varphi_2} = \sigma^W$ for $\sigma = a * W(\varphi_1, \varphi_2)$ by Theorem 3.1.4. We will show that under the given assumptions we have $\sigma \in M^{\infty,1}(\mathbb{R}^{2d})$ The result then follows from Theorem 2.3.1 applied to the Weyl calculus (note that the Weyl calculus is associated to the Wigner distribution, cf. Theorem 2.1.5; this is the bilinear time-frequency distribution with the matrix $\begin{pmatrix} I & \frac{1}{2}I \\ I & -\frac{1}{2}I \end{pmatrix}$, which is left-regular, thus Theorem 2.3.1 is applicable).
In order to show $\sigma \in M^{\infty,1}(\mathbb{R}^{2d})$, we prove a

Lemma 3.1.11 (Convolution Relations for Modulation Spaces I)**.** *We have*

$$M^\infty(\mathbb{R}^d) * M^1(\mathbb{R}^d) \subseteq M^{\infty,1}(\mathbb{R}^d)$$

and

$$\|f * g\|_{M^{\infty,1}} \leq C \cdot \|f\|_{M^\infty} \cdot \|g\|_{M^1}$$

for $f \in M^\infty(\mathbb{R}^d)$ and $g \in M^1(\mathbb{R}^d)$.

Proof. Let $f \in M^\infty(\mathbb{R}^d)$ and $g \in M^1(\mathbb{R}^d)$. Let $\phi \in \mathcal{S}(\mathbb{R}^d)$ and $\Phi = \phi * \phi \in \mathcal{S}(\mathbb{R}^d)$.
Then

$$\begin{aligned} \|f * g\|_{M^{\infty,1}} &\leq C \cdot \|V_\Phi(f * g)\|_{L^{\infty,1}} \\ &= C \cdot \int_{\mathbb{R}^d} \sup_{x \in \mathbb{R}^d} |V_\Phi(f * g)(x, \omega)| \, d\omega. \end{aligned}$$

We use the (easy to prove) identity

$$V_\psi h(x, \omega) = e^{-2\pi i x \cdot \omega}(h * M_\omega \tilde{\psi})(x)$$

for the short-time Fourier transform, where $\tilde{\psi}(t) = \overline{\psi(-t)}$ denotes the usual involution. Then

$$\begin{aligned} |V_\Phi(f * g)(x, \omega)| &= |\big((f * g) * M_\omega \tilde{\Phi}\big)(x)| \\ &= |\big((f * g) * M_\omega(\tilde{\phi} * \tilde{\phi})\big)(x)| \\ &= |\big((f * g) * (M_\omega \tilde{\phi} * M_\omega \tilde{\phi})\big)(x)| \\ &= |\big((f * M_\omega \tilde{\phi}) * (g * M_\omega \tilde{\phi})\big)(x)|. \end{aligned}$$

with help from Lemma A.2.6 and the commutativity of convolutions.
Thus for fixed $\omega \in \mathbb{R}^d$

$$\begin{aligned} \sup_{x \in \mathbb{R}^d} |V_\Phi(f * g)(x, \omega)| &= \|(f * M_\omega \tilde{\phi}) * (g * M_\omega \tilde{\phi})\|_{L^\infty} \\ &\leq \|f * M_\omega \tilde{\phi}\|_{L^\infty} \cdot \|g * M_\omega \tilde{\phi}\|_{L^1} \end{aligned}$$

by Young's Inequality.
The first term can be estimated by

$$\begin{aligned}\|f * M_\omega \tilde{\phi}\|_{L^\infty} &= \sup_{x\in\mathbb{R}^d} |(f * M_\omega \tilde{\phi})(x)| \\ &= \sup_{x\in\mathbb{R}^d} |V_\phi f(x,\omega)| \\ &\leq \|V_\phi f\|_{L^\infty} \\ &\leq C \cdot \|f\|_{M^\infty}\end{aligned}$$

with $C = C(\phi)$. So we find

$$\begin{aligned}&\int_{\mathbb{R}^d} \sup_{x\in\mathbb{R}^d} |V_\Phi(f * g)(x,\omega)|\, d\omega \\ &\qquad\leq C \cdot \|f\|_{M^\infty} \int_{\mathbb{R}^d} \|g * M_\omega \tilde{\phi}\|_{L^1}\, d\omega \\ &\qquad= C \cdot \|f\|_{M^\infty} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |V_\phi g(x,\omega)|\, dx d\omega \\ &\qquad\leq C \cdot \|f\|_{M^\infty} \cdot \|g\|_{M^1} \\ &\qquad< \infty\end{aligned}$$

with some generic constant C.
This concludes the proof of the lemma. □

Now observe that $W(\varphi_1, \varphi_2) \in M^1(\mathbb{R}^{2d})$ since $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$, by Theorem 1.7.2. Thus the lemma yields

$$a * W(\varphi_1, \varphi_2) \in M^\infty(\mathbb{R}^{2d}) * M^1(\mathbb{R}^{2d}) \subseteq M^{\infty,1}(\mathbb{R}^{2d}),$$

so the theorem follows. The norm bound comes from

$$\begin{aligned}\|A_a^{\varphi_1,\varphi_2}\|_{B(M^{p,q})} &= \|(a * W(\varphi_1, \varphi_2))^W\|_{B(M^{p,q})} \\ &\leq C \cdot \|a * W(\varphi_1, \varphi_2)\|_{M^{\infty,1}} \\ &\leq C \cdot \|a\|_{M^\infty} \cdot \|W(\varphi_1, \varphi_2)\|_{M^1} \\ &\leq C \cdot \|a\|_{M^\infty} \cdot \|\varphi_1\|_{M^1} \cdot \|\varphi_2\|_{M^1}\end{aligned}$$

by the respective estimates in Theorem 2.3.1, Lemma 3.1.11 and Theorem 1.7.2. □

Observe that $M^{p,\infty}(\mathbb{R}^{2d}) \subseteq M^{\infty,\infty}(\mathbb{R}^{2d}) = M^{\infty}(\mathbb{R}^{2d})$ for all $1 \leq p \leq \infty$, thus localization operators with symbols in $M^{p,\infty}(\mathbb{R}^{2d})$ are well defined by the preceding theorem. For these operators, the following Schatten class property holds.

Theorem 3.1.12 (Schatten Class for $M^{p,\infty}$-Symbols)**.** *If $a \in M^{p,\infty}(\mathbb{R}^{2d})$, $1 \leq p < \infty$, then $A_a^{\varphi_1,\varphi_2} \in \mathcal{S}^p(L^2(\mathbb{R}^d))$, the Schatten p-class.*

Proof. Assume first that $a \in M^{1,\infty}(\mathbb{R}^{2d})$. We will show that $A_a^{\varphi_1,\varphi_2} \in \mathcal{S}^1(L^2(\mathbb{R}^d))$ is a trace class operator.
Again, we use $A_a^{\varphi_1,\varphi_2} = \sigma^W$ for $\sigma = a * W(\varphi_1, \varphi_2)$ by Theorem 3.1.4. By Theorem 2.2.8, σ^W is trace class if $\sigma \in M^1(\mathbb{R}^{2d})$. For $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$, we have $W(\varphi_1, \varphi_2) \in M^1(\mathbb{R}^{2d})$ by Theorem 1.7.2. Hence the statement is proved once we have shown the

Lemma 3.1.13 (Convolution Relations for Modulation Spaces II)**.** *We have*

$$M^{1,\infty}(\mathbb{R}^d) * M^1(\mathbb{R}^d) \subseteq M^1(\mathbb{R}^d)$$

and

$$\|f * g\|_{M^1} \leq C \cdot \|f\|_{M^{1,\infty}} \cdot \|g\|_{M^1}$$

for $f \in M^{1,\infty}(\mathbb{R}^d)$ and $g \in M^1(\mathbb{R}^d)$.

Proof. The proof is very similar to the proof of Lemma 3.1.11.
Let $f \in M^{1,\infty}(\mathbb{R}^d)$ and $g \in M^1(\mathbb{R}^d)$. Let $\phi \in \mathcal{S}(\mathbb{R}^d)$ and $\Phi = \phi * \phi \in \mathcal{S}(\mathbb{R}^d)$. Then

$$\begin{aligned}\|f * g\|_{M^1} &\leq C \cdot \|V_\Phi(f * g)\|_{L^1} \\ &= C \cdot \iint_{\mathbb{R}^{2d}} |V_\Phi(f * g)(x, \omega)| \, dx d\omega \\ &= C \cdot \iint_{\mathbb{R}^{2d}} \left|\big((f * M_\omega \tilde{\phi}) * (g * M_\omega \tilde{\phi})\big)(x)\right| \, dx d\omega\end{aligned}$$

completely analogous to the argument in Lemma 3.1.11.
Now for fixed $\omega \in \mathbb{R}^d$

$$\begin{aligned}\int_{\mathbb{R}^d} \left|\big((f * M_\omega \tilde{\phi}) * (g * M_\omega \tilde{\phi})\big)(x)\right| \, dx &= \|(f * M_\omega \tilde{\phi}) * (g * M_\omega \tilde{\phi})\|_{L^1} \\ &\leq \|f * M_\omega \tilde{\phi}\|_{L^1} \cdot \|g * M_\omega \tilde{\phi}\|_{L^1}\end{aligned}$$

by Young's Inequality.
The first term yields

$$\begin{aligned}\|f * M_\omega \tilde{\phi}\|_{L^1} &= \int_{\mathbb{R}^d} |\big(f * M_\omega \tilde{\phi}\big)(x)|\, dx \\ &= \int_{\mathbb{R}^d} |V_\phi f(x,\omega|\, dx \\ &\le \sup_{\omega\in\mathbb{R}^d} \int_{\mathbb{R}^d} |V_\phi f(x,\omega|\, dx \\ &= \|V_\phi f\|_{L^{1,\infty}} \\ &\le C \cdot \|f\|_{M^{1,\infty}}\end{aligned}$$

with a constant depending on ϕ. We conclude

$$\begin{aligned}&\iint_{\mathbb{R}^{2d}} |V_\Phi(f * g)(x,\omega)|\, dx d\omega \\ &\qquad \le C \cdot \|f\|_{M^{1,\infty}} \int_{\mathbb{R}^d} \|g * M_\omega \tilde{\phi}\|_{L^1}\, d\omega \\ &\qquad = C \cdot \|f\|_{M^{1,\infty}} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |V_\phi g(x,\omega)|\, d\omega \\ &\qquad \le C \cdot \|f\|_{M^{1,\infty}} \cdot \|g\|_{M^1} \\ &\qquad < \infty\end{aligned}$$

with some generic constant C.
The lemma is thus proved. □

Therefore we have

$$a * W(\varphi_1,\varphi_2) \in M^{1,\infty}(\mathbb{R}^{2d}) * M^1(\mathbb{R}^{2d}) \subseteq M^1(\mathbb{R}^{2d}),$$

hence $A_a^{\varphi_1,\varphi_2} \in \mathcal{S}^1(L^2(\mathbb{R}^d))$ for $a \in M^{1,\infty}(\mathbb{R}^{2d})$.
Now by the preceding Theorem 3.1.10, we also have $A_a^{\varphi_1,\varphi_2} \in B(L^2(\mathbb{R}^d)) = \mathcal{S}^\infty(L^2(\mathbb{R}^d))$ for $A \in M^{\infty,\infty}(\mathbb{R}^{2d}) = M^\infty(\mathbb{R}^{2d})$. By complex interpolation between the Banach spaces $M^{1,\infty}(\mathbb{R}^{2d})$ and $M^{\infty,\infty}(\mathbb{R}^{2d})$ on the one hand and $\mathcal{S}^1(L^2(\mathbb{R}^d))$ and $\mathcal{S}^\infty(L^2(\mathbb{R}^d)) = B(L^2(\mathbb{R}^d))$ on the other hand, we find that $A_a^{\varphi_1,\varphi_2} \in \mathcal{S}^p(L^2(\mathbb{R}^d))$ for symbols $a \in M^{p,\infty}(\mathbb{R}^{2d})$, as claimed. □

Remark: A more general theorem on convolution relations between modulation spaces, that contains both lemmata used above in the proofs of the last two theorems as special cases, can be found in [6]. In this context, also [39] and [40] are of interest.

The following table summarizes this section's results on the mapping properties of the correspondence $\mathcal{A} : a \mapsto A_a^{\varphi_1,\varphi_2}$ from symbols in different classes to localization operators.

Symbol	**Windows**	**Localization Operator**
$\mathcal{S}'(\mathbb{R}^{2d})$	$\mathcal{S}(\mathbb{R}^d)$	ΨDO $\mathcal{S}(\mathbb{R}^d) \to \mathcal{S}'(\mathbb{R}^d)$
$L^\infty(\mathbb{R}^{2d})$	$L^2(\mathbb{R}^d)$	$B(L^2(\mathbb{R}^d))$
$L^p(\mathbb{R}^{2d})$, $1 \leq p < \infty$	$L^2(\mathbb{R}^d)$	$\mathcal{S}^p(L^2(\mathbb{R}^d))$
$M^{\infty,\infty}(\mathbb{R}^{2d})$	$M^1(\mathbb{R}^d)$	$B(M^{p,q}(\mathbb{R}^d))$, $1 \leq p, q \leq \infty$
$M^{p,\infty}(\mathbb{R}^{2d})$, $1 \leq p < \infty$	$M^1(\mathbb{R}^d)$	$\mathcal{S}^p(L^2(\mathbb{R}^d))$

Table 3.1: Localization operators with different symbols and windows

3.2 The Berezin Transform

The Berezin transform is of great importance in complex analysis or more precisely in the theory of Hilbert spaces of analytic functions, e.g. Bergman spaces, see for example [20], or Hardy spaces, see for example [11]. It acts on bounded operators on these spaces. Assume H is a Hilbert space consisting of analytic functions on some open set $\Omega \subseteq \mathbb{C}$, which allows a reproducing kernel, i.e. for every $z \in \Omega$, there is a $K_z \in H$ such that $f(z) = \langle f, k_z \rangle$ for every $f \in H$. Let T be a bounded operator on H. Then the Berezin transform of T is the complex-valued function on Ω defined by

$$\mathcal{B}T(z) = \langle Tk_z, k_z \rangle, \qquad z \in \Omega.$$

We define an appropriate analogous version of the Berezin transform for our purposes, acting on bounded operators on $L^2(\mathbb{R}^d)$.

Definition 3.2.1 (Berezin Transform)**.** *Let $T \in B(L^2(\mathbb{R}^d))$. Let $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^d)$. The* **Berezin transform** *$\mathcal{B}$ maps T to the function on $\mathbb{R}^{2d}$*

$$\mathcal{B}T(z) := \langle T\pi(z)\varphi_2, \pi(z)\varphi_1 \rangle, \qquad z \in \mathbb{R}^{2d}.$$

Note that although the Berezin transform depends on the choice of the window functions $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^d)$, we prefer to omit this dependence notationally by simply writing $\mathcal{B}$ (rather than for example $\mathcal{B}^{\varphi_1,\varphi_2}$). This practice will in general not lead to any confusion.

For convenience and further reference, we give a representation of $\mathcal{B}T$ that is valid for operators that can be written as a strongly convergent series of rank one operators. We will make frequent use of this formula in the following.

Lemma 3.2.2. *Let $T \in B(L^2(\mathbb{R}^d))$ be of the following form: there exist some orthonormal systems $(g_n)_{n\in\mathbb{N}}$ and $(h_n)_{n\in\mathbb{N}}$ in $L^2(\mathbb{R}^d)$ and a (necessarily bounded) sequence of complex numbers $(s_n)_{n\in\mathbb{N}}$ such that*

$$Tf = \sum_{n\in\mathbb{N}} s_n \cdot \langle f, g_n \rangle h_n = \sum_{n\in\mathbb{N}} s_n \cdot (h_n \otimes g_n)(f)$$

for every $f \in L^2(\mathbb{R}^d)$ (where the series is required to converge in $L^2(\mathbb{R}^d)$ for every f).
Then

$$\mathcal{B}T(z) = \sum_{n\in\mathbb{N}} s_n \cdot V_{\varphi_1}h_n(z) \cdot \overline{V_{\varphi_2}g_n(z)}$$

for every $z \in \mathbb{R}^{2d}$.

Proof. We compute

$$\begin{aligned}
\mathcal{B}T(z) &= \langle T\pi(z)\varphi_2, \pi(z)\varphi_1\rangle \\
&= \left\langle \sum_{n\in\mathbb{N}} s_n \cdot \langle \pi(z)\varphi_2, g_n\rangle\, h_n, \pi(z)\varphi_1 \right\rangle \\
&= \sum_{n\in\mathbb{N}} s_n \cdot \langle \pi(z)\varphi_2, g_n\rangle \cdot \langle h_n, \pi(z)\varphi_1\rangle \\
&= \sum_{n\in\mathbb{N}} s_n \cdot \overline{V_{\varphi_2} g_n(z)} \cdot V_{\varphi_1} h_n(z).
\end{aligned}$$

□

Note that the assumptions of the lemma are in particular satisfied for operators belonging to some Schatten p-class, $1 \le p < \infty$, since for these we have the singular value decomposition

$$T = \sum_{n\in\mathbb{N}} s_n \cdot \langle \bullet, g_n\rangle\, h_n$$

with $(g_n)_{n\in\mathbb{N}}$, $(h_n)_{n\in\mathbb{N}}$ orthonormal systems and $(s_n)_{n\in\mathbb{N}}$ the sequence of singular values, $s_n \ge 0$ for all $n \in \mathbb{N}$ and $(s_n) \in \ell^p$ for $T \in \mathcal{S}^p(L^2(\mathbb{R}^d))$.

Theorem 3.2.3. *The Berezin transform* $\mathcal{B}$ *defines a bounded linear operator from* $B(L^2(\mathbb{R}^d))$ *into the space* $C_b(\mathbb{R}^{2d}) \subseteq L^\infty(\mathbb{R}^{2d})$ *of all bounded continuous functions on* $\mathbb{R}^{2d}$*, with norm estimate*

$$\|\mathcal{B}\|_{B(L^2)\to L^\infty} \le \|\varphi_1\| \cdot \|\varphi_2\|.$$

Proof. The function $\mathcal{B}T(z)$ is bounded, since

$$\begin{aligned}
|\mathcal{B}T(z)| &= |\,\langle T\pi(z)\varphi_2, \pi(z)\varphi_1\rangle\,| \\
&\le \|T\|_{B(L^2)} \cdot \|\pi(z)\varphi_2\| \cdot \|\pi(z)\varphi_1\| \\
&= \|T\|_{B(L^2)} \cdot \|\varphi_2\| \cdot \|\varphi_1\|
\end{aligned}$$

for all $z \in \mathbb{R}^{2d}$, hence

$$\|\mathcal{B}T\|_{L^\infty} \le \|T\|_{B(L^2)} \cdot \|\varphi_1\| \cdot \|\varphi_2\|$$

and thus

$$\|\mathcal{B}\|_{B(L^2)\to L^\infty} \leq \|\varphi_1\| \cdot \|\varphi_2\|.$$

It is continuous since for arbitrary $\varphi \in L^2(\mathbb{R}^d)$ the mapping $z \mapsto \pi(z)\varphi$ is continuous from $\mathbb{R}^{2d}$ to $L^2(\mathbb{R}^d)$. □

Now consider the Berezin transform restricted to the set $\mathcal{S}^1(L^2(\mathbb{R}^d))$ of trace class operators.

Theorem 3.2.4. *The Berezin transform $\mathcal{B}$ is a bounded linear operator from $\mathcal{S}^1(L^2(\mathbb{R}^d))$ to $L^1(\mathbb{R}^{2d})$ with operator norm*

$$\|\mathcal{B}\|_{\mathcal{S}^1\to L^1} \leq \|\varphi_1\| \cdot \|\varphi_2\|.$$

Proof. Let $T \in \mathcal{S}^1(L^2(\mathbb{R}^d))$ be a trace class operator. Then T has a spectral representation of the form

$$Tf = \sum_k s_k \langle f, g_k\rangle h_k, \qquad f \in L^2(\mathbb{R}^d),$$

with $(g_k)_{k\in\mathbb{N}}, (h_k)_{k\in\mathbb{N}}$ some orthonormal systems, and $(s_k)_{k\in\mathbb{N}}$ the sequence of singular values of T, $s_k \geq 0$, $\sum_{k\in\mathbb{N}} s_k = \|T\|_{\mathcal{S}^1} < \infty$. The series converges in the norm of $L^2(\mathbb{R}^d)$. Using Lemma 3.2.2, we calculate

$$\begin{aligned}
\iint_{\mathbb{R}^{2d}} |\mathcal{B}T(z)|\, dz &= \iint_{\mathbb{R}^{2d}} |\sum_k s_k V_{\varphi_1} h_k(z) \overline{V_{\varphi_2} g_k(z)}|\, dz \\
&\leq \iint_{\mathbb{R}^{2d}} \sum_k s_k\, |V_{\varphi_1} h_k(z) \overline{V_{\varphi_2} g_k(z)}|\, dz \\
&\underset{\text{Fub.}}{=} \sum_k s_k \iint_{\mathbb{R}^{2d}} |V_{\varphi_1} h_k(z)| \cdot |V_{\varphi_2} g_k(z)|\, dz \\
&\underset{\text{Cauchy-Schw.}}{\leq} \sum_k s_k\, \|V_{\varphi_1} h_k\|_{L^2(\mathbb{R}^{2d})} \cdot \|V_{\varphi_2} g_k\|_{L^2(\mathbb{R}^{2d})} \\
&= \|\varphi_1\| \cdot \|\varphi_2\| \cdot \sum_k s_k \\
&= \|\varphi_1\| \cdot \|\varphi_2\| \cdot \|T\|_{\mathcal{S}^1} < \infty,
\end{aligned}$$

where Fubini's Theorem and the Cauchy-Schwarz Inequality were used at the indicated places. Hence $\mathcal{B}T \in L^1(\mathbb{R}^{2d})$, and $\mathcal{B} : \mathcal{S}^1(L^2(\mathbb{R}^d)) \to L^1(\mathbb{R}^{2d})$ is bounded with the stated norm estimate. □

Once again, we use the preceding two Theorems 3.2.3 and 3.2.4 as the endpoints of a complex interpolation with $L^p = [L^1, L^\infty]_\theta$ and $\mathcal{S}^p = [\mathcal{S}^1, \mathcal{S}^\infty]_\theta = [\mathcal{S}^1, B(L^2)]_\theta$ and find the following theorem:

Theorem 3.2.5. *Let $1 \le p \le \infty$.*
The Berezin transform defines a bounded linear operator

$$\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$$

with operator norm

$$\|\mathcal{B}\|_{\mathcal{S}^p \to L^p} \le \|\varphi_1\| \cdot \|\varphi_2\|.$$

Proof. Complex interpolation yields the result and also the norm:

$$\begin{aligned}\|\mathcal{B}\|_{\mathcal{S}^p \to L^p} &\le \|\mathcal{B}\|^{\theta}_{\mathcal{S}^1 \to L^1} \cdot \|\mathcal{B}\|^{1-\theta}_{\mathcal{S}^\infty \to L^\infty} \\ &\le (\|\varphi_1\| \cdot \|\varphi_2\|)^{\theta} \cdot (\|\varphi_1\| \cdot \|\varphi_2\|)^{1-\theta} \\ &= \|\varphi_1\| \cdot \|\varphi_2\|.\end{aligned}$$

□

If we assume $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$ for the windows, we can considerably strengthen Theorem 3.2.4.

Theorem 3.2.6. *Let $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$. Let $T \in \mathcal{S}^1(L^2(\mathbb{R}^d))$. Then $\mathcal{B}T \in M^1(\mathbb{R}^{2d})$, and the operator $\mathcal{B} : \mathcal{S}^1(L^2(\mathbb{R}^d)) \to M^1(\mathbb{R}^{2d})$ is bounded with norm estimate*

$$\|\mathcal{B}T\|_{M^1} \le C \cdot \|\varphi_1\|_{M^1} \cdot \|\varphi_2\|_{M^1} \cdot \|T\|_{\mathcal{S}^1}$$

for all $T \in \mathcal{S}^1(L^2(\mathbb{R}^d))$.

Proof. Assume $T \in \mathcal{S}^1(\mathbb{R}^d)$.
By Lemma 3.2.2 we have

$$\begin{aligned}\|\mathcal{B}T(z)\|_{M^1} &= \|\sum_{n\in\mathbb{N}} s_n \cdot V_{\varphi_1} h_k(z) \overline{V_{\varphi_2} g_k(z)}\|_{M^1} \\ &\le \sum_{n\in\mathbb{N}} s_n \cdot \|V_{\varphi_1} h_k(z) \overline{V_{\varphi_2} g_k(z)}\|_{M^1},\end{aligned}$$

where $(s_n)_{n\in\mathbb{N}} \in \ell^1$, $s_n \geq 0$ for all $n \in \mathbb{N}$ and $\|T\|_{\mathcal{S}^1} = \sum_{n\in\mathbb{N}} s_n$.
Observe that $L^2(\mathbb{R}^d) = M^2(\mathbb{R}^d)$ with equivalent norms. By Theorem 1.7.3, both the functions $V_{\varphi_1}h$ and $V_{\varphi_2}g_k$ are in the Wiener amalgam space $W(\mathcal{F}L^1, L^2)(\mathbb{R}^{2d})$. Thus their product $V_{\varphi_1}h_k\overline{V_{\varphi_2}g_k}$ is in the Wiener amalgam space $W(\mathcal{F}L^1, L^1)(\mathbb{R}^{2d}) = M^1(\mathbb{R}^{2d})$, by the version of Hölder's Inequality for amalgam spaces, with norm

$$\begin{aligned}
\|V_{\varphi_1}h_k\overline{V_{\varphi_2}g_k}\|_{M^1} &\leq C \cdot \|V_{\varphi_1}h_k\overline{V_{\varphi_2}g_k}\|_{W(\mathcal{F}L^1,L^1)} \\
&\leq C \cdot \|V_{\varphi_1}h_k\|_{W(\mathcal{F}L^1,L^2)} \cdot \|V_{\varphi_2}g_k\|_{W(\mathcal{F}L^1,L^2)} \\
&\leq C \cdot \|h_k\|_{M^2} \cdot \|\varphi_1\|_{M^1} \cdot \|g_k\|_{M^2} \cdot \|\varphi_2\|_{M^1} \\
&\leq C \cdot \underbrace{\|h_k\|_{L^2}}_{=1} \cdot \|\varphi_1\|_{M^1} \cdot \underbrace{\|g_k\|_{L^2}}_{=1} \cdot \|\varphi_2\|_{M^1} \\
&\leq C \cdot \|\varphi_1\|_{M^1} \cdot \|\varphi_2\|_{M^1}
\end{aligned}$$

with a suitable generic constant $C > 0$. So

$$\begin{aligned}
\|\mathcal{B}T(z)\|_{M^1} &\leq \sum_{n\in\mathbb{N}} s_n \cdot \|V_{\varphi_1}h_k(z)\overline{V_{\varphi_2}g_k(z)}\|_{M^1} \\
&\leq C \cdot \|\varphi_1\|_{M^1} \cdot \|\varphi_2\|_{M^1} \cdot \sum_{n\in\mathbb{N}} s_n \\
&= C \cdot \|\varphi_1\|_{M^1} \cdot \|\varphi_2\|_{M^1} \cdot \|T\|_{\mathcal{S}^1} \\
&< \infty.
\end{aligned}$$

So $\mathcal{B}T \in M^1(\mathbb{R}^{2d})$ and the theorem is proved. □

Some more mapping properties of the Berezin transform with windows in $M^1(\mathbb{R}^d)$ will be shown in the next section, where additional tools will be at our disposal.

The relevance of the Berezin transform comes from the following observation.

Theorem 3.2.7. *Let $T \in \mathcal{S}^p(L^2(\mathbb{R}^d))$, $1 \leq p < \infty$, and $a \in L^q(\mathbb{R}^{2d})$, $1 < q \leq \infty$, where p and q denote conjugate exponents $\frac{1}{p} + \frac{1}{q} = 1$.*
Then

$$\langle \mathcal{A}a, T\rangle = \langle a, \mathcal{B}T\rangle\,.$$

(Here the left brackets denote the $\mathcal{S}^q$-$\mathcal{S}^p$ duality, whereas the right brackets denote the L^q-L^p duality.)

Proof. Consider the case $T \in \mathcal{S}^1(L^2(\mathbb{R}^d))$ and $a \in L^\infty(\mathbb{R}^{2d})$ first.
We have

$$\begin{aligned}
\langle \mathcal{A}a, T\rangle &= \langle A_a^{\varphi_1,\varphi_2}, T\rangle \\
&= \operatorname{tr}(T^* \circ A_a^{\varphi_1,\varphi_2}), \qquad \text{where tr denotes the trace,} \\
&= \sum_{n\in\mathbb{N}} \langle (T^* \circ A_a^{\varphi_1,\varphi_2}) e_n, e_n\rangle \\
&= \sum_{n\in\mathbb{N}} \langle A_a^{\varphi_1,\varphi_2} e_n, T e_n\rangle
\end{aligned}$$

for an arbitrary orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of $L^2(\mathbb{R}^d)$ (all such bases yield the same sum).
Now look at the singular value decomposition of T,

$$T = \sum_k s_k \cdot \langle \bullet, g_k\rangle\, h_k,$$

with $(s_k) \in \ell^1$. The orthonormal system $(g_k)_k$ can be completed to an orthonormal basis $(g_n)_{n\in\mathbb{N}}$ of $L^2(\mathbb{R}^d)$. We obviously have $Tg_k = s_k h_k$ for g_k a member of the original collection $(g_k)_k$, whereas $Tg_n = 0$ for all g_n that were joined to the original system $(g_k)_k$ to form a complete basis. Thus, choosing the orthonormal basis $(g_n)_{n\in\mathbb{N}}$ for $(e_n)_{n\in\mathbb{N}}$ above, we get

$$\begin{aligned}
\sum_{n\in\mathbb{N}} \langle A_a^{\varphi_1,\varphi_2} e_n, T e_n\rangle &= \sum_{n\in\mathbb{N}} \langle A_a^{\varphi_1,\varphi_2} g_n, T g_n\rangle \\
&= \sum_k \langle A_a^{\varphi_1,\varphi_2} g_k, T g_k\rangle \\
&= \sum_k \langle A_a^{\varphi_1,\varphi_2} g_k, s_k h_k\rangle \\
&= \sum_k s_k \langle A_a^{\varphi_1,\varphi_2} g_k, h_k\rangle\,.
\end{aligned}$$

The inner product can be written as

$$\begin{aligned}
\langle A_a^{\varphi_1,\varphi_2} g_k, h_k\rangle &= \left\langle V_{\varphi_1}^*(a \cdot V_{\varphi_2} g_k), h_k\right\rangle \\
&= \langle a \cdot V_{\varphi_2} g_k, V_{\varphi_1} h_k\rangle \\
&= \iint_{\mathbb{R}^{2d}} a(z) \cdot V_{\varphi_2} g_k(z) \cdot \overline{V_{\varphi_1} h_k(z)}\, dz.
\end{aligned}$$

Hence we find

$$\langle \mathcal{A}a, T\rangle = \sum_k s_k \cdot \iint_{\mathbb{R}^{2d}} a(z) \cdot V_{\varphi_2} g_k(z) \cdot \overline{V_{\varphi_1} h_k(z)}\, dz.$$

To justify a change of order of summation and integration, we have to check the assumptions of Fubini's Theorem:

$$\begin{aligned}
\sum_k \iint_{\mathbb{R}^{2d}} & s_k \cdot |a(z)| \cdot |V_{\varphi_2} g_k(z)| \cdot |V_{\varphi_1} h_k(z)|\, dz \\
&\leq \sum_k s_k \cdot \|a\|_{L^\infty} \iint_{\mathbb{R}^{2d}} |V_{\varphi_2} g_k(z)| \cdot |V_{\varphi_1} h_k(z)|\, dz \\
&\underset{\text{Cauchy-Schw.}}{\leq} \sum_k s_k \cdot \|a\|_{L^\infty} \cdot \|V_{\varphi_2} g_k\| \cdot \|V_{\varphi_1} h_k\| \\
&= \|a\|_{L^\infty} \cdot \|\varphi_2\| \cdot \underbrace{\|g_k\|}_{=1} \cdot \|\varphi_1\| \cdot \underbrace{\|h_k\|}_{=1} \cdot \sum_k s_k \\
&= \|a\|_{L^\infty} \cdot \|\varphi_2\| \cdot \|\varphi_1\| \cdot \|T\|_{\mathcal{S}^1} \\
&< \infty.
\end{aligned}$$

Thus Fubini's Theorem is applicable and yields

$$\begin{aligned}
\sum_k s_k \cdot \iint_{\mathbb{R}^{2d}} & a(z) \cdot V_{\varphi_2} g_k(z) \cdot \overline{V_{\varphi_1} h_k(z)}\, dz \\
&= \iint_{\mathbb{R}^{2d}} a(z) \cdot \sum_k s_k \cdot V_{\varphi_2} g_k(z) \cdot \overline{V_{\varphi_1} h_k(z)}\, dz \\
&= \iint_{\mathbb{R}^{2d}} a(z) \cdot \overline{\left(\sum_k s_k \cdot V_{\varphi_1} h_k(z) \cdot \overline{V_{\varphi_2} g_k(z)}\right)}\, dz \\
&= \iint_{\mathbb{R}^{2d}} a(z) \cdot \overline{\mathcal{B}T(z)}\, dz \\
&= \langle a, \mathcal{B}T\rangle
\end{aligned}$$

in the sense of L^∞-L^1 duality, by Lemma 3.2.2.
Now consider the case $T \in \mathcal{S}^p(L^2(\mathbb{R}^d))$, $a \in L^q(\mathbb{R}^{2d})$, with $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$.
We cannot immediately repeat the above computation, since now Fubini's

Theorem is no longer allowed at the appropriate spot. However, if we restrict at first to operators $T = \sum_{k=1}^{N} s_k \cdot \langle \bullet, g_k \rangle \, h_k$ of finite rank and functions $a \in \mathcal{S}(\mathbb{R}^{2d}) \subseteq L^q(\mathbb{R}^{2d})$, the computation can be performed completely analogously and yields

$$\begin{aligned}
\langle \mathcal{A}a, T \rangle &= \sum_{k=1}^{N} s_k \cdot \langle A_a^{\varphi_1,\varphi_2} g_k, h_k \rangle \\
&= \sum_{k=1}^{N} s_k \cdot \iint_{\mathbb{R}^{2d}} a(z) \cdot V_{\varphi_2} g_k(z) \cdot \overline{V_{\varphi_1} h_k(z)} \, dz \\
&\underset{\text{Fub.}}{=} \iint_{\mathbb{R}^{2d}} a(z) \cdot \sum_{k=1}^{N} s_k \cdot V_{\varphi_2} g_k(z) \cdot \overline{V_{\varphi_1} h_k(z)} \, dz \\
&= \langle a, \mathcal{B}T \rangle \, .
\end{aligned}$$

Now use the fact that $\mathcal{S}(\mathbb{R}^{2d})$ is dense in $L^q(\mathbb{R}^{2d})$ for $1 < q < \infty$. So for every $a \in L^q(\mathbb{R}^{2d})$ there is a sequence $(a_n)_{n\in\mathbb{N}}$ in $\mathcal{S}(\mathbb{R}^{2d})$ that converges to a in the L^q-norm:

$$\lim_{n\to\infty} \|a_n - a\|_{L^q} = 0.$$

Similarly, every Schatten p-class operator T can be approximated in the $\mathcal{S}^p$-norm by a sequence $(T_m)_{m\in\mathbb{N}}$ of finite rank operators:

$$\lim_{m\to\infty} \|T_m - T\|_{\mathcal{S}^p} = 0.$$

So

$$\begin{aligned}
\langle \mathcal{A}a, T \rangle &= \left\langle \mathcal{A}(\lim_n a_n), \lim_m T_m \right\rangle \\
&= \left\langle \lim_n (\mathcal{A}a_n), \lim_m T_m \right\rangle \\
&= \lim_{n,m} \langle \mathcal{A}a_n, T_m \rangle \\
&= \lim_{n,m} \langle a_n, \mathcal{B}T_m \rangle \\
&= \left\langle \lim_n a_n, \lim_m (\mathcal{B}T_m) \right\rangle \\
&= \langle a, \mathcal{B}T \rangle \, ;
\end{aligned}$$

here we used the boundedness of the operators $\mathcal{A} : L^p(\mathbb{R}^{2d}) \to \mathcal{S}^q(L^2(\mathbb{R}^d))$ and $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$ and the joint continuity of the dual pairings $\langle \bullet, \bullet \rangle$ of $\mathcal{S}^q$ and $\mathcal{S}^p$ resp. L^q and L^p. □

The following is then a simple corollary of Theorem 3.2.7.

Theorem 3.2.8. *Let $1 \leq p < \infty$ and q be the conjugate exponent with $1 < q \leq \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$.*
The operator $\mathcal{A} : L^q(\mathbb{R}^{2d}) \to \mathcal{S}^q(L^2(\mathbb{R}^d))$ is the Banach space adjoint of the operator $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$, i.e. $\mathcal{B}^* = \mathcal{A}$.

Proof. Note the dualities $L^q(\mathbb{R}^{2d}) = (L^p(\mathbb{R}^{2d}))^*$ and $\mathcal{S}^q(L^2(\mathbb{R}^d)) = (\mathcal{S}^p(L^2(\mathbb{R}^d)))^*$ for $1 \leq p < \infty$ (but not for $p = \infty$, $q = 1$!). If $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$, then $\mathcal{B}^* : \mathcal{S}^p(L^2(\mathbb{R}^d)))^* \to (L^p(\mathbb{R}^{2d}))^*$, so $\mathcal{B}^* : L^q(\mathbb{R}^{2d}) \to \mathcal{S}^q(L^2(\mathbb{R}^d))$. The adjoint of $\mathcal{B}$ is the uniquely defined operator $\mathcal{B}^*$ such that

$$\langle \mathcal{B}T, a \rangle = \langle T, \mathcal{B}^* a \rangle$$

for all $T \in \mathcal{S}^p(L^2(\mathbb{R}^d))$ and all $a \in L^q(\mathbb{R}^{2d})$. But by Theorem 3.2.7,

$$\langle \mathcal{B}T, a \rangle = \langle T, \mathcal{A} a \rangle$$

for all $T \in \mathcal{S}^p(L^2(\mathbb{R}^d))$ and all $a \in L^q(\mathbb{R}^{2d})$. Hence

$$\mathcal{B}^* = \mathcal{A}.$$

□

If $1 < p < \infty$, then $L^p(\mathbb{R}^{2d})$ and $\mathcal{S}^p(L^2(\mathbb{R}^d))$ are reflexive spaces, thus in this cases we also have

Theorem 3.2.9. *Let $1 < p < \infty$ and q be the conjugate exponent with $1 < q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$.*
The operator $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$ is the Banach space adjoint of the operator $\mathcal{A} : L^q(\mathbb{R}^{2d}) \to \mathcal{S}^q(L^2(\mathbb{R}^d))$, i.e. $\mathcal{A}^* = \mathcal{B}$.

Proof. The statement is clear by the remark preceding the theorem about the reflexivity of the involved spaces. □

3.3 Auxiliary Results

In this section we present (partly without proofs) some results that will be needed in the sequel.

First, we will need two other versions of the Kernel Theorem 2.2.5. The first one concerns the particular case of Hilbert-Schmidt operators.

Theorem 3.3.1 (Kernel Theorem for Hilbert-Schmidt Operators)**.** *Let $T \in \mathcal{S}^2(L^2(\mathbb{R}^d))$. Then there exists a unique kernel function $\sigma \in L^2(\mathbb{R}^{2d})$ such that*

$$\langle Tf, g\rangle = \langle \sigma, W(g,f)\rangle$$

for all $f, g \in L^2(\mathbb{R}^d)$. ($W(g,f)$ denotes the Wigner distribution of g and f.)

Proof. We first prove existence.
If $T \in \mathcal{S}^2(L^2(\mathbb{R}^d))$, then there exist orthonormal families $(g_n)_{n\in\mathbb{N}}$ and $(h_n)_{n\in\mathbb{N}}$ in $L^2(\mathbb{R}^d)$ and a sequence $(s_n)_{n\in\mathbb{N}}$ in ℓ^2 such that

$$T = \sum_n s_n \langle \bullet, g_n\rangle\, h_n.$$

This implies that for all $f, g \in L^2(\mathbb{R}^d)$

$$\begin{aligned}\langle Tf, g\rangle &= \left\langle \sum_n s_n \langle f, g_n\rangle\, h_n, g \right\rangle \\ &= \sum_n s_n \langle f, g_n\rangle \langle h_n, g\rangle \\ &= \sum_n s_n \left\langle h_n \otimes \overline{g_n}, g \otimes \overline{f}\right\rangle\end{aligned}$$

with the last inner product in $L^2(\mathbb{R}^{2d})$. The family $(h_n \otimes \overline{g_n})_{n\in\mathbb{N}}$ is orthogonal in $L^2(\mathbb{R}^{2d})$, thus the series

$$\sum_n s_n (h_n \otimes \overline{g_n})$$

converges in $L^2(\mathbb{R}^{2d})$ to some function $k \in L^2(\mathbb{R}^{2d})$, because $(s_n) \in \ell^2$. We conclude

$$\begin{aligned}\langle Tf, g\rangle &= \left\langle k, g \otimes \overline{f}\right\rangle \\ &= \left\langle \mathcal{F}_2\mathcal{T}_A k, \mathcal{F}_2\mathcal{T}_A(g \otimes \overline{f})\right\rangle \\ &= \langle \mathcal{F}_2\mathcal{T}_A k, W(g, f)\rangle\end{aligned}$$

for $A = \begin{pmatrix} I & \frac{1}{2}I \\ I & -\frac{1}{2}I \end{pmatrix}$. Denote $\mathcal{F}_2\mathcal{T}_A k$ by σ, then obviously $\sigma \in L^2(\mathbb{R}^{2d})$, since $\mathcal{F}_2$ and $\mathcal{T}_A$ are unitary operators from $L^2(\mathbb{R}^{2d})$ onto $L^2(\mathbb{R}^{2d})$, and

$$\langle Tf, g\rangle = \langle \sigma, W(g, f)\rangle$$

for all $f, g \in L^2(\mathbb{R}^d)$.
Now assume that $\sigma, \tau \in L^2(\mathbb{R}^{2d})$ such that

$$\langle \sigma, W(g, f)\rangle = \langle \tau, W(g, f)\rangle$$

for all $f, g \in L^2(\mathbb{R}^d)$. Then

$$\langle \sigma - \tau, W(g, f)\rangle = 0$$

for all $f, g \in L^2(\mathbb{R}^d)$. By Proposition 1.2.2, the set $\{W(g, f)|\ f, g \in L^2(\mathbb{R}^d)\}$ is a complete subset of $L^2(\mathbb{R}^{2d})$, hence

$$\sigma - \tau = 0 \text{ in } L^2(\mathbb{R}^{2d}).$$

This shows uniqueness. □

The second one is a kernel theorem for modulation spaces. The proof can be found for example in [17], Theorem 14.4.1, or in [13].

Theorem 3.3.2 (Kernel Theorem for Modulation Spaces)**.** *Let $T : M^1(\mathbb{R}^d) \to M^\infty(\mathbb{R}^d)$ be a bounded linear operator. Then there exists a unique $\sigma \in M^\infty(\mathbb{R}^{2d})$ such that*

$$\langle Tf, g\rangle = \langle \sigma, W(g, f)\rangle$$

for all $f, g \in M^1(\mathbb{R}^d)$ (with $W(g, f)$ the Wigner distribution of g and f).
□

Using this theorem, we are now able to complement Theorem 3.2.6 and prove some more mapping properties of the Berezin transform with windows $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$.

Theorem 3.3.3. *Let $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$. Let $T \in B(L^2(\mathbb{R}^d))$. Then $\mathcal{B}T \in M^{\infty,1}(\mathbb{R}^{2d})$ and*

$$\|\mathcal{B}T\|_{M^{\infty,1}} \leq C \cdot \|\varphi_1\|_{M^1} \cdot \|\varphi_2\|_{M^1} \cdot \|T\|_{B(L^2)}.$$

Proof. Since $M^1(\mathbb{R}^d) \hookrightarrow M^2(\mathbb{R}^d) = L^2(\mathbb{R}^d) \hookrightarrow M^\infty(\mathbb{R}^d)$ with continuous embeddings, every $T \in B(L^2(\mathbb{R}^d))$ can be considered as a bounded operator $T : M^1(\mathbb{R}^d) \to M^\infty(\mathbb{R}^d)$. By the Kernel Theorem 3.3.2, there is a unique $\sigma \in M^\infty(\mathbb{R}^{2d})$ such that

$$\langle Tf, g\rangle = \langle \sigma, W(g, f)\rangle$$

for all $f, g \in M^1(\mathbb{R}^d)$. In particular,

$$\begin{aligned}\mathcal{B}T(z) &= \langle T\pi(z)\varphi_2, \pi(z)\varphi_1\rangle \\ &= \langle \sigma, W(\pi(z)\varphi_1, \pi(z)\varphi_2)\rangle\end{aligned}$$

for all $z \in \mathbb{R}^{2d}$. Now a short computation shows

$$W(\pi(z)\varphi_1, \pi(z)\varphi_2) = T_z W(\varphi_1, \varphi_2),$$

hence

$$\begin{aligned}\langle \sigma, W(\pi(z)\varphi_1, \pi(z)\varphi_2)\rangle &= \langle \sigma, T_z W(\varphi_1, \varphi_2)\rangle \\ &= \iint_{\mathbb{R}^{2d}} \sigma(w) \cdot \overline{W(\varphi_1, \varphi_2)(w - z)}\, dw \\ &= \left(\sigma * \widetilde{W(\varphi_1, \varphi_2)}\right)(z)\end{aligned}$$

with

$$\widetilde{W(\varphi_1, \varphi_2)}(w) = \overline{W(\varphi_1, \varphi_2)(-w)}$$

the usual Fourier involution.
By Proposition 1.7.2, $W(\varphi_1, \varphi_2) \in M^1(\mathbb{R}^{2d})$, if $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$, thus Lemma 3.1.11 yields

$$\mathcal{B}T \in M^\infty(\mathbb{R}^{2d}) * M^1(\mathbb{R}^{2d}) \subseteq M^{\infty,1}(\mathbb{R}^{2d})$$

with the stated norm estimate. □

As usual, complex interpolation for

$$\mathcal{S}^p(L^2(\mathbb{R}^d)) = [\mathcal{S}^1(L^2(\mathbb{R}^d)), B(L^2(\mathbb{R}^d))]_\theta = [\mathcal{S}^1(L^2(\mathbb{R}^d)), \mathcal{S}^\infty(L^2(\mathbb{R}^d))]_\theta$$

and

$$M^{p,1}(\mathbb{R}^{2d}) = [M^1(\mathbb{R}^{2d}), M^{\infty,1}(\mathbb{R}^{2d})]_\theta = [M^{1,1}(\mathbb{R}^{2d}), M^{\infty,1}(\mathbb{R}^{2d})]_\theta$$

allows us to conclude the following

Corollary 3.3.4. *Let $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$ and $1 \leq p \leq \infty$. Then $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to M^{p,1}(\mathbb{R}^{2d})$ is a bounded operator with*

$$\|\mathcal{B}T\|_{M^{p,1}} \leq C \cdot \|\varphi_1\|_{M^1} \cdot \|\varphi_2\|_{M^1} \cdot \|T\|_{\mathcal{S}^p}$$

for all $T \in \mathcal{S}^p(L^2(\mathbb{R}^d))$, with some fixed constant $C > 0$.

Proof. The endpoint results for the complex interpolation are given in Theorem 3.2.6 and Theorem 3.3.3. □

Next, we show a version of the famous theorem of Tauber. Whereas the original result (with rather difficult proof) is concerned with functions in L^1, we content ourselves with proving the statement for the much more trivial case of L^2 functions.

Theorem 3.3.5 (Tauberian Theorem for $L^2(\mathbb{R}^d)$)**.** *Let $f \in L^2(\mathbb{R}^d)$. Then $\overline{span\{T_z f|\ z \in \mathbb{R}^d\}} = L^2(\mathbb{R}^d)$ if and only if $\hat{f}(\omega) \neq 0$ for almost all $\omega \in \mathbb{R}^d$.*

Proof. The span of $\{T_z f|\ z \in \mathbb{R}^d\}$ is dense in $L^2(\mathbb{R}^d)$ if and only if

$$\langle g, T_z f\rangle = 0 \text{ for all } z \in \mathbb{R}^d \qquad \text{implies} \qquad g = 0.$$

Now

$$\begin{aligned}\langle g, T_z f\rangle &= \int_{\mathbb{R}^d} g(t) \cdot \overline{f(t-z)}\, dt \\ &= \int_{\mathbb{R}^d} g(t) \cdot \overline{f(-(z-t))}\, dt \\ &= \int_{\mathbb{R}^d} g(t) \cdot \tilde{f}(z-t)\, dt \\ &= (g * \tilde{f})(z)\end{aligned}$$

with $\tilde{f}(t) = \overline{f(-t)}$ the Fourier involution.
Hence

$$\langle g, T_z f\rangle = 0 \text{ for all } z \in \mathbb{R}^d$$

is equivalent to

$$(g * \tilde{f})(z) = 0 \text{ for all } z \in \mathbb{R}^d,$$

which in turn is equivalent to

$$\widehat{g}(\omega) \cdot \widehat{\tilde{f}}(\omega) = 0 \text{ for almost all } \omega \in \mathbb{R}^d$$

by the Fourier convolution theorem.
Thus the span of $\{T_z f |\ z \in \mathbb{R}^d\}$ is dense in $L^2(\mathbb{R}^d)$ if and only if

$$\widehat{g}(\omega) \cdot \widehat{\tilde{f}}(\omega) = 0 \text{ for almost all } \omega \in \mathbb{R}^d \qquad \text{implies} \qquad g = 0.$$

But this is clearly satisfied if and only if

$$\widehat{f}(\omega) \neq 0 \text{ for almost all } \omega \in \mathbb{R}^d :$$

If $\widehat{f}(\omega) \neq 0$ almost everywhere, then $\widehat{g}(\omega) \cdot \widehat{\tilde{f}}(\omega) = 0$ almost everywhere implies $\widehat{g}(\omega) = 0$ almost everywhere, hence $\widehat{g} = 0$ in $L^2(\mathbb{R}^d)$ and so $g = 0$ in $L^2(\mathbb{R}^d)$.
If, on the other hand, $\widehat{f}(\omega) = 0$ on a set of positive measure, then in particular $\widehat{f}(\omega) = 0$ on a set A of positive finite measure, say $0 < |A| < 1$. If we choose g such that $\widehat{g} = \chi_A \in L^2(\mathbb{R}^d)$, the characteristic function of A, then $\widehat{g}(\omega) \cdot \widehat{\tilde{f}}(\omega) = \widehat{g}(\omega) \cdot \overline{\widehat{f}}(\omega) = 0$ for almost every $\omega \in \mathbb{R}^d$, but $\widehat{g} \neq 0$ and so $g \neq 0$ in $L^2(\mathbb{R}^d)$. □

For reference, we state the classic Tauberian Theorem for L^1 without proof. Observe the slightly stronger condition $\hat{f}(\omega) \neq 0$ for *all* (instead of *almost all*) ω in this case.

Theorem 3.3.6 (Classic Tauberian Theorem for $L^1(\mathbb{R}^d)$)**.** *Let $f \in L^1(\mathbb{R}^d)$. Then $\overline{span\{T_z f |\ z \in \mathbb{R}^d\}} = L^1(\mathbb{R}^d)$ if and only if $\hat{f}(\omega) \neq 0$ for all $\omega \in \mathbb{R}^d$.*
□

Finally, we will need an extension of the preceding theorem to the modulation space $M^1(\mathbb{R}^d)$. It turns out that the classic Tauberian Theorem holds unchanged in the case $M^1(\mathbb{R}^d) \subseteq L^1(\mathbb{R}^d)$. The proof (which is once more not given) is based on very general statements about ideals in so-called Segal algebras that can be found in the monograph [33] by Reiter. (Note that the modulation space $M^1(\mathbb{R}^d)$ is a particular case of a Segal algebra in $L^1(\mathbb{R}^d)$.)

Theorem 3.3.7 (Tauberian Theorem for the Modulation Space $M^1(\mathbb{R}^d)$)**.** *Let $f \in M^1(\mathbb{R}^d)$. Then the subspace span$\{T_z f \mid z \in \mathbb{R}^d\} \subseteq M^1(\mathbb{R}^d)$ spanned by all the translates of f is dense in $M^1(\mathbb{R}^d)$ (with respect to the norm topology induced by $\|\cdot\|_{M^1}$) if and only if $\hat{f}(\omega) \neq 0$ for all $\omega \in \mathbb{R}^d$.*

□

3.4 Density Results

This section contains the main results of Chapter 3. We investigate the possibility of approximating a given bounded linear operator on $L^2(\mathbb{R}^d)$ by localization operators. In particular, we are interested in conditions that guarantee the density of the set of localization operators in different topologies.

In order to examine density properties of the set of localization operators with symbols from various classes, we employ some well known results from functional analysis, giving relations between properties of operators on Banach spaces and their adjoints on the respective dual spaces. Precisely, we use the following facts (cf. [5]):
Let X, Y be Banach spaces and $T : X \to Y$ be a bounded operator. Its (Banach space) adjoint operator be denoted by $T^* : Y^* \to X^*$.

- T^* is injective if and only if the range of T is dense in Y with respect to the norm topology on Y.
- T is injective if and only if the range of T^* is dense in X^* with respect to the weak* topology on X^*.

Since $\mathcal{A}$ and $\mathcal{B}$ are essentially adjoint to each other, the following question is of considerable interest:
Under what conditions is the Berezin transform $\mathcal{B}$ a one-to-one operator?

Theorem 3.4.1. *Let $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^d)$ and $1 \leq p \leq 2$. Then the Berezin transform $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$ is a one-to-one operator if and only if $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for almost all $(x, \omega) \in \mathbb{R}^{2d}$.*

Proof. If $T \in \mathcal{S}^p(L^2(\mathbb{R}^d))$ and $1 \leq p \leq 2$, then $T \in \mathcal{S}^2(L^2(\mathbb{R}^d))$ is a Hilbert-Schmidt operator.
The Berezin transform $\mathcal{B}$ is one-to-one if and only if

$$\mathcal{B}T = 0 \text{ in } L^p(\mathbb{R}^{2d}) \qquad \text{implies} \qquad T = 0 \text{ in } \mathcal{S}^p(L^2(\mathbb{R}^d)).$$

But $\mathcal{B}T = 0$ in L^p if and only if $\mathcal{B}T(z) = 0$ for almost every $z \in \mathbb{R}^{2d}$ if and only if $\mathcal{B}T(z) = 0$ for every $z \in \mathbb{R}^{2d}$, since $\mathcal{B}T$ is a continuous function by

Theorem 3.2.3. So, equivalently, the Berezin transform $\mathcal{B}$ is one-to-one if and only if

$$\mathcal{B}T(z) = 0 \text{ for all } z \in \mathbb{R}^{2d} \qquad \text{implies} \qquad T = 0 \text{ in } \mathcal{S}^p(L^2(\mathbb{R}^d)).$$

By the Kernel Theorem 3.3.1

$$\mathcal{B}T(z) = \langle T\pi(z)\varphi_2, \pi(z)\varphi_1\rangle = \langle \sigma, W(\pi(z)\varphi_1, \pi(z)\varphi_2\rangle$$

for a uniquely determined function $\sigma \in L^2(\mathbb{R}^{2d})$. The injectivity of $\mathcal{B}$ is therefore equivalent to the statement, that

$$\langle \sigma, W(\pi(z)\varphi_1, \pi(z)\varphi_2)\rangle = 0 \text{ for all } z \in \mathbb{R}^{2d} \qquad \text{implies} \qquad \sigma = 0 \text{ in } L^2(\mathbb{R}^{2d}).$$

Now for the Wigner distribution we easily verify by direct computation that

$$W(\pi(z)f, \pi(z)g)(x,\omega) = W(f,g)(x - z_1, \omega - z_2) = T_z W(f,g)(x,\omega)$$

for all $f, g \in L^2(\mathbb{R}^d)$, $z = (z_1, z_2) \in \mathbb{R}^{2d}$ and $(x,\omega) \in \mathbb{R}^{2d}$.
So

$$\langle \sigma, W(\pi(z)\varphi_1, \pi(z)\varphi_2\rangle = \langle \sigma, T_z W(\varphi_1, \varphi_2)\rangle = 0 \text{ for all } z \in \mathbb{R}^{2d} \quad \text{implies} \quad \sigma = 0$$

if and only if the subspace spanned by the translates of $W(\varphi_1, \varphi_2)$ is dense in $L^2(\mathbb{R}^d)$. But by Theorem 3.3.5 and Lemma 3.1.2 this is equivalent to

$$\widehat{W(\varphi_1, \varphi_2)}(x,\omega) = e^{-\pi i x\cdot\omega} V(\varphi_1, \varphi_2)(-\omega, x) \neq 0$$

for almost all $(x,\omega) \in \mathbb{R}^{2d}$, that means

$$V(\varphi_1, \varphi_2)(x,\omega) \neq 0$$

for almost all $(x,\omega) \in \mathbb{R}^{2d}$. □

For the remaining values of p, note that the condition for injectivity is slightly stronger.

Theorem 3.4.2. *Let $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$ and $2 < p \leq \infty$. Then the Berezin transform $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$ is a one-to-one operator if and only if $V(\varphi_1, \varphi_2)(x,\omega) \neq 0$ for all $(x,\omega) \in \mathbb{R}^{2d}$.*

Proof. The proof is very similar to the preceding one. Again, we must show that $\mathcal{B}T(z) = 0$ for all $z \in \mathbb{R}^{2d}$ implies $T = 0$ in $\mathcal{S}^p(L^2(\mathbb{R}^d))$ if and only if $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for all $(x, \omega) \in \mathbb{R}^{2d}$. Since $M^1(\mathbb{R}^d) \hookrightarrow M^2(\mathbb{R}^d) = L^2(\mathbb{R}^d) \hookrightarrow M^\infty(\mathbb{R}^d)$ with continuous embeddings, every $T \in \mathcal{S}^p(L^2(\mathbb{R}^d)) \subseteq B(L^2(\mathbb{R}^d))$ can be considered as a bounded operator $T : M^1(\mathbb{R}^d) \to M^\infty(\mathbb{R}^d)$. By the Kernel Theorem 3.3.2, there is a unique $\sigma \in M^\infty(\mathbb{R}^{2d})$ such that

$$\langle Tf, g\rangle = \langle \sigma, W(g, f)\rangle$$

for all $f, g \in M^1(\mathbb{R}^d)$. In particular,

$$\begin{aligned}
\mathcal{B}T(z) &= \langle T\pi(z)\varphi_2, \pi(z)\varphi_1\rangle \\
&= \langle \sigma, W(\pi(z)\varphi_1, \pi(z)\varphi_2)\rangle \\
&= \langle \sigma, T_z W(\varphi_1, \varphi_2)\rangle
\end{aligned}$$

for all $z \in \mathbb{R}^{2d}$. Hence the mapping $\mathcal{B}$ is one-to-one if and only if

$$\langle \sigma, T_z W(\varphi_1, \varphi_2)\rangle = 0 \text{ for all } z \in \mathbb{R}^{2d} \quad \text{implies} \quad \sigma = 0 \text{ in } M^\infty(\mathbb{R}^{2d})$$

(since this is equivalent to $T = 0$ in $\mathcal{S}^p(L^2(\mathbb{R}^d))$, in $B(L^2(\mathbb{R}^d))$ and as operator $M^1(\mathbb{R}^d) \to M^\infty(\mathbb{R}^d)$). But this is in turn equivalent to the subspace

$$\operatorname{span}\{T_z W(\varphi_1, \varphi_2) \mid z \in \mathbb{R}^{2d}\}$$

being dense in $M^1(\mathbb{R}^{2d})$, which, by the Tauberian Theorem 3.3.7, is equivalent to $\widehat{W(\varphi_1, \varphi_2)}(x, \omega) \neq 0$ and thus $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for all $(x, \omega) \in \mathbb{R}^{2d}$. □

The following **example** illustrates that the weaker condition $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ only for *almost all* $x, \omega \in \mathbb{R}^d$ is in general not sufficient for the Berezin transform to be one-to-one:
Assume that there are $x_0, \omega_0 \in \mathbb{R}^d$ such that $V(\varphi_1, \varphi_2)(x_0, \omega_0) = 0$. Then consider the bounded operator $T = \pi(x_0, \omega_0) = M_{\omega_0} T_{x_0} \in B(L^2(\mathbb{R}^d))$. We compute the Berezin transform (with $z = (z_1, z_2)$):

$$\begin{aligned}
\mathcal{B}T(z) &= \langle T\pi(z_1, z_2)\varphi_2, \pi(z_1, z_2)\varphi_1\rangle \\
&= \langle \pi(x_0, \omega_0)\pi(z_1, z_2)\varphi_2, \pi(z_1, z_2)\varphi_1\rangle \\
&= e^{2\pi i(\omega_0 \cdot z_1 - x_0 \cdot z_2)} \cdot \langle \pi(z_1, z_2)\pi(x_0, \omega_0)\varphi_2, \pi(z_1, z_2)\varphi_1\rangle \\
&= e^{2\pi i(\omega_0 \cdot z_1 - x_0 \cdot z_2)} \cdot \langle \pi(x_0, \omega_0)\varphi_2, \varphi_1\rangle \\
&= e^{2\pi i(\omega_0 \cdot z_1 - x_0 \cdot z_2)} \cdot \overline{V(\varphi_1, \varphi_2)(x_0, \omega_0)} \\
&\equiv 0
\end{aligned}$$

(where we have used the commutation relation Lemma A.2.3).
Thus $\mathcal{B}T = 0 \in L^\infty(\mathbb{R}^{2d})$, but $T \neq 0 \in B(L^2(\mathbb{R}^d))$, hence $\mathcal{B}$ is not one-to-one.

Now let us look at localization operators with symbols in $L^\infty(\mathbb{R}^{2d})$, that is $A_a^{\varphi_1,\varphi_2} = \mathcal{A}a$ for $\mathcal{A} : L^\infty(\mathbb{R}^{2d}) \to \mathcal{B}(L^2(\mathbb{R}^d))$.

We have the following negative result on density with respect to the norm topology:

Theorem 3.4.3. *The Fourier transform $\mathcal{F} \in \mathcal{B}(L^2(\mathbb{R}^d))$ is not contained in the norm-closure of the range of $\mathcal{A}$. In particular the set of all localization operators with symbols in $L^\infty(\mathbb{R}^{2d})$ is not dense in $\mathcal{B}(L^2(\mathbb{R}^d))$ with respect to the operator norm.*

Proof. We will show that the Fourier transform $\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ cannot be approximated by localization operators with symbols in $L^\infty(R^{2d})$, that means $\mathcal{F} \notin \overline{ran(\mathcal{A})}$.
For $f \in L^2(\mathbb{R}^d)$ we have

$$\begin{aligned}
\|\mathcal{F}f - A_a^{\varphi_1,\varphi_2} f\|^2 &= \|\hat{f}\|^2 + \|A_a^{\varphi_1,\varphi_2} f\|^2 - 2\,\mathrm{Re}\big(\langle \hat{f}, A_a^{\varphi_1,\varphi_2} f\rangle\big) \\
&\geq \|\hat{f}\|^2 + \|A_a^{\varphi_1,\varphi_2} f\|^2 - 2\,\big|\langle \hat{f}, A_a^{\varphi_1,\varphi_2} f\rangle\big| \\
&= \|\hat{f}\|^2 + \|A_a^{\varphi_1,\varphi_2} f\|^2 - 2\,\big|\langle \hat{f}, V_{\varphi_1}^*(aV_{\varphi_2} f)\rangle\big| \\
&= \|\hat{f}\|^2 + \|A_a^{\varphi_1,\varphi_2} f\|^2 - 2\,\big|\langle V_{\varphi_1}\hat{f}, aV_{\varphi_2} f\rangle\big|.
\end{aligned}$$

Now pick an arbitrary $g \in L^2(\mathbb{R}^d)$, $g \neq 0$, and let $f := M_{\omega_0}T_{x_0}g$. Observe that $\|f\| = \|g\|$ for any x_0, ω_0, since the modulation operator M_{ω_0} and translation operator T_{x_0} are unitary. We will show that

$$\lim_{x_0,\omega_0\to\infty} \big|\langle V_{\varphi_1}(\widehat{M_{\omega_0}T_{x_0}}g), aV_{\varphi_2}\,(M_{\omega_0}T_{x_0}g)\rangle\big| = \lim_{x_0,\omega_0\to\infty} \big|\langle V_{\varphi_1}\hat{f}, aV_{\varphi_2} f\rangle\big| = 0.$$

Let $\varepsilon > 0$ be given.
Choose $R > 0$ such that

$$\iint_{[-R,R]^{2d}} |V_{\varphi_1}\hat{g}(x,\omega)|^2\,dxd\omega \geq (1-\varepsilon)\|g\|^2\|\varphi_1\|^2$$

and

$$\iint_{[-R,R]^{2d}} |V_{\varphi_2}g(x,\omega)|^2\,dxd\omega \geq (1-\varepsilon)\|g\|^2\|\varphi_2\|^2.$$

(Note that it is always possible to find such an R in view of the fact that $\|V_\varphi g\|^2 = \iint_{\mathbb{R}^{2d}} |V_\varphi g(x,\omega)|^2\, dx d\omega = \|g\|^2 \|\varphi\|^2$ for arbitrary $g, \varphi \in L^2(\mathbb{R}^d)$.)
A time-frequency shift of g amounts to a translation of $V_\varphi g$ in the time-frequency plane (up to some phase factor of modulus 1), more precisely

$$|V_\varphi(M_{\omega_0}T_{x_0}g)(x,\omega)| = |V_\varphi g(x - x_0, \omega - \omega_0)|$$

for all $g, \varphi \in L^2(\mathbb{R}^d)$, $x, \omega, x_0, \omega_0 \in \mathbb{R}^d$. This yields

$$|V_{\varphi_2}(M_{\omega_0}T_{x_0}g)(x,\omega)| = |V_{\varphi_2} g(x - x_0, \omega - \omega_0)|$$

and

$$|V_{\varphi_1}(\widehat{M_{\omega_0}T_{x_0}g})(x,\omega)| = |V_{\varphi_1}(T_{\omega_0}M_{-x_0}\hat{g})(x,\omega)| = |V_{\varphi_1}\hat{g}(x - \omega_0, \omega + x_0)|.$$

If we define $U_\varepsilon := \begin{pmatrix}\omega_0 \\ -x_0\end{pmatrix} + [-R,R]^{2d}$ and $V_\varepsilon := \begin{pmatrix}x_0 \\ \omega_0\end{pmatrix} + [-R,R]^{2d}$ then by a change of variable we have

$$\begin{aligned}\iint_{U_\varepsilon} |V_{\varphi_1}(\widehat{M_{\omega_0}T_{x_0}g})(x,\omega)|^2\, dx d\omega &= \iint_{[-R,R]^{2d}} |V_{\varphi_1}\hat{g}(x,\omega)|^2\, dx d\omega \\ &\geq (1-\varepsilon)\|g\|^2\|\varphi_1\|^2 \\ &= (1-\varepsilon)\|M_{\omega_0}T_{x_0}g\|^2\|\varphi_1\|^2\end{aligned}$$

and

$$\begin{aligned}\iint_{V_\varepsilon} |V_{\varphi_2}(M_{\omega_0}T_{x_0}g)(x,\omega)|^2\, dx d\omega &= \iint_{[-R,R]^{2d}} |V_{\varphi_2} g(x,\omega)|^2\, dx d\omega \\ &\geq (1-\varepsilon)\|g\|^2\|\varphi_2\|^2 \\ &= (1-\varepsilon)\|M_{\omega_0}T_{x_0}g\|^2\|\varphi_2\|^2.\end{aligned}$$

By choosing x_0 and ω_0 large enough we can achieve $U_\varepsilon \cap V_\varepsilon = \emptyset$, so that $U_\varepsilon \subset \mathbb{R}^{2d} \setminus V_\varepsilon$ and $V_\varepsilon \subset \mathbb{R}^{2d} \setminus U_\varepsilon$. Then

$$\begin{aligned}\iint_{V_\varepsilon} |V_{\varphi_1}(\widehat{M_{\omega_0}T_{x_0}g})(x,\omega)|^2\, dx d\omega &\leq \iint_{\mathbb{R}^{2d}\setminus U_\varepsilon} |V_{\varphi_1}(\widehat{M_{\omega_0}T_{x_0}g})(x,\omega)|^2\, dx d\omega \\ &\leq \varepsilon\|M_{\omega_0}T_{x_0}g\|^2\|\varphi_1\|^2\end{aligned}$$

and

$$\begin{aligned}\iint_{U_\varepsilon} |V_{\varphi_2}(M_{\omega_0}T_{x_0}g)(x,\omega)|^2\, dx d\omega &\leq \iint_{\mathbb{R}^{2d}\setminus V_\varepsilon} |V_{\varphi_2}(M_{\omega_0}T_{x_0}g)(x,\omega)|^2\, dx d\omega \\ &\leq \varepsilon\|M_{\omega_0}T_{x_0}g\|^2\|\varphi_2\|^2.\end{aligned}$$

Writing $M_{\omega_0}T_{x_0}g =: f$ again, we conclude

$$\begin{aligned}
\big|\langle V_{\varphi_1}\hat{f}, aV_{\varphi_2}f\rangle\big| &= \Big|\iint_{\mathbb{R}^{2d}} V_{\varphi_1}\hat{f}(x,\omega)\overline{a(x,\omega)V_{\varphi_2}f(x,\omega)}\,dxd\omega\Big| \\
&\le \iint_{\mathbb{R}^{2d}} |V_{\varphi_1}\hat{f}(x,\omega)||a(x,\omega)||V_{\varphi_2}f(x,\omega)|\,dxd\omega \\
&\le \|a\|_\infty \iint_{\mathbb{R}^{2d}} |V_{\varphi_1}\hat{f}(x,\omega)||V_{\varphi_2}f(x,\omega)|\,dxd\omega \\
&= \|a\|_\infty \Big[\iint_{U_\varepsilon} |V_{\varphi_1}\hat{f}||V_{\varphi_2}f|\,dxd\omega \\
&\qquad + \iint_{\mathbb{R}^{2d}\setminus U_\varepsilon} |V_{\varphi_1}\hat{f}||V_{\varphi_2}f|\,dxd\omega\Big] \\
&\le \|a\|_\infty \Big[\Big(\iint_{U_\varepsilon} |V_{\varphi_1}\hat{f}|^2\,dxd\omega\Big)^{1/2}\Big(\iint_{U_\varepsilon} |V_{\varphi_2}f|^2\,dxd\omega\Big)^{1/2} \\
&\qquad + \Big(\iint_{\mathbb{R}^{2d}\setminus U_\varepsilon} |V_{\varphi_1}\hat{f}|^2\,dxd\omega\Big)^{1/2}\Big(\iint_{\mathbb{R}^{2d}\setminus U_\varepsilon} |V_{\varphi_2}f|^2\,dxd\omega\Big)^{1/2}\Big] \\
&\le \|a\|_\infty \Big[\|V_{\varphi_1}\hat{f}\|\Big(\iint_{U_\varepsilon} |V_{\varphi_2}f|^2\,dxd\omega\Big)^{1/2} \\
&\qquad + \Big(\iint_{\mathbb{R}^{2d}\setminus U_\varepsilon} |V_{\varphi_1}\hat{f}|^2\,dxd\omega\Big)^{1/2}\|V_{\varphi_2}f\|\Big] \\
&\le \|a\|_\infty\Big[\|f\|\,\|\varphi_1\|\cdot\sqrt{\varepsilon}\,\|f\|\,\|\varphi_2\| + \sqrt{\varepsilon}\,\|f\|\,\|\varphi_1\|\cdot\|f\|\,\|\varphi_2\|\Big] \\
&= 2\,\|a\|_\infty\,\|\varphi_1\|\,\|\varphi_2\|\,\|f\|^2\cdot\sqrt{\varepsilon} \\
&= C\,\|g\|^2\cdot\sqrt{\varepsilon}
\end{aligned}$$

for all sufficiently large x_0 and ω_0, with $C = 2\,\|a\|_\infty\,\|\varphi_1\|\,\|\varphi_2\|$ independent of x_0, ω_0. Thus the claim

$$\lim_{x_0,\omega_0\to\infty}\big|\langle V_{\varphi_1}\hat{f}, aV_{\varphi_2}f\rangle\big| = 0.$$

is proved.

For nonzero $g \in L^2(\mathbb{R}^d)$ choose x_0, ω_0 such that

$$\big|\langle V_{\varphi_1}(\widehat{M_{\omega_0}T_{x_0}}g), aV_{\varphi_2}(M_{\omega_0}T_{x_0}g)\rangle\big| \le \frac{1}{4}\|g\|^2.$$

Then with $f := M_{\omega_0} T_{x_0} g$ we have

$$\begin{aligned} \|\mathcal{F}f - A_a^{\varphi_1,\varphi_2} f\|^2 &\geq \|\hat{f}\|^2 + \|A_a^{\varphi_1,\varphi_2} f\|^2 - 2\left|\langle V_{\varphi_1}\hat{f}, aV_{\varphi_2}f\rangle\right| \\ &\geq \|g\|^2 - 2\,\frac{1}{4}\|g\|^2 \\ &= \frac{1}{2}\|f\|^2 \end{aligned}$$

since $\|f\| = \|\hat{f}\| = \|g\| = \|\hat{g}\|$. This shows

$$\|\mathcal{F} - A_a^{\varphi_1,\varphi_2}\|_{op} \geq \frac{1}{\sqrt{2}} > 0$$

for all $a \in L^\infty(\mathbb{R}^{2d})$, which was to be proved. □

Note that virtually the same argument applies to any operator of the metaplectic representation, since these are exactly the operators that move the time-frequency distribution of functions in $L^2(\mathbb{R}^d)$ in phase space. So none of these operators is contained in the norm-closure of ran$(\mathcal{A}) \subset B(L^2(\mathbb{R}^d))$.

However, the density of the set of localization operators with L^∞-symbols with respect to the weak* topology can be characterized completely.

Theorem 3.4.4. *Let $\mathcal{A} : L^\infty(\mathbb{R}^{2d}) \to B(L^2(\mathbb{R}^d))$ and $\mathcal{B} : \mathcal{S}^1(L^2(\mathbb{R}^d)) \to L^1(\mathbb{R}^{2d})$ be given as before.*
The following conditions are equivalent:

1. *ran$(\mathcal{A})$ is weak* dense in $B(L^2(\mathbb{R}^d))$.*

2. *The Berezin transform $\mathcal{B}$ is one-to-one on $\mathcal{S}^1$.*

3. *The short-time Fourier transform of the windows φ_1, φ_2 is nonzero almost everywhere, i.e. $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for almost all $(x, \omega) \in \mathbb{R}^{2d}$.*

Proof. (1) ⇔ (2): follows from our remarks on functional analysis at the beginning of the section, since $\mathcal{A} = \mathcal{B}^*$.
(2) ⇔ (3): follows from Theorem 3.4.1. □

Now consider $\mathcal{A} : L^q(\mathbb{R}^{2d}) \to \mathcal{S}^q(L^2(\mathbb{R}^d))$ for $2 \leq q < \infty$. Then $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$, where $1 < p \leq 2$ is the conjugate exponent defined by the relation $\frac{1}{p} + \frac{1}{q} = 1$, is the adjoint of $\mathcal{A}$ by Theorem 3.2.9. Thus in this case, the same reasoning as above yields a much stronger statement about norm density.

Theorem 3.4.5. *Let $\mathcal{A} : L^q(\mathbb{R}^{2d}) \to \mathcal{S}^q(L^2(\mathbb{R}^d))$ with $2 \leq q < \infty$ and $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$ with $\frac{1}{p} + \frac{1}{q} = 1$.*
The following conditions are equivalent:

1. *ran($\mathcal{A}$) is norm dense in $\mathcal{S}^q(L^2(\mathbb{R}^d))$.*
2. *The Berezin transform $\mathcal{B}$ is one-to-one on $\mathcal{S}^q$.*
3. *The short-time Fourier transform of the windows φ_1, φ_2 is nonzero almost everywhere, i.e. $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for almost all $(x, \omega) \in \mathbb{R}^{2d}$.*

Proof. The proof is exactly the same as for the preceding theorem. □

We emphasize the special case of $p = q = 2$ in Theorem 3.4.5 as a corollary on its own:

Corollary 3.4.6. *Let $\varphi_1, \varphi_2 \in L^2(\mathbb{R}^d)$ such that $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for almost every $(x, \omega) \in \mathbb{R}^{2d}$.*
Then the set $\{A_a^{\varphi_1,\varphi_2} | a \in L^2(\mathbb{R}^{2d})\} \subseteq \mathcal{S}^2(L^2(\mathbb{R}^d))$ is norm dense in the space $\mathcal{S}^2(L^2(\mathbb{R}^d))$ of all Hilbert-Schmidt operators on $L^2(\mathbb{R}^d)$. □

Finally, for $1 < q < 2$, the analogous statement follows using Theorem 3.4.2.

Theorem 3.4.7. *Let $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$. Let $\mathcal{A} : L^q(\mathbb{R}^{2d}) \to \mathcal{S}^q(L^2(\mathbb{R}^d))$ with $1 < q < 2$ and $\mathcal{B} : \mathcal{S}^p(L^2(\mathbb{R}^d)) \to L^p(\mathbb{R}^{2d})$ with $\frac{1}{p} + \frac{1}{q} = 1$.*
The following conditions are equivalent:

1. *ran($\mathcal{A}$) is norm dense in $\mathcal{S}^q(L^2(\mathbb{R}^d))$.*
2. *The Berezin transform $\mathcal{B}$ is one-to-one on $\mathcal{S}^q$.*
3. *The short-time Fourier transform of the windows φ_1, φ_2 is nonzero everywhere, i.e. $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for all $(x, \omega) \in \mathbb{R}^{2d}$.*

Proof. The proof is again identical to the proof of Theorem 3.4.4, only with Theorem 3.4.1 replaced by Theorem 3.4.2. □

The following lemma allows to extend the previous results to localization operators with symbols in modulation spaces.

Lemma 3.4.8. *The Lebesgue spaces $L^p(\mathbb{R}^d)$ are continuously embedded into the modulation spaces $M^{p,\infty}(\mathbb{R}^d)$:*

$$L^p(\mathbb{R}^d) \hookrightarrow M^{p,\infty}(\mathbb{R}^d)$$

for all $1 \leq p \leq \infty$ and

$$\|f\|_{M^{p,\infty}} \leq C \cdot \|f\|_{L^p}$$

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

Proof. Let $1 \leq p < \infty$ first.
Assume $f, g \in \mathcal{S}(\mathbb{R}^d)$. Then

$$\begin{aligned} \|f\|_{M^{p,\infty}} &\leq C \cdot \|V_g f\|_{L^{p,\infty}} \\ &= C \cdot \sup_{\omega \in \mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V_g f(x,\omega)|^p \, dx \right)^{1/p}. \end{aligned}$$

Now observe that

$$\begin{aligned} V_g f(x,\omega) &= \int_{\mathbb{R}^d} f(t)\overline{g(t-x)} e^{-2\pi i \omega \cdot t} \, dt \\ &= \int_{\mathbb{R}^d} f(s+x)\overline{g(s)} e^{-2\pi i \omega \cdot (s+x)} \, ds \\ &= e^{-2\pi i \omega \cdot x} \cdot \overline{\int_{\mathbb{R}^d} g(s)\overline{f(s+x)} e^{-2\pi i (-\omega)\cdot s} \, ds} \\ &= e^{-2\pi i \omega \cdot x} \cdot \overline{V_f g(-x,-\omega)} \end{aligned}$$

for all $(x,\omega) \in \mathbb{R}^{2d}$. Thus

$$\sup_{\omega \in \mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V_g f(x,\omega)|^p \, dx \right)^{1/p} = \sup_{\omega \in \mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V_f g(-x,-\omega)|^p \, dx \right)^{1/p}.$$

This last integral can be estimated by

$$\begin{aligned}
&\left(\int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} g(t)\overline{f(t+x)}e^{-2\pi i\omega\cdot t}\,dt\right|^p dx\right)^{1/p} \\
&\quad\leq \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |g(t)|^p\cdot|f(t+x)|^p\,dx\right)^{1/p} dt \\
&\quad= \int_{\mathbb{R}^d} |g(t)|\left(\int_{\mathbb{R}^d} |f(t+x)|^p\,dx\right)^{1/p} dt \\
&\quad= \int_{\mathbb{R}^d} |g(t)|\cdot\|f\|_{L^p}\,dt \\
&\quad= \|g\|_{L^1}\cdot\|f\|_{L^p} \\
&\quad< \infty,
\end{aligned}$$

where we have used Minkowski's Inequality for integrals. So

$$\sup_{\omega\in\mathbb{R}^d}\left(\int_{\mathbb{R}^d} |V_gf(x,\omega)|^p\,dx\right)^{1/p} \leq \|g\|_{L^1}\cdot\|f\|_{L^p}$$

and

$$\|f\|_{M^{p,\infty}} \leq C\cdot\|f\|_{L^p}$$

with a constant $C > 0$ depending on g.
Thus the identity mapping $f \mapsto f$ from $\mathcal{S}(\mathbb{R}^d)$ into $M^{p,\infty}(\mathbb{R}^d)$ is continuous with respect to the L^p-norm and the $M^{p,\infty}$-norm. Therefore, it can be extended uniquely to a bounded linear embedding $L^p(\mathbb{R}^d) \hookrightarrow M^{p,\infty}(\mathbb{R}^d)$ by the standard density argument, since $\mathcal{S}(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$ with respect to the L^p-norm.
For the case $p = \infty$, we compute directly

$$\begin{aligned}
\|f\|_{M^\infty} &\leq C\cdot\|V_gf\|_{L^\infty} \\
&= C\cdot\sup_{(x,\omega)\in\mathbb{R}^{2d}} |V_gf(x,\omega)| \\
&\leq C\cdot\sup_{(x,\omega)\in\mathbb{R}^{2d}} \left|\int_{\mathbb{R}^d} f(t)\overline{g(t-x)}e^{-2\pi i\omega\cdot t}\,dt\right| \\
&\leq C\cdot\sup_{(x,\omega)\in\mathbb{R}^{2d}} \int_{\mathbb{R}^d} |f(t)\cdot|g(t-x)|\,dt \\
&\leq C\cdot\|g\|_{L^1}\cdot\|f\|_{L^\infty} \\
&\leq C'\cdot\|f\|_{L^\infty}
\end{aligned}$$

for all $f \in L^\infty(\mathbb{R}^d)$. $\square$

Since $L^p(\mathbb{R}^{2d}) \subseteq M^{p,\infty}(\mathbb{R}^{2d})$ obviously implies

$$\{A_a^{\varphi_1,\varphi_2} |\ a \in L^p(\mathbb{R}^{2d})\} \subseteq \{A_a^{\varphi_1,\varphi_2} |\ a \in M^{p,\infty}(\mathbb{R}^{2d})\} \subseteq \mathcal{S}^p(L^2(\mathbb{R}^d)),$$

we conclude that at least one of the implications of the above density results for localization operators with symbols in L^p also holds for localization operators with symbols in $M^{p,\infty}$. Thus we have

Theorem 3.4.9. *Let $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$ and $\mathcal{A} : M^\infty(\mathbb{R}^{2d}) \to B(L^2(\mathbb{R}^d))$ be given as before.*
If the short-time Fourier transform of the windows φ_1, φ_2 is nonzero almost everywhere, i.e. $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for almost all $(x, \omega) \in \mathbb{R}^{2d}$, then ran($\mathcal{A}$) is weak dense in $B(L^2(\mathbb{R}^d))$.* $\square$

Theorem 3.4.10. *Let $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$ and $\mathcal{A} : M^{q,\infty}(\mathbb{R}^{2d}) \to \mathcal{S}^q(L^2(\mathbb{R}^d))$ with $2 \leq q < \infty$.*
If the short-time Fourier transform of the windows φ_1, φ_2 is nonzero almost everywhere, i.e. $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for almost all $(x, \omega) \in \mathbb{R}^{2d}$, then ran($\mathcal{A}$) is norm dense in $\mathcal{S}^q(L^2(\mathbb{R}^d))$. $\square$

Theorem 3.4.11. *Let $\varphi_1, \varphi_2 \in M^1(\mathbb{R}^d)$ and $\mathcal{A} : M^{q,\infty}(\mathbb{R}^{2d}) \to \mathcal{S}^q(L^2(\mathbb{R}^d))$ with $1 < q < 2$.*
If the short-time Fourier transform of the windows φ_1, φ_2 is nonzero everywhere, i.e. $V(\varphi_1, \varphi_2)(x, \omega) \neq 0$ for all $(x, \omega) \in \mathbb{R}^{2d}$, then ran($\mathcal{A}$) is norm dense in $\mathcal{S}^q(L^2(\mathbb{R}^d))$. $\square$

Appendix

A.1 The Standard Density Argument

Lemma A.1.1. *Let X, Y be Banach spaces, $D \subset X$ a dense linear subspace, and $T : D \to Y$ a bounded linear map. Then T extends uniquely to a bounded operator $\tilde{T} : X \to Y$, i.e. $\tilde{T} \in L(X,Y)$ and $\tilde{T}(x) = T(x)$ for all $x \in D$. We have $||\tilde{T}||_{X \to Y} = ||T||_{D \to Y}$.*

Proof. The subspace D is dense in X, so for every $x \in X$ there is a sequence $(x_n)_{n \in \mathbb{N}}$ in D that converges to x: $\lim_{n\to\infty} ||x_n - x|| = 0$. Define

$$\tilde{T}x := \lim_{n\to\infty} Tx_n.$$

The limit on the r.h.s. exists, since the sequence $(Tx_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in Y:

$$||Tx_n - Tx_m|| = ||T(\underbrace{x_n - x_m}_{\in D})|| \le ||T||_{D\to Y} ||x_n - x_m||;$$

but $||x_n - x_m|| \le \varepsilon$ for $n, m \ge N(\varepsilon)$, since $(x_n)_{n\in\mathbb{N}}$ converges and is therefore a Cauchy sequence in X.
The limit is independent of the approximating sequence:
Assume $x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n$ for two sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in D. Consider the mixed sequence $(z_1, z_1, \ldots) = (x_1, y_1, x_2, y_2, \ldots)$. Obviously $(z_n)_{n\in\mathbb{N}}$ also converges to x, so $\lim_{n\to\infty} Tz_n$ exists. The sequences $(Tx_n)_{n\in\mathbb{N}}$ and $(Ty_n)_{n\in\mathbb{N}}$ are subsequences of $(Tz_n)_{n\in\mathbb{N}}$ and converge therefore to the same limit: $\lim_{n\to\infty} Tx_n = \lim_{n\to\infty} Ty_n$ $(= \lim_{n\to\infty} Tz_n)$. So

$\tilde{T} : X \to Y$ is well defined.
The linearity of $\tilde{T}$ is trivial:
Let $x, y \in X$ and $a, b \in \mathbb{C}$, then choose approximating sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in D for x and y, respectively, and find

$$\tilde{T}(ax + by) = \lim_{n\to\infty} T(ax_n + by_n) = a \lim_{n\to\infty} Tx_n + b \lim_{n\to\infty} Ty_n = a\tilde{T}x + b\tilde{T}y.$$

Boundedness follows from

$$||\tilde{T}x|| = || \lim_{n\to\infty} Tx_n|| = \lim_{n\to\infty} ||Tx_n|| \leq ||T||_{D\to Y} \lim_{n\to\infty} ||x_n|| = ||T||_{D\to Y}||x||,$$

which shows $||\tilde{T}||_{X\to Y} \leq ||T||_{D\to Y}$.
Finally, for $x \in D$ one may choose the constant approximating sequence $(x_1, x_2, \ldots) = (x, x, \ldots)$ in D, which yields

$$\tilde{T}x = \lim_{n\to\infty} Tx = Tx,$$

so $\tilde{T}x = Tx$ for all $x \in D$, and obviously $||\tilde{T}||_{X\to Y} \geq ||T||_{D\to Y}$, since $\tilde{T}$ is an extension of T.
For uniqueness:
Suppose $\tilde{S} : X \to Y$ is another bounded extension of T, then for $x \in X$ choose an approximating sequence $(x_n)_{n\in\mathbb{N}}$ in D with $x = \lim_{n\to\infty} x_n$. By the continuity of $\tilde{S}$ and the definition of $\tilde{T}$ we have

$$\tilde{S}x = \lim_{n\to\infty} \tilde{S}x_n = \lim_{n\to\infty} Tx_n = \tilde{T}x.$$

□

The standard density argument is most often applied in the situation of the following corollary.

Corollary A.1.2. *Let X and Y be Banach spaces, $D \subset X$ a dense linear subspace, and $S : X \to Y$ and $T : X \to Y$ bounded linear operators such that $Sx = Tx$ for all $x \in D$. Then $S = T$.*

Proof. The operator $S - T$ is bounded on all of X and satisfies $Sx - Tx = (S - T)x = 0$ for all $x \in D$, so $S - T$ must be the unique bounded extension of the restriction $(S - T)|_D = 0$ of $S - T$ to the subspace D. But the unique bounded extension of the zero operator (on D) is again the zero operator (on X). Therefore $S - T = 0$. □

A.2 Time-Frequency Shifts

Definition A.2.1 (Translation, Modulation, Time-Frequency Shift)**.** *Let $u, v \in \mathbb{R}^d$ and $\lambda = (u, v) \in \mathbb{R}^{2d}$. Define the following operators on $L^2(\mathbb{R}^d)$:*

- **Translation:**

$$T_u : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d),$$
$$f(x) \mapsto T_u f(x) := f(x-u);$$

- **Modulation:**

$$M_v : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d),$$
$$f(x) \mapsto M_v f(x) := e^{2\pi i v \cdot x} f(x);$$

- **Time-frequency shift:**

$$\pi(\lambda) : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d),$$
$$f(x) \mapsto \pi(\lambda) f(x) := \pi(u, v) f(x) := M_v T_u f(x) = e^{2\pi i v \cdot x} f(x-u).$$

It is easy to see, that these are unitary operators on the Hilbert space $L^2(\mathbb{R}^d)$.

Lemma A.2.2 (Canonical Commutation Relation)**.** *Let $(u, v) \in \mathbb{R}^{2d}$. Then*

$$T_u M_v = e^{-2\pi i v \cdot u} M_v T_u.$$

Proof.

$$T_u M_v f(x) = e^{2\pi i v \cdot (x-u)} f(x-u) = e^{-2\pi i v \cdot u} M_v T_u f(x)$$

for all $f \in L^2(\mathbb{R}^d)$. □

An immediate corollary of this is the

Lemma A.2.3 (Commutation Relation for Time-Frequency Shifts)**.** *Let $\lambda = (x, \omega), \mu = (u, \eta) \in \mathbb{R}^{2d}$. Then*

$$\pi(\lambda)\pi(\mu) = e^{2\pi i (\omega \cdot u - x \cdot \eta)} \pi(\mu)\pi(\lambda).$$

Proof. Using Lemma A.2.2 twice, we find

$$\begin{aligned}\pi(\lambda)\pi(\mu) &= M_\omega T_x M_\eta T_u \\ &= e^{-2\pi i x\cdot\eta} M_\omega M_\eta T_x T_u \\ &= e^{-2\pi i x\cdot\eta} M_\eta M_\omega T_u T_x \\ &= e^{-2\pi i x\cdot\eta} e^{2\pi i \omega\cdot u} M_\eta T_u M_\omega T_x \\ &= e^{2\pi i(\omega\cdot u - x\cdot\eta)}\pi(\mu)\pi(\lambda).\end{aligned}$$

□

Lemma A.2.4. *The following formulas hold for all $f \in L^2(\mathbb{R}^d)$ and $u, v \in \mathbb{R}^d$:*

$$\begin{aligned}\widehat{T_u f} &= M_{-u}\hat{f}, \\ \widehat{M_v f} &= T_u \hat{f}, \\ \widehat{M_v T_u f} &= T_v M_{-u}\hat{f} = e^{2\pi i u\cdot v} M_{-u} T_v \hat{f}.\end{aligned}$$

Proof. Assume that $f \in \mathcal{S}(\mathbb{R}^d)$. Then

$$\begin{aligned}\widehat{T_u f}(\omega) &= \int_{\mathbb{R}^d} f(x-u)e^{-2\pi i\omega\cdot x}\,dx = \int_{\mathbb{R}^d} f(y)e^{-2\pi i\omega\cdot(y+u)}\,dy \\ &= e^{-2\pi i\omega\cdot u}\int_{\mathbb{R}^d} f(y)e^{-2\pi i\omega\cdot y}\,dy = M_{-u}\hat{f}(\omega)\end{aligned}$$

and

$$\begin{aligned}\widehat{M_v f}(\omega) &= \int_{\mathbb{R}^d} e^{2\pi i v\cdot x} f(x)e^{-2\pi i\omega\cdot x}\,dx = \int_{rd} f(x)e^{-2\pi i(\omega-v)\cdot x} \\ &= \hat{f}(\omega - v) = T_v\hat{f}(\omega).\end{aligned}$$

Using these formulas and the commutation relation A.2.2 yields

$$\widehat{M_v T_u f} = T_v \widehat{T_u f} = T_v M_{-u}\hat{f} = e^{2\pi i u\cdot v} M_{-u} T_v \hat{f}.$$

Since $\mathcal{S}(\mathbb{R}^d)$ is a dense subspace of $L^2(\mathbb{R}^d)$, the formulas extend to all of $L^2(\mathbb{R}^d)$ by the standard density argument A.1.1. □

Lemma A.2.5 (Strong Continuity of Time-Frequency Shifts)**.** *Let $\lambda_n = (u_n, v_n)$ converge to $\lambda = (u, v)$ in $\mathbb{R}^{2d}$. Then for any (fixed) $f \in L^2(\mathbb{R}^d)$, $\pi(\lambda_n)f \to \pi(\lambda)f$ in $L^2(\mathbb{R}^d)$.*

Proof. We first show that for fixed $h \in L^2(\mathbb{R}^d)$, $T_{u_n}h \to T_u h$ and $M_{v_n}h \to M_v h$ (i.e. the strong continuity of the unitary group of translations and the unitary group of modulations, respectively). We calculate

$$\begin{aligned} ||M_{v_n}h - M_v h||^2 &= \int_{\mathbb{R}^d} \left| e^{2\pi i v_n \cdot x} h(x) - e^{2\pi i v \cdot x} h(x) \right|^2 \, dx \\ &= \int_{\mathbb{R}^d} \left| e^{2\pi i v \cdot x} (e^{2\pi i (v_n - v)\cdot x} h(x) - h(x)) \right|^2 \, dx \\ &= \int_{\mathbb{R}^d} \left| (e^{2\pi i (v_n - v)\cdot x} - 1) h(x) \right|^2 \, dx. \end{aligned}$$

The integrand satisfies $\lim_{n\to\infty} \left|(e^{2\pi i (v_n - v)\cdot x} - 1)h(x)\right|^2 = 0$ pointwise for almost all $x \in \mathbb{R}^d$, and $\left|(e^{2\pi i (v_n - v)\cdot x} - 1)h(x)\right|^2 \leq 4|h(x)|^2$ for all $n \in \mathbb{N}$. The Dominated Convergence Theorem is applicable and yields

$$\lim_{n\to\infty} ||M_{v_n}h - M_v h||^2 = 0.$$

So $M_{v_n}h \to M_v h$ in $L^2(\mathbb{R}^d)$.
For the strong continuity of the translations, we observe

$$\begin{aligned} ||T_{u_n}h - T_u h|| &= ||(T_{u_n}h - T_u h)^\wedge|| \\ &= ||M_{-u_n}\hat{h} - M_{-u}\hat{h}|| \end{aligned}$$

by Parseval's Formula and Lemma A.2.4. The latter, however, tends to zero for $n \to \infty$, as shown above. So also $T_{u_n}h \to T_u h$ in $L^2(\mathbb{R}^d)$.
We conclude

$$\begin{aligned} ||\pi(\lambda_n)f - \pi(\lambda)f|| &= ||M_{v_n}T_{u_n}f - M_v T_u f|| \\ &= ||M_{v_n}T_{u_n}f - M_{v_n}T_u f + M_{v_n}T_u f - M_v T_u f|| \\ &\leq ||M_{v_n}(T_{u_n}f - T_u f)|| + ||(M_{v_n} - M_v)T_u f|| \\ &\leq \underbrace{||T_{u_n}f - T_u f||}_{\to 0} + \underbrace{||(M_{v_n} - M_v)T_u f||}_{\to 0} \to 0 \end{aligned}$$

for $n \to \infty$. □

Another useful property is given by the

Lemma A.2.6 (Convolution Relations)**.** *Let* $f, g \in L^2(\mathbb{R}^d)$, $x, \omega \in \mathbb{R}^d$. *Then*

$$(T_x f * T_x g)(t) = T_{2x}(f * g)(t)$$

and

$$(M_\omega f * M_\omega g)(t) = M_\omega(f * g)(t).$$

Proof. Compute

$$\begin{aligned}(T_x f * T_x g)(t) &= \int_{\mathbb{R}^d} f(s-x)g(t-s-x)\,ds \\ &= \int_{\mathbb{R}^d} f(y)g(t-y-2x)\,dy \\ &= (f*g)(t-2x) \\ &= T_{2x}(f*g)(t)\end{aligned}$$

with a trivial substitution and

$$\begin{aligned}(M_\omega f * M_\omega g)(t) &= \int_{\mathbb{R}^d} e^{2\pi i\omega\cdot s} f(s) e^{2\pi i \omega\cdot(t-s)} g(t-s)\,ds \\ &= e^{2\pi i\omega\cdot t}\int_{\mathbb{R}^d} f(s)g(t-s)\,ds \\ &= M_\omega(f*g)(t).\end{aligned}$$

□

A.3 Linear Coordinate Transformations

Definition A.3.1 (Coordinate Transformation)**.** *Let $A \in \mathbb{R}^{d\times d}$. Define the* ***coordinate transformation***

$$\mathcal{T}_A : f(x) \mapsto \mathcal{T}_A f(x) := f(Ax).$$

Lemma A.3.2. *If $\det(A) \neq 0$, the coordinate transformation $\mathcal{T}_A$ is an invertible bounded linear operator of $L^2(\mathbb{R}^d)$ onto itself with inverse $(\mathcal{T}_A)^{-1} = \mathcal{T}_{A^{-1}}$ and adjoint $(\mathcal{T}_A)^* = \frac{1}{|\det A|}\mathcal{T}_{A^{-1}}$. If $|\det(A)| = 1$, then $\mathcal{T}_A$ is unitary.*

Proof. We have

$$\begin{aligned} ||\mathcal{T}_A f||^2 &= \int_{\mathbb{R}^d} |\mathcal{T}_A f(x)|^2\, dx = \int_{\mathbb{R}^d} |f(Ax)|^2\, dx \\ &= \frac{1}{|\det(A)|}\int_{\mathbb{R}^d} |f(y)|^2\, dy = \frac{1}{|\det(A)|}||f||^2 \end{aligned}$$

by the change of variables formula, hence $\mathcal{T}_A$ is a bounded operator on $L^2(\mathbb{R}^d)$ for $\det(A) \neq 0$ and an isometry for $|\det(A)| = 1$. It is also invertible, since

$$\mathcal{T}_{A^{-1}}\mathcal{T}_A f = \mathcal{T}_A \mathcal{T}_{A^{-1}} f = f$$

for all $f \in L^2(\mathbb{R}^d)$, so $(\mathcal{T}_A)^{-1} = \mathcal{T}_{A^{-1}}$. Finally, again by a change of variables, we find

$$\begin{aligned} \langle \mathcal{T}_A f, g\rangle &= \int_{\mathbb{R}^d} f(Ax)\overline{g(x)}\, dx = \frac{1}{|\det(A)|}\int_{\mathbb{R}^d} f(y)\overline{g(A^{-1}y)}\, dy \\ &= \left\langle f, \frac{1}{|\det(A)|}\mathcal{T}_{A^{-1}} g\right\rangle, \end{aligned}$$

which proves $(\mathcal{T}_A)^* = \frac{1}{|\det A|}\mathcal{T}_{A^{-1}}$. □

Lemma A.3.3 (Commutation Relations with Time-Frequency Shifts)**.** *Let $f \in L^2(\mathbb{R}^d)$ and $A \in \mathbb{R}^{d\times d}$ be invertible. Then for all $x, \omega \in \mathbb{R}^d$*

$$\begin{aligned} \mathcal{T}_A(T_x f) &= T_{A^{-1}x}(\mathcal{T}_A f), \\ \mathcal{T}_A(M_\omega f) &= M_{A^*\omega}(\mathcal{T}_A f). \end{aligned}$$

Proof. We find

$$\begin{aligned}\mathcal{T}_A(T_x f)(t) &= T_x f(At) = f(At - x) = f(A(t - A^{-1}x)) \\ &= (\mathcal{T}_A f)(t - A^{-1}x) = T_{A^{-1}x}(\mathcal{T}_A f)(t)\end{aligned}$$

and

$$\begin{aligned}\mathcal{T}_A(M_\omega f)(t) &= M_\omega f(At) = e^{2\pi i\omega\cdot At} f(At) \\ &= e^{2\pi i A^*\omega\cdot t} f(At) = M_{A^*\omega}(\mathcal{T}_A f)(t).\end{aligned}$$

□

Lemma A.3.4 (Fourier Transform of a Coordinate Transformation)**.**

$$\widehat{\mathcal{T}_A f} = \frac{1}{|\det A|}\mathcal{T}_{(A^{-1})^*}\widehat{f}$$

holds for all invertible $A \in \mathbb{R}^{d\times d}$, $f \in L^2(\mathbb{R}^d)$.

Proof. For $f \in \mathcal{S}(\mathbb{R}^d)$, a change of variables yields

$$\begin{aligned}\widehat{\mathcal{T}_A f}(\xi) &= \int_{\mathbb{R}^d} f(Ax)e^{-2\pi i x\cdot\xi}\,dx \\ &= \frac{1}{|\det A|}\int_{\mathbb{R}^d} f(y)e^{-2\pi i A^{-1}y\cdot\xi}\,dy \\ &= \frac{1}{|\det A|}\int_{\mathbb{R}^d} f(y)e^{-2\pi i y\cdot(A^{-1})^*\xi}\,dy \\ &= \frac{1}{|\det A|}\widehat{f}((A^{-1})^*\xi) \\ &= \frac{1}{|\det A|}\mathcal{T}_{(A^{-1})^*}\widehat{f}(\xi).\end{aligned}$$

The standard density argument gives the statement for all $f \in L^2(\mathbb{R}^d)$. □

Lemma A.3.5. *Let* $A \in \mathbb{R}^{d\times d}$ *be invertible. Then the coordinate transformation* $\mathcal{T}_A$ *is a continuous mapping from* $\mathcal{S}(\mathbb{R}^d)$ *to* $\mathcal{S}(\mathbb{R}^d)$*.*

Proof. Let $f \in \mathcal{S}(\mathbb{R}^d)$.
We start by considering a single partial derivative of $\mathcal{T}_A f$. For $1 \le j \le d$,

$$\begin{aligned}\partial_j \mathcal{T}_A f(x) &= \partial_j(f(Ax)) \\ &= \partial_j(f(a_{11}x_1 + \ldots + a_{1d}x_d, \ldots, a_{d1}x_1 + \ldots + a_{dd}x_d));\end{aligned}$$

the chain rule gives

$$\partial_j \mathcal{T}_A f(x) = \sum_{k=1}^{d} (\partial_k f)(Ax) \cdot a_{kj} = \mathcal{T}_A(\sum_{k=1}^{d} a_{kj}\partial_k f)(x).$$

Continuing inductively, we get for any multiindex $\beta \in \mathbb{N}_0^d$

$$\partial^\beta \mathcal{T}_A f(x) = \mathcal{T}_A(\sum_{|\gamma|\le|\beta|} c_\gamma \partial^\gamma f)(x)$$

with some appropriate coefficients c_γ.
If $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ is a multiindex, then

$$x^\alpha \mathcal{T}_A f(x) = x^\alpha f(Ax) = (A^{-1}Ax)^\alpha f(Ax) = (A^{-1}y)^\alpha f(y),$$

where we substituted y for Ax. The d components of $A^{-1}y$ are each a linear combination of $y_1, \ldots, y_d$. Denoting $A^{-1} = \left(a'_{ij}\right)_{1\le i,j\le d}$, we have

$$(A^{-1}y)^\alpha = \prod_{i=1}^{d}(\sum_{j=1}^{d} a'_{ij}y_j)^{\alpha_i} = p_\alpha(y)$$

with some polynomial $p_\alpha(y) = \sum_{|\delta|\le|\alpha|} d_\delta y^\delta$ of degree less than or equal $|\alpha|$. The matrix A is invertible, so when x runs through all of $\mathbb{R}^d$, the same is true of $y = Ax$, therefore

$$\sup_{x\in\mathbb{R}^d} |x^\alpha \mathcal{T}_A f(x)| = \sup_{y\in\mathbb{R}^d} |p_\alpha(y) f(y)|.$$

By putting the pieces together, we conclude for any multiindices $\alpha, \beta \in \mathbb{N}_0^d$

$$\begin{aligned}x^\alpha \partial^\beta \mathcal{T}_A f(x) &= x^\alpha \mathcal{T}_A(\sum_{|\gamma|\le|\beta|} c_\gamma \partial^\gamma f) \\ &= \sum_{|\delta|\le|\alpha|} d_\delta y^\delta (\sum_{|\gamma|\le|\beta|} c_\gamma \partial^\gamma f)(y)\end{aligned}$$

and

$$\begin{aligned}\sup_{x\in\mathbb{R}^d} |x^\alpha \partial^\beta \mathcal{T}_A f(x)| &= \sup_{y\in\mathbb{R}^d} |\sum_{|\delta|\le|\alpha|} d_\delta y^\delta (\sum_{|\gamma|\le|\beta|} c_\gamma \partial^\gamma f)(y)| \\ &\le \sum_{|\delta|\le|\alpha|} \sum_{|\gamma|\le|\beta|} |d_\delta| \cdot |c_\gamma| \cdot ||y^\delta \partial^\gamma f(y)||_\infty .\end{aligned}$$

This estimate proves that $\mathcal{T}_A f \in \mathcal{S}(\mathbb{R}^d)$ and $\mathcal{T}_A : \mathcal{S}(\mathbb{R}^d) \to \mathcal{S}(\mathbb{R}^d)$ is continuous. □

A.4 Partial Fourier Transform

Definition A.4.1 (Partial Fourier Transform)**.** *Let* $f = f(x, y) \in L^2(\mathbb{R}^{2d})$, $x, y \in \mathbb{R}^d$.
The ***partial Fourier transform*** *of* f *(with respect to the second argument) is defined as*

$$\mathcal{F}_2 f(x, \omega) := \widehat{f_x}(\omega),$$

where $\widehat{\ }$ *denotes the Fourier transform on* $L^2(\mathbb{R}^d)$ *and*

$$f_x(y) := f(x, y), \qquad y \in \mathbb{R}^d,$$

denotes the cross section of f *for fixed first argument* $x \in \mathbb{R}^d$*.*

By Fubini's Theorem, $f_x \in L^2(\mathbb{R}^d)$ for almost every $x \in \mathbb{R}^d$, thus $\mathcal{F}_2 f$ is well defined.

Lemma A.4.2. $\mathcal{F}_2$ *is a unitary operator from* $L^2(\mathbb{R}^{2d})$ *to* $L^2(\mathbb{R}^{2d})$*. The inverse (=adjoint) is given by*

$$\mathcal{F}_2^{-1} F(x, y) = \mathcal{F}_2^* F(x, y) = \mathcal{F}_2 F(x, -y) = \mathcal{T}_A \mathcal{F}_2 F(x, y)$$

with matrix $A = \left(\begin{smallmatrix} I & 0 \\ 0 & -I \end{smallmatrix}\right)$*.*

Proof. We have

$$\begin{aligned}
\iint_{\mathbb{R}^{2d}} |\mathcal{F}_2 f(x, \omega)|^2 \, dx d\omega &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |\widehat{f_x}(\omega)|^2 \, d\omega \right) dx \\
&= \int_{\mathbb{R}^d} \|\widehat{f_x}\|^2 \, dx \\
&= \int_{\mathbb{R}^d} \|f_x\|^2 \, dx \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f_x(y)|^2 \, dy \right) dx \\
&= \iint_{\mathbb{R}^{2d}} |f(x, y)|^2 \, dx dy
\end{aligned}$$

with Fubini's Theorem and Plancherel's Formula. Thus $\mathcal{F}_2 f \in L^2(\mathbb{R}^{2d})$, $\|\mathcal{F}_2 f\| = \|f\|$, and $\mathcal{F}_2 : L^2(\mathbb{R}^{2d}) \to L^2(\mathbb{R}^{2d})$ is an isometry.
Let $F \in L^2(\mathbb{R}^{2d})$. Set

$$f(x,y) := \mathcal{F}^{-1} F_x(y),$$

where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform. Then

$$\mathcal{F}_2 f(x,\omega) = \mathcal{F}^{-1}\mathcal{F} F_x(\omega) = F_x(\omega) = F(x,\omega),$$

thus $F = \mathcal{F}_2 f$ and $\mathcal{F}_2 : L^2(\mathbb{R}^{2d}) \to L^2(\mathbb{R}^{2d})$ is surjective and thus unitary.
The above also shows that

$$\mathcal{F}_2^{-1} g(x,y) = \mathcal{F}^{-1} F_x(y) = \mathcal{F} F_x(-y) = \mathcal{F}_2 F(x,-y).$$

□

Lemma A.4.3 (Commutation Relations with Time-Frequency Shifts)**.** *Let $F \in L^2(\mathbb{R}^{2d})$. Then*

$$\mathcal{F}_2(T_{\binom{r}{s}} F)(x,\omega) = e^{-2\pi i \omega\cdot s}(\mathcal{F}_2 F)(x-r,\omega) = M_{\binom{0}{-s}} T_{\binom{r}{0}} (\mathcal{F}_2 F)(x,\omega),$$
$$\mathcal{F}_2(M_{\binom{\rho}{\sigma}} F)(x,\omega) = e^{2\pi i \rho\cdot x}(\mathcal{F}_2 F)(x,\omega-\sigma) = M_{\binom{\rho}{0}} T_{\binom{0}{\sigma}} (\mathcal{F}_2 F)(x,\omega)$$

for all $\binom{r}{s}, \binom{\rho}{\sigma} \in \mathbb{R}^{2d}$.

Proof. Assume that $F \in \mathcal{S}(\mathbb{R}^{2d})$. Then a calculation yields

$$\begin{aligned}
\mathcal{F}_2(T_{\binom{r}{s}} F)(x,\omega) &= \int_{\mathbb{R}^d} F(x-r, y-s) e^{-2\pi i \omega\cdot y}\, dy \\
&= \int_{\mathbb{R}^d} F(x-r, t) e^{-2\pi i \omega\cdot (t+s)}\, dt \qquad \text{(Subst. } t = y-s) \\
&= e^{-2\pi i \omega\cdot s}(\mathcal{F}_2 F)(x-r,\omega) \\
&= M_{\binom{0}{-s}} T_{\binom{r}{0}} (\mathcal{F}_2 F)(x,\omega),
\end{aligned}$$

with a substitution at the indicated place.
Analogously, we compute

$$\begin{aligned}
\mathcal{F}_2(M_{\binom{\rho}{\sigma}} F)(x,\omega) &= \int_{\mathbb{R}^d} e^{2\pi i(\rho\cdot x + \sigma\cdot y)} F(x,y) e^{-2\pi i \omega\cdot y}\, dy \\
&= e^{2\pi i \rho\cdot x} \int_{\mathbb{R}^d} F(x,y) e^{-2\pi i(\omega-\sigma)\cdot y}\, dy \\
&= e^{2\pi i \rho\cdot x} (\mathcal{F}_2 F)(x,\omega-\sigma) \\
&= M_{\binom{\rho}{0}} T_{\binom{0}{\sigma}} (\mathcal{F}_2 F)(x,\omega).
\end{aligned}$$

That these formulas are also valid for arbitrary $F \in L^2(\mathbb{R}^{2d})$, follows from the standard density argument. $\square$

Lemma A.4.4. *The partial Fourier transform $\mathcal{F}_2$ is a continuous mapping from $\mathcal{S}(\mathbb{R}^{2d})$ to $\mathcal{S}(\mathbb{R}^{2d})$.*

Proof. Let $f \in \mathcal{S}(\mathbb{R}^{2d})$. Then

$$\mathcal{F}_2 f(x,\omega) = \int_{\mathbb{R}^d} f(x,y) e^{-2\pi i \omega \cdot y}\, dy$$

is well-defined and exists as an integral for all $x, \omega \in \mathbb{R}^d$. We can thus estimate for all $x, \omega \in \mathbb{R}^d$:

$$\begin{aligned}
|\mathcal{F}_2 f(x,\omega)| &= |\int_{\mathbb{R}^d} f(x,y) e^{-2\pi i \omega \cdot y}\, dy| \\
&\leq \int_{\mathbb{R}^d} |f(x,y)|\, dy \\
&= \int_{\mathbb{R}^d} \big((1+|x|^2+|y|^2)^{s/2} \cdot |f(x,y)|\big) \cdot (1+|x|^2+|y|^2)^{-s/2}\, dy \\
&\leq \sup_{(x,y)\in\mathbb{R}^{2d}} \big((1+|x|^2+|y|^2)^{s/2} \cdot |f(x,y)|\big) \cdot \int_{\mathbb{R}^d} (1+|y|^2)^{-s/2}\, dy.
\end{aligned}$$

The supremum is finite for arbitrary $s > 0$, since $f \in \mathcal{S}(\mathbb{R}^{2d})$. The integral is finite if and only if $s > d$. Thus

$$\|\mathcal{F}_2 f\|_\infty \leq C \cdot \|(1+|x|^2+|y|^2)^{s/2} \cdot f(x,y)\|_\infty \tag{$*$}$$

for $s > d$ and $C > 0$ some constant (depending only on s).
For convenience we introduce the following notation: Let $F = F(x,y)$ be a function on $\mathbb{R}^{2d}$ and $\alpha = (\alpha_1, \alpha_2), \beta = (\beta_1, \beta_2) \in \mathbb{N}_0^{2d}$ be multiindices; then write

$$M^\alpha_{(x,y)} := M_x^{\alpha_1} M_y^{\alpha_2} F(x,y) := x^{\alpha_1} y^{\alpha_2} F(x,y)$$

for the multiplication operator and

$$D^\beta_{(x,y)} := D_x^{\beta_1} D_y^{\beta_2} F(x,y) := \partial_x^{\beta_1} \partial_y^{\beta_2} F(x,y)$$

for the partial differential operator.
It is then not hard to prove that the partial Fourier transform satisfies the following rules that are analogous to the full Fourier transform:

$$D_x^{\beta_1} D_\omega^{\beta_2} \mathcal{F}_2 f = (-2\pi i)^{|\beta_2|} \cdot \mathcal{F}_2 (D_x^{\beta_1} M_y^{\beta_2} f)$$

and

$$M_x^{\alpha_1} M_\omega^{\alpha_2} \mathcal{F}_2 f = (\frac{1}{2\pi i})^{|\alpha_2|} \cdot \mathcal{F}_2(M_x^{\alpha_1} D_y^{\alpha_2} f).$$

Hence, by induction, $\mathcal{F}_2 f$ is infinitely differentiable, and for arbitrary multi-indices $\alpha, \beta \in \mathbb{N}_0^{2d}$ we have

$$\begin{aligned} |M_{(x,\omega)}^{\alpha} D_{(x,\omega)}^{\beta} \mathcal{F}_2 f(x,\omega)| &= |M_x^{\alpha_1} M_\omega^{\alpha_2} D_x^{\beta_1} D_\omega^{\beta_2} \mathcal{F}_2 f(x,\omega)| \\ &= |2\pi|^{|\beta_2|-|\alpha_2|} \cdot |\mathcal{F}_2(M_x^{\alpha_1} D_\omega^{\alpha_2} D_x^{\beta_1} M_\omega^{\beta_2} f)(x,\omega)| \\ &\leq C \cdot \|\mathcal{F}_2(M_x^{\alpha_1} D_\omega^{\alpha_2} D_x^{\beta_1} M_\omega^{\beta_2} f)\|_\infty. \end{aligned}$$

The estimate $(*)$ now yields

$$\sup_{(x,\omega)\in\mathbb{R}^{2d}} |M_{(x,\omega)}^{\alpha} D_{(x,\omega)}^{\beta} \mathcal{F}_2 f(x,\omega)| < \infty$$

for all multiindices α, β (hence $\mathcal{F}_2 f \in \mathcal{S}(\mathbb{R}^{2d})$) and continuity of the mapping $\mathcal{F}_2 : \mathcal{S}(\mathbb{R}^{2d}) \to \mathcal{S}(\mathbb{R}^{2d})$. □

A.5 Short-Time Fourier Transform and Modulation Spaces

Definition A.5.1 (Short-Time Fourier Transform)**.** *Let $f, g \in L^2(\mathbb{R}^d)$. The* ***short-time Fourier transform*** *of f with window g (STFT) is defined as*

$$V(f,g)(x,\omega) := \langle f, M_\omega T_x g\rangle = \int_{\mathbb{R}^d} f(y)\overline{g(y-x)}e^{-2\pi i\omega\cdot y}\,dy$$

for $x, \omega \in \mathbb{R}^d$.

Lemma A.5.2. *Let $f, g \in L^2(\mathbb{R}^d)$. Then the following holds for all $x, \omega \in \mathbb{R}^d$:*

$$\begin{aligned} V(f,g)(x,\omega) &= (\widehat{f\cdot T_x\overline{g}})(\omega) \\ &= \left\langle \widehat{f}, T_\omega M_{-x}\widehat{g}\right\rangle \\ &= e^{-2\pi i x\cdot\omega}V(\widehat{f},\widehat{g})(\omega,-x). \end{aligned}$$

Proof. The first expression follows directly from the definition, the second by applying Plancherel's Theorem, and the third from Lemma A.2.3. □

Observe that the STFT is a bilinear time-frequency distribution in the sense considered in this work (with coefficient matrix $A = \left(\begin{smallmatrix} 0 & I \\ -I & I \end{smallmatrix}\right)$, which is right-regular, cf. Definition 1.2.10. Thus all results of the first chapter apply. In particular, one has such things as orthogonality relations or covariance formulas for the STFT.

Definition A.5.3 (Mixed-Norm Spaces)**.** *Let $1 \le p, q \le \infty$. The* ***mixed-norm space*** *$L^{p,q}(\mathbb{R}^{2d})$ is defined as*

$$L^{p,q}(\mathbb{R}^{2d}) := \{F : \mathbb{R}^{2d} \mapsto \mathbb{C} \text{ measurable} : ||F||_{L^{p,q}} < \infty\},$$

with mixed p,q-norm

$$||F||_{L^{p,q}} := \left(\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|F(x,y)|^p\,dx\right)^{q/p}dy\right)^{1/q}.$$

For $p = \infty$ or $q = \infty$, this definition is modified in the expected way, i.e.

$$||F||_{L^{\infty,q}} := \left(\int_{\mathbb{R}^d} \left(\sup_{x \in \mathbb{R}^d} |F(x,y)| \right)^q dy \right)^{1/q}, \quad \text{if } p = \infty,$$

resp.

$$||F||_{L^{p,\infty}} := \sup_{y \in \mathbb{R}^d} \left(\int_{\mathbb{R}^d} |F(x,y)|^p \, dx \right)^{1/p}, \quad \text{if } q = \infty.$$

As usual, we identify functions in $L^{p,q}$ that differ only on a set of Lebesgue measure zero (that is, formally we consider equivalence classes of measurable functions, two functions F and G being equivalent if and only if $\lambda(\{x \in \mathbb{R}^{2d} : F(x) \neq G(x)\}) = 0$, where λ denotes Lebesgue measure on $\mathbb{R}^{2d}$.) With this convention, the mixed p,q-norm becomes indeed a norm (not only a seminorm), and $L^{p,q}(\mathbb{R}^{2d})$ is a Banach space.

Lemma A.5.4 (Transformation Formula for Mixed Norm Spaces)**.** *Let $F \in L^{p,q}(\mathbb{R}^{2d})$, $1 \leq p,q \leq \infty$. Let $A_1, A_2 \in \mathbb{R}^{d \times d}$ be invertible matrices and $b_1, b_2 \in \mathbb{R}^d$ be arbitrary vectors. Consider the function*

$$G(z) = G(z_1, z_2) := F(A_1 z_1 + b_1, A_2 z_2 + b_2), \qquad z = (z_1, z_2) \in \mathbb{R}^{2d}.$$

Then $G \in L^{p,q}(\mathbb{R}^{2d})$ and

$$||G||_{L^{p,q}} = \frac{1}{|\det A_1|^{1/p} \cdot |\det A_2|^{1/q}} ||F||_{L^{p,q}}.$$

(In case $p = \infty$ or $q = \infty$, $\frac{1}{p}$ resp. $\frac{1}{q}$ are understood to be equal to 0.)

Proof. We compute the mixed p,q-norm of G:

$$\begin{aligned} ||G||_{L^{p,q}} &= \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |G(z_1, z_2)|^p \, dz_1 \right)^{q/p} dz_2 \right)^{1/q} \\ &= \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |F(A_1 z_1 + b_1, A_2 z_2 + b_2)|^p \, dz_1 \right)^{q/p} dz_2 \right)^{1/q}. \end{aligned}$$

The inner integral yields

$$\int_{\mathbb{R}^d} |F(A_1 z_1 + b_1, A_2 z_2 + b_2)|^p \, dz_1 = \frac{1}{|\det A_1|} \int_{\mathbb{R}^d} |F(s_1, A_2 z_2 + b_2)|^p \, ds_1$$

by an easy substitution in the first argument, hence

$$\begin{aligned}
||G||_{L^{p,q}} &= \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |G(z_1, z_2)|^p \, dz_1 \right)^{q/p} dz_2 \right)^{1/q} \\
&= \left(\int_{\mathbb{R}^d} \left(\frac{1}{|\det A_1|} \int_{\mathbb{R}^d} |F(s_1, A_2 z_2 + b_2)|^p \, ds_1 \right)^{q/p} dz_2 \right)^{1/q} \\
&= \frac{1}{|\det A_1|^{1/p}} \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |F(s_1, A_2 z_2 + b_2)|^p \, ds_1 \right)^{q/p} dz_2 \right)^{1/q} \\
&= \frac{1}{|\det A_1|^{1/p}} \left(\frac{1}{|\det A_2|} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |F(s_1, s_2)|^p \, ds_1 \right)^{q/p} ds_2 \right)^{1/q} \\
&= \frac{1}{|\det A_1|^{1/p} \cdot |\det A_2|^{1/q}} \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |F(s_1, s_2)|^p \, ds_1 \right)^{q/p} ds_2 \right)^{1/q} \\
&= \frac{1}{|\det A_1|^{1/p} \cdot |\det A_2|^{1/q}} ||F||_{L^{p,q}},
\end{aligned}$$

by another substitution in the second argument.
In case $p = \infty$ or $q = \infty$, the calculations are even simpler. □

Definition A.5.5 (Modulation Spaces)**.** *Let $g \in \mathcal{S}(\mathbb{R}^d)$ be fixed. We define the* ***modulation space***

$$M^{p,q}(\mathbb{R}^d) := \{ f \in \mathcal{S}'(\mathbb{R}^d) : V(f, g) \in L^{p,q}(\mathbb{R}^{2d}) \}.$$

It can be shown that the modulation spaces are Banach spaces whose definition does not depend on the chosen window function g; different windows yield just equivalent norms.

The theory of modulation spaces is explained in detail in the monographs [17] and [14].

A.6 Wiener Amalgam Spaces

This section discusses some properties of Wiener amalgam spaces. These spaces are used as a technical tool at several places in the text. They characterize functions that have certain local and global regularity properties described by a local resp. global component, usually some normed spaces. An excellent survey of Wiener amalgam spaces can be found in [12].

For our purposes, it is not necessary to define Wiener amalgam spaces in full generality. We thus restrict ourselves to the special case where the local component is always the Fourier algebra $\mathcal{F}L^1$ and the global component is some L^p-space. As we will see, such Wiener amalgams are closely related to the short-time Fourier transform and modulation spaces.

Definition A.6.1 (Wiener Amalgam Spaces)**.** *Let $g \in \mathcal{D}(\mathbb{R}^d)$ be a test function (i.e. a C^∞-function with compact support). The* ***Wiener amalgam space*** *$W(\mathcal{F}L^1, L^p)(\mathbb{R}^d)$ (with* ***local component*** *$\mathcal{F}L^1$ and* ***global component*** *L^p, $1 \leq p \leq \infty$) is defined as the space of all functions f on $\mathbb{R}^d$ such that the norm*

$$||f||_{W(\mathcal{F}L^1,L^p)} = \left(\int_{\mathbb{R}^d} ||f \cdot T_z g||^p_{\mathcal{F}L^1} \, dz \right)^{1/p}$$

is finite. (Note that for $p = \infty$ the obvious adjustments need to be made. For the sake of brevity, these will not always be pointed out explicitly in the following.)

If we denote

$$F(z) = F_f(z) = ||f \cdot T_z g||_{\mathcal{F}L^1}, \qquad z \in \mathbb{R}^d,$$

then obviously $f \in W(\mathcal{F}L^1, L^p)(\mathbb{R}^d)$ if and only if $F \in L^p(\mathbb{R}^d)$, and

$$||f||_{W(\mathcal{F}L^1,L^p)} = \left(\int_{\mathbb{R}^d} F(z)^p \, dz \right)^{1/p} = ||F||_{L^p}.$$

It is not hard to see that $W(\mathcal{F}L^1, L^p)(\mathbb{R}^d)$ equipped with the above norm is a Banach space, consisting of continuous functions on $\mathbb{R}^d$.

The next theorem shows that the definition of $W(\mathcal{F}L^1, L^p)(\mathbb{R}^d)$ does not depend on the particular choice of the test function g; different such functions yield the same space with equivalent norms.

Theorem A.6.2 (Equivalence of Norms). *Let $g_1, g_2 \in \mathcal{D}(\mathbb{R}^d)$ be two different test functions. Let f be a continuous function on $\mathbb{R}^d$. Then*

$$||f||_1 = \left(\int_{\mathbb{R}^d} ||f \cdot T_z g_1||^p_{\mathcal{F}L^1}\, dz\right)^{1/p} < \infty$$

if and only if

$$||f||_2 = \left(\int_{\mathbb{R}^d} ||f \cdot T_z g_2||^p_{\mathcal{F}L^1}\, dz\right)^{1/p} < \infty,$$

and in this case there are positive constants $A, B > 0$ independent of f such that

$$A \cdot ||f||_2 \leq ||f||_1 \leq B \cdot ||f||_2.$$

Proof. In order to prove the theorem, we first show a useful

Lemma A.6.3. *Let $f \in \mathcal{F}L^1(\mathbb{R}^d)$ and $\phi \in \mathcal{D}(\mathbb{R}^d)$.*
Then $f \cdot \phi \in \mathcal{F}L^1(\mathbb{R}^d)$ (and in particular $f \cdot T_z\phi \in \mathcal{F}L^1(\mathbb{R}^d)$ for all $z \in \mathbb{R}^d$), and there exists a constant $C = C(\phi) > 0$ independent of f such that

$$||f \cdot T_z\phi||_{\mathcal{F}L^1} \leq C \cdot ||f||_{\mathcal{F}L^1}$$

for all $z \in \mathbb{R}^d$.

Proof. We have $T_z\phi \in \mathcal{D}(\mathbb{R}^d) \subseteq \mathcal{S}(\mathbb{R}^d) = \mathcal{F}\mathcal{S}(\mathbb{R}^d) \subseteq \mathcal{F}L^1(\mathbb{R}^d)$ for all $z \in \mathbb{R}^d$. Now $\mathcal{F}L^1(\mathbb{R}^d)$ is a Banach algebra under pointwise multiplication, thus $f \cdot T_z\phi \in \mathcal{F}L^1$ for all $z \in \mathbb{R}^d$ with

$$||f \cdot T_z\phi||_{\mathcal{F}L^1} \leq ||f||_{\mathcal{F}L^1} \cdot ||T_z\phi||_{\mathcal{F}L^1}.$$

Now, since $\phi \in \mathcal{D}(\mathbb{R}^d) \subseteq \mathcal{S}(\mathbb{R}^d)$, there is a $\psi \in \mathcal{S}(\mathbb{R}^d)$ such that $\phi = \widehat{\psi}$. Then $T_z\phi = T_z\widehat{\psi} = \widehat{M_z\psi}$, hence

$$||T_z\phi||_{\mathcal{F}L^1} = ||M_z\psi||_{L^1} = ||\psi||_{L^1} = ||\phi||_{\mathcal{F}L^1}$$

independent of $z \in \mathbb{R}^d$. So

$$||f \cdot T_z\phi||_{\mathcal{F}L^1} \leq ||\phi||_{\mathcal{F}L^1} \cdot ||f||_{\mathcal{F}L^1}$$

for all $f \in \mathcal{F}L^1(\mathbb{R}^d)$ and $z \in \mathbb{R}^d$. □

Now assume $||f||_2 < \infty$.
Let $\phi = g_2 \cdot \overline{g_2} = |g_2|^2 \in \mathcal{D}(\mathbb{R}^d)$. Then

$$\begin{aligned} ||f \cdot T_z \phi||_{\mathcal{F}L^1} &= ||f \cdot T_z(g_2 \cdot \overline{g_2})||_{\mathcal{F}L^1} \\ &= ||(f \cdot T_z g_2) \cdot T_z \overline{g_2}||_{\mathcal{F}L^1} \\ &\leq C' \cdot ||f \cdot T_z g_2||_{\mathcal{F}L^1} \end{aligned}$$

with a constant $C' = C'(\overline{g_2}) > 0$, by the Lemma. Hence

$$\left(\int_{\mathbb{R}^d} ||f \cdot T_z \phi||^p_{\mathcal{F}L^1} \, dz \right)^{1/p} \leq C' \cdot ||f||_2 < \infty.$$

Observe that $\phi \geq 0$, so we can find a suitable linear combination of translates of ϕ, say

$$\psi = \sum_{j=1}^{n} T_{z_j} \phi \in \mathcal{D}(\mathbb{R}^d),$$

such that $\psi(t) \geq \delta > 0$ for all $t \in \text{supp}(g_1)$, the support of g_1. Then

$$\begin{aligned} ||f \cdot T_z \psi||_{\mathcal{F}L^1} &= ||f \cdot T_z(\sum_{j=1}^{n} T_{z_j} \phi)||_{\mathcal{F}L^1} \\ &= ||f \cdot \sum_{j=1}^{n} T_{z+z_j} \phi||_{\mathcal{F}L^1} \\ &= ||\sum_{j=1}^{n} f \cdot T_{z+z_j} \phi||_{\mathcal{F}L^1} \\ &\leq \sum_{j=1}^{n} ||f \cdot T_{z+z_j} \phi||_{\mathcal{F}L^1}, \end{aligned}$$

so

$$\begin{aligned}\left(\int_{\mathbb{R}^d} ||f \cdot T_z\psi||_{\mathcal{F}L^1}^p \, dz\right)^{1/p} &\leq \left(\int_{\mathbb{R}^d} \sum_{j=1}^{n} ||f \cdot T_{z+z_j}\phi||_{\mathcal{F}L^1} \, dz\right)^{1/p} \\ &\leq \sum_{j=1}^{n} \left(\int_{\mathbb{R}^d} ||f \cdot T_{z+z_j}\phi||_{\mathcal{F}L^1} \, dz\right)^{1/p} \\ &= n \cdot \left(\int_{\mathbb{R}^d} ||f \cdot T_z\phi||_{\mathcal{F}L^1} \, dz\right)^{1/p} \\ &\leq n \cdot C' \cdot ||f||_2 \\ &< \infty,\end{aligned}$$

since the L^p-norm is translation invariant. Finally, choose a test function $\rho \in \mathcal{D}(\mathbb{R}^d)$ such that $\rho(t) = \frac{1}{\psi(t)}$ for $t \in \text{supp}(g_1)$. Then obviously $g_1 = \rho \cdot (\psi \cdot g_1)$. Therefore

$$\begin{aligned}||f \cdot T_z g_1||_{\mathcal{F}L^1} &= ||f \cdot T_z(\rho \cdot \psi \cdot g_1)||_{\mathcal{F}L^1} \\ &= ||(f \cdot T_z\rho) \cdot T_z(\psi \cdot g_1)||_{\mathcal{F}L^1} \\ &\leq C'' \cdot ||f \cdot T_z\rho||_{\mathcal{F}L^1}\end{aligned}$$

with a constant $C'' = C''(\psi \cdot g_1) > 0$, again by the Lemma, and thus

$$\begin{aligned}||f||_1 &= \left(\int_{\mathbb{R}^d} ||f \cdot T_z g_1||_{\mathcal{F}L^1}^p \, dz\right)^{1/p} \\ &\leq C'' \cdot \left(\int_{\mathbb{R}^d} ||f \cdot T_z\rho||_{\mathcal{F}L^1}^p \, dz\right)^{1/p} \\ &\leq C'' \cdot C' \cdot n \cdot ||f||_2 \\ &< \infty.\end{aligned}$$

We conclude that $||f||_2 < \infty$ implies $||f||_1 < \infty$ and

$$||f||_1 \leq B \cdot ||f||_2$$

with $B := C'' \cdot C' \cdot n > 0$.
Finally, by symmetry, we may interchange the roles of g_1 and g_2 in the preceding argument, so the other inequality is valid as well. □

The connection with the short-time Fourier transform is given in

Theorem A.6.4. *Let f be a function (or tempered distribution) on $\mathbb{R}^d$ and $g \in \mathcal{D}(\mathbb{R}^d)$ be a test function. Then*

$$||f||_{W(\mathcal{F}L^1,L^q)} = ||V(f,g)||_{L^{1,q}}$$

(where $L^{1,q}(\mathbb{R}^{2d})$ denotes the mixed-norm space defined in Definition A.5.3). Thus $W(\mathcal{F}L^1, L^q)(\mathbb{R}^d) = M^{1,q}(\mathbb{R}^d)$ with equivalent norms, in particular $W(\mathcal{F}L^1, L^1)(\mathbb{R}^d) = M^1(\mathbb{R}^d)$.

Proof. A direct computation shows

$$\begin{aligned}
||f||_{W(\mathcal{F}L^1,L^q)} &= \left(\int_{\mathbb{R}^d} ||f \cdot T_z g||_{\mathcal{F}L^1}^q \, dz\right)^{1/q} \\
&= \left(\int_{\mathbb{R}^d} ||\widehat{f \cdot T_z g}||_{L^1}^q \, dz\right)^{1/q} \\
&= \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |\widehat{f \cdot T_z g}(\omega)| \, d\omega\right)^q dz\right)^{1/q} \\
&= \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |V(f,\overline{g})(z,\omega)| \, d\omega\right)^q dz\right)^{1/q} \\
&= ||V(f,\overline{g})||_{L^{1,q}} \\
&= C \cdot ||f||_{M^{1,q}}
\end{aligned}$$

with $C > 0$ depending on g. Thus $f \in W(\mathcal{F}L^1, L^q)(\mathbb{R}^d)$ if and only if $||f||_{W(\mathcal{F}L^1,L^q)} < \infty$ if and only if $||f||_{M^{1,q}} < \infty$ if and only if $f \in M^{1,q}(\mathbb{R}^d)$, with equivalent norms. □

Finally, the following theorem gives a version of Hölder's Inequality for Wiener amalgam spaces.

Theorem A.6.5 (Hölder's Inequality for Amalgam Spaces)**.** *Let $f \in W(\mathcal{F}L^1, L^p)(\mathbb{R}^d)$ and $g \in W(\mathcal{F}L^1, L^q)(\mathbb{R}^d)$ with $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leq p, q \leq \infty$, conjugate exponents. Then $f \cdot g \in W(\mathcal{F}L^1, L^1)(\mathbb{R}^d) = M^1(\mathbb{R}^d)$ and*

$$||f \cdot g||_{W(\mathcal{F}L^1,L^1)} \leq C \cdot ||f||_{W(\mathcal{F}L^1,L^p)} \cdot ||g||_{W(\mathcal{F}L^1,L^q)}$$

(with some generic constant $C > 0$).

Proof. Choose a test function $\phi \in \mathcal{D}(\mathbb{R}^d)$. Then $\psi = \phi \cdot \phi \in \mathcal{D}(\mathbb{R}^d)$ is also a test function, and ϕ and ψ generate equivalent norms in all the spaces $W(\mathcal{F}L^1, L^p)(\mathbb{R}^d)$, $1 \leq p \leq \infty$, by Theorem A.6.2. Thus

$$\begin{aligned} ||f \cdot g||_{W(\mathcal{F}L^1,L^1)} &= \int_{\mathbb{R}^d} ||(f \cdot g) \cdot T_z\phi||_{\mathcal{F}L^1}\, dz \\ &\leq C \cdot \int_{\mathbb{R}^d} ||(f \cdot g) \cdot T_z\psi||_{\mathcal{F}L^1}\, dz \\ &= C \cdot \int_{\mathbb{R}^d} ||(f \cdot g) \cdot T_z(\phi \cdot \phi)||_{\mathcal{F}L^1}\, dz \\ &= C \cdot \int_{\mathbb{R}^d} ||(f \cdot T_z\phi) \cdot (g \cdot T_z\phi)||_{\mathcal{F}L^1}\, dz \end{aligned}$$

with a suitable constant $C > 0$ (that depends only on ϕ and ψ). Now the Fourier algebra $\mathcal{F}L^1$ is a Banach algebra, hence

$$||(f \cdot T_z\phi) \cdot (g \cdot T_z\phi)||_{\mathcal{F}L^1} \leq ||f \cdot T_z\phi||_{\mathcal{F}L^1} \cdot ||g \cdot T_z\phi||_{\mathcal{F}L^1},$$

so

$$\begin{aligned} &C \cdot \int_{\mathbb{R}^d} ||(f \cdot T_z\phi) \cdot (g \cdot T_z\phi)||_{\mathcal{F}L^1}\, dz \\ &\qquad \leq C \cdot \int_{\mathbb{R}^d} ||f \cdot T_z\phi||_{\mathcal{F}L^1} \cdot ||g \cdot T_z\phi||_{\mathcal{F}L^1}\, dz. \end{aligned}$$

Hölder's Inequality for L^p-spaces now yields

$$\begin{aligned} &C \cdot \int_{\mathbb{R}^d} ||f \cdot T_z\phi||_{\mathcal{F}L^1} \cdot ||g \cdot T_z\phi||_{\mathcal{F}L^1}\, dz \\ &\qquad \leq C \cdot \left(\int_{\mathbb{R}^d} ||f \cdot T_z\phi||^p_{\mathcal{F}L^1}\, dz \right)^{1/p} \cdot \left(\int_{\mathbb{R}^d} ||g \cdot T_z\phi||^q_{\mathcal{F}L^1}\, dz \right)^{1/q} \\ &\qquad \leq C \cdot ||f||_{W(\mathcal{F}L^1,L^p)} \cdot ||g||_{W(\mathcal{F}L^1,L^p)}. \end{aligned}$$

□

A.7 Trace Class and Hilbert-Schmidt Operators

In this section we collect (mainly without proofs) the most fundamental facts about compact operators on Hilbert space, spectral theory, and trace class and Hilbert-Schmidt operators. These can be found in any standard textbook on functional analysis, e.g. [5] or [35]. Unless otherwise stated, H always denotes a complex separable Hilbert space.

Definition A.7.1 (Compact Operators)**.** *A linear mapping $T : H \to H$ is called* ***compact*** *if T maps bounded subsets of H onto relatively compact sets, i.e. $\overline{T(B)} \subset H$ is compact for all bounded $B \subset H$.*

Since compact subsets of a Hilbert space are bounded, T maps bounded sets onto bounded sets and in particular the closed unit ball onto a bounded set. Therefore there is a $C > 0$ such that $||T(x)|| \leq C$ for all $x \in H$ with $||x|| \leq 1$; so T is automatically a bounded linear operator.

Elementary properties of compact operators are contained in

Theorem A.7.2. *Let S, T, T_n be compact (for $n \in \mathbb{N}$), $\lambda, \mu \in \mathbb{C}$, and $A : H \to H$ be bounded. Then the following holds:*

1. *The operator $\lambda S + \mu T$ is compact.*
2. *The composition of a compact and a bounded operator (in either order) is compact, i.e. both AT and TA are compact.*
3. *If $\lim_{n\to\infty} ||A - T_n||_{H\to H} = 0$, then A is compact: the set of compact operators is closed under limits with respect to the operator norm.*

The statements of the previous theorem may be stated in a concise form by saying: The compact operators form a closed two-sided ideal in the Banach algebra $\mathcal{B}(H \to H)$ of all bounded linear operators on H into itself.

Compact operators may be characterized in several ways:

Theorem A.7.3. *Let $T : H \to H$ be a bounded operator. The following are equivalent:*

1. *T is compact.*
2. *If $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence in H, then its image $(Tx_n)_{n\in\mathbb{N}}$ contains a convergent subsequence.*
3. *If $(x_n)_{n\in\mathbb{N}}$ is weakly convergent, then $(Tx_n)_{n\in\mathbb{N}}$ converges in norm.*

Theorem A.7.4 (Spectral Representation for Compact Self-Adjoint Operators)**.** *Let $T : H \to H$ be compact and self-adjoint. Then there exists a sequence $(\lambda_j)_{j\in\mathbb{N}}$ of real numbers and an orthonormal family $(\phi_j)_{j\in\mathbb{N}}$ of vectors in H such that*

- $\lim_{j\to\infty} \lambda_j = 0$;
- *$T\phi_j = \lambda_j\phi_j$ for all $j \in \mathbb{N}$, i.e. the λ_j are eigenvalues of T with associated eigenvectors ϕ_j;*
- *$Tf = \sum_{j=1}^{\infty} \lambda_j \langle f, \phi_j\rangle \phi_j$ for all $f \in H$, with convergence of the series in the norm of H.*

Definition A.7.5 (Singular Values, Hilbert-Schmidt Operator, Schatten p–Class)**.** *We define:*

1. *The **singular values** of a compact operator $T : H \to H$ are defined as the square-roots of the eigenvalues of the (compact self-adjoint positive) operator T^*T. We write*

$$s_j(T) := \lambda_j(T^*T)^{1/2},$$

where $\lambda_j(\cdot)$ denotes the sequence of eigenvalues of a compact self-adjoint operator in non-increasing order.

2. *A compact operator $T : H \to H$ is called a **Hilbert-Schmidt** operator, if $(s_j)_{j\in\mathbb{N}} \in \ell^2(\mathbb{N})$. The set of all Hilbert-Schmidt operators is denoted by $\mathcal{S}^2$.*
3. *A compact operator $T : H \to H$ belongs to the **Schatten p-class**, if $(s_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$. The set of all Schatten p-class operators is denoted by $\mathcal{S}^p$.*

Theorem A.7.6. *Let $T : H \to H$. The following are equivalent.*

1. *T is a Hilbert-Schmidt operator.*

2. *There exists some orthonormal basis* $(e_k)_{k\in\mathbb{N}}$ *of* H *such that* $\sum_{k\in\mathbb{N}} ||Te_k||^2 < \infty$.

3. $\sum_{k\in\mathbb{N}} ||Tf_k||^2 < \infty$ *holds for* ***all*** *orthonormal bases* $(f_k)_{k\in\mathbb{N}}$.

We have

$$\sum_{k\in\mathbb{N}} ||Te_k||^2 = ||T||^2_{\mathcal{S}^2}$$

independently of the chosen basis $(e_k)_{k\in\mathbb{N}}$.

Lemma A.7.7. *If* $T : H \to H$ *satisfies* $\sum_{k\in\mathbb{N}} ||Te_k||^2 = C < \infty$ *for some orthonormal basis* $(e_k)_{k\in\mathbb{N}}$, *then* T *is bounded. The estimate* $||T||_{H\to H} \leq \sqrt{C}$ *holds.*

Proof. Let $f = \sum_{k=1}^n c_k e_k$. Observe that $||f||^2 = \sum_{k=1}^n |c_k|^2$, since $(e_k)_{k\in\mathbb{N}}$ is orthonormal. Then

$$\begin{aligned} ||Tf||^2 = \langle Tf, Tf \rangle &= \sum_{k=1}^n \sum_{i=1}^n c_k \overline{c_i} \langle Te_k, Te_i \rangle \\ &\leq \sum_{k=1}^n \sum_{i=1}^n |c_k| \cdot |c_i| \cdot |\langle Te_k, Te_i \rangle| \\ &\leq \sum_{k=1}^n \sum_{i=1}^n |c_k| \cdot |c_i| \cdot ||Te_k|| \cdot ||Te_i|| \\ &= \left(\sum_{k=1}^n |c_k| \cdot ||Te_k|| \right)^2 . \end{aligned}$$

This last expression can be estimated further by the Chauchy-Schwarz Inequality:

$$\begin{aligned}\left(\sum_{k=1}^{n}|c_k|\cdot||Te_k||\right)^2 &\leq \left(\sum_{k=1}^{n}|c_k|^2\right)\cdot\left(\sum_{k=1}^{n}||Te_k||^2\right)\\ &\leq ||f||^2\cdot\left(\sum_{k=1}^{n}||Te_k||^2\right)\\ &\leq ||f||^2\cdot\left(\sum_{k=1}^{\infty}||Te_k||^2\right)\\ &= C\cdot||f||^2.\end{aligned}$$

So T is bounded on the dense subspace of finite linear combinations of elements of $(e_k)_{k\in\mathbb{N}}$ with operator norm $||T||_{op}\leq\sqrt{C}$. By the standard density argument, T is bounded on all of H, with the same operator norm. □

Lemma A.7.8. *If $T:H\to H$ satisfies $\sum_{k\in\mathbb{N}}||Te_k||^2<\infty$ for some orthonormal basis $(e_k)_{k\in\mathbb{N}}$, then T is compact.*

Proof. Define the operator T_n as

$$T_n(\sum_{k\in\mathbb{N}}c_ke_k):=\sum_{k=1}^{n}c_kTe_k,$$

i.e. $T_n=TP_n$, where P_n is the projection onto the finite-dimensional space spanned by the basis vectors $e_1,\dots,e_n$. Obviously T_n has finite rank at most n. Then

$$||(T-T_n)f||^2=||\sum_{k>n}c_kTe_k||^2.$$

Using the same estimate as in the previous lemma, we continue

$$\begin{aligned}||\sum_{k>n}c_kTe_k||^2 &\leq \left(\sum_{k>n}|c_k|^2\right)\cdot\left(\sum_{k>n}||Te_k||^2\right)\\ &\leq ||c||_2^2\cdot\left(\sum_{k>n}||Te_k||^2\right)\\ &\leq ||f||^2\cdot\left(\sum_{k>n}||Te_k||^2\right).\end{aligned}$$

This shows that

$$||T - T_n||_{H\to H} \leq \left(\sum_{k>n} ||Te_k||^2\right)^{1/2}.$$

But the last term tends to zero for $n \to \infty$, since the series $\sum_{k\in\mathbb{N}} ||Te_k||^2$ converges. This means that T can be approximated in the operator norm by operators of finite rank. So T is compact. □

The last two theorems in this section give useful criteria to decide whether a bounded operator on $L^2(\mathbb{R}^d)$ is trace class. Proofs can be found in [46].

Theorem A.7.9 (Criterion for Trace Class)**.** *Suppose T is a compact operator on H.*
Then T belongs to the trace class $\mathcal{S}^1(H)$ if and only if

$$\sum_{n\in\mathbb{N}} |\langle Te_n, e_n\rangle| < \infty$$

for every orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of H.
In this case we have

$$||T||_{\mathcal{S}^1} = \sup_{(e_n)_{n\in\mathbb{N}}\ ONB} \sum_{n\in\mathbb{N}} |\langle Te_n, e_n\rangle|.$$

□

Actually, in the preceding theorem one can dispense with the assumption that T be compact in the first place.

Theorem A.7.10 (Weidmann, [42])**.** *Let T be a bounded linear operator on H such that*

$$\sum_{n\in\mathbb{N}} \langle Te_n, e_n\rangle < \infty$$

for every orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of H.
Then T is a trace class operator. □

A.8 Tensor Products

Definition A.8.1 (Tensor Product)**.** *Let $f, g : \mathbb{R}^d \to \mathbb{C}$ be two functions. The* ***tensor product*** *of f and g is defined as the function on $\mathbb{R}^{2d}$*

$$f \otimes g(x, y) := f(x)g(y).$$

Lemma A.8.2. *The tensor product $f \otimes g$ of $f, g \in L^2(\mathbb{R}^d)$ is in $L^2(\mathbb{R}^{2d})$, and*

$$||f \otimes g||_{L^2(\mathbb{R}^{2d})} = ||f|| \cdot ||g||.$$

Proof. If $f, g \in L^2(\mathbb{R}^d)$, then

$$\begin{aligned} \iint_{\mathbb{R}^{2d}} |f \otimes g(x, y)|^2 \, dxdy &= \iint_{\mathbb{R}^{2d}} |f(x)|^2 \cdot |g(y)|^2 \, dxdy \\ &= \int_{\mathbb{R}^d} |f(x)|^2 \, dx \cdot \int_{\mathbb{R}^d} |g(y)|^2 \, dy \\ &= ||f||^2 \cdot ||g||^2 < \infty, \end{aligned}$$

so $f \otimes g \in L^2(\mathbb{R}^{2d})$ and $||f \otimes g||_{L^2(\mathbb{R}^{2d})} = ||f|| \cdot ||g||$. □

Lemma A.8.3. *Let $(ek)_{k\in\mathbb{N}}$ and $(f_l)_{l\in\mathbb{N}}$ be orthonormal bases for $L^2(\mathbb{R}^d)$. Then the family of tensor products $(e_k \otimes f_l)_{(k,l)\in\mathbb{N}\times\mathbb{N}}$ constitutes an orthonormal basis for $L^2(\mathbb{R}^{2d})$.*

Proof. Orthonormality can be shown by direct computation:

$$\begin{aligned} \langle e_k \otimes f_l, e_{k'} \otimes f_{l'} \rangle &= \iint_{\mathbb{R}^{2d}} e_k(x) f_l(y) \cdot \overline{e_{k'}(x) f_{l'}(y)} \, dxdy \\ &= \int_{\mathbb{R}^d} e_k(x) \overline{e_{k'}(x)} \, dx \cdot \int_{\mathbb{R}^d} f_l(y) \overline{f_{l'}(y)} \, dy \\ &= \langle e_k, e_{k'} \rangle \cdot \langle f_l, f_{l'} \rangle = \delta_{k,k'} \delta_{l,l'}. \end{aligned}$$

For completeness, assume that $H = H(x, y) \in L^2(\mathbb{R}^{2d})$ is arbitrary. For $y \in \mathbb{R}^d$, define the auxiliary function

$$H_y(x) := H(x, y).$$

We have

$$\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |H_y(x)|^2 \, dx \right) dy = \iint_{\mathbb{R}^{2d}} |H(x,y)|^2 \, dxdy = ||H||_{L^2(\mathbb{R}^{2d})} < \infty,$$

by Fubini's Theorem, so

$$||H_y||^2 = \int_{\mathbb{R}^d} |H_y(x)|^2 \, dx < \infty$$

and $H_y \in L^2(\mathbb{R}^d)$ for almost every $y \in \mathbb{R}^d$. The (almost everywhere defined) function

$$h : y \mapsto ||H_y|| = \left(\int_{\mathbb{R}^d} |H_y(x)|^2 \, dx \right)^{1/2}$$

is also in $L^2(\mathbb{R}^d)$ and

$$||h||^2 = \int_{\mathbb{R}^d} |h(y)|^2 \, dy = \int_{\mathbb{R}^d} ||H_y||^2 \, dy = \iint_{\mathbb{R}^{2d}} |H(x,y)|^2 \, dxdy = ||H||^2_{L^2(\mathbb{R}^{2d})}.$$

Now, for any $g \in L^2(\mathbb{R}^d)$, set

$$H_g(y) := \langle H_y, g \rangle \, .$$

This function is again defined for almost every $y \in \mathbb{R}^d$ (since $H_y \in L^2(\mathbb{R}^d)$ for almost every $y \in \mathbb{R}^d$), and is in $L^2(\mathbb{R}^d)$ by

$$\begin{aligned}
||H_g||^2 &= \int_{\mathbb{R}^d} | \langle H_y, g \rangle |^2 \, dy \\
&\leq \int_{\mathbb{R}^d} ||H_y||^2 \cdot ||g||^2 \, dy \\
&= ||g||^2 \cdot \int_{\mathbb{R}^d} ||H_y||^2 \, dy \\
&= ||g||^2 \cdot ||h||^2 \\
&= ||g||^2 \cdot ||H||^2_{L^2(\mathbb{R}^{2d})} < \infty,
\end{aligned}$$

with an application of the Cauchy-Schwarz Inequality.

Now we get

$$\begin{aligned}
\langle H, e_i \otimes f_j \rangle_{L^2(\mathbb{R}^{2d})} &= \iint_{\mathbb{R}^{2d}} H(x,y) \cdot \overline{e_i(x) f_j(y)}\, dxdy \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} H(x,y)\overline{e_i(x)}\, dx \right) \overline{f_j(y)}\, dy \\
&= \int_{\mathbb{R}^d} \langle H_y, e_i \rangle \cdot \overline{f_j(y)}\, dy \\
&= \langle H_{e_i}, f_j \rangle
\end{aligned}$$

for all $i, j \in \mathbb{N}$, therefore Parseval's Identity gives

$$\begin{aligned}
\sum_{(i,j)\in\mathbb{N}\times\mathbb{N}} |\, \langle H, e_i \otimes f_j \rangle_{L^2(\mathbb{R}^{2d})}\, |^2 &= \sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} |\, \langle H_{e_i}, f_j \rangle\, |^2 \\
&= \sum_{i\in\mathbb{N}} ||H_{e_i}||^2.
\end{aligned}$$

But

$$||H_{e_i}||^2 = \int_{\mathbb{R}^d} |H_{e_i}(y)|^2\, dy = \int_{\mathbb{R}^d} |\, \langle H_y, e_i \rangle\, |^2\, dy$$

yields

$$\begin{aligned}
\sum_{i\in\mathbb{N}} ||H_{e_i}||^2 &= \sum_{i\in\mathbb{N}} \int_{\mathbb{R}^d} |\, \langle H_y, e_i \rangle\, |^2\, dy \\
&= \int_{\mathbb{R}^d} \sum_{i\in\mathbb{N}} |\, \langle H_y, e_i \rangle\, |^2\, dy \\
&= \int_{\mathbb{R}^d} ||H_y||^2\, dy \\
&= ||h||^2 \\
&= ||H||^2_{L^2(\mathbb{R}^{2d})},
\end{aligned}$$

again by Parseval's Identity (the change of the order of integration and summation is justified by Fubini's Theorem). So we have

$$\sum_{i,j\in\mathbb{N}} |\, \langle H, e_i \otimes f_j \rangle_{L^2(\mathbb{R}^{2d})}\, |^2 = ||H||^2_{L^2(\mathbb{R}^{2d})}$$

for all $H \in L^2(\mathbb{R}^{2d})$, and this is equivalent to the completeness of the orthonormal system $(e_i \otimes f_j)_{(i,j)\in\mathbb{N}\times\mathbb{N}}$. □

Theorem A.8.4. *Let* $T := span\{f \otimes g : f, g \in L^2(\mathbb{R}^d)\} \subset L^2(\mathbb{R}^{2d})$ *be the linear subspace of* $L^2(\mathbb{R}^{2d})$ *spanned by all tensor products of functions in* $L^2(\mathbb{R}^d)$*. Then* T *is dense in* $L^2(\mathbb{R}^{2d})$*.*

Proof. This is an obvious corollary to the preceding lemma. Let $(e_i)_{i\in\mathbb{N}}$ and $(f_j)_{j\in\mathbb{N}}$ be orthonormal bases in $L^2(\mathbb{R}^d)$. If $H \in L^2(\mathbb{R}^{2d})$ such that $H \perp T$, then $H \perp (f \otimes g)$ for all $f, g \in L^2(\mathbb{R}^d)$, in particular $H \perp (e_i \otimes f_j)$ for all $i, j \in \mathbb{N}$. Since $(e_i \otimes f_j)_{(i,j)\in\mathbb{N}\times\mathbb{N}}$ is an orthonormal basis for $L^2(\mathbb{R}^{2d})$, this implies $H = 0$. Therefore $T^\perp = \overline{T}^\perp = \{0\}$ and $\overline{T} = L^2(\mathbb{R}^{2d})$. $\square$

Bibliography

[1] W.O. Amrein and A.M. Berthier. On Support Properties of L^p-Functions and Their Fourier Transforms. *J. Funct. Anal.*, 24(3):258–267, 1977.

[2] M. Benedicks. On Fourier Transforms of Functions Supported on Sets of Finite Lebesgue Measure. *J. Math. Anal. Appl.*, 106, 1985.

[3] F.A. Berezin. Wick and Anti-Wick Symbols of Operators. *Mat. Sb. (N.S.)*, 86(128):578–610, 1971.

[4] Paolo Boggiatto, Elena Cordero, and Karlheinz Gröchenig. Generalized Anti-Wick Operators with Symbols in Distributional Sobolev Spaces. *Integral Equations and Operator Theory*, 48(4):427–442, 2004.

[5] John B. Conway. *A Course in Functional Analysis.* Springer, New York, 2nd edition, 1990.

[6] Elena Cordero and Karlheinz Gröchenig. Time-Frequency Analysis of Localization Operators. *J. Funct. Anal.*, 205(1):107–131, 2003.

[7] Elena Cordero and Karlheinz Gröchenig. Necessary Conditions for Schatten Class Localization Operators. *Proc. Amer. Math. Soc.*, 133(12):3573–3579, 2005.

[8] A. Cordoba and C. Fefferman. Wave Packets and Fourier Integral Operators. *Comm. Partial Diff. Eq.*, 3(11):979–1005, 1978.

[9] I. Daubechies. Time-Frequency Localization Operators: a Geometric Phase Space Approach. *IEEE Trans. Inform. Theory*, 34(4):605–612, July 1988.

[10] Ingrid Daubechies and Thierry Paul. Time-Frequency Localisation Operators - a Geometric Phase Space Approach: II. The Use of Dilations. *Inverse Probl.*, 4(3):661–680, 1988.

[11] M. Engliš. Toeplitz Operators and the Berezin Transform on h^2. *Lin. Alg. Appl.*, 223/224, 1995.

[12] H.G. Feichtinger. Wiener amalgams over Euclidean spaces and some of their applications. 136:123–137, 1992.

[13] H.G. Feichtinger and K. Gröchenig. Gabor Frames and Time-Frequency Analysis of Distributions. *J. Funct. Anal.*, 146(2):464–495, 1997.

[14] H.G. Feichtinger and T. Strohmer, editors. *Gabor Analysis and Algorithms: Theory and Applications.* Appl. and Num. Harm. Anal. Birkhäuser, Boston, 1998.

[15] G.B. Folland. *Harmonic Analysis in Phase Space.* Princeton University Press, Princeton, N.J., 1989.

[16] G. Friedlander and M. Joshi. *Introduction to the Theory of Distributions.* Cambridge University Press, 2nd edition, 2008.

[17] Karlheinz Gröchenig. *Foundations of Time-Frequency Analysis.* Appl. Numer. Harmon. Anal. Birkhäuser Boston, Boston, MA, 2001.

[18] Karlheinz Gröchenig. Uncertainty Principles for Time-Frequency Representations. In H.G. Feichtinger and T. Strohmer, editors, *Advances in Gabor Analysis*, Appl. and Num. Harm. Anal., pages 11–30. Birkhäuser, Boston, 2003.

[19] Karlheinz Gröchenig. A Pedestrian's Approach to Pseudodifferential Operators. In Christopher Heil, editor, *Harmonic Analysis and Applications; In Honor of John J.Benedetto*, Appl. and Num. Harm. Anal., pages 139–169. Birkhäuser, Boston, 2006.

[20] Haakan Hedenmalm, Boris Korenblum, and Kehe Zhu. *Theory of Bergman Spaces.* Springer, 2000.

[21] Lars Hörmander. The Weyl calculus of pseudo-differential operators. *Comm. Pure Appl. Math.*, 32, 1979.

[22] Lars Hörmander. *The Analysis of Linear Partial Differential Operators I.* Classics in Mathematics. Springer, Berlin, Reprint of the second (1990) edition, 2003.

[23] Lars Hörmander. *The Analysis of Linear Partial Differential Operators III.* Classics in Mathematics. Springer, Berlin, Reprint of the 1994 edition, 2007.

[24] R. Howe. Quantum mechanics and partial differential equations. *J. Funct. Anal.*, 38(2):188–254, 1980.

[25] P. Jaming. Principe d'incertitude qualitatif et reconstruction de phase pour la transforme de Wigner. *C. R. Acad. Sci. Paris Sr. I Math*, 327:249–254, 1998.

[26] A. J. E. M. Janssen. Proof of a Conjecture on the Supports of Wigner Distributions. *J. Fourier Anal. Appl.*, 4(6):723–726, 1998.

[27] J.J. Kohn and L. Nirenberg. An Algebra of Pseudo-Differential Operators. *Comm. Pure Appl. Math.*, 18, 1965.

[28] D. Labate. *Time-Frequency Analysis of Pseudodifferential Operators.* PhD thesis, Georgia School of Technology, Atlanta, 2000.

[29] J.E. Moyal. Quantum Mechanics as a Statistical Theory. *Proceeding of the Cambridge Philosophical Society*, 45:99–124, 1949.

[30] J.C.T. Pool. Mathematical Aspects of the Weyl Correspondence. *J. Math. Phys.*, 7:66–76, 1966.

[31] J. Ramanathan and P. Topiwala. Time-Frequency Localization via the Weyl Correspondence. *SIAM J. Math. Anal.*, 24(5):1378–1393, 1993.

[32] Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics I: Functional Analysis.* Academic Press, San Diego, Revised and enlarged edition, 1980.

[33] Hans Reiter. *L^1-Algebras and Segal Algebras*, volume 231 of *Lecture Notes in Mathematics.* Springer, Berlin, Heidelberg, New York, 1971.

[34] R. Rochberg and K. Tachizawa. Pseudodifferential operators, Gabor frames, and local trigonometric bases. In H.G. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms*, Appl. and Num. Harm. Anal., pages 171–192. Birkhäuser, Boston, 1998.

[35] Walter Rudin. *Functional Analysis*. McGraw-Hill Book Company, New York, 1973.

[36] M.A. Shubin. *Pseudodifferential Operators and Spectral Theory*. Springer, Berlin, 2nd edition, 2001.

[37] Barry Simon. Distributions and Their Hermite Expansions. *J. Math. Phys.*, 12(1):140–148, 1971.

[38] Christoph Thiele. Classical Fourier Analysis. Unpublished lecture notes, 2000. *www.math.ucla.edu/ thiele/lecturenotes/fourieranalysis1.ps*.

[39] Joachim Toft. *Convolutions and Embeddings for Weighted Modulation Spaces*, volume 155 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel.

[40] Joachim Toft. Continuity Properties for Modulation Spaces, with Applications to Pseudo-Differential Operators – I. *Journal Funct. Anal.*, 207, 2004.

[41] J. Ville. Théorie et Applications de la Notion de Signal Analytique. *Cables et Transmission*, 2A, 1948.

[42] J. Weidmann. Ein Satz über nukleare Operatoren im Hilbertraum. *Mathematische Annalen*, 158:69–78, 1965.

[43] H. Weyl. *The Theory of Groups and Quantum Mechanics*. Dover, New York, Reprinted 1950 edition, 1931.

[44] E.P. Wigner. On the Quantum Correction for Thermo-Dynamic Equilibrium. *Phys. Rev. Letters*, 40:749–759, 1932.

[45] E. Wilczock. *Zur Funktionalanalysis der Wavelet- und Gabortransformation*. PhD thesis, Technische Universität, München, 1998.

[46] Kehe Zhu. *Operator Theory in Function Spaces*. Mathematical Surveys and Monographs 138. AMS, 2nd edition, 2007.

Curriculum Vitae

Name: Dominik Bayer

Adresse: Institut für Schallforschung
Österreichische Akademie der Wissenschaften
Wohllebengasse 12-14, A-1040 Wien, Austria

Email: bayerd@kfs.oeaw.ac.at

Telefon: 0043-1-51581-2521

Geburtsdatum: 11. Januar 1976

Geburtsort: Weiden i.d.Opf., Deutschland

Schulbildung:
Sept. 1982 - Aug. 1986: Gerhardinger Grundschule, Weiden
Sept. 1986 - Juni 1995: Augustinus-Gymnasium, Weiden

Erlangung der allgemeinen Hochschulreife 1995

Wehrdienst: Juli 1995 - April 1996

Studium: Studium der Mathematik mit Nebenfach Physik

Universität Würzburg, WS 1996/1997
Ludwig-Maximilians Universität München, SS 1997
Technische Universität München, WS 1997/1998 - SS 2003

Diplomarbeit: "Spektraltheorie von Jacobi-Operatoren",
Betreuer: Prof. D. Castrigiano

Abschluss: Dipl. Math., TU München 2003

Bisherige akademische Laufbahn:

- April 2004 - Sept. 2004: Wissenschaftlicher Mitarbeiter an der TU München, Lehrstuhl M7 (Globale Analysis), Prof. Friesecke
- Okt. 2004 - März 2006: Doktorand am Institut für Biomathematik und Biometrie (IBB), GSF Forschungszentrum für Umwelt und Gesundheit, Neuherberg bei München,
  Betreuer: Prof. K. Gröchenig,
  Titel der Dissertation: Bilinear Time-Frequency Distributions and Pseudodifferential Operators
- Okt. 2005 - Dez. 2005: Studienaufenthalt am Technion Institut, Haifa, Israel
- April 2006 - jetzt: Fortsetzung der Dissertation an der Universität Wien,
  Mitarbeit am Projekt "EUCETIFA"
- Feb. 2010 - Aug. 2010: Anstellung (PostDoc) am Institut für Schallforschung, Österreichische Akademie der Wissenschaften, Wien
  Mitarbeit am Project "MULAC"
- Sept. 2010 - Nov. 2010: Anstellung (PostDoc) am Laboratoire d'Analyse, Topologie et Probabilites (LATP), CMI, Université de Provence, Marseille, Frankreich

Bisherige Lehrtätigkeit:

- WS 1999/2000 - SS 2001: Tutor Analysis 1-4 für Mathematiker, Prof. Brokate, TU München
- WS 2001/2002 - SS 2003: Tutor Höhere Mathematik 1-4 für Maschinenwesen, Prof. Vachenauer, TU München

  Preis für die beste Tutorübung in der Fakultät Maschinenwesen, vergeben durch die Fachschaft Maschinenwesen

- WS 2003/2004: Tutor Höhere Mathematik 1 für Bauingenieur- und Vermessungswesen, Prof. Koch, TU München
- SS 2004: Tutor Höhere Mathematik 4 für Elektrotechnik, Dr. Barnerßoi, TU München
- WS 2004/2005: Tutor Höhere Mathematik 3 für Elektro- und Informationstechnik, Dr. Barnerßoi, TU München

  Übungsleiter Ergänzungen zur Mathematik für Physiker, Prof. Castrigiano, TU München
- SS 2005: Übungsleitung Ergänzungen zur Mathematik für Physiker, Prof. Henrichs, TU München

  Hauptseminar (mit Prof. Castrigiano und F. Hofmaier): "Maß und Kategorie"
- WS 2005/2006: Übungsleitung (mit J. Wehrstedt) Analysis 1 für Mathematiker, Prof. Friesecke, TU München

  Hauptseminar (mit Prof. Castrigiano und F. Hofmaier): "Lernen mit Kernen" (Anwendungen der Theorie der reproduzierenden Hilberträume auf die Lerntheorie)
- WS 2007/2008: Tutor Einführung in das mathematische Arbeiten, Prof. Schichl, Universität Wien
- SS 2008: Tutor Einführung in das mathematische Arbeiten, Dr. Kaiblinger, Universität Wien
- WS 2008/2009: Lehrbeauftragter für den Brückenkurs "Mathematik Warm-up" an der FH Technikum Wien

Teilnahme an Konferenzen/Workshops:

- Aug. 2004: Workshop "Quantenmechanische Vielteilchensysteme", Warwick University, Großbritannien
- Januar 2005: Workshop im Rahmen des ESI Sondersemesters "Zeit-Frequenz-Analyse", Erwin-Schrödinger-Institut, Wien, Österreich
- Mai 2005: Internationale Konferenz "Modern Methods in Time Frequency Analysis", Strobl, Österreich
- März 2006: Workshop "Pseudodifferential Operators and Microlocal Analysis", Universität Novi Sad, Serbien

- Sept. 2006: Internationale Konferenz ”HASSIP 06: Multiscale Methods, Sparse Decompositions and Parsimonious Statistics”, Neuherberg bei München, Deutschland
- Juni 2007: Internationale Konferenz ”Trends in Harmonic Analysis”, Strobl, Österreich
- Sept. 2007: Summer School ”New Trends and Directions in Harmonic Analysis, Approximation Theory, and Image Analysis”, Inzell, Deutschland
- Juli 2008: Internationale Konferenz ”Trends in Harmonic Analysis”, Strobl, Österreich
- März 2010: Internationale Konferenz ”DAGA 2010”, 36. Deutsche Jahrestagung für Akustik, Berlin, Deutschland
- April 2010: Workshop ”MULAC Mid-Term Meeting”, Marseille, Frankreich

Publikationen:

- Artikel (mit P. Balazs, F. Jaillet und P. Soendergaard) ”Zeros of the Short-Time Fourier Transform: A Numerical and Analytical Investigation”, to appear in ”EURASIP Journal on Advances in Signal Processing”, special issue on ”Recent Advances in Theory and Methods for Nonstationary Signal Analysis”, 2010

Abstract

Diese Dissertation verfolgt zwei Zielsetzungen.

Erstens wird versucht, wohlbekannte Zeit-Frequenz-Verteilungen wie etwa die Kurzzeit-Fouriertransformation oder die Wigner-Verteilung zu verallgemeinern und in einen einheitlichen Rahmen einzufügen. Insbesondere werden die zugehörigen Pseudodifferentialoperator-Kalküle und deren wesentliche Eigenschaften untersucht und mit bereits bestehenden Kalkülen wie etwa der Kohn-Nirenberg-Korrespondenz oder dem Weyl-Kalkül verglichen. Die Leitfrage besteht darin, ob sich die recht schönen Eigenschaften der erwähnten Kalküle auf die allgemeinere Situation übertragen lassen.

Zweitens wird, basierend auf den Ergebnissen des ersten Teils, ein spezieller Typus von Pseudodifferentialoperatoren, nämlich Zeit-Frequenz-Lokalsationsoperatoren, genauer analysiert. Ihre grundlegenden Eigenschaften, besonders Abbildungseigenschaften des Symbols, werden in einheitlichem Rahmen präsentiert. Der Zusammenhang mit der Berezin-Transformation erlaubt es, neue Dichtheitsresultate für die Menge der Lokalisationsoperatoren als Teilmengen größerer Operatorklassen sowohl bezüglich verschiedener Symbolklassen als auch verschiedener Topologien zu beweisen.

The purpose of this doctoral thesis is twofold.

First, an attempt is made to generalize well-known time-frequency distributions, such as the short-time Fourier transform or the Wigner distribution, and integrate them into a unified framework. In particular, the associated pseudodifferential calculi and their properties are investigated and compared

to already existing calculi, such as the Kohn-Nirenberg correspondence or the Weyl calculus. The guiding question is which of the rather nice properties of the mentioned calculi carry over to the more general situation.

Second, based on the first part, a specific type of pseudodifferential operators, namely the time-frequency localization operators, are analyzed more closely. Their basic properties, in particular mapping properties of the symbol, are reviewed in a unified way. The connection with the Berezin transform allows to prove new density results of the set of localization operators as subsets of larger classes of operators, for different symbol classes and with respect to different topologies.