MA8104 Wavelets

Franz Luef

September 20, 2021

Abstract

These notes is based on a course in applied harmonic analysis at NTNU Trondheim in the fall 2021.

1 Introduction

The role of *Applied Harmonic Analysis* is to provide great tools for representing and processing your data/functions, where these tools are (1) efficient, (2) useful, (3) adapted to the task, and (4) stable and computationally fast.

Concretely, we are going to spend some time on the construction of building blocks $\{g_j\}_{j\in J}$ such that the function/signal f possesses the properties (1)-(4):

$$f = \sum_{j\in J} a_j\, g_j,$$

for an appropriate class of functions.

We are also looking into different ways to represent/analyze a signal f in some space x onto some domain Y via a map T such that $T(f) \in Y$, e.g. the short-time Fourier transform of an audio signal and the wavelet transform of an image. In short, we are often encounter the following strategy in finding a representation of f:

- Pick an appropriate class of functions $\{g_j\}_{j\in J}$.
- Analyse the function f with respect to the $\{g_j\}_{j\in J}$: $f \mapsto \langle f, g_j\rangle$.
- Synthesize / take linear combinations of the building blocks $\{g_j\}_{j\in J}$ such that $\sum_{j\in J} a_j g_j$ is a good approximation or representation of f.
- Find a way to quantify in which sense you would like to have your approximation/representation. In our case this will also amount to the choice of an appropriate Banach space $(X, \|.\|)$.

2 From Fourier analysis to time-frequency representations

2.1 Fourier transform

Let us start describing a setting that is suitable for Fourier analysis. We denote by $L^1(\mathbb{R}^d)$ the space of Lebesgue integrable functions f on $\mathbb{R}^d$ satisfying

$$\|f\|_1 = \int_{\mathbb{R}^d} |f(x)|dx < \infty.$$

For $f \in L^1(\mathbb{R}^d)$ we define the *Fourier transform* by

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

which we might denote by $\mathcal{F}(f)$, when we want to understand the Fourier transform as a linear transform between two function spaces.

Note that for a $f \in L^1(\mathbb{R}^d)$ its Fourier transform $\hat{f}$ is absolutely integrable and it is a continuous and bounded functions. Moreover, one can show that the Fourier transform $\mathcal{F}$ maps $L^1(\mathbb{R}^d)$ into $C_0(\mathbb{R}^d)$, which is the space of continuous functions vanishing at infinity with respect to $\|.\|_\infty$. This elementary statement is due to Riemann and Lebesgue. A non-rigorous way to formulate is that the high-frequency components of an L^1-function vanish.

For $f, g \in L^1(\mathbb{R}^d)$ we define its *convolution* $f \star g$ by

$$f \star g(x) = \int_{\mathbb{R}^d} f(y)g(x-y)dy$$

and it is elementary to show that $f \star g \in L^1(\mathbb{R}^d)$.

Example 1. *Let* $1_{[-\frac{1}{2},\frac{1}{2}]}$ *be the indicator function of the interval* $[-\frac{1}{2},\frac{1}{2}]$*. Then*

$$\widehat{1}_{[-\frac{1}{2},\frac{1}{2}]}(\omega) = \int_{\frac{-1}{2}}^{\frac{1}{2}} e^{-2\pi i x\omega} d\omega = \frac{e^{\pi i x} - e^{-\pi i x}}{2\pi i\omega} = \frac{\sin \pi\omega}{\pi\omega}.$$

The function $\frac{\sin x}{x}$ *is called the sinc-function and denoted by* $\operatorname{sinc}(x)$*. The sinc-function is in* $L^2(\mathbb{R})$ *but not in* $L^1(\mathbb{R})$*.*

Another important example is the Fourier transform of a

Example 2. *The Gaussian function* $g_a(x) = e^{-\pi a x^2}$ *is given by*

$$\widehat{g_a}(\omega) = e^{-\pi\omega^2/a}.$$

Thus, g_1 *is Fourier invariant, i.e.* $\widehat{g_1} = g_1$*.*

The Fourier inversion theorem gives a procedure to reconstruct f from its Fourier transform $\widehat{f}$, which in a heuristic way amounts to

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

The problem is that $\widehat{f}$ might not be in $L^1(\mathbb{R}^d)$. In order to fix this issue, one regularizes $\widehat{f}$ with a smooth cut-off function, such as the Gaussian. We are not going to discuss this in detail. We instead just restrict our attention to the case where $\widehat{f} \in L^1(\mathbb{R}^d)$.

Theorem 2.1. *If f and $\widehat{f}$ are both in $L^1(\mathbb{R}^d)$, then*

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

holds almost everywhere. Hence we have that $f = \mathcal{F}^{-1}(\mathcal{F}(f))$.

We denote by $\check{f}$ the inverse Fourier transform $\mathcal{F}^{-1}(f)$:

$$\check{f} = \int_{\mathbb{R}^d} \widehat{f}(\omega) e^{2\pi i x \omega} d\omega.$$

Thus the Fourier inversion theorem may be expressed as $\check{(\widehat{f})}$.
As a consequence we get that $\mathcal{F}$ is injective on $L^1(\mathbb{R}^d) \cap \mathcal{F}(L^1(\mathbb{R}^d))$.

Theorem 2.2. *If f and $\widehat{f}$ are both in $L^1(\mathbb{R}^d)$ and $\widehat{f} = 0$, then*

$$f = 0 \quad a.e.$$

In order to have f and $\widehat{f} \in L^1(\mathbb{R}^d)$, the Fourier transform has to decay, for example like $|\widehat{f}(\omega)| \leq \frac{C}{(1+\omega)}^{1+\delta}$ for some $\delta > 0$, or $\widehat{f}$ being supported on a compact set, such a function is called *band-limited*.

Recall that decay of the Fourier transform of f is equivalent to smoothness of f.
We close this section with a brief treatment of the Fourier transform on L^2.

As the *sinc* shows, the Fourier transform of an element of $f \in L^2(\mathbb{R}^d)$ might not converge in $L^2(\mathbb{R}^d)$, since $\widehat{f}$ might not be in $L^1(\mathbb{R}^d)$, The way out of this is to find a dense subspace $\mathcal{A}$ of $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. Then we approximate any $f \in L^2(\mathbb{R}^d)$ by a sequence $(f_n) \in X$ and extend the Fourier transform from A to $L^2(\mathbb{R}^d)$. We take for A the space $L^1(\mathbb{R}^d) \cap \mathcal{F}(L^1(\mathbb{R}^d))$ and one can show that it is dense in L^2.

Hence, in order to define the Fourier transform of $f \in L^2(\mathbb{R}^d)$ we may proceed as follows:

- Find a sequence $(f_n) \subset \mathcal{A}$ such that $\|f_n - f\|_2 \to 0$ as $n \to \infty$.

- Since $(f_n) \subset \mathcal{A}$ implies that $(\widehat{f_n}) \subset \mathcal{A}$, we have that $\|\widehat{f_n} - \widehat{f_m}\|_2 = \|f_n - f_m\|_2 \to 0$ as $m, n \to \infty$.

- Hence, $(\widehat{f_n})$ is a Cauchy sequence, then there exists a $g \in L^2(\mathbb{R}^d)$ such that $\|\widehat{f_n} - g\|_2 \to 0$ as $n \to \infty$. We define this limit g to be the Fourier transform of f and denote it by $\widehat{f}$.

Theorem 2.3 (Parseval Theorem). *For any f, g in $L^2(\mathbb{R}^d)$ we have $\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle$. Consequently, for any $f \in L^2(\mathbb{R}^d)$ we have $\|f\|_2 = \|\widehat{F}\|_2$.*

A quite useful tool of Fourier analysis is the *Poisson summation formula* which is just the Fourier series of a periodization of a function.

We associate to $f \in L^1(\mathbb{R})$ the *periodization* of f:

$$P_{\mathbb{Z}}(f)(x) = \sum_{k\in\mathbb{Z}} f(x+k).$$

Lemma 2.4. *For $f \in L^1(\mathbb{R})$ the periodization $P_{\mathbb{Z}}(f)$ is in $L^1(\mathbb{T})$*

Proof. Note that $|P_{\mathbb{Z}}(f)(x)| \leq \sum_{k\in\mathbb{Z}} |f(x+k)|$. Hence, we have to show that the partial sums $\sum_{k=-n}^{n} |f(x+k)|$ converges in $L^1(\mathbb{T})$. Let us set $S_n(x) := \sum_{k=-n}^{n} |f(x+k)|$ and compute $\int_0^1 S_m(x)dx$:

$$\begin{aligned}
\int_0^1 S_m(x)dx &= \int_0^1 \sum_{k=-n}^{n} |f(x+k)|dx \\
&= \sum_{k=-n}^{n} \int_0^1 |f(x+k)|dx \\
&= \sum_{k=-n}^{n} \int_{-k}^{k+1} |f(x)|dx \\
&= \int_{-n}^{n+1} |f(x)|dx \leq \|f\|_1.
\end{aligned}$$

Consequently, $(S_n)_n$ is a monotone-increasing sequence of non-negative functions which by the monotone convergence theorem has a limit $S(x) := \lim_n S_n(x)$ that is measurable and $S \in L^1(\mathbb{T})$, because

$$\int_0^1 S(x)dx = \int_0^1 S_n(d)dx \leq \|f\|_1.$$

The final step is to note that

$$\int_0^1 |P_{\mathbb{Z}}(f)(x)dx| \leq \int_0^1 S(x)dx \leq \|f\|_1 < \infty.$$

$\square$

In order to make the steps in the proof of Poisson summation rigorous we have to add a few assumptions on the class of admissible functions.

Theorem 2.5 (Poisson summation)**.** *For functions $f \in C_0(\mathbb{R}) \cap L^1(\mathbb{R})$ satisfying*

$$(1) \sum_{k\in\mathbb{Z}} \|f(\cdot + k)\|_\infty < \infty,$$
$$(2) \sum_{k\in\mathbb{Z}} |\widehat{f}(k)| < \infty,$$

the Fourier coefficients of $P_{\mathbb{Z}}(f)$ are $(\widehat{f}(k))_{k\in\mathbb{Z}}$ and consequently

$$\sum_{k\in\mathbb{Z}} f(x+k) = \sum_{k\in\mathbb{Z}} \widehat{f}(k)e^{2\pi ikx}.$$

Proof. Condition (1) yields that $P_{\mathbb{Z}}(f)$ is continuous and the series $P_{\mathbb{Z}}(f)$ converges uniformly. Thus we obtain

$$\begin{aligned} c_k(P_{\mathbb{Z}}(f)) &= \int_0^1 \sum_{k\in\mathbb{Z}} f(x+k)e^{-2\pi ikx}dx \\ &= \sum_{k\in\mathbb{Z}} \int_0^1 f(x+k)e^{-2\pi ikx}dx \\ &= \int_{\mathbb{R}} f(x)e^{-2\pi ikx}dx = \widehat{f}(k). \end{aligned}$$

By Condition (2) the Fourier series of $P_{\mathbb{Z}}(f)$ converges uniformly to $P_{\mathbb{Z}}(f)$:

$$P_{\mathbb{Z}}(f) = \sum_{k\in\mathbb{Z}} \widehat{f}(k)e^{2\pi ikx}$$

the desired assertion. $\square$

For $x = 0$ the Poisson summation gives a useful formula:

$$\sum_{k\in\mathbb{Z}} f(k) = \sum_{k\in\mathbb{Z}} \widehat{f}(k).$$

3 Time-frequency representations and wavelet transform

In the first section We introduce time-frequency representations, like the short-time Fourier transform (STFT), Wigner distribution and ambiguity function, and discuss some of their basic properties, including the orthogonality relations for the STFT and a reconstruction formula for the STFT. While in the second section we give a similar presentation for the wavelet transform.

3.1 Time-frequency representations

We introduce several time-frequency transforms and start with the *Short-Time Fourier Transform (STFT)* aka *Sliding window Fourier transform.* For a given $g \in L^2(\mathbb{R}^d)$ we define the short-time Fourier transform of a function f by

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

Note that the STFT maps a 1D signal in the time domain to a 2D function on the time-frequency plane. There is a general principle: The values $V_g f$ and $V_g f(x',\omega')$ for $(x,\omega) \neq (x',\omega')$ are intimately coupled and thus cannot be fixed arbitrarily. The strength of this dependence is given by the reproducing kernel and will be discussed in detail later. The STFT is a overcomplete representation of a function containing a lot of redundant information and it was D. Gabor's insight to take advantage of this fact and sample the STFT. We come back to this point of view in our treatment of Gabor frames.
Let us compute the $V_{g_a} g_a$ to get a better idea about the STFT.

Lemma 3.1. $V_{g_a} g_a(x,\omega) = e^{\pi i x\omega}e^{-\pi(ax^2+\omega^2/a)}$

Proof. This is an application of completing squares

$$\begin{aligned}
V_{g_a} g_a(x,\omega) &= \int_{\mathbb{R}} e^{-\pi a t^2}e^{-\pi a(t-x)^2}e^{-2\pi i t\omega}dt \\
&= \int_{\mathbb{R}} e^{-\pi a(2t^2-2tx+x^2)}e^{-2\pi i t\omega}dt \\
&= \int_{\mathbb{R}} e^{-\pi 2a(t^2-2tx/2+x^2/2)}e^{-2\pi i t\omega}dt \\
&= \int_{\mathbb{R}} e^{-\pi a(t-x/2)^2}e^{-\pi a x^2/2}e^{-2\pi i t\omega}dt \\
&= e^{-\pi i x\omega}e^{-\pi a x^2/2}\int_{\mathbb{R}} e^{-\pi i a s^2}e^{-2\pi i s\omega}ds \quad = e^{-\pi i x\omega}e^{-a\pi x^2/2}e^{-\pi\omega^2/a}.
\end{aligned}$$

$\square$

There is a closely related object, the *ambiguity function* in the field of radar/sonar technology:

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

If one changes the variable t in $V_g f$ by $t + x/2$, then one gets that

$$A(f)(x,\omega) = e^{\pi i x\omega} V_f f(x,\omega).$$

Another prominent time-frequency representation is the *Wigner distribution* of g:

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

Wigner's motivation for introducing $W(g)$ was to have a time-frequency representation that has the marginal properties:

$$\int_{\mathbb{R}^d} W(g)(x,\omega)d\omega = |g(x)|^2 \quad\text{and}\quad \int_{\mathbb{R}^d} W(g)(x,\omega)dx = |\widehat{g}(\omega)|^2.$$

Since we have

$$\begin{aligned}
\int_{\mathbb{R}^d} W(g)(x,\omega)dx &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g(x+\tfrac{t}{2})\overline{g(x-\tfrac{t}{2})}e^{-2\pi i t\cdot\omega}\,dtdx \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g(y)\overline{g(y-t)}e^{-2\pi i t\cdot\omega}\,dtdy \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} g(y)\overline{g(s)}e^{2\pi i s\cdot\omega}e^{-2\pi i y\cdot\omega}\,dsdy \\
&= |\widehat{g}(\omega)|^2.
\end{aligned}$$

and the second assertion may be deduced in a similar manner.

Let us rephrase the definition of the STFT:

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

where $\pi(x,\omega)g(t) = e^{2\pi i\omega t}g(t-x)$ denotes the **time-frequency shift** of g by $z=(x,\omega)$. As the name suggests, is $\pi(x,\omega)$ the composition $M_omega\circ T_x$ of a translation/time shift $T_x g(t) = g(t-x)$ and a modulation $M_\omega g(t) = e^{2\pi i t\cdot\omega}g(t)$ of $g\in L^2(\mathbb{R}^d)$.

Here are a few expressions for the STFT that display different aspects of this time-frequency representation.

Lemma 3.2. *For f, g in $L^2(\mathbb{R}^d)$ we have*

$$V_g f(x,\omega) = (f \cdot T_x \overline{g})\widehat{\ } = (f * M_\omega \check{g}) = \langle \widehat{f}, T_\omega M_{-x} \widehat{g} \rangle,$$

where we denote the reflection at the origin by $\check{g}(x) = g(-x)$.

Proof. Left as an exercise. Hint: Use Parseval's Theorem. $\square$

Let us collect a few properties of translation and modulation operators.

Lemma 3.3. *We have the following identities:*

1. $\widehat{T_x f} = M_{-\omega} \widehat{f}$ *and* $\widehat{M_\omega f} = T_\omega \widehat{f}$
2. $M_\omega T_x = e^{2\pi i x\omega} T_x M_\omega$

The commutation relation $M_\omega T_x = e^{2\pi i x\omega} T_x M_\omega$ is the basic identity of time-frequency analysis and we will see many of its consequences in our discussion of the STFT and Gabor frames.

Proof. The stated assertions are elementary computations:

1.

$$\begin{aligned}
\widehat{T_x f}(\omega) &= \int_{\mathbb{R}^d} f(t-x) e^{-2\pi i t\omega} dt \\
&= \int_{\mathbb{R}^d} f(s) e^{-2\pi i (x+s)\omega} dt \\
&= e^{-2\pi i x\omega} \widehat{f}(\omega) \\
&= M_{-\omega} \widehat{f}(\omega).
\end{aligned}$$

The second assertion may be deduced in a similar way:

$$\begin{aligned}
\widehat{M_\omega} f(\eta) &= \int_{\mathbb{R}^d} e^{2\pi i \omega t} f(t) e^{-2\pi i t\eta} dt \\
&= \int_{\mathbb{R}^d} f(t) e^{-2\pi i t(\eta-\omega)} dt \\
&= \widehat{f}(\eta - \omega) \\
&= T_\omega \widehat{f}(\eta).
\end{aligned}$$

2.

$$\begin{aligned} T_x M_\omega f(t) &= T_x(e^{2\pi i t\omega} f)(t) \\ &= e^{2\pi i (t-x)\omega} f(t-x) \\ &= e^{-2\pi i x\omega} e^{2\pi t\omega} f(t-x) \\ &= e^{-2\pi i x\omega} M_\omega T_x f(t). \end{aligned}$$

□

Lemma 3.4. *If f, g are in $L^2(\mathbb{R}^d)$, then $V_g f$ is uniformly continuous on $\mathbb{R}^{2d}$.*

Proof. This is left as an exercise. You might want to use that for all $f \in L^2(\mathbb{R}^d)$ we have $\lim_{x\to 0} \|T_x f - f\|_2 = 0$ and $\lim_{\omega\to 0} \|M_\omega f - f\|_2 = 0$. □

Let us use the commutation relation to relate the STFT of f with the STFT of its Fourier transfrom $\widehat{f}$:

$$V_{\widehat{g}} \widehat{f}(x,\omega) = \langle \widehat{f}, T_\omega M_{-x} \widehat{g} \rangle = e^{-2\pi i x\cdot\omega} V_{\widehat{g}} \widehat{f}(\omega, -x). \tag{1}$$

We close the treatment of elementary properties with a computation of the STFT of a time-frequency shift of a signal:

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

For later reference we have obtained that

$$V_g(\pi(y,\eta)f)(x,\omega) = e^{-2\pi i y\cdot(\omega-\eta)} V_g f(x-y, \omega-\eta). \tag{2}$$

Let us take a step back and distill some of the concepts underlying the STFT.

- We pick a window function g.
- We associate to each $f \in L^2(\mathbb{R})$ the time-frequency representation $V_g f$, i.e. this induces a map $\mathcal{V}_g : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^{2d})$.

In the remainder of this section we are going to address these evident questions:

1. Do the values of the STFT $V_g f(z)$ for all points z in the time-frequency plane $\mathbb{R}^{2d}$ determine f?
2. What are the properties of the mapping $\mathcal{V}_g$?
3. Is there a decomposition of the signal f based on the values of the STFT $\{V_g f(z)\}_{z\in\mathbb{R}^{2d}}$?

The answers to all these questions might be deduced from an identity that at first seems not to be of relevance in this context: *Orthogonality relations for the STFT* aka *Moyal's identity.*

Proposition 3.5 (Moyal's identity)**.** *For functions f_1, f_2 in $L^2(\mathbb{R}^d)$ and window functions g_1, g_2 in $L^2(\mathbb{R}^d)$ we have that*

$$\iint_{\mathbb{R}^{2d}} V_{g_1} f_1(z)\overline{V_{g_2} f_2(z)}\, dz = \langle f_1, f_2\rangle\overline{\langle g_1, g_2\rangle}. \tag{3}$$

Proof. It suffices to show that this holds for functions $f_1, f_2.g_1, g_2$ in the Schwartz class $\mathscr{S}(\mathbb{R}^d)$. Since $\mathscr{S}(\mathbb{R}^d)$ is stable under pointwise multiplication and Fourier transform, we have that $V_{g_i} f_i$ is in $\mathscr{S}(\mathbb{R}^{2d})$ since $V_{g_i} f_i(x,\omega) = \mathcal{F}(f_i \cdot T_x\overline{g_i})(\omega)$ for $i = 1, 2$.

$$\begin{aligned}
\iint_{\mathbb{R}^{2d}} V_{g_1} f_1(z)\overline{V_{g_2} f_2(z)}\, dz &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} (\widehat{f_1 \cdot T_x\overline{g_1}})(\omega)\overline{(\widehat{f_2 \cdot T_x\overline{g_2}})(\omega)}d\omega\right)dx \\
&= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f_1(t)\overline{g_1(t-x)}\overline{f_2(t)}g_2(t-x)dt\right)dx \\
&= \int_{\mathbb{R}^d} f_1(t)\overline{f_2(t)}\left(\int_{\mathbb{R}^d} \overline{g_1(t-x)}g_2(t-x)dx\right)dt \\
&= \langle f_1, f_2\rangle\overline{\langle g_1, g_2\rangle},
\end{aligned}$$

where we used Parseval's Theorem in the second step and Fubini's Theorem in the third step, which is justified for functions in the Schwartz class. □

As consequences of the orthogonality relations for the STFT we provide affirmative answers to the questions stated above.

Proposition 3.6. *For a non-zero $g \in L^2(\mathbb{R}^d)$ and a $f \in L^2(\mathbb{R}^d)$ we assume that $V_g f(z) = 0$ for all $z \in \mathbb{R}^{2d}$. Then $f = 0$.*

Proof. By assumption $V_g f(z) = 0$ and thus $\iint_{\mathbb{R}^{2d}} |V_g f(z)|^2 dz = 0$. By the orthogonality relations we have $\iint_{\mathbb{R}^{2d}} |V_g f(z)|^2 dz = \|f\|_2^2 \|g\|_2^2$, which implies that $f = 0$ since $g \neq 0$. $\square$

Let us give an equivalent formulation of the preceding assertion.

Proposition 3.7. *For a non-zero $g \in L^2(\mathbb{R}^d)$ we have that the closed linear span of $\{\pi(z)g\}_{z \in \mathbb{R}^{2d}}$ is dense in $L^2(\mathbb{R}^d)$.*

Proof. We denote by X the closed linear span of $\{\pi(z)g\}_{z \in \mathbb{R}^{2d}}$ in $L^2(\mathbb{R}^d)$. Then the desired assertion is equivalent to show that the orthogonal complement $X^\perp$ of X is the trivial subspace $\{0\}$. Recall that $X^\perp$ consists of all $f \in L^2(\mathbb{R}^d)$ such that $\langle f, \pi(z)g \rangle = 0$ for all $z \in \mathbb{R}^{2d}$, which by the orthogonality relations for the STFT implies that $f = 0$. $\square$

The reconstruction formula of $f \in L^2(\mathbb{R}^d)$ is given in terms of two functions $g, h \in L^2(\mathbb{R}^d)$.

Proposition 3.8. *Given $g, h \in L^2(\mathbb{R}^d)$ satisfying $\langle g, h \rangle \neq 0$. Then for any $f \in L^2(\mathbb{R}^d)$ we have*

$$f = \frac{1}{\langle h, g \rangle} \iint_{\mathbb{R}^{2d}} V_g f(z) \pi(z) h \, dz,$$

where we interpret the vector-valued integral weakly.

Proof. Let us rewrite $\iint_{\mathbb{R}^{2d}} V_g f(z) \overline{V_h k(z)} \, dz$ as follows:

$$\begin{aligned}
\iint_{\mathbb{R}^{2d}} V_g f(z) \overline{V_h f(z)} \, dz &= \iint_{\mathbb{R}^{2d}} \langle f, \pi(z) g \rangle \langle \pi(z) h, k \rangle \, dz \\
&= \Big\langle \iint_{\mathbb{R}^{2d}} \langle f, \pi(z) g \rangle \pi(z) h \, dz, k \Big\rangle,
\end{aligned}$$

which by the orthogonality relations yields the identity:

$$\Big\langle \iint_{\mathbb{R}^{2d}} \langle f, \pi(z) g \rangle \pi(z) h \, dz, k \Big\rangle = \Big\langle \langle h, g \rangle f, k \Big\rangle$$

that holds for all $k \in \mathscr{S}(\mathbb{R}^d)$ and hence we have the claimed reconstruction formula:

$$\iint_{\mathbb{R}^{2d}} \langle f, \pi(z) g \rangle \pi(z) h \, dz = \langle h, g \rangle f.$$

$\square$

Consequently, any non-zero $g \in L^2(\mathbb{R}^d)$ with $\|g\|_2 = 1$ yields a reconstruction formula for $f \in L^2(\mathbb{R}^d)$:

$$f = \iint_{\mathbb{R}^{2d}} V_g f(z) \pi(z) g \, dz.$$

3.2 Wavelet transform

For a non-zero real number a we define the dilation operator D_a acting on $f \in L^2(\mathbb{R})$ by $D_a f(x) = |a|^{-1/2} f(\frac{x}{a})$, where we think of a as the scale, which has as Fourier transform $\widehat{D_a f}(\omega) = |a|^{1/2} f(ax) = |a|^{1/2} D_{1/a} f(x)$. Note that a dilation $D_a f$ has the same shape as f, but its support differs from the one of f. More concretely, if the support of f is contained in a subset X of $\mathbb{R}$, then the support of $D_a f$ is contained in aX.

The *wavelet transform* of a function f with respect to the wavelet g is given by

$$\mathcal{W}_g f(x,a) = |a|^{-1/2} \int_{\mathbb{R}} f(t)\overline{g(\tfrac{t-x}{a})} dt,$$

which we might rewrite in a similar way as the STFT:

$$\mathcal{W}_g f(x,a) = \langle f, T_x D_a g \rangle,$$

where the dilation operator D_a takes its place instead of the modulation operator M_ω.

A different expression for the wavelet transform is in terms of convolution:

$$\mathcal{W}_g f(x,a) = f * D_a \check{g}.$$

Thus for a given scale a we have that the Fourier transform of $W_g f(x,a)$ is of the form

$$\widehat{\mathcal{W}_g f}(\omega, a) = |a|^{1/2} \widehat{f}(\omega) \widehat{g}(a\omega).$$

For a fixed $a > 0$, the wavelet transform can be interpreted as an approximation of f that sees only details of size a and smooths out smaller details. While for a fixed x and values of a tending to 0 the wavelet transform $\mathcal{W}_g f(x,a)$ zooms in to the point x and resolves local details at x. This is in contrast to the STFT, where the support of g remains fixed and prevents a study of the local properties of f at a given point.

Before we state and proof the key result about the wavelet transform, the orthogonality relations, we state some of its basic properties.

- $\mathcal{W}_g(T_y f)(x,a) = \mathcal{W}_g f(x, a-y)$;
- For $c \neq 0$ we have $\mathcal{W}_g(D_c f)(x,a) = \mathcal{W}_g f(\frac{x}{x}, \frac{a}{c})$;
- $\mathcal{W}_{\check{g}} \check{f}(x,a) = \mathcal{W}_g f(x, -a)$;
- $\mathcal{W}_g f(x,a) = \mathcal{W}_f g(\frac{1}{x}, -\frac{a}{x})$.

These useful identities are straightforward consequences of the definition of the wavelet transform.

Proposition 3.9 (Orthogonality relations for the wavelet transform)**.** *For $f_1, f_2 \in L^2(\mathbb{R})$ and g_1, g_2 in $L^2(\mathbb{R})$ satisfying the condition $\mathbb{C}_{g_1,g_2} := int_{\mathbb{R}}\widehat{g}_1(a\omega)\overline{\widehat{g}_2}(a\omega)\frac{da}{a} < \infty$ we have that*

$$\iint_{\mathbb{R}^2} W_{g_1}(f_1)(x,a)\overline{W_{g_2}(f_2)(x,a)}\frac{dxda}{a^2} = \langle f_1, f_2\rangle \int_{\mathbb{R}} \widehat{g}_1(a\omega)\overline{\widehat{g}_2}(a\omega)\frac{da}{a}. \tag{4}$$

Proof. It suffices to prove the orthogonality relations for f_1, f_2 in $L^1 \cap L^2(\mathbb{R})$. Since $\mathcal{W}_g f(x,a) = f * D_a\check{g}$, we have that for a fixed a the Fourier transform of the wavelet transform is

$$\widehat{\mathcal{W}_g f}(\omega, a) = |a|^{1/2}\widehat{f}(\omega)\overline{\widehat{g}(a\omega)},$$

which we are going to use in our argument below.

$$\begin{aligned}
\int_{\mathbb{R}}\left(\int_{\mathbb{R}} \mathcal{W}_{g_1}f_1(x,a)\overline{\mathcal{W}_{g_2}f_2(x,a)}dx\right)\frac{da}{a^2} &= \int_{\mathbb{R}}\left(\int_{\mathbb{R}} |a|\widehat{f}_1(\omega)\overline{\widehat{g}_1(a\omega)}\overline{\widehat{f}_2(\omega)}\widehat{g}_2(a\omega)\frac{da}{|a|}\right)d\omega \\
&= \int_{\mathbb{R}} \widehat{f}_1(\omega)\overline{\widehat{f}_2(\omega)}\left(\int_{\mathbb{R}} \overline{\widehat{g}_1(a\omega)}\widehat{g}_2(a\omega)\frac{da}{|a|}\right)d\omega \\
&= \langle f_1, f_2\rangle \int_{\mathbb{R}} \overline{\widehat{g}_1(\eta)}\widehat{g}_2(\eta)\frac{d\eta}{|\eta|},
\end{aligned}$$

where we have used that the Theorem of Fubini is applicable due to $\widehat{f}_1\widehat{f}_1 \in L^1(\mathbb{R})$ and that $\int_{\mathbb{R}} \overline{\widehat{g}_1(a\omega)}\widehat{g}_2(a\omega)\frac{da}{|a|}$ is actually independent of ω as one sees by substituting $\eta = a\omega$ in the integral. $\square$

In the case that for $g \in L^2(\mathbb{R})$ we have $C_g = \int_{\mathbb{R}} \frac{|\widehat{g}(\omega)|^2}{|\omega|}d\omega < \infty$, we say that g satisfies the *admissibility condition* for the (mother) wavelet g. If g is continuous, then the admissibility condition implies that $\widehat{g}(0) = \int_{\mathbb{R}} g(t)dt = 0$.

As for the STFT, the orthogonality relations yield to reconstruction formulas.

Proposition 3.10. *For $g, h \in L^2(\mathbb{R})$ satisfying $\int_{\mathbb{R}} \overline{\widehat{g}(\eta)}\,\widehat{h}(\eta)\frac{d\eta}{|\eta|} < \infty$ we have that any f may be expressed in the form*

$$f = C_{g_1,g_2}^{-1}\iint_{\mathbb{R}^2} W_g f(x,a) T_x D_a h \frac{dxda}{a^2}.$$

In particular, for an admissible g this yields

$$f = C_g^{-1}\iint_{\mathbb{R}^2} W_g f(x,a) T_x D_a g \frac{dxda}{a^2}.$$

The proof is left as an exercise to the reader.

4 RKHS and applications to the STFT and the wavelet transform

4.1 Reproducing kernel Hilbert spaces – Basics

Let us make a small detour into a class of Hilbert spaces of functions, known as *reproducing kernel Hilbert spaces*, that have some attractive features not shared by general Hilbert spaces.

Definition 4.1. *For a set X we denote by $\mathcal{F}(X)$ the space of real or complex-valued functions on X. Then a reproducing kernel Hilbert space (rkhs) is a subspace $\mathcal{H}$ of $\mathcal{F}(X)$ equipped with an inner product $\langle .,. \rangle$ such that $(\mathcal{H}, \langle .,. \rangle)$ is a Hilbert space satisfying that the evaluation functional E_x, given by $E_x(f) = f(x)$, is bounded: $|E_x(f)| \leq C\|f\|_{\mathcal{H}}$ for all $f \in \mathcal{H}$.*

The Riesz representation theorem yields that there exists for each $x \in X$ an element $k_x \in \mathcal{H}$ such that

$$E_x(f) = \langle f, k_x \rangle.$$

We refer to k_x as the *reproducing kernel at x*. Since k_x is an element of $\mathcal{H}$ the values of $k_x(y)$ are expressible in the form: $k_x(y) = \langle k_y, k_x \rangle$.
We call $k(x,y) := \langle k_y, k_x \rangle$ the *reproducing kernel* of $\mathcal{H}$, we will often denote a rkhs by $(\mathcal{H}, k)$.
Since $E_x(k_x) = \langle k_x, k_x \rangle = k(x,x) = \|k_x\|^2$ we obtain that the operator norm of E_x equals $\sqrt{k(x,x)}$, i.e. $\|E_x\| = \sqrt{k(x,x)}$. The *normalized reproducing kernel* $\tilde{k}(x,y) = k(x,y)/k(x,x)$. By the properties of the inner product and the definition of the reproducing kernel we obtain that reproducing kernels possess a symmetry property:

$$k(x,y) = \overline{k(y,x)} \quad \text{for any} \quad x, y \in X.$$

The short argument goes as follows: Since $k_x \in \mathcal{H}$ we have that $k_x(y) = \langle k_y, k_x \rangle$ and thus

$$k(x,y) = k_x(y) = \langle k_y, k_x \rangle = \overline{\langle k_x, k_y \rangle} = \overline{k(y,x)}.$$

Let us state some basic properties of RKHS.

Proposition 4.2. *Let $(\mathcal{H}, k)$ be a RKHS on the set X. Then the linear span of $\{k_y(.) : y \in X\}$ is dense in $\mathcal{H}$.*

Proof. The statement is equivalent to $\{k_y(.) : y \in X\}^\perp = \{0\}$ which holds if and only if $\langle f, k_y \rangle = 0$ for every $y \in X$. Since $\langle f, k_y \rangle = f(y)$, we have that $f(y) = 0$ for every $y \in X$. Hence we deduce that $f = 0$, which is the desired assertion. $\square$

The next result is elementary and we are going to apply it several times.

Lemma 4.3. *Let $(\mathcal{H}, k)$ be a RKHS on X and let $\{f_n\}$ be a sequence of $\mathcal{H}$. If (f_n) converges in norm to $f \in \mathcal{H}$, then it converges moreover also point-wise: $f(x) = \lim_n f_n(x)$ for every $x \in X$.*

Proof. Since $f_n(x) - f(x) \in \mathcal{H}$ we have that $f_n(x) - f(x) = \langle f_n - f, k_x \rangle$ and by an application of Cauchy-Schwarz:

$$|f_n(x) - f(x)| = |\langle f_n - f, k_x \rangle| \leq\leq \|f_n - f\| \|k_x\|.$$

Consequently, $\lim_n \|f_n - f\| = 0$ implies that $\lim_n |f_n(x) - f(x)| = 0$ for every $x \in X$. $\square$

One consequence is an expression for the reproducing kernel k of a RKHS.

Proposition 4.4. *Let $(\mathcal{H}, k)$ be a RKHS on X. If $\{e_j\}_{j \in J}$ is an orthonormal basis for $\mathcal{H}$, then for $x, y \in X$*

$$k(x,y) = \sum_{j \in J} \overline{e_j(x)} e_j(y) \tag{5}$$

converges pointwise.

Proof. For any $y \in X$ we have that the generalized Fourier series expansion $k_y = \sum_{j \in J} \langle k_y, e_j \rangle e_j$ which converges in norm in $\mathcal{H}$. Hence we have that

$$k(x,y) = k_y(x) = \sum_{j \in J} \langle k_y, e_j \rangle e_j(x)$$

converges pointwise by the preceding lemma. The final step is to note that the Fourier coefficient $\langle k_y, e_j \rangle = \overline{\langle e_j, k_y \rangle} = \overline{e_j(y)}$ by the definition of a reproducing kernel and thus

$$k(x,y) = \sum_{j \in J} \overline{e_j(y)} e_j(x).$$

$\square$

Another application of the preceding lemma is the fact that a reproducing kernel determines the RKHS uniquely.

Proposition 4.5. *Let* $(\mathcal{H}_a, k_a)$ *and* $(\mathcal{H}_b, k_b)$ *be two RKHSs on* X*. If* $k_a(x,y) = k_b(x,y)$ *fpr all* $x, y \in X$*, then* $(\mathcal{H}_a, \|.\|_a) = (\mathcal{H}_b, \|.\|_b)$*.*

Proof. See [4, Prop. 2.3] for a proof of this statement. □

There are numerous examples of RKHS from which three are of relevance for us: (i) the Paley-Wiener space, (ii) Bergman spaces, and (iii) Bargmann-Fock spaces.

4.2 Examples

Paley-Wiener spaces: For a fixed $A > 0$ we define the *Paley-Wiener space* $\mathrm{PW_A}$ to be the set of all those Fourier transforms of L^2-functions with support in $[-A, A]$: $\mathrm{PW_A} := \{\widehat{f} : f \in L^2[-A, A]\}$. Since $L^2[-A, A] \subset L^1[-A, A]$ we have that $\mathrm{PW_A}$ is a space of continuous functions on the real line. Furthermore $\mathrm{PW_A}$ is a closed subspace pf $L^2[-A, A]$ and thus a Hilbert space.

If $F \in \mathrm{PW_A}$, then we have that there exists a unique $f \in L^2[-A, A]$ such that

$$F(x) = \int_{-A}^{A} f(t) e^{-2\pi i t x} \, dt,$$

since the Fourier transform from $L^2[-A, A]$ to $\mathrm{PW_A}$ is a Hilbert space isomorphism. This is a consequence of the fact that $\{e^{-2\pi i n t/A}\}_{n \in \mathbb{Z}}$ is an ONB for $L^2[-A, A]$ and thus the Fourier coefficients $(\widehat{f}(\frac{n}{A}))_{n \in \mathbb{Z}}$ uniquely determine $f \in L^2[-A, A]$, but the Fourier coefficients $\widehat{f}(n/A)$ are actually equal to $F(n/A)$ and hence also uniquely determine $F \in L^2[-A, A]$ uniquely. Therefore, we may define a norm on $\mathrm{PW_A}$ by $\|F\|_{\mathrm{PW_A}} := \|f\|_2$. In other words, the Fourier transform provides us with an identification of $L^2[-A, A]$ with $\mathrm{PW_A}$ which yields that for $F, G \in \mathrm{PW_A}$ we use their representatives $f, g \in L^2[-A, A]$ to define the inner product of F and G via the Fourier transforms of $\widehat{f}$ and $\widehat{g}$:

$$\langle F, G \rangle_{\mathrm{PW_A}} := \langle f, g \rangle = \int_{-A}^{A} f(t) \overline{g(t)} \, dt.$$

For $F \in \mathrm{PW_A}$ we have that the evaluation functional

$$|F(x)| = \left| \int_{-A}^{A} f(t) e^{-2\pi i t x} \, dt \right| \leq \|f\|_2 \Big(\int_{-A}^{A} |e^{-2\pi i t x}|^2 \, dt \Big)^{1/2} = \sqrt{2A} \|f\|_2$$

is bounded and so $\mathrm{PW_A}$ is a RKHS.
The Paley-Wiener space $\mathrm{PW_A}$ is a RKHS and we might determine its reproducing kernel either using the ONB $\{e^{-2\pi i n t/A}\}_{n \in \mathbb{Z}}$ via the identity $k(x, y) =$

$\sum_{n\in\mathbb{Z}} e^{-2\pi inxt}e^{2\pi iyt}$ or by the definition of the reproducing kernel to represent the evaluation functional on $\mathrm{PW_A}$.
For $F \in \mathrm{PW_A}$ we want to determine the function $k_y \in \mathrm{PW_A}$ such that

$$F(y) = \langle f, k_y \rangle,$$

but $F(y) = \int_{-A}^{A} f(t)e^{-2\pi iyt}\,dt = \langle f, e_y\rangle$ where $e_y \in L^2[-A,A]$ denotes $e_y(t) = e^{-2\pi iyt}$. Consequently, $k_y(x) = \widehat{e_y}(x)$:

$$k_y(x) = k(x,y) = \int_{-A}^{A} e^{2\pi it(y-x)}\,dt,$$

which may be explicitly written as

$$k(x,y) = \begin{cases} \frac{\sin(2\pi A(x-y))}{\pi(x-y)} & \text{for } x \neq y \\ 2A & \text{for } x = y \end{cases}.$$

Hardy spaces of the unit disc: We denote by $H^2(\mathbb{D})$ the *Hardy space* of the unit disc $\mathbb{D} = \{z \in \mathbb{C} \,|\, |z| < 1\}$, which consists of all holomorphic functions f on $\mathbb{D}$ with power series expansion:

$$f(z) = \sum_{n\geq 0} a_n z^n$$

such that $\sum_{n\geq 0} |a_n|^2 < \infty$. We define a norm on $H^2(\mathbb{D})$ by $\|f\|_{H^2} := \|(a_n)\|_2$ and its inner product by polarization.

We identify $H^2(\mathbb{D})$ with ℓ^2 via $f \mapsto (a_0, a_1, a_2, ...)$ and thus $H^2(\mathbb{D})$ is a Hilbert space with respect to the norm $\|f\|_{H^2}$.

Let us show that the evaluation functional on $H^2(\mathbb{D})$ is bounded: For any $z \in \mathbb{D}$ we have

$$|f(z)| \leq \left| \sum_{n\geq 0} a_n z^n \right| \leq \sum_{n\geq 0} |a_n||z|^n,$$

which by an application of the Cauchy-Schwarz inequality yields

$$|f(z)| \leq \left(\sum_{n\geq 0} |a_n|^2\right)^{1/2} \left(\sum_{n\geq 0} |z|^{2n}\right)^{1/2} = \|f\|_{H^2} \frac{1}{\sqrt{1-|z|^2}}.$$

In summary, we have that $H^2(\mathbb{D})$ is a RKHS. The reproducing kernel at $w \in \mathbb{D}$ satisfies

$$f(w) = \langle f, k_w \rangle \quad \text{for some} \quad k_w \in H^2(\mathbb{D}).$$

But by definition of $H^2(\mathbb{D})$ this means that

$$f(w) = \sum_{n\geq 0} a_n W^n,$$

i.e. $\langle f, k_w\rangle = \sum_{n\geq 0} a_n \overline{b_n}$, for some $g(z) = \sum_{n\geq 0} b_n z^n$. Consequently, we obtain that $b_n = \overline{w}^n$ and

$$k_w(z) = \sum_{n\geq 0} \overline{w}^n z^n = \frac{1}{1-\overline{w}z}$$

and for the kernel function $k(w,z) = (1-\overline{w}z)^{-1}$.

Bargmann-Fock spaces: A well-studied class of RKHS is given by the **Bargmann-Fock space** $\mathcal{F}$, which consists of all holomorphic functions on $\mathbb{C}$ with power series expansions

$$f(z) = \sum_{n\geq 0} a_n z^n \qquad \text{satisfying} \quad \sum_{n\geq 0} n!|a_n|^2 < \infty.$$

We identify $\mathcal{F}$ with

$$\ell_v^2 = \{a = (a_n) : \sum_{n\geq 0} n!|a_n|^2 < \infty\}$$

via $f \mapsto (a_0, a_1, ...)$. Hence $\mathcal{F}$ is a Hilbert space with respect to the inner product:

$$\langle f, g\rangle_{\mathcal{F}} = \sum_{n\geq 0} n! a_n \overline{b_n},$$

for $f(z) = \sum_{n\geq 0} a_n z^n$ and $g(z) = \sum_{n\geq 0} b_n z^n$.
The evaluation functional is once more bounded by the Cauchy-Schwarz inequality:

$$|f(z)| \leq \sum_{n\geq 0} \sqrt{n!}|a_n| \frac{|z|^n}{\sqrt{n!}} \leq \sum_{n\geq 0} n!|a_n|^2 \sum_{n\geq 0} \frac{|z|^{2n}}{n!} = \|f\|_{\mathcal{F}} e^{|z|^2}.$$

Consequently, $\mathcal{F}$ is a RKHS.
The reproducing kernel k of $\mathcal{F}$ is given by $k(w,z) = e^{\overline{w}z}$. The argument follows along the lines of the one for the Hardy space on the unit disc, and is left as an exercise to the reader.

Exercise: Show that $\mathcal{F}$ is isomorphic to the space of all holomorphic functions f on $\mathbb{C}$ such that $\int_{\mathbb{C}} |f(z)|^2 e^{-|z|^2}\, dz < \infty$, and the square root of the later expression is equal to $\|f\|_{\mathcal{F}}$.

Bergman spaces: The final example is the *Bergman space* on the unit disc $\mathbb{D}$, denoted by $B^2(\mathbb{D})$, consisting of all holomorphic functions on $\mathbb{D}$ satisfying a square-integrability condition:

$$\|f\|_{B^2} := \frac{1}{\pi}\left(\iint |f(x+iy)|^2\, dxdy\right)^{1/2} < \infty.$$

The proof the $B^2(\mathbb{D})$ is a Hilbert space requires some results from complex analysis that are not part of the prerequisites of MA8004, see [4], also the argument of showing the boundedness of the evaluation functional. Let us just state that $B^2(\mathbb{D})$ is a RKHS with kernel $k(z,w) = \frac{1}{1-wz}$ for $w, z \in \mathbb{D}$, see [4].

4.3 Gabor spaces

For a normalized $g \in L^2(\mathbb{R})$ the STFT V_g induces an isometry from $L^2(\mathbb{R})$ to $L^2(\mathbb{R}^2)$, and we refer to its range $V_g(L^2(\mathbb{R}))$ as *Gabor space*, which we will further on denote by $V_g(L^2)$. Note that $V_g(L^2)$ is a proper closed subspace of $L^2(\mathbb{R}^2)$. By definition $F \in V_g(L^2)$ if and only if there exists an $f \in L^2(\mathbb{R})$ such that $F = V_g f$.
In order to gain a better understanding of Gabor spaces we have to look a bit more into the mapping $f \mapsto V_g f$ for a fixed normalized $g \in L^2(\mathbb{R})$.

Lemma 4.6. *For a fixed normalized $g \in L^2(\mathbb{R})$ the mapping $V_g : L^2(\mathbb{R}) \to L^2(\mathbb{R}^2)$ is bounded and its adjoint mapping is given by*

$$V_g^* F = \iint_{\mathbb{R}^2} F(z)\pi(z)\, dz, \quad F \in L^2(\mathbb{R}).$$

Proof. For any $F \in L^2(\mathbb{R}^2)$ and $f \in L^2(\mathbb{R})$ we have

$$\begin{aligned}
\langle V_g f, F\rangle &= \iint_{\mathbb{R}^2} V_g f(z)\overline{F(z)}\, dz \\
&= \iint_{\mathbb{R}^2} \langle f, \pi(z)g\rangle \overline{F(z)}\, dz \\
&= \langle f, \iint_{\mathbb{R}^2} F(z)\pi(z)g\, dz\rangle \\
&= \langle f, V_g^\star F\rangle,
\end{aligned}$$

and thus the adjoint of V_g is of the form:

$$V_g^\star F = \iint_{\mathbb{R}^2} F(z)\pi(z)g\,dz.$$

$\square$

For $g \in L^2(\mathbb{R})$ with $\|g\|_2 = 1$ the composition of $V_g^\star V_g$ is

$$V_g^\star V_g f = \iint_{\mathbb{R}^2} V_g f(z)\pi(z)\,dz = f$$

by the reconstruction formula for the STFT and thus equivalent to the statement

$$V_g^\star V_g = I.$$

Proposition 4.7. *For $g \in L^2(\mathbb{R})$ with $\|g\|_2 = 1$ the Gabor space $V_g(L^2)$ is a RKHS with reproducing kernel k_z^g at $z \in \mathbb{R}^2$ given by $k_z^g = V_g(\pi(z)g)$ and its reproducing kernel is $k^g(z,z') = V_g(\pi(z')g)(z) = \langle \pi(z')g, \pi(z)g\rangle$.*

5 Supplementary resources

The standard text on time-frequency analysis is Gröchenig's book [3], Daubechies's book is the all-time classic [2]. The theory of frames and Riesz bases is well-explained in Christensens book [1] and for reproducing kernel Hilbert spaces we refer to [4].

References

[1] O. Christensen. *An Introduction to Frames and Riesz Bases. 2nd edition.* Applied and Numerical Harmonic Analysis. Birkhäuser, Boston, 2016.

[2] I. Daubechies. *Ten Lectures on Wavelets.*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* SIAM, Philadelphia, PA, 1992.

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

[4] V. I. Paulsen and M. Raghupathi. *An introduction to the theory of reproducing kernel Hilbert spaces*, volume 152. Cambridge university press, 2016.