universität wien

Diplomarbeit

Titel der Diplomarbeit

Banach-Gelfand Triples and Applications in Time-Frequency Analysis

Angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Verfasser:	Severin Bannert
Matrikelnummer:	0205038
Studienkennzahl lt. Studienblatt:	A405
Studienrichtung lt. Studienblatt:	Mathematik
Betreuer:	Prof. Dr. Hans Georg Feichtinger

Wien, December 2, 2010

Contents

List of Figures

1 Abstract

A recurrent theme in mathematics is extension. The rational numbers, for example, were extended to the real numbers which in turn were extended to the complex numbers when it became clear that the respective number sets were too small to solve particular problems in. In this work we highlight a related theme in the context of time-frequency analysis.

The natural domain of the Fourier transform, which is the fundamental operation in time-frequency analysis, is $\boldsymbol{L}^1$. This is somewhat limited but luckily one can extend its domain to $\boldsymbol{L}^2$ by taking limits, so we are able to work in a Hilbert space, which is a very nice place to be, amongst others because many concepts of linear algebra can be carried over to this setting. But Hilbert spaces too, are limited. As will be pointed out in the beginning of Section 3, it is easy to pose questions which lead outside the Hilbert space setting and call for distribution spaces.

In this work we will introduce *Banach-Gelfand triples*, a concept which encompasses the "niceness" of Hilbert spaces and the generality of Banach spaces of distributions. A Banach-Gelfand triple is a triple of spaces, consisting of a Hilbert space $\mathcal{H}$ which contains a smaller Banach space $\boldsymbol{B}$ and is itself contained in the dual space $\boldsymbol{B}'$. Our aim is to emphasize the role of Banach-Gelfand triples as a background to modern time-frequency analysis. Thus the subsequent section (Section 2) will give a short introduction to the key concepts of time-frequency analysis. The Fourier transform, translation and modulation operators, the short-time Fourier transform and Gaussian windows will be introduced and important properties will be proved. In Section 2.2 we will introduce *Gabor frames* which enable us to discretize the short-time Fourier transform.

Section 3 and the subsequent sections encompass the main part of this work. We will introduce the general concept of Banach-Gelfand triples and subsequently in Section 4 its specific incarnation $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')$, which plays a key role in time-frequency analysis as presented here, consequently its discussion will make up a major part of this paper. Section 5 is dedicated to various classes of operator representation. Different methods to identify operators with functions on the time-frequency plane are introduced there.

2 Introduction to Time-Frequency Analysis

2.1 Some Important Operations

Let f be a function from $\mathbb{R}^d$ to $\mathbb{C}$. We will interpret f as signal, be it a sound, an image or the vibration of a steel bridge etc. It is the goal of time-frequency analysis to describe f in terms of its behaviour in time and frequency *simultaneously*, i.e. we want an operator F, which takes a function f on $\mathbb{R}^d$ that is only dependent on time t and maps it to a function on the time-frequency plane $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$, that is, $F(f)$ is a function of time *and* frequency.

Before we can start an attempt to construct F, we will talk a little bit about the structure of the time frequency plane $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$. Therefore we have to introduce some group theory.

Definition 1. Let G be an arbitrary abelian group, then the *dual group* $\widehat{G}$ of G, also called the group of *characters*, is defined as the set

$$\widehat{G} := \{\nu \,|\, \nu(x+y) = \nu(x)\nu(y), |\nu(x)| = 1 \text{ where } x, y \in G\}.$$

$\widehat{G}$ is again a group under pointwise multiplication.

If we take $G = \mathbb{R}^d$, and interpret $\mathbb{R}^d$ as an additive group then the characters are the functions $\nu(x) = e^{2\pi i x\cdot\nu}$, the so-called *pure frequencies*. With this in mind we can define the most often used mathematical tool to map a function from the time domain $\mathbb{R}^d$ to the frequency domain $\widehat{\mathbb{R}}^d$: the Fourier transformation.

Definition 2. (The Fourier Transform) Let $f \in \boldsymbol{L}^1(\mathbb{R}^d)$, then the *Fourier Transform* $\mathcal{F}f$ is defined as

$$\mathcal{F}f(\omega) = \hat{f}(\omega) = \int_{\mathbb{R}^d} f(t)e^{-2\pi i t\cdot\omega}dt. \tag{2.1}$$

If we switch back to the more general character notation this reads $\hat{f}(\omega) = \int_{\mathbb{R}^d} f(t)\overline{\omega(t)}dt$. This formula looks suspiciously like an inner product $\langle f, \omega\rangle$, which suits the interpretation that the Fourier transform gives us the "energy" of f at frequency ω – but of course it is not. A priori the Fourier Transform is only defined for functions $f \in \boldsymbol{L}^1(\mathbb{R}^d)$, so we can't legally interpret it as inner product. But (2.1) can be extended to $\boldsymbol{L}^2(\mathbb{R}^d)$ using Plancherel's Theorem which states that the Fourier transform is energy preserving for functions $f \in \boldsymbol{L}^1 \cap \boldsymbol{L}^2(\mathbb{R}^d)$.

Theorem 1. *(Plancherel's Theorem) If $f \in \boldsymbol{L}^1 \cap \boldsymbol{L}^2(\mathbb{R}^d)$ then*

$$\|f\|_2 = \|\hat{f}\|_2 \tag{2.2}$$

Before we can prove this theorem we need some more theory. In particular we will have to introduce the inverse operator of $\mathcal{F}$ (for what its worth, an operator like $\mathcal{F}$ would not be very useful without its inverse $\mathcal{F}^{-1}$). This is done in Theorem

3. With Plancherel's Theorem in place we can reason that, since $\boldsymbol{L}^1 \cap \boldsymbol{L}^2(\mathbb{R}^d)$ is dense in $\boldsymbol{L}^2(\mathbb{R}^d)$, an arbitrary function $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ is the limit of some sequence $(f_n)_n \in \boldsymbol{L}^1 \cap \boldsymbol{L}^2(\mathbb{R}^d)$. Equation (2.2) implies that $(\hat{f}_n)_n$ is a Cauchy sequence in $\boldsymbol{L}^2(\mathbb{R}^d)$, hence it has a limit and we define $\hat{f} = \lim_{n\to\infty} \hat{f}_n$.

Another consequence of Plancherel's Theorem is *Parseval's formula* which states that:

Corollary 1. *If* $f, g \in \boldsymbol{L}^2(\mathbb{R}^d)$ *then*

$$\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle. \tag{2.3}$$

The drawback of the Fourier transformation is, that we loose all time information of f. We know the frequencies present in the signal and their respective energies, but we have no information on when a frequency occurs in time, which frequencies occur together and so on. The basic idea to overcome this problem is to divide the support of f into small sections and then apply the Fourier transform piecewise. To accomplish this we multiply f with some compactly supported (or at least rapidly decreasing) function g, the so-called *window function* (or just *window*), which we "slide" along the timeline to obtain a mapping from complex valued functions on $\mathbb{R}^d$ (the signals) to complex valued functions on the time-frequency plane $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ (e.g. spectrograms). To make this more precise we will introduce some fundamental operations.

Definition 3. (Translation and Modulation) Let T_x denote the *translation* (aka time shift) operator and M_ω the *modulation* (aka frequency shift) operator. Their respective definitions are

$$\begin{aligned} T_x f(t) &= f(t-x), \\ M_\omega f(t) &= e^{2\pi i \omega \cdot t} f(t), \quad x, \omega \in \mathbb{R}^d. \end{aligned}$$

One can easily calculate the following *commutation relations* for T_x and M_ω.

$$T_x M_\omega = e^{-2\pi i x \cdot \omega} M_\omega T_x \tag{2.4}$$

Thus T_x and M_ω commute if and only if $x \cdot \omega \in \mathbb{Z}$, otherwise there is always the phasefactor $e^{-2\pi i x \cdot \omega}$ which has to be taken into account. If we combine translation and modulation we call the resulting operator *time-frequency shift* and denote it by $\pi(\lambda)$:

$$\pi(\lambda) f = M_\omega T_x f, \quad \lambda = (x, \omega) \in \mathbb{R}^{2d}.$$

There is also a time frequency shift on the operator level.

Definition 4. Let K be an operator on $\boldsymbol{L}^2$, then we define the *time frequeny shift of operators* as

$$(\pi \otimes \pi^*)(\lambda) K = \pi(\lambda) K \pi^*(\lambda).$$

Definition 5. A *dilation* operator acts on a function f by stretching or squeezing its support. We define two kinds of dilation:

- *mass preserving* dilation: $St_\rho f(x) = \rho f(\rho x)$
- *value preserving* dilation: $D_\rho f(x) = f(\rho x)$

Now we can further tackle the problem of a joint time-frequency mapping. The apparent choice for a window g would be to simply take the characteristic function $\mathcal{X}_I$ over some interval I but due to its discontinuities at the edges of I it introduces oscillations under the Fourier transform (note that $\mathcal{F}\mathcal{X}_I$ is the *sinc*-function). So we will choose a smooth window and the most popular choice in literature is the *Gaussian.*

Definition 6. (The Gaussian) Let

$$\varphi_\alpha(x) = e^{-\pi x^2/\alpha}$$

denote the non-normalized *Gaussian function* with parameter $\alpha > 0$, where x is in $\mathbb{R}^d$ (see Figure 1).

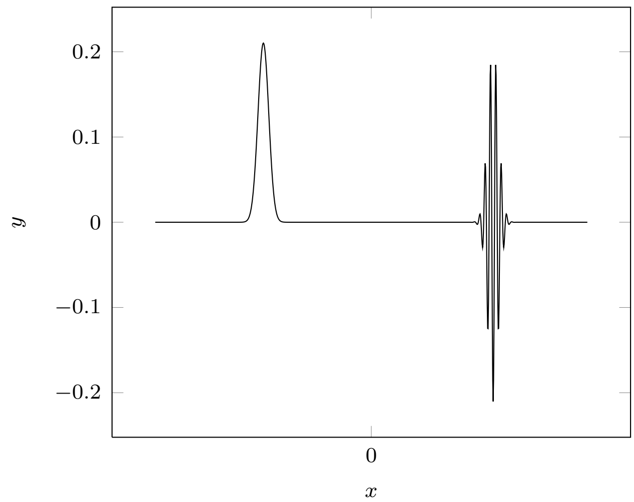


Figure 1: A Gaussian and its time-frequency shift.

As a small finger exercise we compute the $\boldsymbol{L}^1$-norm of the Gaussian.

Lemma 1. *(The $\boldsymbol{L}^1$-norm of φ_α)*

$$||\varphi_\alpha||_{\boldsymbol{L}^1(\mathbb{R}^d)} = (\sqrt{\alpha})^d$$

Proof. We first prove the lemma for $d = 1$.

$$\begin{aligned}
||\varphi_\alpha||^2_{\boldsymbol{L}^1(\mathbb{R})} &= \left(\int_{\mathbb{R}} |e^{-\frac{\pi}{\alpha}x^2}|dx\right)^2 \\
&= \int_{\mathbb{R}} |e^{-\frac{\pi}{\alpha}x^2}|dx \int_{\mathbb{R}} |e^{-\frac{\pi}{\alpha}x^2}|dx \\
&= \int_{\mathbb{R}} |e^{-\frac{\pi}{\alpha}x^2}|dx \int_{\mathbb{R}} |e^{-\frac{\pi}{\alpha}y^2}|dy \\
&= \int_{\mathbb{R}^2} |e^{-\frac{\pi}{\alpha}(x^2+y^2)}|dxdy \\
&= \int_0^{2\pi} \int_0^{\infty} re^{-\frac{\pi}{\alpha}r^2} drd\varphi \\
&= 2\pi \frac{(-\alpha)}{2\pi} e^{-\frac{\pi}{\alpha}r^2}|_{r=0}^{\infty} \\
&= \alpha
\end{aligned}$$

The result for $d > 1$ follows by induction. □

Of utter importance for the theory developed in the sequel of this work is the behaviour of the Gaussian under the Fourier transform. It turns out that it has the very convenient property that the Fourier transform of a Gaussian is again a Gaussian.

Lemma 2. *(The Fourier Transform of the Gaussian)*

$$\mathcal{F}\varphi_a(\omega) = a^{\frac{d}{2}} \varphi_{\frac{1}{a}}(\omega)$$

where a is positive.

Proof. We will give a proof which uses completion of the square as the central trick. For a very elegant alternate proof which uses differential equations instead, see [13], Lemma 1.5.1. W.l.o.g. let $d = 1$, then

$$\begin{aligned}
\widehat{\varphi_a}(\omega) &= \int_{\mathbb{R}} e^{-\frac{\pi}{a}x}\, e^{-2\pi i x\omega} dx \\
&= e^{-\pi a\omega^2} \int_{\mathbb{R}} e^{-(\sqrt{\frac{\pi}{a}}x + i\sqrt{a\pi}\omega)^2} dx \\
&= e^{-\pi a\omega^2} \int_{\mathbb{R}} e^{-\frac{\pi}{a}(x+ia\omega)^2} dx.
\end{aligned}$$

Now we set $y = x + ia\omega$ and get

$$\widehat{\varphi_a}(\omega) = e^{-\pi a\omega^2} \int_{\mathbb{R}} e^{-\frac{\pi}{a}y^2} dy = \sqrt{a}\, e^{-\pi a\omega^2} = \sqrt{a}\, \varphi_{1/a}(\omega),$$

which can easily be seen by a closer look at the proof of the preceding lemma. For $d > 1$ the result follows from the fact that the Fourier transform on $\mathbb{R}^d$ respects products, i.e.

$$\mathcal{F}\left(\prod_{n=1}^{d} f_n\right) = \prod_{n=1}^{d} \hat{f}_n,$$

and the d-dimensional Gaussian factors as $\varphi_a(x) = \prod_{n=1}^{d} e^{-\frac{\pi}{a}x_n^2}$. $\square$

Remark 1. We see that for $\alpha = 1$ the Gaussian is invariant under the Fourier transformation.

Next we will work towards introducing the inverse Fourier transform $\mathcal{F}^{-1}$. The following result is known as the *Fundamental Relation of the Fourier Transform* or alternatively the *Multiplication Formula*.

Theorem 2. *Let $f, g \in \boldsymbol{L}^1(\mathbb{R}^d)$, then*

$$\int_{\mathbb{R}^d} \hat{f}(x)g(x)dx = \int_{\mathbb{R}^d} f(x)\hat{g}(x)dx. \tag{2.5}$$

Proof. The proof is a straight forward calculation. With the help of Fubini's Theorem we get

$$\begin{aligned}\int_{\mathbb{R}^d} \hat{f}(x)g(x)dx &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f(t)e^{-2\pi itx}dt\right)g(x)dx \\ &= \int_{\mathbb{R}^d} f(t)\left(\int_{\mathbb{R}^d} g(x)e^{-2\pi itx}dx\right)dt \\ &= \int_{\mathbb{R}^d} f(t)\hat{g}(t)dt.\end{aligned}$$

$\square$

Now we have all the results in place to prove the inverse Fourier transform.

Theorem 3. *Let $f, \hat{f} \in \boldsymbol{L}^1(\mathbb{R}^d)$, then the* inverse Fourier transform $\mathcal{F}^{-1}$ *is given by*

$$f(x) = \mathcal{F}^{-1}\hat{f}(x) = \int_{\mathbb{R}^d} \hat{f}(\omega)e^{2\pi ix\omega}d\omega \qquad \forall x \in \mathbb{R}^d. \tag{2.6}$$

Proof. First, let $f \in \boldsymbol{C}_c(\mathbb{R}^d)$, the space of continuous functions with compact support. Let $\varphi(x) = e^{-\pi x^2}$ be the normalized Gaussian, then $\|St_\rho\varphi\|_1 = \|\varphi\|_1 = 1$. Set $\psi_\rho = D_{1/\rho}\varphi$, then $\widehat{\psi_\rho} = St_\rho\widehat{\varphi} = St_\rho\varphi$ by Lemma 2. By Theorem 2 it follows that

$$\begin{aligned}\int_{\mathbb{R}^d} \hat{f}(\omega)e^{2\pi ix\omega}\psi_\rho(\omega)d\omega &= \int_{\mathbb{R}^d} \hat{f}(\omega)M_\omega\psi_\rho(\omega)d\omega \\ &= \int_{\mathbb{R}^d} f(\omega)T_x\widehat{\psi_\rho}(\omega)d\omega\end{aligned} \tag{2.7}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^d} f(\omega) T_x St_\rho \varphi(\omega) d\omega \\
&= \int_{\mathbb{R}^d} f(\omega) St_\rho \varphi(\omega - x) d\omega \\
&= f * St_\rho \check{\varphi}(x),
\end{aligned}
$$

where $\check{\varphi}(x) = \varphi(-x)$. Next we have to show that $f * St_\rho \check{\varphi}(x) \to f(x)$ for $\rho \to \infty$. This results from the following considerations:

$$
\begin{aligned}
f * St_\rho \check{\varphi}(x) - f(x) &= \int_{\mathbb{R}^d} \left(f(x-t) - f(x)\right) St_\rho \check{\varphi}(t) dt \\
&= \int_{\mathbb{R}^d} \left(f(x - \rho^{-1} y) - f(y)\right) \check{\varphi}(y) dy.
\end{aligned}
$$

Since by assumption f in $\boldsymbol{C}_c(\mathbb{R}^d)$, the claim is proved and therefore $f * St_\rho \check{\varphi} \to f$ for $\rho \to \infty$. By Lebesgue's Theorem of Dominated Convergence the left-hand side of 2.7 converges to $\int_{\mathbb{R}^d} \hat{f}(\omega) e^{2\pi i x \omega} d\omega$ and thus 2.6 holds for all $f \in \boldsymbol{C}_c(\mathbb{R}^d)$ but since this is a dense subspace of $\boldsymbol{L}^1(\mathbb{R}^d)$ it holds on all of $\boldsymbol{L}^1(\mathbb{R}^d)$ and thus the proof is finished.

□

With $\mathcal{F}^{-1}$ at hand we can prove Plancherel's Theorem.

Proof of Theorem 1. The proof is a straight forward calculation.

$$
\begin{aligned}
\|\hat{f}\|_2 &= \int_{\mathbb{R}^d} |\hat{f}(\omega)|^2 d\omega \\
&= \int_{\mathbb{R}^d} \hat{f}(\omega) \overline{\hat{f}(\omega)} d\omega \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(x) e^{-2\pi i x \omega} dx \int_{\mathbb{R}^d} \overline{f(t)} e^{2\pi i t \omega} dt \right) d\omega \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{f(t)} e^{2\pi i (t-x)\omega} dx dt d\omega \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x) \overline{f(t)} \delta(t-x) dx dt \\
&= \int_{\mathbb{R}^d} f(x) \overline{f(x)} dx \\
&= \|f\|_2
\end{aligned}
$$

□

Now we are able to obtain a joint time-frequency representation of a signal f by "sliding" a window along the time axis and taking Fourier transforms. The following operator lies at the center of time-frequency analysis, it is the main building block for the subsequent theory.

Definition 7. The *Short Time Fourier Transform (STFT)* of a function $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ with respect to a window $g \in \boldsymbol{L}^2(\mathbb{R}^d)$ is defined as

$$\begin{aligned} V_g f(x,\omega) &= \int_{\mathbb{R}^d} f(t)\overline{g(t-x)}e^{-2\pi i t\cdot\omega}dt \\ &= \int_{\mathbb{R}^d} f(t)\overline{M_\omega T_x g(t)}dt \\ &= \langle f, M_\omega T_x g\rangle \end{aligned} \tag{2.8}$$

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

Again, the question arises immediately if there is an inverse operator and how to construct it. This, again, requires some more results before we can proceed.

Theorem 4. *(Orthogonality relations of the STFT)*
Let $f_1, f_2, g_1, g_2 \in \boldsymbol{L}^2(\mathbb{R}^d)$, then

$$\langle V_{g_1} f_1, V_{g_2} f_2\rangle = \langle f_1, f_2\rangle \overline{\langle g_1, g_2\rangle}. \tag{2.9}$$

Furthermore $V_{g_k} f_k \in \boldsymbol{L}^2(\mathbb{R}^{2d})$ for $k = 1, 2$.

Proof. We will follow the proof given in [13], p.42. At its heart lies a clever application of Parseval's formula.

$$\begin{aligned} \langle V_{g_1} f_1, V_{g_2} f_2\rangle &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V_{g_1} f_1(x,\omega)\overline{V_{g_2} f_2(x,\omega)}d\omega dx \\ &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} \mathcal{F}(f_1 \cdot T_x\bar{g}_1)(\omega)\overline{\mathcal{F}(f_2 \cdot T_x\bar{g}_2)(\omega)}d\omega\right) dx \\ &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f_1(t)\overline{f_2(t)}\,\overline{g_1(t-x)}g_2(t-x)dt\right) dx \\ &= \int_{\mathbb{R}^d} f_1(t)\overline{f_2(t)}\left(\int_{\mathbb{R}^d} \overline{g_1(t-x)}g_2(t-x)dx\right) dt \\ &= \langle f_1, f_2\rangle\overline{\langle g_1, g_2\rangle}. \end{aligned}$$

□

The orthogonality relations immediatly lead to the following corollary.

Corollary 2. *(Moyal's Formula) Let $f, g \in \boldsymbol{L}^2(\mathbb{R}^d)$, then*

$$\|V_g f\|_{\boldsymbol{L}^2(\mathbb{R}^d\times\widehat{\mathbb{R}}^d)} = \|g\|_{\boldsymbol{L}^2(\mathbb{R}^d)}\|f\|_{\boldsymbol{L}^2(\mathbb{R}^d)} \tag{2.10}$$

Remark 2. If in particular $\|g\|_{\boldsymbol{L}^2} = 1$ (for example if g is the normed Gaussian), then by (2.10) the STFT is an isometry, $V_g : \boldsymbol{L}^2(\mathbb{R}^d) \to \boldsymbol{L}^2(\mathbb{R}^{2d})$, and thus injective, i.e. each $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ is uniquely determined by its STFT.

Next we will find a way to reconstruct f from its STFT. With the previous results we have the right tools at hand to formulate and prove the inversion formula.

Theorem 5. *(The Inversion formula of the STFT)*
Let $g \in \boldsymbol{L}^2(\mathbb{R}^d)$ with $\|g\|_{\boldsymbol{L}^2} = 1$, then for $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ we have

$$f = \int_{\mathbb{R}^d \times \widehat{\mathbb{R}}^d} V_g f(x, \omega) M_\omega T_x g \, d\omega dx \tag{2.11}$$

Proof. Since $\|g\|_2 = 1$, (2.9) implies that

$$\langle V_g f_1, V_g f_2 \rangle = \langle f_1, f_2 \rangle \qquad \forall f_1, f_2 \in \boldsymbol{L}^2(\mathbb{R}^d),$$

which leads to

$$\langle V_g^* V_g f_1, f_2 \rangle = \langle f_1, f_2 \rangle, \tag{2.12}$$

which in turn implies $V_g^* V_g = \mathrm{id}$, where V_g^* denotes the adjoint operator of V_g. What we have to show now is that

$$V_g^* F = \int_{\mathbb{R}^d \times \mathbb{R}^d} F(x, \omega) M_\omega T_x g \, dx d\omega, \tag{2.13}$$

for $F \in \boldsymbol{L}^2(\mathbb{R}^{2d})$. This is done by the following computations:

$$\begin{aligned}
\langle V_g f, F \rangle &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} V_g f(x, \omega) \overline{F(x, \omega)} \, dx d\omega \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f(t) \overline{T_x g(t)} e^{-2\pi i t \omega} dt \right) dx d\omega \\
&= \int_{\mathbb{R}^d} f(t) \overline{\left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F(x, \omega) M_\omega T_x g(t) \, dx d\omega \right)} dt \\
&= \langle f, V_g^* F \rangle.
\end{aligned}$$

Hence, by setting $F = V_g f$ and with the help of (2.12) the proof is complete. □

Remark 3. We can omit the assumption that $\|g\|_{\boldsymbol{L}^2} = 1$ in the previous theorem. The inversion formula then reads

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

(2.11) can also be generalized to a situation where we use two different windows g and $\gamma \in \boldsymbol{L}^2$ for analysis and synthesis respectively. We only have to make sure that $\langle \gamma, g \rangle \neq 0$, then

$$f = \frac{1}{\langle \gamma, g \rangle} \int_{\mathbb{R}^d \times \mathbb{R}^d} V_g f(x, \omega) M_\omega T_x \gamma \, d\omega dx.$$

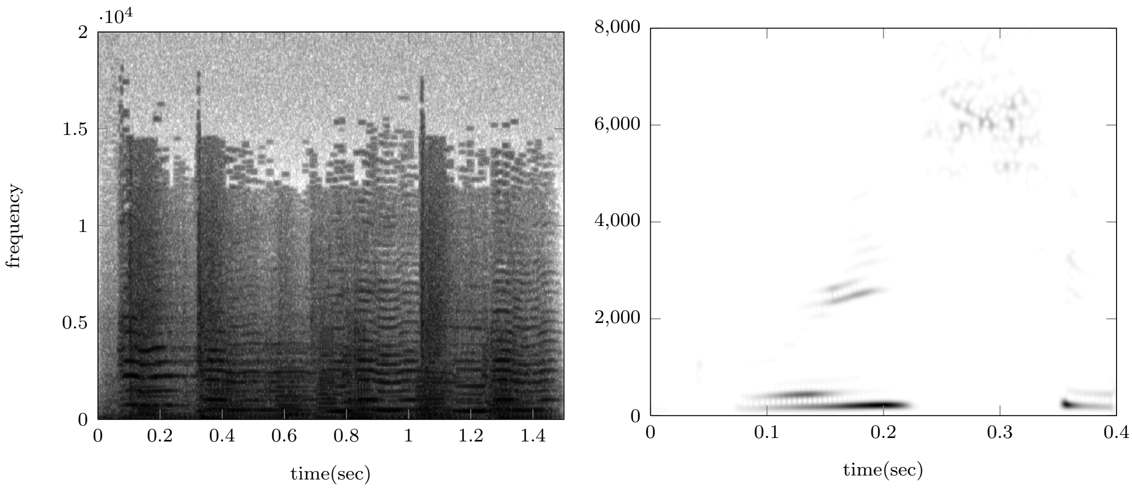


Figure 2: Two spectrograms. Left: spectrogram of a short piece of Klezmer music. One can clearly identify the drum beats (vertical lines) and a closer inspection reveals the melody (far right, two alternating notes). Right: spectrogram of the spoken word "greasy". Note the frequency pattern in the upper right: that is the sibilant sound of the "s" phoneme.

With the *STFT* we have a tool to determine the behaviour of a signal in the frequency domain over time. A typical usecase are spectrograms, which are just *STFTs* with the negative frequencies left out (see Figure 2).

One is tempted to use smaller time intervals to get more exact frequency localization, but there are uncertainty principles in place which prevent us from obtaining infinitely precise time frequency localization. Practically that means that better resolution in the time domain yields worse resolution in the frequency domain and vice versa (see Figure 3). For a more thorough discussion of uncertainty principles see [13]. As an example we state one particular uncertainty principle.

Proposition 1. *Let $U \subseteq \mathbb{R}^{2d}$ and $f, g \in \boldsymbol{L}^2$ with $\|f\|_2 = \|g\|_2 = 1$. Then for $\varepsilon \geq 0$*

$$\int\int_U |V_g f(x,\omega)|^2 \, dx d\omega \geq 1 - \varepsilon$$

implies that $|U| \geq 1 - \varepsilon$.

Proof. By the Cauchy-Schwartz inequality we have $|V_g f(x,\omega)| \leq \|f\|_2 \|g\|_2 = 1$. This yields

$$1 - \varepsilon \leq \int\int_U |V_g f(x,\omega)|^2 \, dx d\omega \leq \|V_g f\|_\infty^2 |U| \leq |U|.$$

□

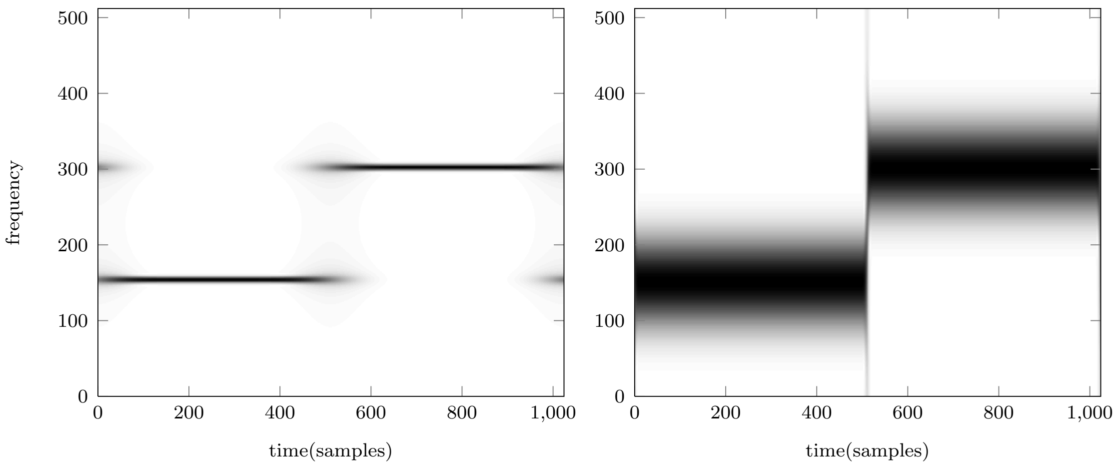


Figure 3: An extreme example of how the choice of the analyzing window influences the STFT. The signal is a simple sine wave at two different frequencies. On the left, a wide window was used, resulting in very good frequency localization, on the right, a very short window was used, which yields good time localization but very bad frequency localization.

2.2 Gabor Frames

In the previous section we got to know the *STFT* as our main tool that yields a joint time-frequency representation of a signal $f \in \boldsymbol{L}^2(\mathbb{R}^d)$. But a quick review of the underlying mechanism shows, that this representation is redundant. If we look closer at what happens in the time-frequency plane, we see that we let the window g move around continuously while taking inner products of the signal f with $T_x M_\omega g$. But since the essential supports of the time-frequency shifted versions $M_\omega T_x g$ of g overlap, we get highly redundant information, i.e. many of the coefficients $\langle f, M_\omega T_x g\rangle$ in the inversion formula of the *STFT* carry essentially the same information. We need a way to thin out this abundance. Discretization is the way to go here. We sample the *STFT* at equidistant points in the time frequency plane, then, by (2.8) we should be able to represent f by the following expansion:

$$f = \sum_{m \in \mathbb{Z}^d} \sum_{n \in \mathbb{Z}^d} \langle f, M_{nb} T_{ma} g\rangle M_{nb} T_{ma} \gamma \tag{2.14}$$

Where $a, b \in \mathbb{R}$ are constants defining the spacing of the sample points in the time-frequency plane and $g, \gamma \in \boldsymbol{L}^2(\mathbb{R}^d)$ are suitable windows for analysis and synthesis respectively. Of course there remain some questions. Under which conditions does the sum in the previous equation converge, and does it converge at all? How do we have to choose the windows g and γ?

Definition 8. A discrete subgroup $\Lambda \in \mathbb{R}^{2d}$ of the form $\Lambda = A\mathbb{Z}^{2d}$ for some $2d \times 2d$-matrix A with entries in $\mathbb{R}$ is called a *lattice* . We will only work with *time-frequency lattices*, which are lattices of the above form in the time-frequency plane $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$. A lattice of the form $\Lambda = a\mathbb{R}^d \times b\widehat{\mathbb{R}}^d$ with $a, b \in \mathbb{R}$ is called *separable lattice*.

Definition 9. A *frame* is a sequence $(g_j)_{j\in J}$ in some Hilbert space $\mathcal{H}$, where J is a countable index set, for which there exists constants $A, B > 0$ such that for all $f \in \mathcal{H}$

$$A\|f\|_{\mathcal{H}}^2 \le \sum_{j\in J} |\langle f, g_j\rangle|^2 \le B\|f\|_{\mathcal{H}}^2. \tag{2.15}$$

Note that the left-hand inequality ensures injectivity of the sampling operation $f \mapsto \sum_{j\in J} |\langle f, g_j\rangle|^2$ while the right-hand inequality takes care of its continuity. A frame is called a *tight frame* if $A = B$.

Definition 10. Associated with every frame are the following operators.

- The *analysis operator* $C : \mathcal{H} \to \boldsymbol{\ell}^2$ is given by

$$C : f \mapsto (\langle f, g_j\rangle)_{j\in J}. \tag{2.16}$$

- The *synthesis operator* $D : \boldsymbol{\ell}^2 \to \mathcal{H}$ is given by

$$D : c \mapsto \sum_{j\in J} c_j g_j. \tag{2.17}$$

- The *frame operator* $S : \mathcal{H} \to \mathcal{H}$ is given by

$$S : f \mapsto \sum_{j\in J} \langle f, g_j\rangle g_j, \tag{2.18}$$

  i.e. $S = DC$.

Remark 4. It is easy to see that

$$\begin{aligned}\langle Cf, c\rangle &= \sum_{j\in J} \langle f, g_j\rangle \overline{c_j} \\ &= \sum_{j\in J} \int f(t)\overline{c_j g_j(t)}dt \\ &= \int f(t)\overline{\sum_{j\in J} c_j g_j(t)}dt \\ &= \langle f, Dc\rangle,\end{aligned}$$

and thus D and C are adjoint to each other. Furthermore the frame operator $S = C^*C = DD^*$ is self-adjoint.

Lemma 3. *Let $\{g_j : j \in J\}$ be a frame with frame bounds $A, B > 0$, then the set $\{S^{-1}g_j : j \in J\}$ is also a frame, the so called* dual frame *with frame bounds A^{-1}, B^{-1}.*

Proof. For the proof we refer the reader to [13], Corollary 5.1.3. □

Lemma 4. *Let $\{g_j : j \in J\}$ be a frame and $\{\gamma_j = S^{-1}g_j : j \in J\}$ its dual frame, then every $f \in \mathcal{H}$ can be written as non-orthogonal expansions of the form*

$$f = \sum_{j \in J} \langle f, g_j \rangle \gamma_j \tag{2.19}$$

and

$$f = \sum_{j \in J} \langle f, \gamma_j \rangle g_j. \tag{2.20}$$

Both sums converge unconditionally in $\mathcal{H}$.

Proof. We only sketch the proof which basically consists of a straight forward calculation. We only show (2.19):

$$f = S^{-1}(Sf) = \sum_{j \in J} \langle f, g_j \rangle S^{-1} g_j = \sum_{j \in J} \langle f, g_j \rangle \gamma_j.$$

□

In a frame expansion of the form (2.19) the coefficients are in general not unique. Since frames are redundant we can drop certain coefficients and still be able to reconstruct f from the remaining ones. This property is very useful for example in wireless communication because it enables us to reconstruct a signal from incomplete transmissions. If we want unique coefficients, we loose the redundancy as shown in the following theorem which introduces *Riesz bases.*

Theorem 6. *(Riesz basis)*
Let $(g_i)_{i \in I}$ be a frame in some Hilbert space $\mathcal{H}$ then $(g_i)_{i \in I}$ is called a Riesz basis *if one (and therefore all) of the following equivalent conditions hold:*

1. *The coefficients $\langle f, \gamma_i \rangle$ are unique.*
2. *$C : \boldsymbol{L}^2 \to \boldsymbol{\ell}^2$ is surjective.*
3. *There exists constants $A, B > 0$ such that for arbitrary sequences $c \in \ell(I)$*

$$A\|c\|_2 \leq \| \sum_{i \in I} c_i g_i \|_{\mathcal{H}} \leq B\|c\|_2. \tag{2.21}$$

4. *$(g_i)_{i \in I}$ is the image of an orthonormal basis $(e_i)_{i \in I}$ under an invertible operator $T \in \mathcal{B}(\mathcal{H})$.*
5. *The Gram matrix $(G_{nk})_{nk}$, where $G_{nk} = \langle g_n, g_k \rangle$ defines a positive, invertible operator on $\boldsymbol{\ell}^2$.*

Proof. The equivalencies are shown in [13], p. 90. □

For a thorough discussion of frames and Riesz bases see [1] and [13]. Now we will apply the machinery of frames to our problem at hand: discretizing the *STFT* in order to reduce the redundancy and thus achieving a more efficient joint time-frequency representation of a given signal f. The idea is simple: move a window $g \in \boldsymbol{L}^2(\mathbb{R}^d)$ across the time-frequency plane and take inner products at equidistant points. The following definition makes this precise.

Definition 11. A *Gabor system* or *Gabor family* is a set of functions which is generated by shifting some window function $g \in \boldsymbol{L}^2(\mathbb{R}^d)$ along a time-frequency lattice $\Lambda = a\mathbb{Z}^d \times b\mathbb{Z}^d \subset \mathbb{R}^d \times \widehat{\mathbb{R}}^d$. Formally, we write

$$G(g, \Lambda) = \{\pi(\lambda)g, \lambda \in \Lambda\}.$$

The generating function g is called *Gabor atom*. If $G(g, \Lambda)$ is a frame, then it is called *Gabor frame*. Consequently the *Gabor frame operator* is given as

$$Sf = \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)g \rangle \pi(\lambda)g. \tag{2.22}$$

What we are really interested in is a frame expansion of the form (2.14), i.e. we need a dual window. With Gabor frames this becomes pleasantly easy.

Proposition 2. *The Gabor frame operator commutes with time-frequency shifts, i.e.*

$$(\pi(\lambda))^{-1} S \pi(\lambda) f = Sf. \tag{2.23}$$

Thus the dual frame is again a Gabor frame, generated by the dual atom $\gamma = S^{-1}g$.

Proof. The proof is a straight forward calculation. Let Λ be a lattice in $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$, then

$$\begin{aligned}
(\pi(\lambda)))^{-1} S \pi(\lambda) f &= \sum_{\lambda' \in \Lambda} \langle \pi(\lambda) f, \pi(\lambda')g \rangle (\pi(\lambda))^{-1} \pi(\lambda')g \\
&= \sum_{\lambda' \in \Lambda} \langle f, \pi(\lambda' - \lambda)g \rangle \pi(\lambda' - \lambda)g \\
&= Sf
\end{aligned}$$

□

From this result we can finally derive our Gabor expansion.

Corollary 3. *Let f be a function in $\boldsymbol{L}^2(\mathbb{R}^d)$ and $G(g, a, b)$ a Gabor frame, then f has the expansions*

$$\begin{aligned}
f &= \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)g \rangle \pi(\lambda)\gamma \\
&= \sum_{\lambda \in \Lambda} \langle f, \pi(\lambda)\gamma \rangle \pi(\lambda)g.
\end{aligned} \tag{2.24}$$

Remark 5. We succeeded in "thinning out" the information needed to reconstruct a function $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ from its *STFT* by introducing Gabor frames and the Gabor expansion. But we have been somehow vague about when this expansion is really possible. Up to now we use *suitable* windows in $\boldsymbol{L}^2(\mathbb{R}^d)$ but which ones are they?

3 Banach-Gelfand Triples

The results from the preceding section give rise to some questions. Until now we have defined the STFT, our main tool in time frequency analysis, only for functions in $\boldsymbol{L}^2(\mathbb{R}^d)$ with *suitable* windows $g \in \boldsymbol{L}^2(\mathbb{R}^d)$. In this chapter we will elaborate on what this means.

A particular problem arises when we ask for the eigenvectors of the unitary operators T_x or M_ω. From linear algebra we know the fact that every self-adjoint operator on a finite dimensional euclidian vector space has a complete set of eigenvectors. In the infinite dimensional case this isn't necessarily true anymore. Assume that $T_x f(t) = cf(t)$ for $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ and $c \in \mathbb{R}$, then an application of the Fourier transform yields $M_{-x}\hat{f}(\omega) = c\hat{f}(\omega)$ almost everywhere, but this holds true only if $\hat{f} = 0$ almost everywhere except for the set of all x such that $e^{-2\pi i x\omega} = c$, a set which has measure zero, thus $\hat{f}$ must be 0 almost everywhere and consequently $f \equiv 0$. Hence the operator T_x has no eigenvectors in $\boldsymbol{L}^2(\mathbb{R}^d)$. But if we apply T_x to a pure frequency $\chi_\omega(t) = e^{2\pi i t\omega}$ we find that

$$T_x \chi_\omega(t) = e^{-2\pi i x\omega} e^{2\pi i t\omega}$$

and thus we can consider the pure frequencies χ_ω as eigenfunctions of the operator T_x to the eigenvalue $e^{-2\pi i x\omega}$. But χ_ω does not belong to $\boldsymbol{L}^2(\mathbb{R}^d)$. The solution to this problem will be the interpretation of the pure frequencies as functionals on a suitable (Banach-) space of test functions as we will see in Section 4. This idea evolves further into the concept of *generalized eigenvectors* as introduced in [12] and [2].

Another problem arises if we want to generalize the STFT. In Equation (2.8) we saw that the STFT can be considered as an inner product, which only makes sense in $\boldsymbol{L}^2$, but what if we want to apply the STFT to more general objects like functions from $\boldsymbol{L}^\infty$ or distributions. Here the concept of duality comes to rescue. If we could find a subspace $U \subset \boldsymbol{L}^2(\mathbb{R}^d)$, we could consider its dual U' and extend the definition of the STFT in analogy to the Hilbert space case as

$$V_g f(x, \omega) = \langle f, M_\omega T_x g \rangle$$

where $f \in U'$ and $g \in U$. It turns out that this is the correct path to take and we will even find a suitable Banach space $U = \boldsymbol{S}_0$, leading to a situation where we have a Banach space embedded into a Hilbert space which in turn is embedded in the dual of the small space. This is one possible motivation for introducing *Banach-Gelfand triple*.

Remark 6. The following definitions assume a basic familiarity with the concept of *weak*-convergence* and the *weak*-topology* on the dual space B' of a Banach space B. A deeper discussion of this concepts can be found in Section 4.6.

Definition 12. Let $\boldsymbol{B}_1, \boldsymbol{B}_2$ be Banach spaces. An operator $U : \boldsymbol{B}_1' \to \boldsymbol{B}_2'$ is called *weak*-weak* continuous* or *w*-w* continuous* if it maps bounded weak*-convergent sequences in $\boldsymbol{B}_1'$ to bounded weak*-convergent sequences in $\boldsymbol{B}_2'$, i.e.

$$\langle \sigma_n - \sigma, f \rangle_{\boldsymbol{B}_1} \to 0 \Rightarrow \langle U\sigma_n - U\sigma, g \rangle_{\boldsymbol{B}_2} \to 0 \quad \forall f \in \boldsymbol{B}_1, g \in \boldsymbol{B}_2, \sigma_n, \sigma \in \boldsymbol{B}_1'.$$

Now we can give the central definition in this work.

Definition 13. (Banach-Gelfand-Triple) Let $\mathcal{H}$ be a Hilbert space and let $\boldsymbol{B}$ be a Banach space which is continuously and densely embedded into $\mathcal{H}$. If $\mathcal{H}$ in turn is weak*-continuously and densely embedded into $\boldsymbol{B}'$, the dual space of $\boldsymbol{B}$, we call the triple $(\boldsymbol{B}, \mathcal{H}, \boldsymbol{B}')$ a *Banach-Gelfand-Triple.*

Lemma 5. $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)$ *is a Banach-Gelfand-Triple.*

Proof. Let c_{00} be the space of finite sequences

$$c_{00} := \{(a_n)_{n \geq 0} | \ \exists k : a_n = 0 \ \ \forall n > k\}$$

First we will prove that c_{00} is dense in $\boldsymbol{\ell}^1$ and $\boldsymbol{\ell}^2$. Let $f \in \boldsymbol{\ell}^1$ and $(c_n)_n \in c_{00}$ such that

$$c_{n_k} = \begin{cases} f_k & k \leq n \\ 0 & k > n \end{cases}$$

Then we have

$$||f - c_n|| = \sum_{k>n} |f_k| \to 0 \quad \text{for } n \to \infty$$

The same argument holds for $\boldsymbol{\ell}^2$ and it follows that c_{00} is dense in $\boldsymbol{\ell}^1 \cap \boldsymbol{\ell}^2$. Now for $f \in \boldsymbol{\ell}^2$ there exists a sequence $(g_n)_n \in \boldsymbol{\ell}^1$ such that

$$||f - g_n||_2 = ||f - c_n + c_n - g_n||_2 \leq ||f - c_n||_2 + ||c_n - g_n||_2 \to 0$$

since c_{00} is dense in the intersection of $\boldsymbol{\ell}^1$ and $\boldsymbol{\ell}^2$ and therefore $\boldsymbol{\ell}^1$ is dense in $\boldsymbol{\ell}^2$. Now we show that $\boldsymbol{\ell}^2$ is weak*-dense in $\boldsymbol{\ell}^\infty$. Weak*-convergence in $\boldsymbol{\ell}^\infty$ amounts to coordinate-wise convergence since

$$|\langle g_n, f \rangle - \langle g, f \rangle| = \left| \sum_{k \geq 0} g_{n_k} f_k - \sum_{k \geq 0} g_k f_k \right| \to 0 \iff g_{n_k} \to g_k \text{ for } n \to \infty$$

where $(g_n)_n$ is a sequence in $\boldsymbol{\ell}^\infty$, $g \in \boldsymbol{\ell}^\infty$ and $f \in \boldsymbol{\ell}^1$. Now let $(g_n)_n \in \boldsymbol{\ell}^\infty$ be defined as

$$g_{n_k} = \begin{cases} g_k & k \leq n \\ 0 & k > n \end{cases}$$

then for $f \in \boldsymbol{\ell}^1 \cap \boldsymbol{\ell}^2$

$$|\sum_{k\geq 0} g_{n_k} f_k - \sum_{k\geq 0} g_k f_k| = |\sum_{k>n} g_k f_k| \leq \max_{k>n}(g_k) \sum_{k>n} |f_k| \to 0 \text{ for } n \to \infty.$$

Since $\boldsymbol{\ell}^1 \cap \boldsymbol{\ell}^2$ is dense in $\boldsymbol{\ell}^2$, this holds for all $f \in \boldsymbol{\ell}^2$. □

Definition 14. (Banach-Gelfand triple Homomorphism) Let $(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}'_1)$ and $(\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}'_2)$ be two Banach-Gelfand triples, then a *Banach-Gelfand triple homomorphism* is a linear mapping T which satisfies the following conditions:

- There exists a constant $C_{\boldsymbol{B}} \in \mathbb{R}$ such that the operatornorm $\|T\|_{\boldsymbol{B}_1 \to \boldsymbol{B}_2} \leq C_{\boldsymbol{B}}$.
- There exists a constant $C_{\mathcal{H}} \in \mathbb{R}$ such that $\|T\|_{\mathcal{H}_1 \to \mathcal{H}_2} \leq C_{\mathcal{H}}$.
- There exists a constant $C_{\boldsymbol{B}'} \in \mathbb{R}$ such that $\|T\|_{\boldsymbol{B}'_1 \to \boldsymbol{B}'_2} \leq C_{\boldsymbol{B}'}$.
- T is a w*-w*-continuous mapping from $\boldsymbol{B}'_1$ to $\boldsymbol{B}'_2$, i.e. if a net $(\sigma_\alpha)_\alpha \in \boldsymbol{B}'_1$ satisfies
$$\langle \sigma_\alpha, g \rangle \to \langle \sigma, g \rangle \quad \forall g \in \boldsymbol{B}_1,$$
then
$$\langle T\sigma_\alpha, h \rangle \to \langle T\sigma, h \rangle \quad \forall h \in \boldsymbol{B}_2.$$

Remark 7. Of course, by taking $C = \max\{C_{\boldsymbol{B}}, C_{\mathcal{H}}, C_{\boldsymbol{B}'}\}$, we can abbreviate the above definition by saying that a mapping T between $(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}'_1)$ and $(\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}'_2)$ is a BGT-homomorphism if it is bounded on all three "layers" of the Gelfand triple, i.e.

$$\|T\|_{(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}'_1) \to (\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}'_2)} \leq C$$

and T is w*-w*-continuous from $\boldsymbol{B}'_1$ to $\boldsymbol{B}'_2$.

Definition 15. (Banach-Gelfand triple Isomorphism) Let $(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}'_1)$ and $(\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}'_2)$ be two Banach-Gelfand triples. A linear mapping U is called a (unitary) Banach-Gelfand triple isomorphism if

- U is an isomorphism between $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$.
- U is an isomorphism (resp. U is an unitary operator) between $\mathcal{H}_1$ and $\mathcal{H}_2$.
- U extends to a weak*-isomorphism and a norm-to-norm continuous isomorphism between $\boldsymbol{B}'_1$ and $\boldsymbol{B}'_2$.

Definition 16. For an operator U we can express the fact that U is a unitary Banach-Gelfand triple isomorphism with the following *Gelfand bracket* notation

$$\langle f_1, f_2 \rangle_{(\boldsymbol{B},\mathcal{H},\boldsymbol{B}')} = \langle Uf_1, Uf_2 \rangle_{(\boldsymbol{B},\mathcal{H},\boldsymbol{B}')}. \tag{3.1}$$

This notation extends the inner product notation for Hilbert spaces to the functional brackets of Banach spaces.

The following theorem is of utter importance and immediately shows how useful the concept of Banach-Gelfand triples is.

Theorem 7. *(Extension of Operators)*
Let U be a unitary mapping $U : \mathcal{H}_1 \to \mathcal{H}_2$. Then U extends to a Banach-Gelfand triple isomorphism from $(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}_2')$ to $(\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}_2')$ if and only if U, as well as its adjoint U', restricted to $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ respectively, are bounded linear operators. In other words if and only if there exists a constant $C > 0$ such that

$$\|Uf\|_{\boldsymbol{B}_2} \leq C\|f\|_{\boldsymbol{B}_1} \quad \forall f \in \boldsymbol{B}_1 \tag{3.2}$$

as well as

$$\|U'g\|_{\boldsymbol{B}_1} \leq C\|g\|_{\boldsymbol{B}_2} \quad \forall g \in \boldsymbol{B}_2 \tag{3.3}$$

Proof. One direction of the proof is clear. If U extends to a Banach-Gelfand triple isomorphism, then U and U' are bounded on the innermost level of the Gelfand triple. On the other hand, if (3.3) holds, then we can define a mapping $\overline{U}$ by

$$\langle \overline{U}g, f \rangle = \langle g, U'f \rangle$$

where $g \in \boldsymbol{B}_1'$ and $f \in \boldsymbol{B}_2$. This mapping is bounded on $\boldsymbol{B}_1'$ because

$$\|\overline{U}g\|_{\boldsymbol{B}_2'} = \max_{\|f\|_{\boldsymbol{B}_2} \leq 1} |\langle \overline{U}g, f \rangle| = \max_{\|f\|_{\boldsymbol{B}_2} \leq 1} |\langle g, U'f \rangle| < C\|g\|_{\boldsymbol{B}_1'}.$$

It thus extends the unitary mapping $U : \mathcal{H}_1 \to \mathcal{H}_2$, and since $\mathcal{H}_1$ is weak*-dense in $\boldsymbol{B}_1'$, this extension is unique as a function mapping weak*-convergent sequences in $\boldsymbol{B}_1'$ on weak*-convergent sequences in $\boldsymbol{B}_2'$. Similarly, we define

$$\langle \overline{U}^{-1} g, f \rangle = \langle g, Uf \rangle, \quad g \in \boldsymbol{B}_2', f \in \boldsymbol{B}_1.$$

This is again a bounded operator, thus $\overline{U}$ is an isomorphism from $\boldsymbol{B}_1'$ to $\boldsymbol{B}_2'$. Note that the bijectivity of U restricted to $\boldsymbol{B}_1$ follows from (3.2) and (3.3) and the bijectivity of $U : \mathcal{H}_1 \to \mathcal{H}_2$ (in particular $U : \boldsymbol{B}_1 \to \boldsymbol{B}_2$ has a right-inverse and is thus surjective, the injectivity is trivial).

□

As a corollary we get immediately

Corollary 4. *Let U be an isomorphism from $\boldsymbol{B}_1$ to $\boldsymbol{B}_2$, then U extends to an isomorphism on $(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}_1')$ to $(\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}_2')$ if and only if*

$$\langle f, g\rangle_{\mathcal{H}_1} = \langle Uf, Ug\rangle_{\mathcal{H}_2} \quad \forall f, g \in \boldsymbol{B}_1 \tag{3.4}$$

Proof. If U is an BGT isomorphism between $(\boldsymbol{B}_1, \mathcal{H}_1, \boldsymbol{B}_1')$ and $(\boldsymbol{B}_2, \mathcal{H}_2, \boldsymbol{B}_2')$, then it is clear that (3.4) holds. On the other hand it follows from (3.4) that

$$\|f\|^2_{\mathcal{H}_1} = \langle f, f\rangle_{\mathcal{H}_1} = \langle Uf, Uf\rangle_{\mathcal{H}_2} = \|Uf\|^2_{\mathcal{H}_2} \qquad \forall f \in \boldsymbol{B}_1,$$

and thus U extends to an isometry from $\mathcal{H}_1$ to $\mathcal{H}_2$ with dense range $U(\boldsymbol{B}_1) = \boldsymbol{B}_2$. Thus $U : \mathcal{H}_1 \to \mathcal{H}_2$ is a unitary mapping. This implies that the restriction of $U' = U^{-1}$ to $\boldsymbol{B}$ is bounded, thus Theorem 7 applies. □

Remark 8. Corollary 4 gives us a straight forward tool to determine if a given operator is a Banach-Gelfand triple isomorphism. All we have to do is to check if Equation (3.4) holds.

4 The Gelfand Triple (S_0, L^2, S_0')

In this section we will introduce the key players of this survey: the Banach space $\boldsymbol{S}_0$ and its topological dual $\boldsymbol{S}_0'$. Together with $\boldsymbol{L}^2$ they form the most prominent and important Banach-Gelfand triple in modern time-frequency analysis. Extensive further information on the topics in this section can be found in [11], [4], [13] and [5].

First we define two very general classes of spaces which provide the background for the results in the sequel of this work.

Definition 17. (Mixed-norm spaces)
Let $1 \le p, q < \infty$, then the *mixed norm space* $L^{p,q}(\mathbb{R}^{2d})$ is defined as the space of all measureable functions on $\mathbb{R}^{2d}$ such that

$$\|f\|_{L^{p,q}(\mathbb{R}^{2d})} = \left(\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|f(x,\omega)|^p dx\right)^{q/p} d\omega\right)^{1/q} < \infty.$$

with the usual modification if $p = \infty$ or $q = \infty$.

Remark 9. Mixed-norm spaces are a generalization of the usual $\boldsymbol{L}^p$-spaces to the $2d$-dimensional plane. They use separate $\boldsymbol{L}^p$-norms for each “direction” or “domain” (e.g. time and frequency).

The next definition expresses the idea that lies at the core of the theory of time-frequency analysis as developed in this work: To define function spaces by the behaviour of their members under the *STFT*.

Definition 18. (Modulation Spaces) A function (or distribution) f belongs to the *modulation space* $\boldsymbol{M}^{p,q}(\mathbb{R}^d)$ if

$$\|f\|_{\boldsymbol{M}^{p,q}} = \|V_g f\|_{\boldsymbol{L}^{p,q}} < \infty,$$

for some suitable window g.

Remark 10. Membership in a modulation space gives us information on the overall behaviour of a function on the time-frequency plane. It tells us if it is well concentrated (e.g. bandlimited functions) or spread out (e.g. noise signal).

The following section introduces a class of spaces which will turn out to be very useful in the daily work of the time-frequency analyst (see also [5] and [14]).

4.1 Wiener amalgam spaces

The idea behind this special class of spaces stems from the desire to describe a function jointly in terms of its local and global behaviour. Therefore we will first decompose a given function f into compactly supported pieces and then measure the behaviour of each piece (the local behaviour of f) as well as the behaviour of all the pieces together (the global behaviour of f). To decompose f we need the following definitions.

Definition 19. Let $\boldsymbol{A} = \mathcal{F}\boldsymbol{L}^1$ be the image under the Fourier transform of the space $\boldsymbol{L}^1$. It is equipped with the norm

$$\|\hat{f}\|_{\boldsymbol{A}} = \|f\|_{\boldsymbol{L}^1}. \tag{4.1}$$

Its topological dual is the space $\boldsymbol{A}' = \mathcal{F}\boldsymbol{L}^\infty$, the Fourier image of $\boldsymbol{L}^\infty$ (we don't know yet how this might be interpreted). The space $\boldsymbol{A}_c$ is the space of all band-limited, absolutely integrable functions.

Definition 20. Let φ be a function in $\boldsymbol{A}_c(\mathbb{R}^d)$. Then φ generates a *Bounded Uniform Partition of Unity* (BUPU) if

$$\sum_{k\in\mathbb{Z}^d} T_k\varphi(x) = \sum_{k\in\mathbb{Z}^d} \varphi(x-k) = 1 \quad \forall x \in \mathbb{R}^d.$$

Definition 21. A sequence space S is called *BK-space* if it is a Banach space with respect to the topology of pointwise convergence, i.e. a sequence $(s_k)_k \in S$ converges to $s \in S$ if $s_k^n \to s^n$ for $k \to \infty$. A BK-space S is called *solid* if for every sequence $s = (s^k)_k$ there exists a sequence $t = (t^k)_k \in S$ such that $|s^n| \leq |t^n|$ implies $s \in S$ and $\|s\|_S \leq \|t\|_S$.

By using Bounded Uniform Partitions of Unity we can define Wiener Amalgam spaces as follows.

Definition 22. Let $\varphi \in \boldsymbol{A}_c$ generate a BUPU. Let $\boldsymbol{Y}$ be a solid BK-space and $\boldsymbol{X}$ some translation invariant Banach space of functions or distributions satisfying $\boldsymbol{A} \cdot \boldsymbol{X} \subseteq \boldsymbol{X}$, then a function $f \in \boldsymbol{X}$, belongs to the *Wiener Amalgam Space* $\boldsymbol{W}(\boldsymbol{X}, \boldsymbol{Y})$ if

$$||f||_{\boldsymbol{W}(\boldsymbol{X},\boldsymbol{Y})} = \| ||f\ T_k\varphi||_{\boldsymbol{X}} \|_{\boldsymbol{Y}(k)} < \infty, \tag{4.2}$$

where $\boldsymbol{Y}(k)$ emphasizes that we take the $\boldsymbol{Y}$-norm with respect to k.

Remark 11. The assumption that $\boldsymbol{Y}$ is solid is necessary to show that (4.2) actually defines a norm. To verify the triangle inequation consider $f, g \in \boldsymbol{W}(\boldsymbol{X}, \boldsymbol{Y})$, then

$$\begin{aligned}\|f+g\|_{\boldsymbol{W}(\boldsymbol{X},\boldsymbol{Y})} &= \|\|(f+g)T_k\varphi\|_{\boldsymbol{X}}\|_{\boldsymbol{Y}(y)} \\ &= \|\|fT_k\varphi + gT_k\varphi\|_{\boldsymbol{X}}\|_{\boldsymbol{Y}(k)}.\end{aligned}$$

Now a priori we don't know if $\|fT_k\varphi + gT_k\varphi\|_{\boldsymbol{X}} \in \boldsymbol{Y}$, but

$$\|fT_k\varphi + gT_k\varphi\|_{\boldsymbol{X}} \leq \|fT_k\varphi\|_{\boldsymbol{X}} + \|gT_k\varphi\|_{\boldsymbol{X}},$$

by the triangle equation in $\boldsymbol{X}$. But $\|fT_k\varphi\|_{\boldsymbol{X}} + \|gT_k\varphi\|_{\boldsymbol{X}} \in \boldsymbol{Y}$ since

$$\begin{aligned}\|\|fT_k\varphi\|_{\boldsymbol{X}} + \|gT_k\varphi\|_{\boldsymbol{X}}\|_{\boldsymbol{Y}(k)} &\leq \|\|fT_k\varphi\|_{\boldsymbol{X}}\|_{\boldsymbol{Y}(k)} + \|\|gT_k\varphi\|_{\boldsymbol{X}}\|_{\boldsymbol{Y}(k)} \\ &= \|f\|_{\boldsymbol{W}(\boldsymbol{X},\boldsymbol{Y})} + \|g\|_{\boldsymbol{W}(\boldsymbol{X},\boldsymbol{Y})}.\end{aligned}$$

Since $\boldsymbol{Y}$ was assumed to be solid this implies $\|fT_k\varphi + gT_k\varphi\|_{\boldsymbol{X}} \in \boldsymbol{Y}$ and the triangle equation is verified.

Remark 12. In this work we will always use $\boldsymbol{Y} = \boldsymbol{\ell}^p, 1 \leq p \leq \infty$. A function $f \in \boldsymbol{W}(\boldsymbol{X}, \boldsymbol{\ell}^p)$ then has the norm

$$\|f\|_{\boldsymbol{W}(\boldsymbol{X},\boldsymbol{\ell}^p)} = \left(\sum_{k\in\mathbb{Z}} \|f\, T_k\varphi\|_{\boldsymbol{X}}^p\right)^{1/p}$$

The definition of Wiener amalgam spaces captures the local behaviour of a function f by applying some Banach space norm and at the same time describes its global membership in some Banachspace $\boldsymbol{Y}$. Thus it allows for a more finer grained description of f, e.g. two functions with the same $\boldsymbol{L}^p$-norm do not necessarily have the same Wiener norm.

Wiener amalgam spaces obey to the following important convolution relations which will be needed later on.

Theorem 8. *Let $X_1 * X_2 \subseteq X_3$ and $Y_1 * Y_2 \subseteq Y_3$, then*

$$(X_1, Y_1) * W(X_2, Y_2) \subseteq W(X_3, Y_3).$$

Proof. The proof largely follows [5]. Let $f_1 \in W(X_1, Y_1)$ and $f_2 \in W(X_2, Y_2)$. Let ψ_1 and ψ_2 generate BUPUs for $W(X_1, Y_1)$ and $W(X_2, Y_2)$ respectively. Then we know that $f_r T_k \psi_r$ is a function in $X_r, r = 1, 2$. Let $\tau_r = \chi_{\operatorname{supp} T_i\psi_r}$ denote the characteristic function of $\operatorname{supp} T_i\psi_r$. First consider the function

$$\nu(x) = \sum_{i\in I} \|f\, T_i\psi_r\|_{X_r} \tau_r(x), \tag{4.3}$$

where I is some index set. We will show that $\nu(x) \in Y_r$. Let $I_x := \{i : x \in \operatorname{supp} T_i\psi_r\}$ and let n be such that

$$\|f\, T_k\psi_r\|_{X_r} = \sup_{i\in I_x} \|f\, T_i\psi_r\|_{X_r}.$$

I_x is finite since ψ_r is compactly supported, hence

$$\begin{aligned}\nu(x) &= \sum_{i\in I} \|f\,T_i\psi_r\|_{X_r}\tau_r(x)\\ &= \sum_{i\in I_x} \|f\,T_i\psi_r\|_{X_r}\\ &\le C\|f\,T_k\psi_r\|_{X_r}\end{aligned}$$

and therefore $\|\nu\|_{Y_r} \le C\|f\|_{W(X_r,Y_r)}$ by (4.2). Now choose $g \in \boldsymbol{A}_c$ with $\operatorname{supp} g \subset \operatorname{supp}\psi_1$ as well as $\operatorname{supp} g \subset \operatorname{supp}\psi_2$ generating a BUPU for $W(X_3, Y_3)$. Then we can write

$$f_1 * f_2 = \sum_k (f_1 * f_2)T_k g = \sum_k (f_1 * f_2)g_k.$$

First we see that

$$(f_1 * f_2)g_k = \sum_{i,j\in I_k} g_k(f_1T_i\psi_1 * f_2T_j\psi_2)$$

where the sum is taken over the set I_k of all index pairs (i, j) such that $\operatorname{supp}(g_k) \cap \operatorname{supp}(T_i\psi_1 * T_j\psi_2) \neq \emptyset$. Then

$$\|(f_1 * f_2)g_k\|_{X_3} \le \sum_{(i,j)\in I_k} \|g_k\|_{\boldsymbol{A}}\|f_1T_i\psi_1\|_{X_1}\|f_2T_j\psi_2\|_{X_2}.$$

Now we have to bring in convolution on the global spaces Y_1 and Y_2. Observe that

$$(\tau_1 * \tau_2)(x) \ge \|\chi_{\operatorname{supp} g_k}\|_1$$

if $x \in \operatorname{supp}(g_k)$ as above. Now consider the function $\eta : k \mapsto \|(f_1 * f_2)g_k\|_{X_3}$, then for $x \in \operatorname{supp} g_k$ fixed

$$\begin{aligned}\eta(k) &= \|(f_1 * f_2)g_k\|_{X_3}\\ &\le \frac{\|g_k\|_{\boldsymbol{A}}}{\|\operatorname{supp} g_k\|_1}\left(\sum_i \|f_1T_i\psi_1\|_{X_1}\tau_1\right) * \left(\sum_j \|f_2T_j\psi_2\|_{X_2}\tau_2\right)(x)\\ &=: C_k(F_i^1 * F_j^2)(x).\end{aligned}$$

By definition of the Wiener amalgam norm, the functions F_i^1 and F_j^2 are elements of the spaces Y_1, Y_2 respectively. Thus the second convolution relation in the assumption applies and yields $F_i^1 * F_j^2 \in Y_3$. Since this holds true for all $x \in \operatorname{supp} g_k$, the preceding formula implies that $k \mapsto \|(f_1 * f_2)g_k\|_{X_3} \in Y_3$ and thus $f_1 * f_2 \in W(X_3, Y_3)$. □

In the context of Gabor frames the subsequent theorem will prove to be helpful. It shows that sampling of "well behaved" functions leads to equally nice behaviour of the resultant sequences of sampling points.

Theorem 9. *(Sampling estimate for Wiener amalgam spaces)*
Let Λ be a lattice for the time frequency plane $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$, then for a continuous function f and $p \in [1, \infty)$ there exists a constant C_Λ such that

$$\|f|_\Lambda\|_{\boldsymbol{\ell}^p(\Lambda)} \leq C_\Lambda \|f\|_{\boldsymbol{W}(\boldsymbol{L}^\infty, \boldsymbol{\ell}^p)} \tag{4.4}$$

Proof. Since $\psi \in \boldsymbol{A}_c$ we know that there exists some $R > 0$ such that $\operatorname{supp} \psi \subseteq B_R(0)$, where $B_R(0)$ denotes the ball with radius R around 0 in $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$. Let N denote the number of lattice points in $\Lambda \cap B_R(n)$ then there exists a constant C_Λ with $N \leq C_\Lambda$, hence

$$\begin{aligned}
\sum_{\lambda \in \Lambda} |f(\lambda)|^p &\leq \sum_{\lambda \in \Lambda} C_\Lambda^{p-1} \sum_{n \in \mathbb{Z}^{2d}} |f(\lambda) T_n \psi(\lambda)| \\
&= C_\Lambda^{p-1} \sum_{n \in \mathbb{Z}^{2d}} \sum_{\lambda \in \Lambda \cap B_R(n)} |f(\lambda) T_n \psi(\lambda)| \\
&\leq C_\Lambda^p \sum_{n \in \mathbb{Z}^{2d}} \|f \, T_n \psi\|_\infty^p = C_\Lambda^p \|f\|_{\boldsymbol{W}(\boldsymbol{L}^\infty, \boldsymbol{\ell}^p)}.
\end{aligned}$$

□

Remark 13. The theorem also holds for $p = \infty$ with slightly modified proof (see [13], Proposition 11.14).

4.2 The space S_0 and its dual

The STFT gives rise to a special Banach space which proved to be optimally suited for time-frequency analysis. The idea is to define a space of functions based on the properties of their STFTs. In particular we are interested in functions with integrable STFT, i.e. functions that are in a sense bounded or well concentrated in the time frequency plane. The resulting space is known as *Feichtinger's algebra* and is denoted by $\boldsymbol{S}_0$.

Definition 23. (The Banach Space $\boldsymbol{S}_0(\mathbb{R}^d)$)
Let g be the Gaussian. A function f belongs to the space $\boldsymbol{S}_0(\mathbb{R}^d)$ if

$$||f||_{\boldsymbol{S}_0(\mathbb{R}^d)} = ||V_g f||_{\boldsymbol{L}^1} = \int_{\mathbb{R} \times \widehat{\mathbb{R}}} |V_g f(x, \omega)| dx d\omega < \infty \tag{4.5}$$

It can be shown that the definition of $\boldsymbol{S}_0(\mathbb{R}^d)$ does not depend on the window g (see Corollary 11 and [7]). Any non-zero $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ is suitable and different windows yield equivalent norms. $\boldsymbol{S}_0(\mathbb{R}^d)$ contains the Schwartz space $\boldsymbol{\mathcal{S}}(\mathbb{R}^d)$. Furthermore $\boldsymbol{S}_0(\mathbb{R}^d)$ is a special member of the class of *Modulation spaces* (see Definition 18) where it corresponds to the space $M_0^{1,1}(\mathbb{R}^d) = M^1(\mathbb{R}^d)$.

Theorem 10. *(Important Properties of $\boldsymbol{S}_0$)*
Let g denote the Gaussian.

1. *$\boldsymbol{S}_0(\mathbb{R}^d)$ is a time-frequency homogeneous Banach space, i.e. for every $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ we have $\pi(\lambda)f \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\|\pi(\lambda)f\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}$.*

2. *$\boldsymbol{S}_0(\mathbb{R}^d)$ is a dense subspace of $\boldsymbol{L}^2(\mathbb{R}^d)$, furthermore it is the smallest time-frequency homogeneous Banach space containing g.*

3. *$\boldsymbol{S}_0(\mathbb{R}^d)$ is invariant under the Fourier transform, i.e. if $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ then $\hat{f} \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\|\hat{f}\|_{\boldsymbol{S}_0(\mathbb{R}^d)} = \|f\|_{\boldsymbol{S}_0(\mathbb{R}^d)}$.*

Proof. First we observe that the $STFT$ of a time-frequency shifted version of $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ is just a time-frequency shift of the $STFT$ on the time-frequency plane:

$$\begin{aligned} V_g(M_\xi T_u f)(x,\omega) &= \int_{\mathbb{R}^d} M_\xi T_u f(t)\overline{g(x-t)}e^{-2\pi i t\omega}dt \\ &= \int_{\mathbb{R}^d} f(t-u)\overline{g(x-t)}e^{-2\pi i t(\omega-\xi)}dt \\ &= \int_{\mathbb{R}^d} f(z)\overline{g(x-u-z)}e^{-2\pi i z(x-\xi)}e^{-2\pi i u(\omega-\xi)}dz \\ &= e^{2\pi i u\xi}M_{(0,-u)}T_{(u,\xi)}V_g f(x,\omega). \end{aligned}$$

Thus $\|M_\xi T_u f\|_{\boldsymbol{S}_0} = \|f\|_{\boldsymbol{S}_0}$. To show that $\boldsymbol{S}_0(\mathbb{R}^d)$ is a Banach space, we have to verify that $\sum_{n\in\mathbb{N}} \|f\|_{\boldsymbol{S}_0} < \infty$ implies $f = \sum_{n\in\mathbb{N}} f_n$ is in $\boldsymbol{S}_0(\mathbb{R}^d)$. Indeed,

$$\begin{aligned} \|f\|_{\boldsymbol{S}_0} &= \iint |V_g f(x,\omega)|dxd\omega \\ &\leq \sum_{n\in\mathbb{N}} \iint |V_g f_n(x,\omega)|dxd\omega \\ &= \sum_{n\in\mathbb{N}} \|f_n\|_{\boldsymbol{S}_0} < \infty \end{aligned}$$

and thus finishes the proof of the first statement. To prove the second one, we need that

$$\|V_g f\|_\infty = \sup_{(x,\omega)\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |\langle f, M_\omega T_x g\rangle| \leq \|f\|_2\|g\|_2,$$

because in combination with (2.10) this yields

$$\begin{aligned} \|f\|_2^2 &= \|g\|_2^{-2}\|V_g f\|_2^2 \\ &\leq \|g\|_2^{-2}\|V_g f\|_\infty\|V_g f\|_1 \\ &\leq \|g\|_2^{-1}\|f\|_2\|f\|_{\boldsymbol{S}_0}, \end{aligned}$$

which implies that $\boldsymbol{S}_0(\mathbb{R}^d) \subseteq \boldsymbol{L}^2(\mathbb{R}^d)$. Now let $(g_n)_n \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\mathrm{supp}(g_n)$ compact for all $n \in \mathbb{Z}$ generate a BUPU in $\boldsymbol{L}^2(\mathbb{R}^d)$, then $fg_n \in \boldsymbol{S}_0(\mathbb{R}^d)$ for $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ and

$$f = \lim_{k\to\infty} \sum_{|n|\leq k} fg_n,$$

thus for every $f \in \boldsymbol{L}^2(\mathbb{R}^d)$ there is a sequence of functions $\psi_k = \sum_{|n|\leq k} fg_n, \psi_k \in \boldsymbol{S}_0(\mathbb{R}^d)$ converging to f and hence $\boldsymbol{S}_0(\mathbb{R}^d)$ is dense in $\boldsymbol{L}^2(\mathbb{R}^d)$.

To show that $\boldsymbol{S}_0$ is the smallest Banach space of its kind we observe that for any time-frequency homogeneous Banach space $\boldsymbol{B}$ containing g we can write $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ as

$$f = \frac{1}{\langle \gamma, g\rangle} \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} V_\gamma f(x,\omega) M_\omega T_x g \, d\omega dx$$

for an analyzing window $\gamma \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\langle g, \gamma\rangle \neq 0$. From this follows

$$\|f\|_{\boldsymbol{B}} \leq \frac{1}{|\langle g, \gamma\rangle|} \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |V_\gamma f(x,\omega)| \|M_\omega T_x g\|_{\boldsymbol{B}} \, d\omega dx \leq \frac{1}{|\langle g, \gamma\rangle|} \|g\|_{\boldsymbol{B}} \|f\|_{\boldsymbol{S}_0}.$$

Thus $\boldsymbol{S}_0(\mathbb{R}^d)$ is embedded in every time-frequency homogeneous Banach space $\boldsymbol{B}$ containing g.

For the third claim we use Parseval's formula to obtain

$$\begin{aligned} V_g f(x,\omega) &= \langle f, M_\omega T_x g\rangle \\ &= \langle \hat{f}, T_\omega M_{-x} \hat{g}\rangle \\ &= \langle \hat{f}, e^{2\pi i x\omega} M_{-x} T_\omega \hat{g}\rangle \\ &= e^{-2\pi i x\omega} V_{\hat{g}} \hat{f}. \end{aligned}$$

If we let g be a Gaussian window, which is invariant under the Fourier transform, it thus follows that

$$\|\hat{f}\|_{\boldsymbol{S}_0} = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |V_g \hat{f}(\omega, x)| \, d\omega dx = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |V_g f(-x,\omega)| \, dx d\omega = \|f\|_{\boldsymbol{S}_0}.$$

□

When talking about $\boldsymbol{S}_0(\mathbb{R}^d)$ we will also have to consider the space of all linear functionals on $\boldsymbol{S}_0(\mathbb{R}^d)$. Let $\boldsymbol{S}_0'(\mathbb{R}^d)$ denote the dual space of $\boldsymbol{S}_0(\mathbb{R}^d)$. First we will extend the Fourier transform to $\boldsymbol{S}_0'(\mathbb{R}^d)$ using *Parseval's formula* (2.3).

Definition 24. For $f \in \boldsymbol{S}_0'(\mathbb{R}^d)$ and $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ the Fourier transform of f is defined as $\hat{f} \in \boldsymbol{S}_0'(\mathbb{R}^d)$ such that

$$\langle \hat{f}, g\rangle = \langle f, \hat{g}\rangle.$$

Lemma 6. *The Fourier transform is a Banach Gelfand triple isomorphism.*

Proof. By Theorem 10 the Fourier transform is an isomorphism on $\boldsymbol{S}_0(\mathbb{R}^d)$, which by Plancherel's Theorem (Theorem 1) extends to an unitary mapping on $\boldsymbol{L}^2(\mathbb{R}^d)$. An application of Corollary 4 finishes the proof. □

With the previous definition we can interpret the STFT, which can be expressed as an inner product on $\boldsymbol{L}^2(\mathbb{R}^d)$, as the application of an element of $\boldsymbol{S}_0'(\mathbb{R}^d)$ to an element of $\boldsymbol{S}_0(\mathbb{R}^d)$, namely a time-frequency shifted version of $g \in \boldsymbol{S}_0(\mathbb{R}^d)$.

Definition 25. If $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d), g \in \boldsymbol{S}_0(\mathbb{R}^d)$ then

$$V_g\sigma(x,\omega) = \langle \sigma, M_\omega T_x g\rangle.$$

This is a uniformly continuous function on $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$.

Remark 14. The (uniform) continuity can be seen as follows. An easy consequence of Theorem 10 is that

$$\lim_{(x,\omega)\to(0,0)} \|M_\omega T_x g - g\|_{\boldsymbol{S}_0} = 0, \quad \forall g \in \boldsymbol{S}_0(\mathbb{R}^d).$$

Hence $M_\omega T_x$ acts continuously on $\boldsymbol{S}_0(\mathbb{R}^d)$. Thus for $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d), g \in \boldsymbol{S}_0(\mathbb{R}^d)$

$$\langle \sigma, M_\omega T_x g\rangle \to \langle \sigma, g\rangle$$

for $(x,\omega) \to 0$, which shows the continuity of $V_g\sigma$.

Lemma 7. *With the previous result in mind we see that $\boldsymbol{S}_0'(\mathbb{R}^d)$ can be characterized in terms of the STFT. Let $\boldsymbol{\mathcal{S}}'(\mathbb{R}^d)$ denote the dual of the Schwartz space of rapidly decreasing functions (we could also use the even bigger space $D'(\mathbb{R}^d)$, the dual space of the space of infinitely differentiable functions on $\mathbb{R}^d$ with compact support), then*

$$\boldsymbol{S}_0'(\mathbb{R}^d) = \left\{ \sigma \in \boldsymbol{\mathcal{S}}'(\mathbb{R}^d) : ||\sigma||_{\boldsymbol{S}_0'(\mathbb{R}^d)} = ||V_g\sigma||_{\boldsymbol{L}^\infty} = \sup_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |V_g\sigma(x,\omega)| < \infty \right\}$$

Example 1. The δ-distribution belongs to $\boldsymbol{S}_0'(\mathbb{R}^d)$. This follows from

$$\begin{aligned}
\|\delta\|_{\boldsymbol{S}_0'(\mathbb{R}^d)} &= \sup_{(x,\omega)\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |V_g\delta(x,\omega)| \\
&= \sup_{(x,\omega)\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |\langle \delta, M_\omega T_x g\rangle| \\
&= \sup_{(x,\omega)\in\mathbb{R}^d\times\widehat{\mathbb{R}}^d} |g(-x)| \\
&= \|g\|_\infty \le \|g\|_{\boldsymbol{S}_0} < \infty
\end{aligned}$$

Remark 15. With Definition 25 we have extended the inner product notation, which is only meaningful on a Hilbert space, to the Banach space $\boldsymbol{S}_0$ and its dual. Summing up, we can write for $f \in (\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$:

$$V_g f(x,\omega) = \langle f, M_\omega T_x g\rangle_{(\boldsymbol{S}_0,\boldsymbol{L}^2,\boldsymbol{S}_0')}$$

Now we can reassess the problem of eigenvalues of the shift operator T_x as mentioned on page 20. The trick is to consider the pure frequencies as functionals on $\boldsymbol{S}_0(\mathbb{R}^d)$, indeed we have

$$\begin{aligned}
\|\chi_\omega\|_{\boldsymbol{S}_0'}(\mathbb{R}^d) &= \sup_{(x,\nu)\in\mathbb{R}\times\widehat{\mathbb{R}}} |V_g\chi_\omega(x,\nu)| \\
&= \sup_{(x,\nu)\in\mathbb{R}\times\widehat{\mathbb{R}}} |\int_{\mathbb{R}^d} e^{2\pi it\omega}\overline{g(t-x)}e^{-2\pi it\nu}dt| \\
&\leq \sup_{(x,\nu)\in\mathbb{R}\times\widehat{\mathbb{R}}} \int_{\mathbb{R}^d} |e^{2\pi it(\omega-\nu)}||\overline{g(t-x)}|dt \\
&\leq \|g\|_{\boldsymbol{L}^1(\mathbb{R}^d)} \leq \|g\|_{\boldsymbol{S}_0} < \infty,
\end{aligned}$$

thus the operators T_x and M_ω have eigenvectors in $\boldsymbol{S}_0'(\mathbb{R}^d)$, namely χ_x and δ_ω respectively.

Next we will also extend the inverse STFT to $\boldsymbol{S}_0'(\mathbb{R}^d)$.

Theorem 11. *Let $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d), g \in \boldsymbol{S}_0(\mathbb{R}^d)$, then*

$$\sigma = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} V_g\sigma(x,\omega)M_\omega T_x\gamma\, dxd\omega, \quad \gamma \in \boldsymbol{S}_0(\mathbb{R}^d), \tag{4.6}$$

where the integral is to be interpreted in the weak sense, i.e.

$$\langle\sigma, f\rangle = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} V_g\sigma(x,\omega)\langle M_\omega T_x\gamma, f\rangle \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d).$$

Proof. For $\gamma \in \boldsymbol{S}_0(\mathbb{R}^d)$ define $\tilde{\sigma} = \int_{\mathbb{R}^{2d}} V_g\sigma(x,\omega)M_\omega T_x\gamma\, dxd\omega$, then for every $f \in \boldsymbol{S}_0(\mathbb{R}^d)$

$$\begin{aligned}
\langle\tilde{\sigma}, f\rangle &= \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} V_g\sigma(x,\omega)\langle M_\omega T_x\gamma, f\rangle\, dxd\omega \\
&= \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} \langle\sigma, M_\omega T_x g\rangle\overline{V_\gamma f(x,\omega)}\, dxd\omega \\
&= \langle\sigma, f\rangle,
\end{aligned}$$

since $f = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} V_\gamma f(x,\omega)M_\omega T_x g\, dxd\omega$ is valid for $f \in \boldsymbol{S}_0(\mathbb{R}^d)$, thus $\tilde{\sigma} = \sigma$. □

One of the most important characterizations of $\boldsymbol{S}_0$ is as Wiener amalgam space. Using this alternative viewpoint may lead to very accessible proofs as will be seen in the sequel. The following theorem again emphasizes the "niceness" of $\boldsymbol{S}_0$, identifying it locally with the Fourier image (thus a subspace of C_0) of all integrable functions which are piecewise summable. So once more we can appreciate how well concentrated $\boldsymbol{S}_0$-functions are on the time-frequency plane.

Theorem 12. *Let* $\boldsymbol{A} = \mathcal{F}\boldsymbol{L}^1$ *with* $\|f\|_{\boldsymbol{A}} := \|h\|_{\boldsymbol{L}^1}$ *for* $f = \hat{h}$*, and* $\boldsymbol{A}' = \mathcal{F}\boldsymbol{L}^\infty$*, then*

$$\boldsymbol{S}_0(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1) \tag{4.7}$$

and

$$\boldsymbol{S}_0'(\mathbb{R}^d) = \boldsymbol{W}(\boldsymbol{A}', \boldsymbol{\ell}^\infty). \tag{4.8}$$

Proof. We begin with the first equation and have to show that $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ if and only if $\sum_n \|f\, T_n\psi\|_{\boldsymbol{A}} < \infty$. First assume that $f \in \boldsymbol{S}_0(\mathbb{R}^d)$, then the invariance of $\boldsymbol{S}_0(\mathbb{R}^d)$ under the Fourier transform implies that there exists a function $h \in \boldsymbol{S}_0(\mathbb{R}^d)$ such that $f = \hat{h}$. Let $\psi \in \boldsymbol{A}_c \cap \boldsymbol{S}_0(\mathbb{R}^d)$ (for example let ψ be a triangular function) generate a BUPU. By the same argument as above there exists a $\varphi \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $\psi = \widehat{\varphi}$. We have to show, that $\sum_n \|f T_n\psi\|_{\boldsymbol{A}} = \sum_n \|h * M_n\varphi\|_1 < \infty$. We compute that

$$(h * M_n\varphi)(x) = \int h(t)\varphi(x-t)e^{2\pi i n(x-t)}dt = e^{2\pi i n x} V_{\overline{\varphi}^\vee} h(x,n),$$

where $\overline{\varphi}^\vee(x) = \overline{\varphi}(-x)$. Thus

$$\sum_n \|h * M_n\varphi\|_1 = \sum_n \int |V_{\overline{\varphi}^\vee} h(x,n)||e^{2\pi i n x}|dx = \int \sum_n |V_{\overline{\varphi}^\vee} h(x,n)|dx < \infty, \tag{4.9}$$

since $h \in \boldsymbol{S}_0(\mathbb{R}^d)$. Conversely assume $f \in \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)$, then $f = \sum_n f\, T_n\psi$ and $\sum_n \|f\, T_n\psi\|_A < \infty$. Now let $g \in \boldsymbol{S}_0 \cap \boldsymbol{A}_c$, then a quick calculation yields

$$\begin{aligned} \|f\, T_n\psi\|_{\boldsymbol{S}_0(\mathbb{R}^d)} &= \| \|f\, T_n\psi\, T_x\overline{g}\|_{\boldsymbol{A}}\|_1 \\ &\le \|f\, T_n\psi\|_{\boldsymbol{A}} \|g\|_{\boldsymbol{A}} |\operatorname{supp} g \cap \operatorname{supp} f\, T_n\psi|^d \\ &< \infty, \end{aligned} \tag{4.10}$$

since both functions in question are compactly supported and thus $f \in \boldsymbol{S}_0(\mathbb{R}^d)$. □

Remark 16. Equation (4.10) also shows that $\boldsymbol{A}_c \subseteq \boldsymbol{S}_0$, in fact it is even a dense subspace of $\boldsymbol{S}_0$.

Theorem 13. $\boldsymbol{S}_0(\mathbb{R}^d)$ *is a dense subspace of* $\boldsymbol{L}^2(\mathbb{R}^d)$ *and weak*- densely embedded in* $\boldsymbol{S}_0'(\mathbb{R}^d)$*, i.e.* $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ *is a Banach Gelfand triple.*

Proof. That $\boldsymbol{S}_0(\mathbb{R}^d)$ is a dense subspace of $\boldsymbol{L}^2(\mathbb{R}^d)$ was already shown in Theorem 10. What remains to be shown is the weak* density of $\boldsymbol{S}_0(\mathbb{R}^d)$ in $\boldsymbol{S}_0'(\mathbb{R}^d)$. But before we proceed, we need the following result.

Lemma 8. *Let* $g > 0$ *and* h *be in* $\boldsymbol{S}_0(\mathbb{R}^d)$*, then for every* $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ *we have*

$$\lim_{(\rho,\tau)\to(\infty,\infty)} \|St_\rho g * (D_\tau h\, f) - f\|_{\boldsymbol{S}_0(\mathbb{R}^d)} = 0 \tag{4.11}$$

as well as

$$\lim_{(\rho,\tau)\to(\infty,\infty)} \|D_\rho h(St_\tau g * f) - f\|_{\boldsymbol{S}_0(\mathbb{R}^d)} = 0. \tag{4.12}$$

Proof. We will only sketch the proof. Therefore we show that $St_\rho g * f \in \boldsymbol{S}_0(\mathbb{R}^d)$ for $f, g \in \boldsymbol{S}_0(\mathbb{R}^d)$ with $g > 0$. Observe that for a window $\gamma \in \boldsymbol{S}_0(\mathbb{R}^d)$

$$M_\omega \gamma * (St_\rho g * f)(t) = e^{2\pi i t\omega} V_\gamma (St_\rho g * f)(t)),$$

thus

$$|M_\omega \gamma * (St_\rho g * f)(t)| = |V_\gamma (St_\rho g * f)(t, \omega)|.$$

On the other hand,

$$\begin{aligned} |M_\omega \gamma * (St_\rho g * f)(t)| &= |(St_\rho g * M_\omega \gamma * f)(t)| \\ &\leq (|St_\rho g| * |V_\gamma f(\cdot, \omega)|(t)). \end{aligned}$$

Now choose $g > 0 \in \boldsymbol{S}_0(\mathbb{R}^d)$ such that $\|g\|_1 \leq 1$. Now it follows that

$$\begin{aligned} \int\int |V_\gamma (St_\rho g * f)(t,\omega)| dt d\omega &\leq \int\int |St_\rho g| * |V_\gamma f(\cdot, \omega)|(t) dt d\omega \\ &= \int \left(\int |St_\rho g(t)| dt \right) \left(\int |V_\gamma f(t, \omega)| dt \right) d\omega \\ &\leq \int\int |V_\gamma f(t,\omega)| dt d\omega, \end{aligned}$$

hence

$$\lim_{\rho\to\infty} \|St_\rho g * f - f\|_{\boldsymbol{S}_0} = 0.$$

This result is equivalent under the Fourier transform to

$$\lim_{\rho\to 0} \|(D_{\frac{1}{\rho}} \hat{g})\hat{f} - \hat{f}\|_{\boldsymbol{S}_0(\mathbb{R}^d)} = 0. \tag{4.13}$$

Since both St_ρ and D_ρ are bounded operators on $\boldsymbol{S}_0(\mathbb{R}^d)$, we can combine them and thus finish the proof. □

Continuation of the proof of Theorem 13. First observe that $\boldsymbol{\ell}^1 * \boldsymbol{\ell}^\infty \subseteq \boldsymbol{\ell}^\infty$ and $\boldsymbol{A}' * \boldsymbol{A} \subseteq \boldsymbol{A}$. The latter relation follows from the fact that the Fourier transform maps convolution to multiplication and vice versa and the relation $\boldsymbol{L}^1 \cdot \boldsymbol{L}^\infty \subseteq \boldsymbol{L}^1$, which can easily be verified. Using (4.7) and Theorem 8 we see that

$$\boldsymbol{S}_0' * \boldsymbol{S}_0 = \boldsymbol{W}(\boldsymbol{A}', \boldsymbol{\ell}^\infty) * \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1) \subseteq \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^\infty). \tag{4.14}$$

Proceeding with $\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1) \cdot \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^\infty) \subseteq \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)$ we arrive at

$$\boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1) \cdot (\boldsymbol{W}(\boldsymbol{A}', \boldsymbol{\ell}^\infty) * \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)) \subseteq \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^1)$$

or, more compactly,

$$\boldsymbol{S}_0 \cdot (\boldsymbol{S}_0' * \boldsymbol{S}_0) \subseteq \boldsymbol{S}_0. \tag{4.15}$$

Now we reevaluate Lemma 8. Since the limit holds true for the $\boldsymbol{S}_0$-norm, it also applies for $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ in the weak* sense, i.e. if we let g and h as in Lemma 8, Equation (4.11) and (4.12) applied to $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ hold true (with absolute value instead of the norm). Let's verify this for Equation (4.12), if $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ we have

$$\begin{aligned}(D_\rho h(St_\tau g * \sigma))(f) &= \int f(x) D_\rho h(x) \sigma(T_x(St_\tau g)^\vee) dx \\ &= \iint f(x) D_\rho h(x) St_\tau g(x-t)\sigma(t) dt dx \\ &= \int (St_\tau \breve{g} * (D_\rho h\, f))(t)\sigma(t) \\ &= \sigma(St_\tau \breve{g} * (D_\rho h\, f),\end{aligned}$$

where $\breve{g}(x) = g(-x)$. Thus

$$|(D_\rho h(St_\tau g * \sigma))(f) - \sigma(f)| = |\sigma(St_\tau \breve{g} * (D_\rho h\, f - f))| \leq \varepsilon \|\sigma\|_{\boldsymbol{S}_0'},$$

for suitable τ, ρ according to (4.11). Now Equation (4.15) implies that $\boldsymbol{S}_0(\mathbb{R}^d)$ is weak*-dense in $\boldsymbol{S}_0'(\mathbb{R}^d)$. □

In the preceding proof we made heavy use of a technique called *regularization* which will be the topic of the following subsection.

4.3 Regularization

It is a standard technique in mathematics to approximate objects which live in general spaces by objects that live in smaller spaces which are well known and easier to handle. In the setting of time-frequency analysis such spaces are typically $\boldsymbol{S}_0'(\mathbb{R}^d)$ and $\boldsymbol{S}_0(\mathbb{R}^d)$ respectively. To approximate $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ we will start by "taming" it in the time domain and subsequently in the frequency domain. Here "taming" is to be understood as localizing the support of σ in the time frequency plane. This can be done in various ways, three of which we will discuss here.

Of course we need to make sure that our localization procedure converges to the original function if we let the localization parameter go to infinity (or zero as the case may be).

Definition 26. A regularizing sequence is a sequence of operators $(A_n)_n$ with kernels $(K_n)_n \in \boldsymbol{S}_0(\mathbb{R}^{2d})$, i.e. A_n maps $\boldsymbol{S}_0'(\mathbb{R}^d)$ into $\boldsymbol{S}_0(\mathbb{R}^d)$, where

- each A_n is a Banach-Gelfand triple homomorphism on $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$
- $\|A_n f - f\|_{\boldsymbol{S}_0} \to 0$ for $n \to \infty$ and $f \in \boldsymbol{S}_0$.

Regularization via PC and CP operators

Definition 27. A product-convolution operator (PC operator) is an operator of the form

$$A_{g,h}f = g(h * f)$$

and conversely a convolution-product operator (CP operator) is of the form

$$B_{g,h}f = g * (hf).$$

The following proposition reconsiders the proof of Theorem 13 from the angle of regularization.

Proposition 3. *Let $A_\alpha : f \mapsto D_\alpha g(St_\alpha h * f)$, then A_α is a regularizing sequence for $\alpha \to \infty$, more precisely*

1. $A_\alpha : \boldsymbol{S}_0'(\mathbb{R}^d) \to \boldsymbol{S}_0(\mathbb{R}^d)$
2. $A_\alpha f \to f \quad \forall f \in \boldsymbol{S}_0(\mathbb{R}^d)$
3. $A_\alpha \sigma \to \sigma \quad \forall \sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ *in the weak* sense.*

Proof. The proof is identical to the proof of Theorem 13. □

Remark 17. This result shows that parameterized PC operators with $g, h \in \boldsymbol{S}_0(\mathbb{R}^d)$ are suitable for constructing regularizing sequences. An analoguous result holds for CP operators. By analyzing what a PC operator with $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ actually does, we find that it works by applying lowpass filtering (convolution with g) to its argument, which amounts to localization in the frequency domain, followed by localization in time via multiplication with h and vice versa for CP operators.

Regularization via Fourier Transform

The same effect, as described in the previous remark, can be achieved by multiplying with g on the time side followed by the same operation on the Fourier transform side. In the following we will w.l.o.g. let $g = h \in \boldsymbol{S}_0(\mathbb{R}^d)$.

Let's define an operator $A_\alpha f(t) = f(t)g_\alpha(t)$ where $g_\alpha(t) = e^{-\alpha\pi t^2}$ is the Gaussian. Obviously, the action of A_α is localization in time. Now consider the operator

$$T_\alpha := A_\alpha \circ \mathcal{F} \circ A_\alpha.$$

The integral kernel of T_α is

$$\begin{aligned} K_\alpha(\omega, t) &= e^{-\alpha\pi t^2} e^{-2\pi i t\omega} e^{-\alpha\pi\omega^2} \\ &= e^{-\pi(\alpha t^2 + 2it\omega + \alpha\omega^2)}. \end{aligned}$$

As we can see, the action of this operator on a function $f \in (\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ is twofold. First it gently applies localization in the time domain before transferring the result to the Fourier transform side and localizing again. The adjoint operator of T_α is

$$T'_\alpha = A_\alpha \circ \mathcal{F}^{-1} \circ A_\alpha$$

with integral kernel

$$\begin{aligned} K'_\alpha(t,\omega) &= e^{-\alpha\pi\omega^2} e^{2\pi i t\omega} e^{-\alpha\pi t^2} \\ &= e^{-\pi(\alpha\omega^2 - 2it\omega + \alpha t^2)}. \end{aligned}$$

As α approaches zero, $T'_\alpha \circ T_\alpha$ converges to $\mathcal{F}^{-1} \circ \mathcal{F} = \mathrm{Id}$. From this we might conclude that the kernels $K'_\alpha \cdot K_\alpha$ converge to $\delta(t-x)$ since

$$\int f(t)\delta(t-x)dt = f(x)$$

and so $\delta(t-x)$ can be interpreted as the kernel of the identity. Indeed, for $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ we compute

$$\begin{aligned} (T'_\alpha \circ T_\alpha)f(x) &= T'_\alpha \left(\int\limits_{\mathbb{R}^d} K_\alpha(\omega,t) f(t)dt \right)(x) \\ &= \int\limits_{\mathbb{R}^d}\int\limits_{\mathbb{R}^d} K'_\alpha(x,\omega)K_\alpha(\omega,t)f(t)dtd\omega \\ &= \int\limits_{\mathbb{R}^d}\int\limits_{\mathbb{R}^d} f(t)e^{-\pi(\alpha t^2 + 2it\omega + \alpha\omega^2)} e^{-\pi(\alpha\omega^2 - 2ix\omega + \alpha x^2)} dtd\omega \\ &= \int\limits_{\mathbb{R}^d}\int\limits_{\mathbb{R}^d} f(t)e^{-\pi(\alpha t^2 + 2it\omega + 2\alpha\omega^2 - 2ix\omega + \alpha x^2)} dtd\omega \\ &= \int\limits_{\mathbb{R}^d}\int\limits_{\mathbb{R}^d} f(t)e^{-\pi\alpha t^2} e^{-2\pi i\omega(t-x)} e^{2\pi\alpha\omega^2} e^{-\pi\alpha x^2} dtd\omega \\ &= \int\limits_{\mathbb{R}^d}\int\limits_{\mathbb{R}^d} f(t)e^{-\alpha\pi(t^2+x^2)} e^{-2\pi\alpha\omega^2} e^{-2\pi i\omega(t-x)} dtd\omega \\ &= \int\limits_{\mathbb{R}^d} f(t)e^{-\alpha\pi(t^2+x^2)} \int\limits_{\mathbb{R}^d} e^{-2\pi\alpha\omega^2} e^{-2\pi i\omega(t-x)} d\omega dt \\ &= \int\limits_{\mathbb{R}^d} f(t)e^{-\alpha\pi(t^2+x^2)}\, \mathcal{F}\varphi_{\frac{1}{2\alpha}}(t-x)dt \\ &= \frac{\sqrt{2}}{2\sqrt{\alpha}} \int\limits_{\mathbb{R}^d} f(t)e^{-\alpha\pi(t^2+x^2)} e^{-\frac{1}{2\alpha}\pi(t-x)^2} dt, \end{aligned}$$

where we assumed $d = 1$. In the last integral we set $y = (t-x)/\sqrt{\alpha}$ and write $t^2 + x^2$ in the first exponent as $(t-x)^2 + 2tx$. This yields

$$\frac{\sqrt{2}}{2\sqrt{\alpha}} \int\limits_{\mathbb{R}^d} f(x+\sqrt{\alpha}y) e^{-\alpha\pi(\alpha y^2 + 2(\sqrt{\alpha}y + x)x)} e^{-\frac{1}{2\alpha}\pi\alpha y^2} \sqrt{\alpha} dy,$$

which reduces to

$$\frac{1}{\sqrt{2}} \int\limits_{\mathbb{R}^d} f(x+\sqrt{\alpha}y) e^{-\alpha\pi(\alpha y^2 + 2(\sqrt{\alpha}y + x)x)} e^{-\frac{1}{2}\pi y^2} dy.$$

Now we take the limit $\alpha \to 0$ and conclude

$$\begin{aligned}
&\lim_{\alpha\to 0} \frac{1}{\sqrt{2}} \int\limits_{\mathbb{R}^d} f(x+\sqrt{\alpha}y) e^{-\alpha\pi(\alpha y^2 + 2(\sqrt{\alpha}y + x)x)} e^{-\frac{1}{2}\pi y^2} dy \\
&= \frac{1}{\sqrt{2}} \int\limits_{\mathbb{R}^d} \lim_{\alpha\to 0} f(x+\sqrt{\alpha}y) e^{-\alpha\pi(\alpha y^2 + 2(\sqrt{\alpha}y + x)x)} e^{-\frac{1}{2}\pi y^2} dy \\
&= \frac{1}{\sqrt{2}} f(x) \int\limits_{\mathbb{R}^d} e^{-\frac{1}{2}\pi y^2} dy \\
&= \frac{1}{\sqrt{2}} f(x) \sqrt{2} \\
&= f(x).
\end{aligned}$$

To justify the interchanging of limit and integral one could either use Lebesgue's theorem of dominated convergence or argue that, since we are in $\boldsymbol{S}_0$, the integrand decreases rapidly enough such that there are virtually no contributions from "far away" and hence we can interpret the integral as a "quasi finite" Riemann integral.

A simple duality argument shows that $(T'_\alpha \circ T_\alpha)\sigma \to \sigma$ for $\sigma \in \boldsymbol{S}'_0(\mathbb{R}^d)$. Observe that

$$(T'_\alpha \circ T_\alpha)' = T'_\alpha \circ (T'_\alpha)' = T'_\alpha \circ T_\alpha,$$

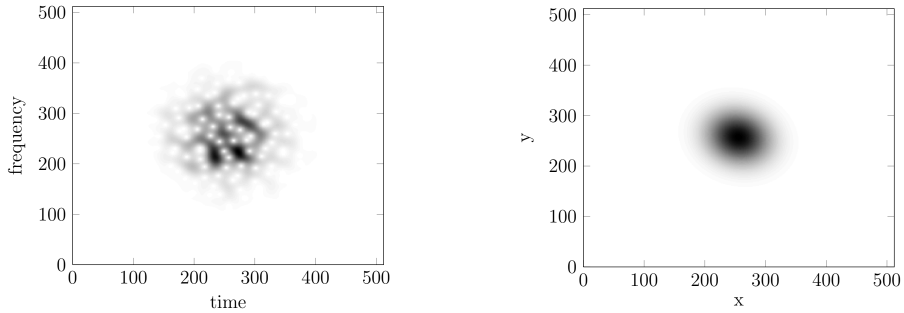
i.e. $T'_\alpha \circ T_\alpha$ is self-adjoint, then

$$\langle (T'_\alpha \circ T_\alpha)\sigma, g \rangle = \langle \sigma, (T'_\alpha \circ T_\alpha) g \rangle,$$

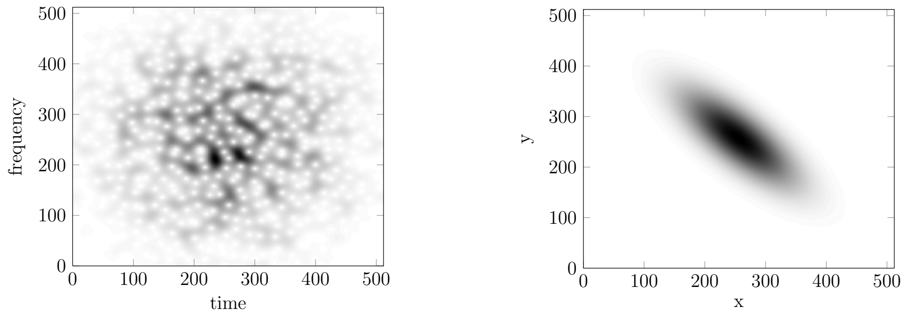
for every $g \in \boldsymbol{S}_0(\mathbb{R}^d)$. Figure 4 is a visualization of this calculation. Note how the kernel approaches the identity matrix.

Regularization via Gabor sums

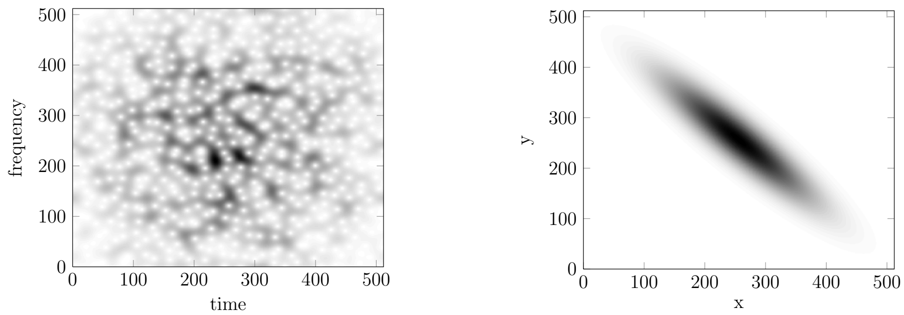
The third regularization method we will briefly mention is regularization via partial Gabor sums.



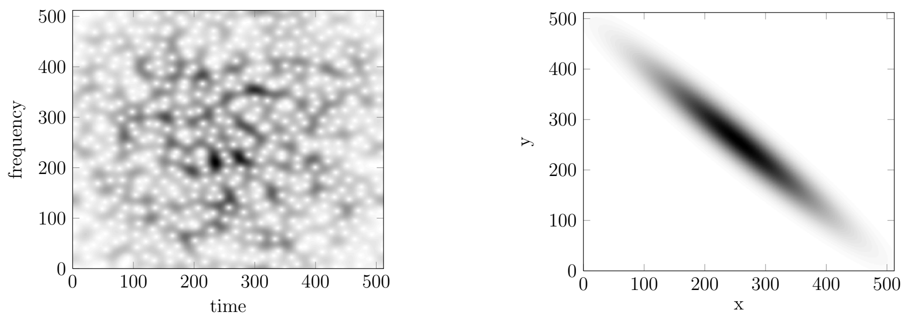
(a) $\alpha = 2$



(b) $\alpha = 0.4$



(c) $\alpha = 0.22$



(d) $\alpha = 0.15$

Figure 4: The action of $T'_\alpha \circ T_\alpha$ *on a random noise signal. On the left is the Gabor coefficient matrix, on the right the operator kernel.*

Lemma 9. *Let $\Lambda \subset \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ be a lattice with elements $\lambda = (\lambda_1, \lambda_2)$, let $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ be a Gabor atom with dual atom $\tilde{g}$, then the operators*

$$A_N : f \mapsto \sum_{\lambda_1^2+\lambda_2^2 \leq N^2} \langle f, g_\lambda \rangle \tilde{g}_\lambda$$

form a regularizing sequence for $N \to \infty$.

Proof. It is clear from Corollary 6 that $A_N : \boldsymbol{S}_0' \to \boldsymbol{S}_0$. The convergence $A_N f \to f$ is also obvious for the whole Gelfand triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$. □

Remark 18. Of course we are not limited to circular subsets of $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ only. The preceding lemma is also true for sums over rectangular subsets of the time-frequency plane or, more general, for any finite subset of $\Lambda \subset \mathbb{R}^d \times \widehat{\mathbb{R}}^d$.

4.4 $(\mathbf{S_0}, \mathbf{L^2}, \mathbf{S_0'})$ and Gabor frames

Now that we have shown that $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ is a full-fledged Banach-Gelfand triple we will investigate some of its properties regarding Gabor frames as introduced in Section 2.2. We start with an inequality.

Lemma 10. *Let $g \in \boldsymbol{S}_0(\mathbb{R}^d)$, $\|g\|_2 = 1$ and $f \in \boldsymbol{S}_0'(\mathbb{R}^d)$, then*

$$|V_g f(x,\omega)| \leq (|V_g f| * |V_g g|)(x,\omega) \tag{4.16}$$

Proof. The key to the proof is the following calculation. Let $F \in (\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^{2d})$, then

$$\begin{aligned} V_g(V_g^{-1}(F))(x,\omega) &= \langle V_g^{-1} F, M_\omega T_x g \rangle \\ &= \langle F, V_g(M_\omega T_x g) \rangle \\ &= \int\int F(t,\nu) \overline{V_g(M_\omega T_x g)(t,\nu)} dt d\nu \\ &= \int\int F(t,\nu) e^{-2\pi i x(\nu-\omega)} V_g g(x-t, \omega-\nu) dt d\nu, \end{aligned}$$

from which it follows that

$$|V_g V_g^{-1} F| \leq (|F| * |V_g g|)(x,\omega). \tag{4.17}$$

The result follows by taking $F = V_g f$. □

Remark 19. Combining the preceding lemma with (4.14) yields that $V_g f \in \boldsymbol{W}(\boldsymbol{A}, \boldsymbol{\ell}^\infty)$ if $f \in \boldsymbol{S}_0'(\mathbb{R}^d)$.

Theorem 14. *Let $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $f \in (\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$, then it follows that $V_g f \in (\boldsymbol{W}(\boldsymbol{L}^\infty, \boldsymbol{\ell}^1), \boldsymbol{W}(\boldsymbol{L}^\infty, \boldsymbol{\ell}^2), \boldsymbol{W}(\boldsymbol{L}^\infty, \boldsymbol{\ell}^\infty))$ and the subsequent inequality holds:*

$$\|V_g f\|_{\boldsymbol{W}(\boldsymbol{L}^\infty, \boldsymbol{\ell}^p)} \leq C \|V_g g\|_{\boldsymbol{W}(\boldsymbol{L}^\infty, \boldsymbol{\ell}^1)} \|f\|_{(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')}. \tag{4.18}$$

Proof. The proof uses *Young's inequality* for modulation spaces, which states that

$$\|f * g\|_{\boldsymbol{W}(\boldsymbol{L}^\infty,\boldsymbol{\ell}^p)} \leq C\|f\|_{\boldsymbol{L}^p}\|g\|_{\boldsymbol{W}(\boldsymbol{L}^\infty,\boldsymbol{\ell}^1)} \qquad 1 \leq p \leq \infty. \tag{4.19}$$

For an exhaustive proof of this formula we refer the reader to [13], Theorem 11.1.5. We apply (4.19) to (4.16), which yields

$$\|V_g f\|_{\boldsymbol{W}(\boldsymbol{L}^\infty,\boldsymbol{\ell}^p)} \leq C\|V_g f\|_{\boldsymbol{L}^p}\|V_g g\|_{\boldsymbol{W}(\boldsymbol{L}^\infty,\boldsymbol{\ell}^1)} \qquad p \in \{1, 2, \infty\},$$

thus the proof is finished.

□

With these technicalities at hand we will prove that the analysis operator and consequently the synthesis operator are bounded.

Theorem 15. *Let $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ and Λ a lattice in $\mathbb{R}^{2d}$, then the analysis operator $C_g : f \mapsto (\langle f, \pi(\lambda)g\rangle)_{\lambda\in\Lambda}$ is bounded from $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')$ into $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)$ and we have*

$$\|C_g\|_{op} \leq C_\Lambda\|V_g g\|_{\boldsymbol{W}(\boldsymbol{L}^\infty,\boldsymbol{\ell}^1)} \tag{4.20}$$

Proof. First we observe that $C_g f(x, \omega) = \langle f, M_\omega T_x g\rangle = V_g f|_\Lambda$. Since $V_g f$ is continuous we can apply Theorem 9 and thus by (4.18) get

$$\begin{aligned}
\|C_g f\|_{\boldsymbol{\ell}^p} &= \|V_g f|_\Lambda\|_{\boldsymbol{\ell}^p} \\
&\leq C_\Lambda\|V_g f\|_{\boldsymbol{W}(\boldsymbol{L}^\infty,\boldsymbol{\ell}^p)} \\
&\leq \tilde{C}_\Lambda\|V_g g\|_{\boldsymbol{W}(\boldsymbol{L}^\infty,\boldsymbol{\ell}^1)}\|f\|_{(\boldsymbol{S}_0,\boldsymbol{L}^2,\boldsymbol{S}_0')}.
\end{aligned}$$

□

With a straight forward duality argument we get the same result for the synthesis operator $D_g : (c_\lambda)_{\lambda\in\Lambda} \mapsto \sum_{\lambda\in\Lambda} c_\lambda\pi(\lambda)g$.

Theorem 16. *Let $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ and Λ a lattice in $\mathbb{R}^{2d}$, then the synthesis operator D_g is a bounded mapping from $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)$ to $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')$.*

Proof. Since $D_g = C_g^*$, we have

$$\begin{aligned}
|\langle D_g c, f\rangle| &= |\langle c, C_g f\rangle| \\
&\leq \|c\|_{(\boldsymbol{\ell}^1,\boldsymbol{\ell}^2,\boldsymbol{\ell}^\infty)}\|C_g f\|_{(\boldsymbol{\ell}^1,\boldsymbol{\ell}^2,\boldsymbol{\ell}^\infty)} \\
&\leq \|c\|_{(\boldsymbol{\ell}^1,\boldsymbol{\ell}^2,\boldsymbol{\ell}^\infty)}\|C_g\|_{op}\|f\|_{(\boldsymbol{S}_0,\boldsymbol{L}^2,\boldsymbol{S}_0')}.
\end{aligned}$$

We consider D_g as adjoint operator, hence

$$\begin{aligned}
\|D_g c\|_{(\boldsymbol{S}_0,\boldsymbol{L}^2,\boldsymbol{S}_0')} &= \sup_{\|f\|_{(\boldsymbol{S}_0,\boldsymbol{L}^2,\boldsymbol{S}_0')}\leq 1} |\langle D_g c, f\rangle| \\
&\leq \|c\|_{(\boldsymbol{\ell}^1,\boldsymbol{\ell}^2,\boldsymbol{\ell}^\infty)}\|C_g\|_{op}
\end{aligned}$$

□

Corollary 5. *Let $g, \gamma \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\Lambda \subset \mathbb{R}^{2d}$ a lattice, then the* generalized *Gabor frame operator*

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

is a Banach-Gelfand triple isomorphism. Furthermore in analogy to (2.19) and (2.20), every $f \in (\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')$ can be expanded as

$$\begin{aligned} f &= \sum_{\lambda\in\Lambda} \langle f, \pi(\lambda)g\rangle \pi(\lambda)\tilde{g} \\ &= \sum_{\lambda\in\Lambda} \langle f, \pi(\lambda)\tilde{g}\rangle \pi(\lambda)g, \end{aligned}$$

where $\tilde{g} = S_{g,g}^{-1}g$ is the dual window of g.

Remark 20. The preceding theorems are also valid for general modulation spaces, i.e.

$$\begin{aligned} C_g &: M^{p,q}(\mathbb{R}^{2d}) \to \ell^{p,q}(\mathbb{Z}^{2d}), \\ D_g &: \ell^{p,q}(\mathbb{Z}^{2d}) \to M^{p,q}(\mathbb{R}^{2d}) \end{aligned}$$

and the Gabor frame operator

$$S_g : M^{p,q}(\mathbb{R}^{2d}) \to M^{p,q}(\mathbb{R}^{2d}).$$

See [13], Theorem 12.2.3. and 12.2.4.

Corollary 5 immediatly leads to another quite important characterization of $\boldsymbol{S}_0(\mathbb{R}^d)$. It implies that every function in $\boldsymbol{S}_0(\mathbb{R}^d)$ is a superposition of time-frequency shifted versions of a single atom g_0.

Corollary 6. *(Atomic characterization of $\boldsymbol{S}_0(\mathbb{R}^d)$)*
Let $g_0 \in \boldsymbol{S}_0(\mathbb{R}^d), g_0 \neq 0$, then

$$\boldsymbol{S}_0(\mathbb{R}^d) = \left\{ f = \sum_{n=1}^{\infty} c_n M_{\omega_n} T_{x_n} g_0, \sum_{n=0}^{\infty} |c_n| < \infty \right\}. \tag{4.21}$$

Lemma 11. *The last corollary implies that the definition of $\boldsymbol{S}_0$ does not depend on the choice of g.*

Proof. Assume that $g, g' \in \boldsymbol{S}_0(\mathbb{R}^d)$, then $g' = \sum_{n=0}^{\infty} c_n M_{\omega_n} T_{x_n} g$. This allows us to make the following (rough) calculation:

$$\begin{aligned} V_{g'}f(x,\omega) &= \langle f, M_\omega T_x g'\rangle \\ &= \langle f, M_\omega T_x \sum_n c_n M_{\omega_n} T_{x_n} g\rangle \end{aligned}$$

$$
\begin{aligned}
&= \sum_n c_n \langle f, M_\omega T_x M_{\omega_n} T_{x_n} g\rangle \\
&= \sum_n c_n \langle f, M_{\nu_n} T_{y_n} g\rangle \\
&= \sum_n c_n V_g f(y_n, \nu_n).
\end{aligned}
$$

From this it follows that

$$
\|V_{g'} f\|_1 \leq \sum_n |c_n| \|V_g f\|_1 < \infty,
$$

thus the independence of g in the definition of $\boldsymbol{S}_0$ is proved. □

4.5 Wilson bases

In this section we will introduce a basis for the Banach-Gelfand triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')$. To get there we will make use of so-called *mixed-norm spaces*, a generalization of $\boldsymbol{L}^p$-spaces.The key result, Wilson bases, sprung from a simple but profound modification of a Gabor system. Instead of time frequency shifted versions $M_\omega T_x g$ of a Gabor atom g, which are concentrated at (x, ω), we consider a system which is symmetrically concentrated at (x, ω) and $(x, -\omega)$. We will define Wilson bases for dimension $d = 1$ and then generalize the concept to higher dimensions which is easily done by taking tensor products. For more on Wilson bases see for example [13] or [8]. First, we will define mixed-norm spaces.

Definition 28. (Mixed-norm spaces)
Let $1 \leq p, q < \infty$, then the *mixed norm space* $L^{p,q}(\mathbb{R}^{2d})$ is defined as the space of all measureable functions on $\mathbb{R}^{2d}$ such that

$$
\|f\|_{L^{p,q}(\mathbb{R}^{2d})} = \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(x,\omega)|^p dx \right)^{q/p} d\omega \right)^{1/q} < \infty.
$$

with the usual modification if $p = \infty$ or $q = \infty$.

Definition 29. Let $\mathcal{G}(g, \frac{1}{2}, 1)$ be a Gabor system of redundancy 2 in $\boldsymbol{L}^2(\mathbb{R})$, then the associated *Wilson system* $\mathcal{W}(g)$ consists of the functions

$$
\psi_{kn} = c_n T_{\frac{k}{2}} (M_n + (-1)^{k+n} M_{-n}) g, \quad (k,n) \in \mathbb{Z} \times \mathbb{Z}^+, \tag{4.22}
$$

where $c_0 = 1$, $c_n = \frac{1}{\sqrt{2}}$ for $n \geq 1$ and

$$
\psi_{2k,0} = T_k g, \quad k \in \mathbb{Z}. \tag{4.23}
$$

As seen in this definition a member function of a Wilson system consists of the sum of two time shifted versions of the atom g which are modulated with opposite sign. For the Wilson system $\mathcal{W}(g)$ to be a basis we need some more structure of the underlying Gabor system.

Theorem 17. *If $\mathcal{G}(g, \frac{1}{2}, 1)$ is a tight frame for $\boldsymbol{L}^2(\mathbb{R})$ and $\|g\|_2 = 1$ as well as $g(x) = g^*(x) = \overline{g(-x)}$. Then $\mathcal{W}(g)$ is an orthonormal basis of $\boldsymbol{L}^2(\mathbb{R}^d)$*

Proof. The proof of this theorem is quite involved and comprehensive. Therefore we will leave it to the interested reader to study it on his own. A real in-depth discussion and proof can be found in ([13], p.168f). □

For a Wilson basis $\mathcal{W}(g)$ we define the associated analysis operator C_ψ and synthesis operator D_ψ as

$$C_\psi f = (\langle f, \psi_{kn} \rangle)_{(k,n) \in \mathbb{Z} \times \mathbb{Z}^+} \tag{4.24}$$

$$D_\psi c = C_\psi^* f = \sum_{n \in \mathbb{N}} \sum_{k \in \mathbb{Z}} c_{kn} \psi_{kn}. \tag{4.25}$$

We will need these operators to prove the following theorem where we will establish an isomophism between modulation spaces of functions or distributions and discrete *mixed-norm spaces*. As a corollary we will establish a Banach-Gelfand triple basis for $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$.

Theorem 18. *(Isomorphism Theorem)*
Let $\mathcal{W}(g)$ with $g \in \boldsymbol{S}_0(\mathbb{R})$ be an orthonormal Wilson basis, then the analysis operator C_ψ establishes an isomorphism between the spaces $\boldsymbol{M}^{p,q}(\mathbb{R}^d)$ and $\boldsymbol{\ell}^{p,q}(\mathbb{Z} \times \mathbb{Z}^+)$.

For the proof of this theorem we need some more facts about Wilson bases which are collected in the following theorem.

Theorem 19. *Let $\mathcal{W}(g)$ with $g \in \boldsymbol{S}_0(\mathbb{R}^d)$ be an orthonormal basis for $\boldsymbol{L}^2(\mathbb{R}^d)$, then there exists a constant $C \geq 1$, such that*

$$\frac{1}{C} \|f\|_{\boldsymbol{M}^{p,q}} \leq \left(\sum_{n \in \mathbb{N}} \left(\sum_{k \in \mathbb{Z}} |\langle f, \psi_{kn} \rangle|^p \right)^{q/p} \right)^{1/q} \leq C \|f\|_{\boldsymbol{M}^{p,q}}, \tag{4.26}$$

which implies that C_ψ is one-to-one from $\boldsymbol{M}^{p,q}(\mathbb{R}^d)$ into $\boldsymbol{\ell}^{p,q}(\mathbb{Z} \times \mathbb{Z}^+)$. Furthermore the coefficients $\langle f, \psi_{kn} \rangle$ in the orthogonal expansion

$$f = \sum_{n \in \mathbb{N}} \sum_{k \in \mathbb{Z}} \langle f, \psi_{kn} \rangle \psi_{kn} \tag{4.27}$$

are unique and the sum converges unconditionally in the $\boldsymbol{M}^{p,q}$-norm for $p, q < \infty$ and weak otherwise. To put it differently, if $f = D_\psi c$, then $c = C_\psi f$.*

Proof. We will only sketch the proof, for the details we refer the reader to [13], p.265-266. Let us write the identity operator in $\boldsymbol{L}^2(\mathbb{R}^d)$ as $\mathrm{Id} = D_\psi C_\psi$, which is possible because $\mathcal{W}(g)$ is an orthonormal basis for $\boldsymbol{L}^2(\mathbb{R}^d)$. Then we can verify (4.26) if we can show the continuity of the operators D_ψ and C_ψ. The key to achieve this, is to reduce the statement about Wilson bases to a statement about the synthesis and analysis operator for frames respectively. If for an arbitrary sequence $(c_{k,n})_{(k,n)\in\mathbb{Z}\times\mathbb{Z}^+}$ we define the sequence $(\tilde{c}_{k,n})_{(k,n)\in\mathbb{Z}\times\mathbb{Z}^+}$ as

$$\tilde{c}_{k,n} = \begin{cases} \frac{1}{\sqrt{2}} c_{k,n} & n > 0 \\ \frac{(-1)^{k+n}}{\sqrt{2}} c_{k,-n} & n < 0 \\ c_{k,0} & n = 0, \end{cases}$$

it can be shown that $\|\tilde{c}\|_{\ell^{p,q}} \leq \sqrt{2}\|c\|_{\ell^{p,q}}$ and $D_\psi c = D_g \tilde{c}$. By Theorem 16 it follows that

$$\|D_\psi c\|_{M^{p,q}} = \|D_g \tilde{c}\|_{M^{p,q}} \leq C\|\tilde{c}\|_{\ell^{p,q}} \leq \sqrt{2}C\|c\|_{\ell^{p,q}},$$

and thus D_ψ is continuous. The continuity of C_ψ follows in a similar fashion from Theorem (4.20) and so (4.26) follows by

$$\begin{aligned} \|f\|_{M^{p,q}} &= \|D_\psi C_\psi\|_{M^{p,q}} \\ &\leq \|D_\psi\|_{op}\|C_\psi f\|_{\ell^{p,q}} \\ &\leq \|D_\psi\|_{op}\|C_\psi\|_{op}\|f\|_{M^{p,q}} \end{aligned}$$

if we take $C = \max \|D_\psi\|_{op}\|C_\psi\|_{op}$. The rest of the theorem is an analogy to Corollary 5. □

Proof of Theorem 18. Admittedly we have more or less outsourced the proof of Theorem 18 to the preceding Theorem 19 and the only thing left to do now is to sum up. C_ψ maps $\boldsymbol{M}^{p,q}$ one-to-one into $\boldsymbol{\ell}^{p,q}(\mathbb{Z}\times\mathbb{Z}^+)$ and conversely for $c \in \boldsymbol{\ell}^{p,q}$ we find $f = D_\psi c$ with $c = C_\psi f$ and so C_ψ is onto $\boldsymbol{\ell}^{p,q}(\mathbb{Z}\times\mathbb{Z}^+)$. Its inverse is the operator D_ψ mapping $\boldsymbol{\ell}^{p,q}(\mathbb{Z}\times\mathbb{Z}^+)$ onto $\boldsymbol{M}^{p,q}$. □

Before we can go on we need to extend the notion of Wilson bases to higher dimensions $d > 1$. We accomplish this by taking tensor products as seen without proof (for the proof see [13], p.270) in the following lemma.

Lemma 12. *Let $\mathcal{W}(g)$ be an orthonormal Wilson basis for $\boldsymbol{L}^2(\mathbb{R})$, then the functions*

$$\psi_{rs} = \prod_{j=1}^{d} \psi_{r_j s_j}(x_j) \quad r, s \in \mathbb{Z}^d, s \geq 0 \tag{4.28}$$

define an unconditional orthonormal basis for $\boldsymbol{L}^2(\mathbb{R}^d)$.

Up until now we have introduced Wilson bases only as bases for $\boldsymbol{L}^2(\mathbb{R}^d)$ but Theorem 18 suggests that they in fact can act as basis for the whole Gelfand triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$.

Lemma 13. *Let $(\psi_{kn})_{(k,n)\in\mathbb{Z}^d\times\mathbb{N}^d}$ be an orthonormal Wilson basis for $\boldsymbol{L}^2(\mathbb{R}^d)$, then*

- *$(\psi_{kn})_{(k,n)\in\mathbb{Z}^d\times\mathbb{N}^d}$ is a bounded, absolute basis of $\boldsymbol{S}_0(\mathbb{R}^d)$.*

- $(\psi_{kn})_{(k,n)\in\mathbb{Z}^d\times\mathbb{N}^d}$ *is a weak* basis of* $\boldsymbol{S}_0'(\mathbb{R}^d)$.

Consequently the analysis operator $C_\psi f = (\langle f, e_i\rangle)_{i\in I}$ *establishes an unitary Banach Gelfand triple isomorphism between* $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ *and* $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\mathbb{Z}^d \times \mathbb{N}^d)$.

Proof. If we recall that $\boldsymbol{S}_0(\mathbb{R}^d)$ coincides with the modulation space $M_0^{1,1}(\mathbb{R}^d)$ we can apply Theorem 18 and find that C_ψ maps $\boldsymbol{S}_0(\mathbb{R}^d)$ isomorphically to $\boldsymbol{\ell}^{1,1}(\mathbb{Z}^d \times \mathbb{N}^d) = \boldsymbol{\ell}^1(\mathbb{Z}^d \times \mathbb{N}^d)$. Now Equation (4.26) states that $(\psi_{kn})_{(k,n)\in\mathbb{Z}^d\times\mathbb{N}^d}$ is a bounded, absolute basis of $\boldsymbol{S}_0(\mathbb{R}^d)$. For $f \in \boldsymbol{S}_0'(\mathbb{R}^d)$ the expansion

$$f = \sum_{k\in\mathbb{Z}^d}\sum_{n\in\mathbb{N}^d} \langle f, \psi_{kn}\rangle \psi_{kn}$$

converges unconditionally in the weak* sense according to Theorem (19), i.e.

$$\begin{aligned}\langle f, g\rangle &= \left\langle \sum_{k\in\mathbb{Z}^d}\sum_{n\in\mathbb{N}^d} \langle f, \psi_{kn}\rangle \psi_{kn}, g \right\rangle \\ &= \sum_{k\in\mathbb{Z}^d}\sum_{n\in\mathbb{N}^d} \langle f, \psi_{kn}\rangle \langle \psi_{kn}, g\rangle < \infty, \quad \forall g \in \boldsymbol{S}_0(\mathbb{R}^d).\end{aligned}$$

This shows the second part of the lemma. □

4.6 Weak*-topology

In this section we will have a thorough look at the weak*-topology, especially on $\boldsymbol{S}_0'(\mathbb{R}^d)$. We start off with some topological theory.

Definition 30. Let X be a set. A topology $\mathcal{T}$ on X is a collection of subsets of X (called the *open sets*) which satisfy

1. $X \in \mathcal{T}$ and $\emptyset \in \mathcal{T}$
2. If $U_1 \in \mathcal{T}$ and $U_2 \in \mathcal{T}$ then $U_1 \cap U_2 \in \mathcal{T}$
3. For any index set I, $U_i \in \mathcal{T}$ implies that $\bigcup_{i\in I} U_i \in \mathcal{T}$

Let X be a set and $F \neq \emptyset$ a family of mappings $f : X \to Y_f$ for some topological space Y_f. Let $\mathcal{T}_X$ denote the union of all finite intersections of $f^{-1}(V)$, where V is an open subset of Y_f, then $\mathcal{T}_X$ is a topology on X. If we say a topology $\mathcal{T}_1$ is weaker than a topology $\mathcal{T}_2$, if every $\mathcal{T}_1$-open set is also $\mathcal{T}_2$-open, in symbols: $\mathcal{T}_1 \subset \mathcal{T}_2$, then $\mathcal{T}_X$ is the weakest topology on X such that every $f : X \mapsto Y_f$ is continuous.

Now let X be a topological vector space and X' its dual. Every $x \in X$ defines a linear functional on X' via

$$f_x(x') := \langle x, x'\rangle \qquad x' \in X', \tag{4.29}$$

where the brackets are to be understood in a functional sense, i.e. $\langle x, x' \rangle$ is to be read as $x'(x)$. This simply means that $X \subseteq (X')'$. A topological space for which $X'' = X$ is called *reflexive*.

Now we are in a situation similar to the one described above, where X' is the argument space and $\mathbb{C}$ the target space for all functions $f_x : X' \to \mathbb{C}$. Hence the functions f_x induce a topology on X', the so-called *weak*-topology* (cf. [15]).

A slightly less abstract way to introduce the weak*-topology is via seminorms on X' defined by (4.29) as

$$p_x(x') = |f_x(x')| = |\langle x, x' \rangle|. \tag{4.30}$$

Let P be the collection of all seminorms on X' and let $F \subset P$ be a *finite* subset. For $\varepsilon > 0$ we then define the sets

$$K_{F,\varepsilon} := \{x' \in X' : p(x') \leq \varepsilon \quad \forall p \in F\} \tag{4.31}$$

and

$$\mathscr{K} := \{K_{F,\varepsilon} : F \subset P, F \text{ as above and } \varepsilon > 0\}. \tag{4.32}$$

The following definition puts this in a broader topological context.

Definition 31. Let (X, τ) be a topological space. A set $U \subseteq X$ is a *neighbourhood* of $x \in X$ if there exists a $T \in \tau$ such that $x \in T \subset U$. A collection $\mathscr{U}$ of neighbourhoods of x is called a *local base* at x if for every neighbourhood V of x there exists a $U \subset \mathscr{U}$ with $U \subset V$.

Remark 21. With this definition we see that (4.31) defines neighbourhoods of 0 and the set $\mathscr{K}$ in (4.32) is a local base at 0. The notion of a neighbourhood of some point x in a topological space generalizes the concept of the ball with radius ε in a metric space (Y, d), $B_\varepsilon(x) := \{y \in Y : d(x, y) \leq \varepsilon\}$. Also note that in a metric space Y $(B_{1/n}(x))_n$ is a local base for $x \in Y$.

Now we can use (4.31) and (4.32) to define the weak*–topology on X'. It consists of all sets $U \subset X'$ which are open in the sense that for all $x' \in U$ there exists a set $K_{x'} \in \mathscr{K}$ such that $x' + K_{x'} \subset U$, in other words, a set $U \subset X'$ is open if for all $x' \in U$ there is a whole neighbourhood $K_{x'} \in \mathscr{K}$ of x with $K_{x'} \subset U$ (see [17] for details and proofs).

Next we will develop the notion of convergence in the weak*-topology. Therefore we will define *nets* which are a generalization of sequences, because the latter demand more structure than a general topological space may provide.

Definition 32. A *directed set* is a set S together with an reflexive and transitive relation $\preceq$ such that for every $a, b \in S$ there exists an element $c \in S$ such that $a \preceq c$ and $b \preceq c$. We write $(S, \preceq)$ to denote a directed set.

Definition 33. Let X be some topological space. A *net* x is a mapping from a directed set $(S, \preceq)$ to X, $x : s \mapsto x_s$.

Remark 22. Nets generalize the concept of sequences by allowing more general index-sets. A sequence in X is just a net with domain $\mathbb{N}$ and range X.

So why do we need nets? Recall that a sequence $(x_n)_n \in X$ converges by definition to a point $x \in X$ if for all $N \in \mathbb{N}$ there exists an $M \in \mathbb{N}$ such that $x_n \in B_{1/N}(x)$ for all $n \geq M$. So the convergence of sequences depends on the availability of countable local bases, but in general topological spaces they might not exist. Topological spaces which *do* possess countable local bases are called *AA2*-spaces. To define the convergence of nets we can use general local bases by saying that a net $(x_\alpha)_\alpha$ converges to x if for every $U \subset \mathscr{U}_x$, where $\mathscr{U}_x$ is a local base of x, there exists an index α_U such that $x_\alpha \in U$ for all $\alpha \succeq \alpha_U$.

To define convergence in the weak* topology we will again use seminorms as defined in (4.30) instead of the very abstract definition above.

Definition 34. (Weak*-convergence) Let B be a Banach space and B' its topological dual. A net $(\sigma_\alpha)_\alpha \in B'$ is said to converge *weak** to $\sigma \in B'$ if and only if

$$p_f(\sigma_\alpha - \sigma) \to 0$$

for all seminorms p_f on B', $\alpha \to \infty$, or stated differently,

$$|\sigma_\alpha(f) - \sigma(f)| \to 0$$

for all $f \in B$ and $\alpha \to \infty$, i.e. $(\sigma_\alpha)_\alpha \in B'$ converges pointwise to $\sigma \in B'$.

Remark 23. This definition matches our abstract definition of net-convergence, because for every finite set $F \subset B$ and $\varepsilon > 0$ there exists an α_0 such that $p_f(\sigma_\alpha - \sigma) < \varepsilon$ for all $\alpha > \alpha_0$. Thus, reconsidering (4.31) leads directly to the definition of net-convergence given in remark 22.

Lemma 14. *A Banach space B is reflexive with respect to the weak* topology on B', i.e. $(B', w^*)' = B$.*

Proof. It is clear that every $g \in B$ defines an element in $(B')'$ by the definition

$$\sigma''_g(\sigma) := \langle g, \sigma \rangle, \quad \sigma \in B'. \tag{4.33}$$

Thus B is a subspace of $(B')'$. Since $(B', w^*)'$ is the space of all linear functionals on B' which are continuous with respect to its weak*-topology, every $\sigma'' \in (B', w^*)'$ satisfies

$$\langle \sigma_n, \sigma'' \rangle \to \langle \sigma, \sigma'' \rangle.$$

if $\sigma_n \to \sigma$ in the weak* sense. By (4.33) these are exactly the elements of B. Thus $B = (B', w^*)'$. □

According to the following lemma, pointwise convergence from Definition 34 is equivalent to uniform convergence on compact sets.

Lemma 15. *Let B be a vector space and $(\sigma_\alpha)_\alpha$ a net in B'. σ_α converges weak* to $\sigma \in B'$ if and only if for all $\varepsilon > 0$ and for all compact sets $K \subset B$, there exists an index α_0 such that for all $\alpha \succeq \alpha_0$*

$$|\sigma_\alpha(f) - \sigma(f)| \leq \varepsilon \quad \forall f \in K.$$

Now lets turn our attention to $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$. As we have seen in Corollary 6, $\boldsymbol{S}_0(\mathbb{R}^d)$ can be characterized as the space of all linear combinations of time-frequency shifted copies of a single atom g_0 (e.g. the Gaussian). Since the definition of $\boldsymbol{S}_0(\mathbb{R}^d)$ does not depend on a particular choice of atom, every atom in $\boldsymbol{S}_0(\mathbb{R}^d)$ is suitable for said characterization, which implies that the set $\{M_\omega T_x g, (x,\omega) \in \mathbb{R}^d \times \mathbb{R}^d\}$ is a total subset of $\boldsymbol{S}_0(\mathbb{R}^d)$. This leads to the following characterization of weak*-convergence on $\boldsymbol{S}_0'(\mathbb{R}^d)$.

Theorem 20. *(Weak-* convergence in $\boldsymbol{S}_0'(\mathbb{R}^d)$) Let $(\sigma_\alpha)_\alpha$ be a net in $\boldsymbol{S}_0'(\mathbb{R}^d)$. $(\sigma_\alpha)_\alpha$ converges weak* to $\sigma \in \boldsymbol{S}_0'(\mathbb{R}^d)$ if for every $\varepsilon > 0$ and every $\varrho > 0$ there exists an index α_0 such that for all $\alpha \succeq \alpha_0$*

$$\begin{aligned}|\sigma_\alpha(M_\omega T_x g) - \sigma(M_\omega T_x g)| &= |\langle \sigma_\alpha, M_\omega T_x g\rangle - \langle \sigma, M_\omega T_x g\rangle| \\ &= |V_g\sigma_\alpha(x,\omega) - V_g\sigma(x,\omega)| \\ &\leq \varepsilon,\end{aligned}$$

where $|(x,\omega)| \leq \varrho$.

Remark 24. Theorem 20 shows that weak* convergence of σ_α to σ on $\boldsymbol{S}_0'(\mathbb{R}^d)$ can be described as pointwise resp. uniform convergence over compact sets of the time frequency plane of $V_g\sigma_\alpha$ to $V_g\sigma$. This can be depicted as the spectrograms of σ_α resembling more and more the spectrogram of σ.

Example 2. As an example where we have weak*-convergence on $\boldsymbol{S}_0'(\mathbb{R}^d)$ but not norm convergence consider the δ-distribution. Let $x_n \to x_0 \in \mathbb{R}^d$, if we use the standard norm on $\boldsymbol{S}_0'(\mathbb{R}^d)$, we have

$$\|\delta_{x_n} - \delta_{x_0}\|_\infty = \max_{\|f\|_{\boldsymbol{S}_0} \leq 1} |f(x_n) - f(x_0)| = 2,$$

on the other hand, using the weak*topology, we have

$$\begin{aligned}p_f(\delta_{x_n} - \delta_{x_0}) &= |\langle f, \delta_{x_n} - \delta_{x_0}\rangle \\ &= |\langle f, \delta_{x_n}\rangle - \langle f, \delta_{x_0}\rangle| \\ &= |f(x_n) - f(x_0)| \to 0,\end{aligned}$$

for all $f \in \boldsymbol{S}_0(\mathbb{R}^d)$.

As another example consider the pure frequencies $\chi_\omega = \omega(\cdot) = e^{2\pi i \omega \cdot} \in \boldsymbol{S}_0'(\mathbb{R}^d)$. In the weak*-topology we have

$$p_f(\chi_{\omega_n} - \chi_{\omega_0}) = |\langle f, \chi_{\omega_n} - \chi_{\omega_0}\rangle|$$

$$
\begin{aligned}
&= |\langle f, \chi_{\omega_n}\rangle - \langle f, \chi_{\omega_0}\rangle| \\
&= |\hat{f}(\omega_n) - \hat{f}(\omega_0)| \to 0,
\end{aligned}
$$

where $\omega_n \to \omega_0 \in \widehat{\mathbb{R}^d}$ and $f \in \boldsymbol{S}_0(\mathbb{R}^d)$ arbitrary. In the norm-topology we have again that

$$
\begin{aligned}
\|\chi_{\omega_n} - \chi_{\omega_0}\|_\infty &= \max_{\|f\|_{\boldsymbol{S}_0} \le 1} |\hat{f}(\omega_n) - \hat{f}(\omega_0)| \\
&= \max_{\|\hat{f}\|_{\boldsymbol{S}_0} \le 1} |\hat{f}(\omega_n) - \hat{f}(\omega_0)| = 2,
\end{aligned}
$$

since the Fourier transform is an isometry on $\boldsymbol{S}_0(\mathbb{R}^d)$ (see Theorem 10).

Proposition 4. *A net σ_α is weak* convergent in $\boldsymbol{S}_0'(\mathbb{R}^d)$ if and only if it converges on the finite dimensional subspace $V = \operatorname{span}(g_1, \dots, g_n) \subset \boldsymbol{S}_0(\mathbb{R}^d)$, where $g_1, g_2, \dots, g_n$ are arbitrary elements of $\boldsymbol{S}_0(\mathbb{R}^d)$.*

Proof. Assume that σ_α is a weak* convergent net in $\boldsymbol{S}_0'(\mathbb{R}^d)$, $\langle \sigma_\alpha, g\rangle \to \langle \sigma_0, g\rangle, \forall g \in \boldsymbol{S}_0(\mathbb{R}^d)$. By using the Gram-Schmidt process we can construct a biorthogonal system $\{f_1, \dots, f_n\}$ from $g_1, \dots, g_n$. Consider the projection operator

$$
P_V \sigma_\alpha = \sum_{k=1}^{n} \langle \sigma_\alpha, f_k\rangle f_k,
$$

then by assumption $P_V\sigma_\alpha$ converges pointwise to $P_V\sigma_0 = \sum_{k=1}^{n} \langle \sigma_0, f_k\rangle f_k$, it even converges in the $\boldsymbol{S}_0$-norm. Now for the other direction: We know that σ_α converges on every finite dimensional subspace of $\boldsymbol{S}_0'(\mathbb{R}^d)$, thus in particular on the one dimensional space $V = \{cf | f \in \boldsymbol{S}_0(\mathbb{R}^d) \text{ fixed }, c \in \mathbb{R}\}$. Then $P_V\sigma_\alpha = \langle \sigma_\alpha, f\rangle f$. If we let w.l.o.g. $\|f\|_2 = 1$, it follows that $\|P_V\sigma_\alpha - P_V\sigma_0\|_2 = |\langle \sigma_\alpha, f\rangle - \langle \sigma_0, f\rangle|$, which tends to zero by assumption. Since this is true for any $f \in \boldsymbol{S}_0(\mathbb{R}^d)$, the proof is complete. □

5 Operator Gelfand Triple

The underlying theme in this section is the search for methods to identify operators acting on the Gelfand triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ with corresponding functions on the time-frequency plane. This approach is known as *operator quantization*. In the following sections we will discuss three different takes on this theme: Identification of an operator with its integral kernel, the Kohn-Nirenberg correspondence and the spreading representation. The material covered here is largely based on [9].

5.1 Integral Operators

In a finite dimensional world, problems which are described by linear systems amount to the action of matrices on vectors by multiplication. But if we step into the infinite dimensional world of function spaces, we have to adapt our thinking. An infinite dimensional linear system is described by some linear operator K which acts on a function f. This action can be described – in analogy to matrix multiplication but using integration instead of summation – as

$$Kf(x) = \int_{\mathbb{R}^d} k(x,y) f(y) dy.$$

The function k, which plays the role of the matrix, is called the *kernel* of the operator K. Operators of this kind are known as *integral operators*. A particularly nice example are the operators with kernels in $\boldsymbol{L}^2(\mathbb{R}^d \times \mathbb{R}^d)$, the so-called *Hilbert-Schmidt* operators. The space of Hilbert-Schmidt operators is denoted by $\mathcal{HS}$. It is equipped with an inner product via

$$\langle K, L \rangle_{\mathcal{HS}} = \langle \kappa(K), \kappa(L) \rangle_{\boldsymbol{L}^2(\mathbb{R}^d \times \mathbb{R}^d)} \quad F, G \in \mathcal{HS}$$

which turns $\mathcal{HS}$ into a Hilbert space with the derived norm

$$\|K\|_{\mathcal{HS}} = \sqrt{\langle K, K \rangle}.$$

These results instantly raise the question how a general operator K is related to its kernel κ or which properties of the operator can be encoded in its kernel. This leads us to the so-called *Kernel Theorems* which identify operator spaces with corresponding kernel spaces.

Definition 35. For $f, g \in \boldsymbol{S}_0(\mathbb{R}^d)$ let $g \otimes \bar{f}$ denote the mapping

$$g \otimes \bar{f} : h \mapsto \langle h, f \rangle g, \quad \forall h \in \boldsymbol{S}_0'(\mathbb{R}^d). \tag{5.1}$$

This is a rank-one operator. Its kernel is given by $\kappa(g \otimes \bar{f})(x, y) = g(x)\bar{f}(y)$.

Theorem 21. *Let $\mathcal{B}' = \mathcal{L}(\boldsymbol{S}_0, \boldsymbol{S}_0')$ denote the space of all continuous linear operators mapping $\boldsymbol{S}_0(\mathbb{R}^d)$ into $\boldsymbol{S}_0'(\mathbb{R}^d)$. These operators are in one-to-one correspondence with their kernels in $\boldsymbol{S}_0'(\mathbb{R}^d \times \mathbb{R}^d)$, i.e. every distribution $\kappa(K) \in \boldsymbol{S}_0'(\mathbb{R}^d \times \mathbb{R}^d)$ defines a bounded linear operator K from $\boldsymbol{S}_0(\mathbb{R}^d)$ to $\boldsymbol{S}_0'(\mathbb{R}^d)$ by*

$$\langle Kf, g \rangle = \langle \kappa(K), g \otimes \bar{f} \rangle \quad f, g \in \boldsymbol{S}_0(\mathbb{R}^d), \tag{5.2}$$

and conversely for every operator K which maps $\boldsymbol{S}_0(\mathbb{R}^d)$ into $\boldsymbol{S}_0'(\mathbb{R}^d)$ there exists a unique kernel $\kappa(K) \in \boldsymbol{S}_0'(\mathbb{R}^d \times \mathbb{R}^d)$ such that the action of K is described by 5.2.

Proof. For a given $\kappa(K) \in \boldsymbol{S}_0'(\mathbb{R}^d \times \mathbb{R}^d)$ we have

$$\begin{aligned} |\langle Kf, g \rangle| &= |\langle \kappa(K), \kappa(g \otimes \bar{f}) \rangle| \\ &\le \|\kappa(K)\|_{\boldsymbol{S}_0'} \|\kappa(g \otimes \bar{f})\|_{\boldsymbol{S}_0} \\ &= \|\kappa(K)\|_{\boldsymbol{S}_0'} \|g\|_{\boldsymbol{S}_0} \|\bar{f}\|_{\boldsymbol{S}_0}, \end{aligned}$$

which is valid for all $g \in \boldsymbol{S}_0(\mathbb{R}^d)$. Therefore we conclude that $Kf \in \boldsymbol{S}_0'(\mathbb{R}^d)$ and hence $K \in \mathcal{B}'$ with operator norm $\|K\| \leq \|\kappa(K)\|_{\boldsymbol{S}_0'}$.

Conversely let K be an operator in $\mathcal{B}'$ and let $(\psi_{kn})_{(k,n)\in\mathbb{Z}^d\times\mathbb{N}^d}$ be a Wilson basis, then we define the matrix corresponding to K as $K_{l,n,r,s} = (k_{ln,rs})_{(l,n),(r,s)\in\mathbb{Z}^{2d}}$ where

$$k_{ln,rs} = \langle K\psi_{rs}, \psi_{ln}\rangle, \quad l,n,r,s \in \mathbb{Z}^d, n,s \geq 0. \tag{5.3}$$

The second part of Lemma 13 implies that $|k_{ln,rs}| < \infty$ and thus $K_{l,n,r,s} \in \boldsymbol{\ell}^\infty(\mathbb{Z}^{4d})$. But in fact $K_{l,n,r,s}$ is even $\boldsymbol{\ell}^1$ because $(\psi_{kn})_{(k,n)\in\mathbb{Z}^d\times\mathbb{N}^d}$ is a basis for $\boldsymbol{S}_0$. Now we define

$$\kappa(K) = \sum_{l,n,r,s\in\mathbb{Z}^d} k_{ln,rs}\ \psi_{ln} \otimes \overline{\psi}_{rs}. \tag{5.4}$$

The sum converges absolutely since

$$\begin{aligned}&\sum_{l,n,r,s\in\mathbb{Z}^d} |k_{ln,rs}\, \psi_{ln} \otimes \overline{\psi}_{rs}| \\ &\leq \|\psi_{ln}\|_{\boldsymbol{S}_0(\mathbb{R}^d)}\, \|\psi_{rs}\|_{\boldsymbol{S}_0(\mathbb{R}^d)} \sum_{l,n,r,s\in\mathbb{Z}^d} |k_{ln,rs}| \\ &\leq \infty.\end{aligned}$$

By Lemma 12 $\psi_{ln} \otimes \overline{\psi}_{rs}$ is an orthonormal Wilson basis for $\boldsymbol{L}^2(\mathbb{R}^{2d})$ and hence a Gelfand triple basis by Lemma 13. Thus we can verify that

$$\begin{aligned}\langle Kf, g\rangle &= \left\langle K\left(\sum_{r,s\in\mathbb{Z}^d} \langle f, \psi_{rs}\rangle\psi_{rs})\right), \sum_{l,n\in\mathbb{Z}^d} \langle g, \psi_{ln}\rangle\psi_{ln}\right\rangle \\ &= \sum_{r,s,l,n\in\mathbb{Z}^d} \langle f, \psi_{rs}\rangle\overline{\langle g, \psi_{ln}\rangle}\langle K\psi_{rs}, \psi_{ln}\rangle \\ &= \sum_{r,s,l,n\in\mathbb{Z}^d} k_{ln,rs}\langle \psi_{ln}, g\rangle\overline{\langle \psi_{rs}, f\rangle} \\ &= \sum_{r,s,l,n\in\mathbb{Z}^d} k_{ln,rs}\langle \psi_{ln} \otimes \overline{\psi}_{rs}, g \otimes \bar{f}\rangle \\ &= \langle \kappa(K), g \otimes \bar{f}\rangle\end{aligned} \tag{5.5}$$

To prove the uniqueness of $\kappa(K)$ assume that there exists a $\kappa'(K)$ with $\langle \kappa(K), \psi_{ln} \otimes \overline{\psi}_{rs}\rangle = \langle K\psi_{rs}, \psi_{ln}\rangle = \langle \kappa'(K), \psi_{ln} \otimes \overline{\psi}_{rs}\rangle$. Then we can calculate

$$\begin{aligned}\langle \kappa(K), \psi_{ln} \otimes \overline{\psi}_{rs}\rangle &= \sum_{i,j,g,h\in\mathbb{Z}^d} k_{ij,gh}\langle \psi_{ij} \otimes \overline{\psi}_{gh}, \psi_{ln} \otimes \overline{\psi}_{rs}\rangle \\ &= \sum_{i,j,g,h\in\mathbb{Z}^d} k_{ij,gh}\ \delta_{ijgh,lnrs} \\ &= k_{ln,rs}.\end{aligned}$$

Now it follows that $k_{ln,rs} = k'_{ln,rs}$, hence $\kappa(K)$ is unique. □

Theorem 22. *Let $\mathcal{B}$ be the space of bounded linear operators from $\boldsymbol{S}_0'(\mathbb{R}^d)$ to $\boldsymbol{S}_0(\mathbb{R}^d)$, which map weak*-convergent sequences $(g_n)_{n\in\mathbb{Z}^d} \in \boldsymbol{S}_0'(\mathbb{R}^d)$ to Cauchy sequences in $\boldsymbol{S}_0(\mathbb{R}^d)$. Then there exists a unique kernel $\kappa(K) \in \boldsymbol{S}_0(\mathbb{R}^d \times \mathbb{R}^d)$, i.e. there is a one-to-one correspondence between the space $\boldsymbol{S}_0(\mathbb{R}^d \times \mathbb{R}^d)$ and the space $\mathcal{B}$.*

Proof. Let $\kappa(K)$ be a function in $\boldsymbol{S}_0(\mathbb{R}^{2d})$, let $(\psi_{rs})_{(r,s)\in\mathbb{Z}^d\times\mathbb{N}^d}$ be a BGT basis. In particular $\kappa(K) \in \boldsymbol{S}_0(\mathbb{R}^{2d}) \subseteq \boldsymbol{L}^2(\mathbb{R}^{2d})$ which implies that there exists a Hilbert-Schmidt operator K with a corresponding matrix defined by

$$\langle K\psi_{rs}, \psi_{lm}\rangle_{\boldsymbol{L}^2(\mathbb{R}^{2d})} = \langle \kappa(K), \psi_{rs} \otimes \psi_{lm}\rangle_{\boldsymbol{L}^2(\mathbb{R}^{2d})}.$$

Since $\psi_{rs} \otimes \psi_{lm}$ is in particular a basis for $\boldsymbol{S}_0(\mathbb{R}^{2d})$ we see that this matrix is in $\boldsymbol{\ell}^1(\mathbb{Z}^{2d})$ and therefore describes the action of an operator from $\boldsymbol{\ell}^\infty(\mathbb{Z}^d)$ to $\boldsymbol{\ell}^1(\mathbb{Z}^d)$. According to Theorem 18 we can now conclude that K is an operator from $\boldsymbol{S}_0'(\mathbb{R}^d)$ to $\boldsymbol{S}_0(\mathbb{R}^d)$.

Conversely assume that K is bounded and linear from $\boldsymbol{S}_0'(\mathbb{R}^d)$ to $\boldsymbol{S}_0(\mathbb{R}^d)$. Then we can define a matrix representation of K by

$$k_{lm,rs} = \langle K\psi_{lm}, \psi_{rs}\rangle.$$

According to Lemma 13 this is an $\boldsymbol{\ell}^1$-matrix and K acts on $f = \sum_{(r,s)\in\mathbb{Z}^d\times\mathbb{N}^d} c_{rs}\psi_{rs}$, $f \in \boldsymbol{S}_0'(\mathbb{R}^d), c \in \boldsymbol{\ell}^\infty$ as

$$Kf = \sum_{l,m,r,s\in\mathbb{Z}^d} k_{lm,rs}\, c_{rs}\, \psi_{lm}.$$

A priori this sum only converges in $\boldsymbol{S}_0(\mathbb{R}^d)$ if f lies in the closed linear span of $(\psi_{rs})_{(r,s)\in\mathbb{Z}^d\times\mathbb{N}^d}$ in $\boldsymbol{S}_0'(\mathbb{R}^d)$ but thanks to the extra condition that K maps weak* convergent sequences to Cauchy sequences the expansion is valid for all $f \in \boldsymbol{S}_0'(\mathbb{R}^d)$. Now we can define our kernel as in (5.4)

$$\kappa(K) = \sum_{l,m,r,s\in\mathbb{Z}^d} k_{lm,rs}\, \psi_{lm} \otimes \overline{\psi}_{rs}\,,$$

hence $\kappa(K)$ is in the space $\boldsymbol{S}_0(\mathbb{R}^{2d})$. The uniquness of the kernel is shown just like in the preceding theorem. □

Remark 25. Note that $\mathcal{B}'$ is the dual space of $\mathcal{B}$. If we combine the preceding theorems with the fact that each $\boldsymbol{L}^2$-kernel defines a Hilbert-Schmidt operator, we notice that the Banach-Gelfand triple $(\mathcal{B}, \mathcal{HS}, \mathcal{B}')$ is isomorphic to the triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \mathbb{R}^d)$.

5.2 Pseudodifferential Operators

In this section we will present a different concept to identify operators with functions on the time-frequency plane: the so-called *Kohn-Nirenberg* correspondence.

Definition 36. For a function or distribution $\sigma \in \boldsymbol{S}_0'(\mathbb{R} \times \widehat{\mathbb{R}})$ we define the corresponding *pseudodifferential operator* with symbol σ as

$$Kf(x) = \int_{\mathbb{R}^d} \sigma(K)(x,\omega)\hat{f}(\omega)e^{2\pi i x\cdot\omega}d\omega. \tag{5.6}$$

The function σ is called the *Kohn-Nirenberg symbol* of K and the mapping $\sigma \mapsto K$ is referred to as the *Kohn-Nirenberg* correspondence. We will often write $\sigma(K)$ to emphasize the relation between the symbol σ and its corresponding operator K.

Before we go any further a motivation for Equation (5.6) is due, because at first sight it seems rather odd. A closer look reveals that the action of K on f is defined as weighted inverse Fourier transform of $\hat{f}$.

The concept of pseudodifferential operators stems from the theory of partial differential equations. The preceding definition is motivated by the following considerations (see [13] p.302). Let $f \in \boldsymbol{C}_c^\infty$ be a smooth function with compact support. A differential operator A acts on f as

$$Af(x) = \sum_{|\alpha|\leq N} \sigma_\alpha(x)D^\alpha f(x) \tag{5.7}$$

where $\sigma_\alpha \in C^\infty$ are non constant coefficients and N is the order of A. Then we can calculate

$$\begin{aligned}
\mathcal{F}(D^\alpha f)(\omega) &= \int_{\mathbb{R}^d} (D^\alpha f)(t)e^{-2\pi i t\cdot\omega}dt \\
&= f(t)e^{-2\pi i t\cdot\omega}\Big|_{-\infty}^{\infty} - \int_{\mathbb{R}^d} (-2\pi i\omega)^\alpha f(t)e^{-2\pi i t\cdot\omega}dt \\
&= (2\pi i\omega)^\alpha \hat{f}(\omega),
\end{aligned}$$

which is true due to f having compact support. Now we substitute in (5.7) and get

$$\begin{aligned}
D^\alpha f(t) &= \mathcal{F}^{-1}(\mathcal{F}(D^\alpha f)(\omega)) \\
&= \int_{\mathbb{R}^d} \hat{f}(\omega)(2\pi i\omega)e^{2\pi i t\cdot\omega}d\omega
\end{aligned}$$

and thus

$$Af(t) = \int_{\mathbb{R}^d} \left(\sum_{|\alpha|\leq N} \sigma_\alpha(t)(2\pi i\omega)^\alpha \right) \hat{f}(\omega)e^{2\pi i t\cdot\omega}d\omega.$$

Now we set $\sigma(A)(x,\omega) = \sum_{|\alpha|\leq N} \sigma_\alpha(x)(2\pi i\omega)^\alpha$ yielding a special case of (5.6).

Example 3. If the symbol σ only depends on ω, i.e. $\sigma(x,\omega) = \hat{h}(\omega)$, then

$$\begin{aligned}
Kf(x) &= \int_{\mathbb{R}^d} \hat{h}(\omega)\hat{f}(\omega)e^{2\pi i x\cdot\omega}d\omega \\
&= \mathcal{F}^{-1}(\hat{h}\hat{f})(x) \\
&= (h * f)(x).
\end{aligned}$$

So K acts as a convolution operator.

Example 4. If σ is of the form $\sigma(x,\omega) = u(x)\hat{v}(\omega)$, then

$$\begin{aligned} Kf(x) &= \int_{\mathbb{R}^d} u(x)\hat{v}(\omega)e^{2\pi i x\cdot\omega}d\omega \\ &= v(x)(v * f)(x), \end{aligned}$$

and so K acts as product convolution (PC) operator.

In the following lemma we quickly introduce the partial Fourier transform, which will be needed later on. We will omit the proof but the interested reader is referred to [9] Lemma 7.3.6.

Lemma 16. *Let f be a function or distribution in $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \mathbb{R}^d)$ then the* partial Fourier transform *with respect to the second variable of f is defined as*

$$(\mathcal{F}_2 f)(x,\omega) := \int_{\mathbb{R}^d} f(x,t)e^{-2\pi i t\cdot\omega}dt.$$

This is a unitary Banach–Gelfand triple isomorphism between $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \mathbb{R}^d)$ and $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R} \times \widehat{\mathbb{R}})$, hence

$$\langle \mathcal{F}_2 f, \mathcal{F}_2 g\rangle_{(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}\times\widehat{\mathbb{R}})} = \langle f, g\rangle_{(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d\times\mathbb{R}^d)}.$$

Theorem 23. *The Kohn-Nirenberg symbol $\sigma(K) \in \boldsymbol{S}_0(\mathbb{R} \times \widehat{\mathbb{R}})$ of an operator K can be expressed by $\kappa(K) \in \boldsymbol{S}_0(\mathbb{R}^{2d})$ as*

$$\sigma(K)(x,\omega) = \int_{\mathbb{R}^d} \kappa(K)(x, x-t)e^{-2\pi i t\cdot\omega}dt \tag{5.8}$$

On the other hand the kernel $\kappa(K)$ can be written by means of σ as

$$\kappa(K)(x,t) = \int_{\widehat{\mathbb{R}^d}} \sigma(K)(x,\omega)e^{2\pi i(x-t)\cdot\omega}d\omega. \tag{5.9}$$

In other words, there is a one-to-one correspondence between the Kohn-Nirenberg symbols $\sigma(K)$ on $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ of operators and their kernels $\kappa(K)$ on $\mathbb{R}^{2d}$ and thus the Kohn-Nirenberg correspondence establishes a unitary Banach-Gelfand triple isomorphism between $(\mathcal{B}, \mathcal{HS}, \mathcal{B}')$ and $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$.

Proof. In (5.6) we substitute the integral formula for $\hat{f}$ and get

$$\begin{aligned} K_\sigma f(x) &= \int_{\mathbb{R}^d} \sigma(K)(x,\omega)\hat{f}(\omega)e^{2\pi i x\cdot\omega}d\omega \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \sigma(K)(x,\omega)e^{2\pi i(x-t)\cdot\omega}f(t)d\omega dt \qquad (5.10) \\ &= \int_{\mathbb{R}^d} \kappa(K)(x,t)f(t)dt, \end{aligned}$$

where $\kappa(K)(x,t) = \int_{\mathbb{R}^d} \sigma(K)(x,\omega)e^{2\pi i(x-t)\cdot\omega}d\omega$. Now let $\mathcal{F}_2$ denote the partial Fourier transform in the second variable and define the shift operator in the second variable as

$$T_x^2 f(v,w) := f(v, v-w). \tag{5.11}$$

This is a self-inverse automorphism on $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \mathbb{R}^d)$. With this notation we can write κ as

$$\kappa(K)(x,t) = T_x^2\, \mathcal{F}_2^{-1}\sigma(K)(x,t),$$

from which it follows that

$$\begin{aligned}\sigma(x,\omega) &= \mathcal{F}_2 T_x^2 \kappa(K)(x,\omega)\\ &= \int_{\mathbb{R}^d} \kappa(x, x-t)e^{-2\pi i t\cdot\omega}dt.\end{aligned}$$

Thus, according to Lemma 16 we have an isomorphism between $\boldsymbol{S}_0(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$ and $\boldsymbol{S}_0(\mathbb{R}^{2d})$, furthermore this also implies an isomorphism between $\boldsymbol{S}_0(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$ and $\mathcal{B}$ according to Theorem 22. In order to show that we can expand this isomorphism to a Banach–Gelfand triple isomporphism it remains to verify that for two operators $K, L \in \mathcal{B}$ the following equation holds:

$$\langle \sigma(K), \sigma(L)\rangle_{\boldsymbol{L}^2(\mathbb{R}\times\widehat{\mathbb{R}})} = \langle \kappa(K), \kappa(L)\rangle_{\boldsymbol{L}^2(\mathbb{R}^{2d})}$$

But this follows immediately as is seen by

$$\begin{aligned}\langle \sigma(K), \sigma(L)\rangle_{\boldsymbol{L}^2(\mathbb{R}\times\widehat{\mathbb{R}})} &= \langle \mathcal{F}_2 T_x^2\, \kappa(K), \mathcal{F}_2 T_x^2\, \kappa(L)\rangle_{\boldsymbol{L}^2(\mathbb{R}\times\widehat{\mathbb{R}})}\\ &= \langle T_x^2\, \kappa(K), T_x^2\, \kappa(L)\rangle_{\boldsymbol{L}^2(\mathbb{R}^{2d})}\\ &= \langle \kappa(K), \kappa(L)\rangle_{\boldsymbol{L}^2(\mathbb{R}^{2d})}.\end{aligned}$$

□

Now that we got to know the KNS, we would like to explore what it is good for. The next theorem states that the KNS can be used to write an operator as sum of time-frequecy shifted versions of a prototype operator P_0.

Theorem 24. *Let the prototype operator P_0 be an operator in $\mathcal{B}'$ with kernel $\kappa(P_0) = \delta_0 \otimes 1$ and $K \in \mathcal{B}$, then*

$$K = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} \sigma(K)(\lambda)(\pi\otimes\pi^*)(\lambda)P_0\, d\lambda \tag{5.12}$$

where $\sigma(K)(\lambda) = \langle K, (\pi\otimes\pi^*)(\lambda)P_0\rangle_{\mathcal{B}}$.

Proof. Let $\lambda = (t,\nu)$. We first verify the latter equation. Therefore we need the kernel of $(\pi\otimes\pi^*)(\lambda)P_0$ and after a simple calculation, we find that

$$\kappa((\pi\otimes\pi^*)(\lambda)P_0)(x,y) = \delta_t(x)e^{2\pi i(x-y)\nu}.$$

With this in mind it now follows easily that

$$\begin{aligned}\langle K, (\pi\otimes\pi^*)(\lambda)P_0\rangle_{\mathcal{B}} &= \langle \kappa(K), \kappa((\pi\otimes\pi^*)(\lambda)P_0)(x,y)\rangle\\ &= \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} \kappa(K)(x,y)\delta_t(x)e^{-2\pi i\nu(x-y)}dxdy\\ &= \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} \kappa(K)(t,y)e^{-2\pi i\nu(t-y)}dy\\ &= \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} \kappa(K)(t,t-z)e^{-2\pi i\nu z}dz\\ &= \sigma(K)(t,\nu).\end{aligned}$$

To show that the integral in (5.12) really represents the operator K, we look at $\kappa(K)$ and find that

$$\begin{aligned}\kappa(K)(x,y) &= \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} \sigma(K)(t,\nu)\delta_t(x)e^{2\pi i\nu(x-y)}dtd\nu\\ &= \int_{\widehat{\mathbb{R}}^d} \sigma(K)(x,\nu)e^{2\pi i\nu(x-y)}d\nu\\ &= \int_{\widehat{\mathbb{R}}^d} \sigma(K)(x,\nu)e^{2\pi i\nu(x-y)}d\nu.\end{aligned}$$

The last line in this equation is just the inversion formula for the KNS (5.9). □

Example 5. As an illustration we will calculate the Kohn-Nirenberg symbol of the rank-one operator $f\otimes g^*$:

$$\begin{aligned}\sigma(f\otimes g^*)(x,\omega) &= \int_{\mathbb{R}^d} \kappa(f\otimes g^*)(x,x-t)e^{-2\pi it\cdot\omega}dt\\ &= \int_{\mathbb{R}^d} f(x)\overline{g(x-t)}e^{-2\pi it\cdot\omega}dt\\ &= f(x)\int_{\mathbb{R}^d} \overline{g(z)}e^{-2\pi i(x-z)\cdot\omega}dz\\ &= f(x)e^{-2\pi ix\omega}\int_{\mathbb{R}^d} \overline{g(z)}e^{2\pi iz\cdot\omega}dz\\ &= f(x)\,\overline{\widehat{g}(\omega)}e^{-2\pi ix\cdot\omega}.\end{aligned}$$

This is the so called Rihaczek distribution of f against g.

A very convenient property of the KNS is derived in the next lemma. It states that time-frequency shifting of the operator translates to time-frequency shifting of its KNS.

Lemma 17. *(Covariance Lemma)*
Let $K\in(\mathcal{B},\mathcal{HS},\mathcal{B}')$, then the action of $\pi\otimes\pi^*(\lambda)$ on K translates to a time-frequency shift of the Kohn-Nirenberg symbol $\sigma(K)$, i.e.

$$\sigma((\pi\otimes\pi^*)(\lambda)K) = T_\lambda\sigma(K).$$

Proof. Note that $(\pi\otimes\pi^*)(\lambda)(K)$ is an automorphism on $(\mathcal{B},\mathcal{HS},\mathcal{B}')$ and thus with the help of (5.12) we get

$$\begin{aligned}(\pi\otimes\pi^*)(\lambda)(K) &= \int_{\mathbb{R}\times\widehat{\mathbb{R}}} \sigma(K)(\mu)(\pi\otimes\pi^*)(\lambda)(\pi\otimes\pi^*)(\mu)P_0\,d\mu\\ &= \int_{\mathbb{R}\times\widehat{\mathbb{R}}} \sigma(K)(\mu)(\pi\otimes\pi^*)(\lambda+\mu)P_0\,d\mu\\ &= \int_{\mathbb{R}\times\widehat{\mathbb{R}}} T_\lambda\sigma(K)(\nu)(\pi\otimes\pi^*)(\nu)P_0\,d\nu.\end{aligned}$$

□

5.3 Spreading representation

We start this final chapter on operator quantization with the definition of the *symplectic Fourier transform* which is an operator on functions which live on the time-frequency plane $\mathbb{R}^d\times\widehat{\mathbb{R}}^d$.

Definition 37. (Symplectic Fourier Transform) Let $\mathcal{I}$ denote an isomorphism from $\mathbb{R}^d\times\widehat{\mathbb{R}}^d$ to $\widehat{\mathbb{R}}^d\times\mathbb{R}^d$ given by

$$\mathcal{I} := \begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}, \qquad \mathcal{I}^* = \begin{pmatrix} 0 & -1\\ 1 & 0\end{pmatrix}. \tag{5.13}$$

The *symplectic fourier transform* on $\mathbb{R}^d\times\widehat{\mathbb{R}}^d$ is defined as

$$\mathcal{F}_s f(\lambda) = \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} f(\eta)e^{-2\pi i\lambda\cdot\mathcal{I}^*\eta}\,d\eta \quad \lambda\in\mathbb{R}\times\widehat{\mathbb{R}}, \tag{5.14}$$

or, if we set $\lambda=(x,\omega)$ and $\eta=(t,\xi)$

$$\begin{aligned}\mathcal{F}_s f(x,\omega) &= \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} f(t,\xi)e^{-2\pi i\langle(x,\omega),(-\xi,t)\rangle}dtd\xi\\ &= \int_{\mathbb{R}^d\times\widehat{\mathbb{R}}^d} f(t,\xi)e^{-2\pi i(t\omega-x\xi)}dtd\xi \qquad x,t\in\mathbb{R}^d,\ \xi,\omega\in\widehat{\mathbb{R}}^d.\end{aligned} \tag{5.15}$$

Remark 26. Note that this means, that the symplectic Fourier transform is just the usual Fourier transform on the time-frequency plane rotated by 90°, i.e

$$\mathcal{F}_s f(x,\omega) = \mathcal{F}f(\omega,-x) = \mathcal{F}(\mathcal{I}(x,\omega)).$$

This implies that the symplectic Fourier transform is a Banach-Gelfand triple automorphism on the Gelfand triple $(\boldsymbol{S}_0,\boldsymbol{L}^2,\boldsymbol{S}_0')(\mathbb{R}^d\times\widehat{\mathbb{R}}^d)$. Furthermore we observe, that $\mathcal{F}_s$ is self-inverse since

$$\begin{aligned}\mathcal{F}_s(\mathcal{F}_s f)(x,\xi) &= \mathcal{F}(\mathcal{F}_s f)(\xi,-x)\\ &= \mathcal{F}^{-1}(\mathcal{F}_s f)(-\xi,x)\\ &= \mathcal{F}^{-1}\mathcal{F}f(x,\xi)\\ &= f(x,\xi).\end{aligned}$$

At the moment we already know two different "views" on pseudodifferential operators. On the one hand we have the definition of such an operator K with symbol $\sigma(K)$, on the other hand we can think of them as integral operators with some kernel $\kappa(K)$ which can be calculated from $\sigma(K)$.

The symplectic Fourier transform allows us to introduce yet another "view" on pseudodifferental operators which turns out to be very natural in the context of time-frequency analysis: the so called *spreading representation* of K, which allows us to describe an operator as superposition of weighted time-frequency shifts. This is motivated by the following calculation:

$$\begin{aligned}\kappa(K)(x,y) &= \mathcal{F}_2^{-1}\sigma(x,x-y)\\ &= \mathcal{F}_1\mathcal{F}_1^{-1}\mathcal{F}_2^{-1}\sigma(x,x-y)\\ &= \mathcal{F}_1\mathcal{F}^{-1}\sigma(x,x-y)\\ &= \mathcal{F}_1\mathcal{F}\sigma(-x,y-x)\\ &= \mathcal{F}_1^{-1}\hat{\sigma}(x,y-x).\end{aligned} \tag{5.16}$$

Switching to integrals we thus get

$$\begin{aligned}\kappa(K)(x,y) &= \int_{\mathbb{R}^d} \hat{\sigma}(\nu,y-x)e^{2\pi i\nu\cdot x}d\nu \qquad (5.17)\\ &= \int_{\mathbb{R}^d} \mathcal{F}_s\sigma(x-y,\nu)e^{2\pi i\nu\cdot x}d\nu.\end{aligned}$$

Now we define $\eta(K)(x,\nu) := \mathcal{F}_s\sigma(x,\nu)$ and write K as integral operator substituting (5.17) for $\kappa(K)$ which yields

$$\begin{aligned}Kf(x) &= \int_{\mathbb{R}^d} \kappa(K)(x,y)f(y)dy\\ &= \int_{\mathbb{R}^d}\int_{\widehat{\mathbb{R}}^d} \eta(K)(x-y,\nu)e^{2\pi i\nu\cdot x}f(y)d\nu dy\\ &= \int_{\mathbb{R}^d}\int_{\widehat{\mathbb{R}}^d} \eta(K)(t,\nu)f(x-t)e^{2\pi i\nu\cdot x}d\nu dt\\ &= \int_{\mathbb{R}^d}\int_{\widehat{\mathbb{R}}^d} \eta(K)(t,\nu)M_\nu T_t f(x)dtd\nu.\end{aligned} \tag{5.18}$$

The function $\eta(K)$ is the so-called *spreading function* of K. Equation (5.18) also gives a nice link to applications. Assume we want to send the signal f from A to B. Depending on the distance between A and B and the medium through which we send f, the signal will undergo some amount of transformation. So, instead of f the receiver B will get the distorted signal Kf. The above equation states that the action of K can be considered as a superposition of time-frequency shifts.

Theorem 25. *Let* $K \in (\mathcal{B},\mathcal{HS},\mathcal{B}')$*, then the* spreading function $\eta(K)$ *is defined on* $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ *as*

$$\eta(K) := \mathcal{F}_s(\sigma(K)). \tag{5.19}$$

$\eta(K)$ is related to the kernel $\kappa(K)$ of K via

$$\eta(K)(x,\nu) = \int_{\mathbb{R}^d} \kappa(K)(t,t-x)e^{-2\pi i t\cdot\nu}dt \tag{5.20}$$

and

$$\kappa(K)(x,y) = \int_{\widehat{\mathbb{R}}^d} \eta(K)(x-y,\nu)e^{2\pi i\nu\cdot x}d\nu. \tag{5.21}$$

Consequently the mapping $K \mapsto \eta(K)$ is a unitary Banach-Gelfand triple isomorphism between $(\mathcal{B}, \mathcal{HS}, \mathcal{B}')$ and $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d \times \widehat{\mathbb{R}}^d)$.

Proof. First we need to recall that the Kohn-Nirenberg Symbol of an operator K can be written as $\sigma(K)(t,\nu) = \mathcal{F}_2 T_x^2 \kappa(K)(t,\nu)$. Then a straight forward calculation yields (5.19):

$$\begin{aligned}\mathcal{F}_s(\sigma(K))(x,\nu) &= \mathcal{F}_s\mathcal{F}_2 T_x^2\kappa(K)(x,\nu)\\ &= \mathcal{F}\mathcal{F}_2 T_x^2\kappa(K)(\nu,-x)\\ &= \mathcal{F}_1^{-1} T_x^2\kappa(K)(\nu,-x)\end{aligned}$$

and thus

$$\begin{aligned}\eta(K)(x,\nu) &= \int_{\mathbb{R}^d} T_x^2\kappa(K)(t,x)e^{-2\pi i\nu t}dt\\ &= \int_{\mathbb{R}^d} \kappa(K)(t,t-x)e^{-2\pi i\nu t}dt.\end{aligned}$$

The inverse formula was already derived in (5.17). And the Banach-Gelfand triple isomorphism follows from the properties of the symplectic Fourier transform. □

Example 6. We compute the spreading function of a rank-one operator $f \otimes g^*$,

$$\begin{aligned}\eta(f\otimes g^*) &= \int_{\mathbb{R}^d} \kappa(f\otimes g^*)(t,t-x)e^{-2\pi i t\cdot\nu}dt\\ &= \int_{\mathbb{R}^d} f(t)\overline{g(t-x)}e^{-2\pi i t\cdot\nu}dt\\ &= V_g f(x,\nu)\end{aligned}$$

and find that it is simply the STFT of f with respect to window g.

5.4 Gabor Multiplier

In this last section on operators we will tap into the Gabor setting by defining Gabor multipliers. The action of this kind of operators is given by multiplying the coefficient sequence in the Gabor expansion of a function with some sequence on the lattice and thus manipulating its time-frequency content, an operation which is commonly known in signal processing as *filtering of a signal.* For a closer discussion of the topics in this section we refer the reader to [3], [6], [10] and [16]. We begin by formally defining Gabor multipliers.

Definition 38. Let g_1 and g_2 be functions in $\boldsymbol{L}^2(\mathbb{R}^d)$, let Λ be a time-frequency lattice in $\mathbb{R}^d \times \widehat{\mathbb{R}}^d$ and $(m(\lambda))_{\lambda\in\Lambda}$ a complex-valued sequence on Λ. Then the *Gabor multiplier* with *upper symbol* $\mathbf{m}(\lambda)$ is defined as

$$G_m(f) = G_{g_1,g_2,\Lambda,m}(f) = \sum_{\lambda\in\Lambda} m(\lambda)\langle f, \pi(\lambda)g_1\rangle \pi(\lambda)g_2.$$

If $g_1 = g = g_2$ we will simply write $G_{g,\Lambda,\mathbf{m}}$, or, even shorter, $G_{\mathbf{m}}$. With this assumption we can also define $P_\lambda := \pi(\lambda)g \otimes \pi(\lambda)g^*$, this is the projection operator on the one dimensional subspace spanned by time-frequency shifted versions of g. With this notation $G_{\mathbf{m}}$ now looks like

$$G_{\mathbf{m}} = \sum_{\lambda\in\Lambda} m(\lambda)P_\lambda.$$

It is clear from the definition that the choice of the analysis window g_1, the synthesis window g_2 and the symbol $\mathbf{m}(\lambda)$ will determine the mapping properties of $G_{g,\Lambda,\mathbf{m}}$ between different spaces. In the case where $g_1 = g = g_2$, we see that $G_{g,\Lambda,\mathbf{m}}$ is just a linear combination of projection operators P_λ with coefficients $\mathbf{m}(\lambda)$ onto one-dimensional subspaces of $\boldsymbol{L}^2(\mathbb{R}^d)$. The following theorem sums up the basic mapping properties of $G_{\mathbf{m}}$ in dependence on the symbol $\mathbf{m}$.

Theorem 26. *(Properties of $G_{g,\Lambda,\mathbf{m}}$)*
Let $g_1, g_2 \in \boldsymbol{S}_0(\mathbb{R}^d)$ and $\Lambda \in \mathbb{R}^d \times \widehat{\mathbb{R}}^d$ be some lattice, then

1. *$G_{\mathbf{m}}$ is a bounded operator on $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')$ whenever $\mathbf{m} \in \boldsymbol{\ell}^\infty(\Lambda)$.*

2. *$G_{\mathbf{m}}$ is in $\mathcal{HS}$ whenever $\mathbf{m} \in \boldsymbol{\ell}^2(\Lambda)$.*

3. *$G_{\mathbf{m}}$ is in $\mathcal{B} = \mathcal{L}(\boldsymbol{S}_0', \boldsymbol{S}_0)$ whenever $\mathbf{m} \in \boldsymbol{\ell}^1(\Lambda)$.*

Or, reformulated in a more compact fashion, the mapping $(\mathbf{m}(\lambda))_{\lambda\in\Lambda} \mapsto G_{\mathbf{m}}$ maps the Gelfand-Triple $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda)$ into the Gelfand-Triple $(\mathcal{B}, \mathcal{HS}, \mathcal{B}')$.

Proof. To prove this theorem we write

$$G_{\mathbf{m}} = D_{g_2} \circ M \circ C_{g_1},$$

where C_{g_1} is the analysis operator with respect to g_1, D_{g_2} is the synthesis operator with respect to g_2 and M is defined as $M : c \mapsto (m(\lambda)c(\lambda))_{\lambda\in\Lambda}$ for sequences $c = (c(\lambda))_{\lambda\in\Lambda}$. $g_1, g_2 \in \boldsymbol{S}_0(\mathbb{R}^d)$ in particular implies that they are Bessel atoms, i.e. establishing a Bessel sequence for any TF-lattice Λ and thus $G_{\mathbf{m}} = D_{g_2} \circ M \circ C_{g_1}$ is bounded whenever M is bounded. For $\mathbf{m}$ in the various sequence spaces we have that

- $M : (\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda) \longrightarrow (\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda)$ whenever $\mathbf{m} \in \boldsymbol{\ell}^\infty(\Lambda)$ and $\|M\| \leq \|\mathbf{m}\|_\infty$

- $M : (\boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda) \longrightarrow (\boldsymbol{\ell}^2, \boldsymbol{\ell}^1)(\Lambda)$ whenever $\mathbf{m} \in \boldsymbol{\ell}^2(\Lambda)$ and $\|M\| \leq \|\mathbf{m}\|_2$
- $M : \boldsymbol{\ell}^\infty(\Lambda) \longrightarrow \boldsymbol{\ell}^1(\Lambda)$ whenever $\mathbf{m} \in \boldsymbol{\ell}^1(\Lambda)$ and $\|M\| \leq \|\mathbf{m}\|_1$,

which follows from basic inequalities concerning pointwise products of sequences. The rest of the theorem follows from the mapping properties of D_g and C_g between $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')$ and $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)$ (see Theorem 15 and 16). □

Definition 39. Let $(G\mathcal{M}_1, G\mathcal{M}_2, G\mathcal{M}_\infty)$ denote the Gelfand triple of Gabor multipliers with upper symbol $\mathbf{m} \in (\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda)$ respectively. Then the preceding theorem can be summarized as $(G\mathcal{M}_1, G\mathcal{M}_2, G\mathcal{M}_\infty) \subseteq (\mathcal{B}, \mathcal{HS}, \mathcal{B}')$.

With these results in mind the question arises, wether it is possible to retrieve the symbol $\mathbf{m}$ from a given Gabor multiplier $G_{\mathbf{m}}$. This is covered in the following theorem.

Theorem 27. *(Best approximation by Gabor multipliers)*

1. *$(P_\lambda)_{\lambda\in\Lambda}$ is a Riesz basis for $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^d)$ if and only if the Λ- Fourier transform of $(|V_g g(\lambda)|^2)_{\lambda\in\Lambda}$ is free of zeros.*
2. *There exists a canonical biorthogonal family $(Q_\lambda)_{\lambda\in\Lambda} \subseteq span((P_\lambda)_{\lambda\in\Lambda})$, where $Q_\lambda = \pi(\lambda) Q \pi^*(\lambda)$ and Q is an operator with kernel in $\boldsymbol{S}_0(\mathbb{R}^d)$.*
3. *For any operator $T \in (\boldsymbol{B}, \mathcal{H}, \boldsymbol{B}')$ the best approximation by Gabor multipliers in $(G\mathcal{M}_1, G\mathcal{M}_2, G\mathcal{M}_\infty)$ is given by*

$$PG(T) = \sum_{\lambda\in\Lambda} \langle T, Q_\lambda \rangle P_\lambda. \tag{5.22}$$

In particular the mapping $\mathbf{m}(\lambda) \mapsto G_{\mathbf{m}}$ is invertible between $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda)$ and $(\mathcal{M}G_1, \mathcal{M}G_2, \mathcal{M}G_\infty) \subseteq (\mathcal{B}, \mathcal{HS}, \mathcal{B}')$. Explicitly, the symbol $\mathbf{m}$ is given by

$$(m(\lambda))_{\lambda\in\Lambda} = (\langle G_{\mathbf{m}}, Q_\lambda \rangle)_{\lambda\in\Lambda}.$$

Proof. To prove the first part of the theorem we have to verify that for constants $A, B > 0$

$$A\|c\|_2^2 \leq \| \sum_{\lambda\in\Lambda} c(\lambda) P_\lambda \|_{\mathcal{HS}}^2 \leq B\|c\|_2^2.$$

Since g is in $\boldsymbol{S}_0(\mathbb{R}^d)$ and thus $(g_\lambda)_{\lambda\in\Lambda}$ a Bessel sequence, the upper bound exists. For the lower bound we first notice that

$$\| \sum_{\lambda\in\Lambda} c(\lambda) P_\lambda \|_{\mathcal{HS}}^2 = \sum_{\lambda\in\Lambda} \sum_{\lambda'\in\Lambda'} c(\lambda) \overline{c(\lambda')} \langle P_\lambda, P_{\lambda'} \rangle \tag{5.23}$$

and after a quick calculation (see [16] for details) we find that

$$\langle P_\lambda, P_{\lambda'} \rangle = V_g g(\lambda' - \lambda). \tag{5.24}$$

If we insert this into (5.23) writing $\varphi(\lambda) = |V_g g(\lambda)|^2$ we get

$$\begin{aligned}\|\sum_{\lambda\in\Lambda} c(\lambda)P_\lambda\|_{\mathcal{HS}} &= \sum_{\lambda\in\Lambda} c(\lambda) \sum_{\lambda'\in\Lambda'} \overline{c(\lambda')}\varphi(\lambda'-\lambda)\\ &= \langle c *_\Lambda \varphi, c\rangle_{\boldsymbol{\ell}^2(\Lambda)}\\ &= \langle \hat{c}\cdot\hat{\varphi}, \hat{c}\rangle_{\boldsymbol{\ell}^2(\Lambda)}\\ &= \int_\Lambda \hat{\varphi}(\lambda)|\hat{c}(\lambda)|^2 d\lambda.\end{aligned}$$

This last expression is bounded below by $A\|c\|_2^2$ if and only if $\hat{\varphi}(\lambda) = \mathcal{F}_\Lambda(|V_g g|^2) \geq A > 0$, since by Plancherel's Theorem (2.2) $\int_\Lambda |\hat{c}(\lambda)|^2 d\lambda = \|c\|_2^2$.

Now for the second statement of the theorem: Since $(P_\lambda)_{\lambda\in\Lambda}$ is a Riesz basis, there exists a biorthogonal system $(Q_\lambda)_{\lambda\in\Lambda}$ and we only need to show that it is of the claimed form $Q_\lambda = \pi(\lambda)\otimes Q\otimes\pi^*(\lambda)$ with generating operator $Q \in \mathcal{B}$, i.e. $\kappa(Q) \in \boldsymbol{S}_0(\mathbb{R}^d)$. The last claim follows from Theorem 3.6 in [6] whenever $g \in \boldsymbol{S}_0(\mathbb{R}^d)$. With the help of Theorem 17 then follows that

$$\begin{aligned}\langle \pi(\lambda)\otimes Q\otimes\pi^*(\lambda), P_{\lambda'}\rangle &= \langle T_\lambda\sigma(Q), T_{\lambda'}\sigma(P)\rangle\\ &= \langle \sigma(Q), T_{\lambda'-\lambda}\sigma(P)\rangle\\ &= \langle Q, P_{\lambda'-\lambda}\rangle\\ &= \delta_{\lambda,\lambda'}\end{aligned}$$

Since the biorthogonal basis is unique, this proves the second part of the theorem.

Now for $T \in \mathcal{HS}$ the best approximation in the space $G\mathcal{M}_2$ is given by the orthogonal projection onto that space, i.e. by

$$PG(T) = \sum_{\lambda\in\Lambda}\langle T, Q_\lambda\rangle P_\lambda = \sum_{\lambda\in\Lambda}\langle T, P_\lambda\rangle Q_\lambda. \tag{5.25}$$

Since P_λ and Q_λ are in $\mathcal{B}$, this equation can be extended to the whole Gelfand triple $(\mathcal{B}, \mathcal{HS}, \mathcal{B}')$. Finally, by Theorem 15, the analysis operator C_g maps the operator kernels from $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')(\mathbb{R}^{2d})$ to $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda)$ and thus follows the last claim of the theorem.

□

Corollary 7. *If we define the mapping* $\beta : T \mapsto (\langle T, Q_\lambda\rangle)_{\lambda\in\Lambda}$*, then* β *is a surjective and bounded mapping from* $(\mathcal{B}, \mathcal{HS}, \mathcal{B}')$ *onto* $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda)$*.*

Proof. The surjectivity follows from the fact that every $\mathbf{m}$ in $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda)$ defines a Gabor multiplier in $(G\mathcal{M}_1, G\mathcal{M}_2, G\mathcal{M}_\infty) \subseteq (\mathcal{B}, \mathcal{HS}, \mathcal{B}')$. □

If we review Equation (5.25) once more, we find that we can rewrite the coefficients of the last sum as

$$\langle T, P_\lambda\rangle = \int\int_{\mathbb{R}^{2d}} \kappa(T)(x,y) g_\lambda(x)\overline{g_\lambda(y)}dxdy = \langle Tg_\lambda, g_\lambda\rangle. \tag{5.26}$$

We name this result in the following definition:

Definition 40. (Lower symbol) The *lower symbol* of an arbitrary operator $T \in (\mathcal{B}, \mathcal{HS}, \mathcal{B}')$ with respect to window g and lattice Λ is given by

$$\sigma_L(T)(\lambda) = \langle T g_\lambda, g_\lambda \rangle.$$

Lemma 18. *The lower symbol $\sigma_L(G)$ of a Gabor multiplier G is related to the upper symbol $\mathbf{m}$ via a bounded, invertible linear mapping on $(\boldsymbol{\ell}^1, \boldsymbol{\ell}^2, \boldsymbol{\ell}^\infty)(\Lambda)$,*

$$\sigma_L = |V_g g|^2 *_\Lambda \mathbf{m}.$$

Proof. Let $G = \sum_{\lambda \in \Lambda} m(\lambda) P_\lambda$, then

$$\begin{aligned}
\sigma_L(G)(\lambda') &= \left\langle \sum_{\lambda \in \Lambda} m(\lambda) P_\lambda g_{\lambda'}, g_{\lambda'} \right\rangle \\
&= \sum_{\lambda \in \Lambda} m(\lambda) \langle P_\lambda g_{\lambda'}, g_{\lambda'} \rangle \\
&= \sum_{\lambda \in \Lambda} m(\lambda) \langle P_\lambda, P_{\lambda'} \rangle \\
&= \sum_{\lambda \in \Lambda} |V_g g(\lambda' - \lambda)|^2 m(\lambda) \\
&= (|V_g g|^2 *_\Lambda \mathbf{m})(\lambda').
\end{aligned}$$

□

A look at Figure 5 sums up the results of this section.

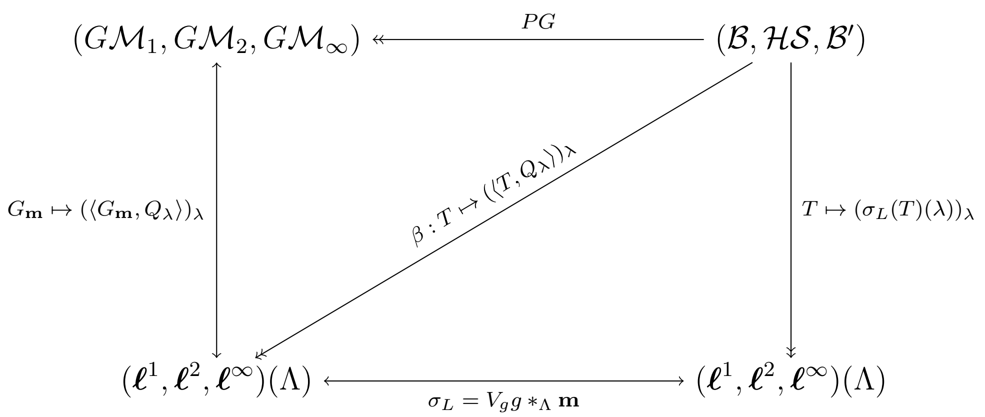


Figure 5: The relationships between the different Gelfand triples involved when we talk about Gabor multipliers.

5.5 Composition of operators

Finally we will briefly discuss the composition of operators. The previous discussion in this section suggests that operator composition corresponds to some sort of composition of their kernels (or symbols). Indeed, if the operators in question both have kernels in $\boldsymbol{S}_0(\mathbb{R}^d)$ then their composition amounts to the continuous analogon of matrix multiplication.

Lemma 19. *Let T_1 and T_2 be operators from $\boldsymbol{S}_0'(\mathbb{R}^d)$ to $\boldsymbol{S}_0(\mathbb{R}^d)$, then the kernel of the composition $T_2 \circ T_1$ is given by*

$$K(x,y) = \int_{\mathbb{R}^d} K_2(x,s)K_1(s,y)ds, \tag{5.27}$$

where $K_1, K_2 \in \boldsymbol{S}_0(\mathbb{R}^{2d})$ are the kernels of T_1, T_2 respectively.

Proof. The proof follows easily by a short calculation. □

Remark 27. Formula (5.27) also holds true if one of the kernels is an element of $\boldsymbol{L}^\infty(\mathbb{R}^{2d})$ since $\boldsymbol{S}_0 \cdot \boldsymbol{L}^\infty \subseteq \boldsymbol{S}_0$.

When dealing with more general operators, e.g. with kernels in $\boldsymbol{L}^2$ or distributional kernels, we cannot expect to get away with Equation (5.27), since this product need not be properly defined. Instead we will apply regularization techniques as discussed in Section 4.3 to get approximations of the operators in question before calculating the kernel of the composite mapping.

Lemma 20. *Let T_1, T_2 be linear mappings on $\boldsymbol{S}_0'(\mathbb{R}^d)$, then for each regularizing sequence A_n the operators $A_n \circ T_1$ and $A_n \circ T_2$ map $\boldsymbol{S}_0'(\mathbb{R}^d)$ in $\boldsymbol{S}_0(\mathbb{R}^d)$. Thus they can be composed at the kernel level as in Lemma 19 such that the kernel of the composition $A_n \circ T_2 \circ A_n \circ T_1$ is weak* convergent to the kernel of $T_2 \circ T_1$ and the resulting operator converges pointwise to the action of $T_2 \circ T_1$.*

Proof. Let K_n denote the kernel of the composite operator $A_n \circ T_2 \circ A_n \circ T_1$, K the kernel of $T_2 \circ T_1$ and K_1, K_2 the kernels of T_1, T_2 respectively. As $n \to \infty$

$$\|T_2(T_1\sigma) - A_n(T_2(A_n(T_1\sigma)))\|_{\boldsymbol{S}_0'} \to 0.$$

This shows the pointwise convergence of the operators which implies the weak* convergence at the kernel level. □

We will close this work with the diagram shown in Figure 6, which gives a rough overview of the relationships between the spaces discussed within the preceding pages.

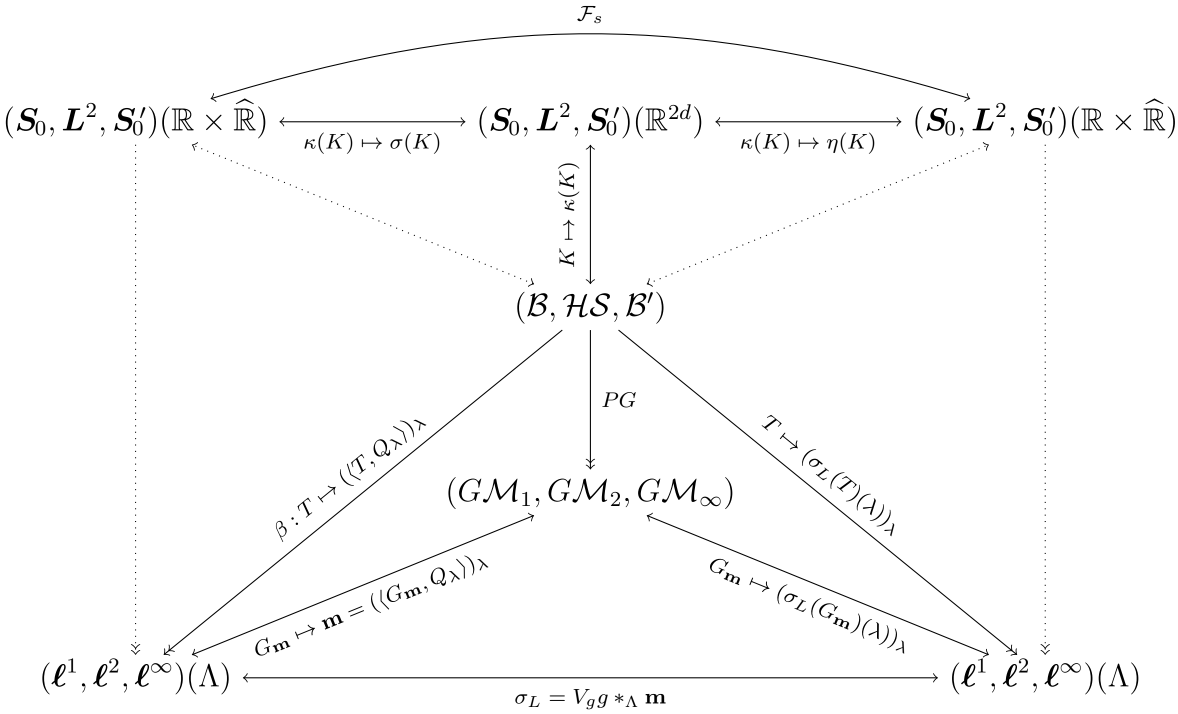


Figure 6: The big picture.

References

[1] O. Christensen. *Frames and Bases. An Introductory Course.* Applied and Numerical Harmonic Analysis. Basel Birkhäuser, 2008.

[2] E. Cordero, H. G. Feichtinger, and F. Luef. Banach Gelfand triples for Gabor analysis. In *'Pseudo-Differential Operators, Quantization and Signals', C.I.M.E. 2006*, Lecture Notes in Mathematics. Springer, 2007.

[3] M. Dörfler. *Gabor Analysis for a Class of Signals called Music.* PhD thesis, 2002.

[4] H. G. Feichtinger. On a new Segal algebra. *Monatsh. Math.*, 92:269–289, 1981.

[5] H. G. Feichtinger. Banach convolution algebras of Wiener type. In *Proc. Conf. on Functions, Series, Operators, Budapest 1980*, volume 35 of *Colloq. Math. Soc. Janos Bolyai*, pages 509–524. North-Holland, Amsterdam, Eds. B. Sz.-Nagy and J. Szabados. edition, 1983.

[6] H. G. Feichtinger. Spline-type spaces in Gabor analysis. In D. X. Zhou, editor, *Wavelet Analysis: Twenty Years Developments Proceedings of the International Conference of Computational Harmonic Analysis, Hong Kong, China, June 4–8, 2001*, volume 1 of *Ser. Anal.*, pages 100–122. World Sci.Pub., River Edge, NJ, 2002.

[7] H. G. Feichtinger and K. Gröchenig. A unified approach to atomic decompositions via integrable group representations. *Lect. Notes in Math.*, 1302:52–73, 1988.

[8] H. G. Feichtinger, K. Gröchenig, and D. F. Walnut. Wilson bases and modulation spaces. *Math. Nachr.*, 155:7–17, 1992.

[9] H. G. Feichtinger and W. Kozek. Quantization of TF lattice-invariant operators on elementary LCA groups. In H. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms. Theory and Applications.*, Applied and Numerical Harmonic Analysis, pages 233–266, 452–488, Boston, MA, 1998. Birkhäuser Boston.

[10] H. G. Feichtinger and K. Nowak. A first survey of Gabor multipliers. In H. Feichtinger and T. Strohmer, editors, *Advances in Gabor Analysis*, Appl. Numer. Harmon. Anal., pages 99–128. Birkhäuser, 2003.

[11] H. G. Feichtinger and G. Zimmermann. A Banach space of test functions for Gabor analysis. In H. Feichtinger and T. Strohmer, editors, *Gabor Analysis and Algorithms: Theory and Applications*, Applied and Numerical Harmonic Analysis, pages 123–170, Boston, MA, 1998. Birkhäuser Boston.

[12] I. M. Gel'fand and N. Y. Vilenkin. *Generalized Functions, Vol. 4: Applications of Harmonic Analysis.* Academic Press, New York and London, 1964.

[13] K. Gröchenig. *Foundations of Time-Frequency Analysis.* Appl. Numer. Harmon. Anal. Birkhäuser Boston, Boston, MA, 2001.

[14] C. Heil. An introduction to weighted Wiener amalgams. In M. Krishna, R. Radha, and S. Thangavelu, editors, *Wavelets and their Applications (Chennai, January 2002)*, pages 183–216. Allied Publishers, New Delhi, 2003.

[15] W. Rudin. *Functional Analysis.* McGraw-Hill Book Company, New York, 1973.

[16] K. Schnass. Gabor Multipliers. A self-contained survey. Master's thesis, University of Vienna, 2004.

[17] D. Werner. *Funktionalanalysis. 6., korrigierte Auflage.* Springer Verlag, 2007.

Index

A Deutsche Zusammenfassung

Ziel der vorliegenden Arbeit ist es, *Banach-Gelfand Triple* im Kontext der Zeitfrequenzanalyse vorzustellen. Banach-Gelfand triple verbinden die Eigenschaften von Hilberträumen mit denen von allgemeinen Banachräumen von Distributionen. Ein Banach-Gelfand Triple ist ein Triple von Räumen, bestehend aus einem Hilbertraum $\mathcal{H}$ in den ein (kleinerer) Banachraum $\boldsymbol{B}$ eingebettet liegt. $\mathcal{H}$ selbst wiederum ist enthalten in dem Banachraum $\boldsymbol{B}'$, dem Dualraum von $\boldsymbol{B}$. Wir werden die Rolle aufzeigen, die Banach-Gelfand Triple als grundlegendes Konzept in der Zeit-Frequenz Analyse spielen.

Kapitel 2 gibt eine kurze Einführung in das Themengebiet. Fouriertransformation, Modulations- und Verschiebungsoperator, die Kurzzeitfouriertransformation sowie Gaussfunktionen werden vorgestellt und deren, für diese Arbeit wichtigen, Eigenschaften bewiesen. In Kapitel 2.2 werden Gabor Frames eingeführt und diskutiert. Sie erlauben eine Diskretisierung der Kurzzeitfouriertransformation.

Kapitel 3 leitet den Hauptteil dieser Arbeit ein. Banach-Gelfand Triple werden allgemein definiert und wichtige Eigenschaften bewiesen. Kapitel 4 widmet sich im Anschluss daran dem Gelfand Triple $(\boldsymbol{S}_0, \boldsymbol{L}^2, \boldsymbol{S}_0')$, das eine Schlüsselrolle in der hier vorgestellten Form der Zeit-Frequenz Analyse spielt. In Kapitel 5 werden schließlich unterschiedliche Klassen der Operatordarstellung besprochen. Es werden Methoden vorgestellt, um Operatoren mit Funktionen auf der Zeit-Frequenz Ebene zu identifizieren.

B Curriculum Vitae

Severin Bannert

Geburtsdatum:	22. Dezember 1982
Geburtsort:	Wien, Österreich
Staatsangehörigkeit:	Österreich
Eltern:	Dr. Herbert Bannert und Dr. Margaretha Bannert
Kontakt:	severin.bannert@gmail.com

2003 – 2010	Diplomstudium Mathematik, Universität Wien.
2002 – 2003	Diplomstudium Molekularbiologie, Universität Wien.
2001	Matura.
1993 – 2001	GRG 23, Gymnasium.
1989 – 1993	Volksschule Alt-Erlaa.

C Acknowledgments

First and foremost I want to thank my girlfriend and partner in life Gloria for her love, patience and steady support. I also thank my parents for giving me the financial possibilities to pursue my studies and of course Prof. Feichtinger for offering me substantial advice in the course of this work. Last but not least, many thanks to my friends.