Chapter 3

The Short-Time Fourier Transform

We have seen that ideal time-frequency analysis faces a fundamental obstacle in the form of the uncertainty principle. Nevertheless, the example of the musical score indicates that a reasonable and useful form of time-frequency analysis should still be possible and realizable.

In this and the following chapters we will discuss several ideas for joint time-frequency representations. We first look at a linear and continuous representation, the short-time Fourier transform. The reader may think of the short-time Fourier transform as the mathematical analogue of the musical score. We will build the theory of time-frequency analysis almost entirely on the short-time Fourier transform because most other time-frequency representations can be expressed in terms of the short-time Fourier transform.

3.1 Elementary Properties of the Short-Time Fourier Transform

The idea of the short-time Fourier transform is implicit in the discussion of Chapter 2. In order to obtain information about local properties of f, in particular about some "local frequency spectrum," we restrict f to an interval and take the Fourier transform of this restriction. Since a sharp cut-off introduces artificial discontinuities and can create unwanted problems, we choose a smooth cut-off function as a "window."

Definition 3.1.1. Fix a function $g \neq 0$ (called the *window function*). Then the short-time Fourier transform (STFT) of a function f with respect to g is defined as

$$V_g f(x, \omega) = \int_{\mathbb{R}^d} f(t)\, \overline{g(t-x)}\, e^{-2\pi i t\cdot\omega}\, dt\,, \qquad \text{for } x, \omega \in \mathbb{R}^d\,. \tag{3.1}$$

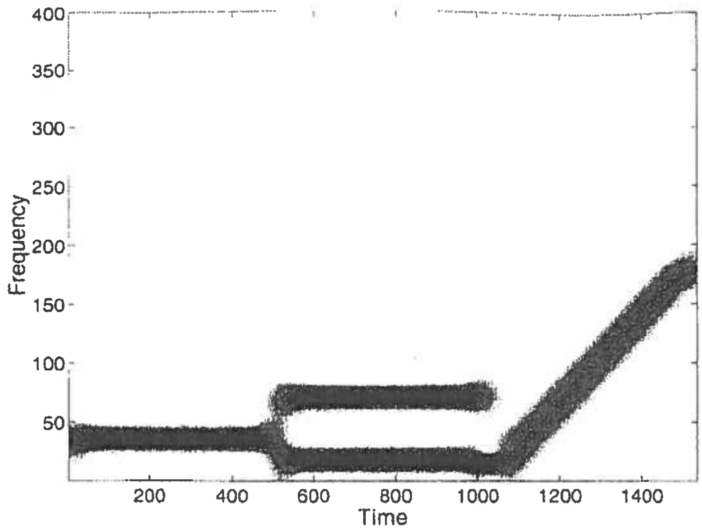

Figure 3.1: The short-time Fourier transform of a multi-component signal.

Figure 3.1 shows the absolute values of the short-time Fourier transform of the multi-component signal of Figure 2.1. The window was chosen to be a bump function similar to the one in Figure 1.1(a). The gray level is proportional to the magnitude of $|V_g f|$, so that dark regions indicate the main time-frequency concentration of f. The short-time Fourier transform separates clearly the three components of f: the first segment corresponds to the pure frequency, the two bands in the second segment indicate the superposition of two frequencies, and the third segment represents the chirp with its linearly increasing frequency.

REMARKS: 1. If g is compactly supported with its support centered at the origin, then $V_g f(x, \cdot)$ is the Fourier transform of a segment of f centered in a neighborhood of x. As x varies, the window slides along the x-axis to different positions. For this reason the STFT is often called the "sliding window Fourier transform" (Figure 3.2). With some reservations, $V_g f(x, \omega)$ can be thought of as a measure for the amplitude of the frequency band near ω at time x. In this sense $V_g f(x, \cdot)$ is a substitute for the impossible "instantaneous frequency spectrum" at x.

2. In signal analysis, at least in dimension $d = 1$, $\mathbb{R}^{2d}$ is called the *time-frequency plane*, and in physics $\mathbb{R}^{2d}$ is called the *phase space*.

3. The STFT is linear in f and conjugate linear in g. Usually the window g is kept fixed, and $V_g f$ is considered a linear mapping from functions on

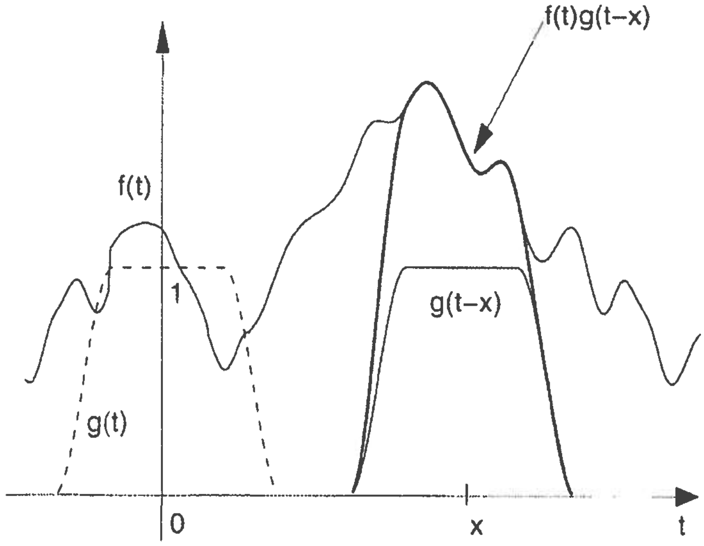

Figure 3.2: The short-time Fourier transform.

$\mathbb{R}^d$ to functions on $\mathbb{R}^{2d}$. Clearly the function $V_g f$ and the properties of the mapping $f \rightarrow V_g f$ depend crucially on the choice of the window g. We will return later to the question of how the STFT depends on the window. Chapters 12.1 and 13 are devoted entirely to the investigation of good window classes. For window optimization see also [180, 181].

We will spend some time becoming acquainted with the basic properties of the STFT. The next lemma lists several useful equivalent forms of the STFT. Recall that * is the involution $g^*(x) = \overline{g(-x)}$.

Lemma 3.1.1. *If $f, g \in L^2(\mathbb{R}^d)$, then $V_g f$ is uniformly continuous on $\mathbb{R}^{2d}$, and*

$$
\begin{aligned}
V_g f(x,\omega) &= (f \cdot T_x \bar{g})\hat{}\,(\omega) &(3.2)\\
&= \langle f, M_\omega T_x g\rangle &(3.3)\\
&= \langle \hat{f}, T_\omega M_{-x} \hat{g}\rangle &(3.4)\\
&= e^{-2\pi i x\cdot\omega} (\hat{f} \cdot T_\omega \bar{\hat{g}})\hat{}\,(-x) &(3.5)\\
&= e^{-2\pi i x\cdot\omega} V_{\hat{g}} \hat{f}(\omega, -x) &(3.6)\\
&= e^{-2\pi i x\cdot\omega} (f * M_\omega g^*)(x) &(3.7)\\
&= (\hat{f} * M_{-x} \hat{g}^*)(\omega) &(3.8)\\
&= e^{-\pi i x\cdot\omega} \int_{\mathbb{R}^d} f(t + \tfrac{x}{2})\, \bar{g}(t - \tfrac{x}{2})\, e^{-2\pi i t\cdot\omega}\, dt\,. &(3.9)
\end{aligned}
$$

Proof. The identities are mostly a matter of notation and are left as exercises. The key ingredients are Parseval's formula (1.2), the commutation relations (1.7), and formulas (1.8), (1.9), and (1.17). The uniform continuity of $V_g f$ follows from the continuity of the operator groups $\{T_x\}$ and $\{M_\omega\}$, that is, the facts

$$\lim_{x\to 0} \|T_x f - f\|_2 = 0\,,$$

and

$$\lim_{\omega\to 0} \|M_\omega f - f\|_2 = \lim_{\omega\to 0} \|T_\omega \hat{f} - \hat{f}\|_2 = 0\,.$$

■

These formulas contain different faces of the short-time Fourier transform. In (3.2) and in (3.5) the STFT is written as a (local) Fourier transform of f and $\hat{f}$, according to the main idea for its definition, whereas in (3.7) and in (3.8) the STFT is written as a convolution. In (3.3) and (3.4) $V_g f$ is written as an inner product of f with a time-frequency shift. This form is most convenient for formal manipulations and reveals some of the deeper structures of the STFT (see in particular Chapter 9). The symmetric form $\int f(t+\frac{x}{2})\,\bar{g}(t-\frac{x}{2})e^{-2\pi i t\cdot\omega}\,dt$ is often called the *cross-ambiguity function*. It plays an important role in radar and in optics [56, 258]. Except for the phase factor $e^{-\pi i x\cdot\omega}$, which can be frequently neglected, it coincides with the STFT. See also Chapter 4.2.

Formula (3.6), namely,

$$V_g f(x,\omega) = e^{-2\pi i x\cdot\omega} V_{\hat{g}} \hat{f}(\omega, -x) \tag{3.10}$$

is the *fundamental identity of time-frequency analysis*. It combines both f and $\hat{f}$ into a joint time-frequency representation. In this representation the Fourier transform amounts to a rotation of the time-frequency plane by an angle of $\pi/2$.

In Lemma 3.1.1 we have emphasized the linearity of the STFT in the case of a fixed window g. Alternatively, the STFT may be considered as the sesquilinear form $(f,g) \longmapsto V_g f$. Let $f\otimes g$ be the (tensor) product $f\otimes g(x,t) = f(x)g(t)$, let $\mathcal{T}_a$ be the asymmetric coordinate transform

$$\mathcal{T}_a F(x,t) = F(t, t-x)\,, \tag{3.11}$$

and let $\mathcal{F}_2$ be the partial Fourier transform

$$\mathcal{F}_2 F(x,\omega) = \int_{\mathbb{R}^d} F(x,t)\, e^{-2\pi i t\cdot\omega}\, dt \tag{3.12}$$

of a function F on $\mathbb{R}^{2d}$. Using this notation, Definition 3.1.1 can be reformulated in terms of a factorization of the STFT.

Lemma 3.1.2. *If* $f, g \in L^2(\mathbb{R}^d)$, *then*

$$V_g f = \mathcal{F}_2 \mathcal{T}_a(f \otimes \bar{g}) \,. \tag{3.13}$$

The Domain of the Short-Time Fourier Transform. In Definition 3.1.1 we have been unprofessionally sloppy and have not specified a domain for f and g. Clearly, if $f, g \in L^2(\mathbb{R}^d)$, then $f \cdot T_x\bar{g} \in L^1(\mathbb{R}^d)$, and $V_g f(x,\omega) = (f \cdot T_x \bar{g})\hat{}\,(\omega)$ is defined pointwise. Similarly, if $g \in L^p(\mathbb{R}^d)$ and $f \in L^{p'}(\mathbb{R}^d)$, then by Hölder's inequality $f \cdot T_x\bar{g} \in L^1(\mathbb{R}^d)$ and again the STFT is defined pointwise.

Writing the STFT as the inner product $V_g f(x,\omega) = \langle f, M_\omega T_x g\rangle$ is useful for extending it to situations when the integral is no longer defined. As a rule of thumb, we may consider the STFT, whenever the bracket $\langle \cdot, \cdot \rangle$ is well defined by some form of duality. For example, if B is a Banach space contained in $\mathcal{S}'(\mathbb{R}^d)$ that is invariant under time-frequency shifts, then the STFT is defined when $f \in B$, $g \in B^*$ or $f \in B^*, g \in B$. More generally, $V_g f$ is well-defined for all tempered distributions $f \in \mathcal{S}'(\mathbb{R}^d)$, provided that $g \in \mathcal{S}(\mathbb{R}^d)$. The detailed study of the STFT on $\mathcal{S}'$ and the time-frequency analysis of tempered distributions will be pursued in Chapters 11 and 12.

With Lemma 3.1.2, the domain of the STFT can be extended even further. Note first that both operators $\mathcal{T}_a$ and $\mathcal{F}_2$ are isomorphisms on $\mathcal{S}'(\mathbb{R}^{2d})$. If $f, g \in \mathcal{S}'(\mathbb{R}^d)$, then $f \otimes \bar{g} \in \mathcal{S}'(\mathbb{R}^{2d})$, and consequently $V_g f = \mathcal{F}_2\mathcal{T}_a(f \otimes \bar{g}) \in \mathcal{S}'(\mathbb{R}^{2d})$ as well. Thus $V_g f$ is a well-defined tempered distribution, whenever $f, g \in \mathcal{S}'(\mathbb{R}^d)$.

The next property is sometimes called the *covariance property* of the STFT.

Lemma 3.1.3. *Whenever* $V_g f$ *is defined, we have*

$$V_g(T_u M_\eta f)(x,\omega) = e^{-2\pi i u\cdot\omega} V_g f(x-u, \omega-\eta) \tag{3.14}$$

for $x, u, \omega, \eta \in \mathbb{R}^d$. *In particular,*

$$|V_g(T_u M_\eta f)(x,\omega)| = |V_g f(x-u, \omega-\eta)| \,.$$

Proof. We substitute the commutation relation $M_{-\eta}T_{-u}M_\omega T_x = e^{2\pi i u\cdot\omega} M_{\omega-\eta}T_{x-u}$ into the definition and obtain

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

■

3.2 Orthogonality Relations and Inversion Formula

The STFT enjoys several properties similar to those possessed by the ordinary Fourier transform. The following theorem on inner products of STFT's corresponds to Parseval's formula (1.2), and will be used frequently.

Theorem 3.2.1 (Orthogonality relations for STFT). *Let $f_1, f_2, g_1, g_2 \in L^2(\mathbb{R}^d)$; then $V_{g_j} f_j \in L^2(\mathbb{R}^{2d})$ for $j = 1, 2$, and*

$$\langle V_{g_1} f_1, V_{g_2} f_2 \rangle_{L^2(\mathbb{R}^{2d})} = \langle f_1, f_2 \rangle \overline{\langle g_1, g_2 \rangle}\,. \tag{3.15}$$

Proof. We first assume that the windows g_j are in $L^1 \cap L^\infty(\mathbb{R}^d) \subseteq L^2(\mathbb{R}^d)$, so that $f_j \cdot T_x \bar{g}_j \in L^2(\mathbb{R}^d)$ for all $x \in \mathbb{R}^d$. Therefore Parseval's formula applies to the ω-integral and yields

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

Here $f_1 \bar{f}_2 \in L^1(\mathbb{R}^d, dt)$ and $\overline{g_1}\, g_2 \in L^1(\mathbb{R}^d, dx)$, therefore Fubini's theorem (Appendix A.13) allows us to interchange the order of integration. We continue as follows:

$$\begin{aligned}
\langle V_{g_1} f_1, V_{g_2} f_2 \rangle_{L^2(\mathbb{R}^{2d})} &= \int_{\mathbb{R}^d} f_1(t) \overline{f_2(t)} \left(\int_{\mathbb{R}^d} \overline{g_1(t-x)}\, g_2(t-x)\, dx \right) dt \\
&= \langle f_1, f_2 \rangle \overline{\langle g_1, g_2 \rangle}\,.
\end{aligned}$$

The extension to general $g_j \in L^2(\mathbb{R}^d)$ is done by a standard density argument (see Appendix A.1). With $g_1 \in L^1 \cap L^\infty$ fixed, the mapping $g_2 \longmapsto \langle V_{g_1} f_1, V_{g_2} f_2 \rangle_{L^2(\mathbb{R}^{2d})}$ is a linear functional that coincides with $\langle f_1, f_2 \rangle \langle g_2, g_1 \rangle$ on the dense subspace $L^1 \cap L^\infty$. It is therefore bounded and extends to all $g_2 \in L^2(\mathbb{R}^d)$. In the same way, for arbitrary f_1, f_2 and $g_2 \in L^2(\mathbb{R}^d)$, the conjugate linear functional $g_1 \longmapsto \langle V_{g_1} f_1, V_{g_2} f_2 \rangle_{L^2(\mathbb{R}^{2d})}$ equals $\langle f_1, f_2 \rangle \overline{\langle g_1, g_2 \rangle}$ on $L^1 \cap L^\infty$ and extends to all of L^2.

The orthogonality relations are therefore established for all $f_j, g_j \in L^2(\mathbb{R}^d)$. ■

Second Proof of the Orthogonality Relations. We use the factorization (3.13) of the STFT. Since on $L^2(\mathbb{R}^{2d})$ both operators $\mathcal{F}_2$ and $\mathcal{T}_a$ are unitary, we deduce the orthogonality relations as follows:

$$\begin{aligned}\langle V_{g_1} f_1, V_{g_2} f_2\rangle_{L^2(\mathbb{R}^{2d})} &= \langle \mathcal{F}_2 \mathcal{T}_a(f_1 \otimes \bar{g}_1), \mathcal{F}_2 \mathcal{T}_a(f_2 \otimes \bar{g}_2)\rangle \\ &= \langle f_1 \otimes \bar{g}_1, f_2 \otimes \bar{g}_2\rangle_{L^2(\mathbb{R}^{2d})} \\ &= \langle f_1, f_2\rangle \, \overline{\langle g_1, g_2\rangle} \,.\end{aligned}$$

■

Corollary 3.2.2. *If $f, g \in L^2(\mathbb{R}^d)$, then*

$$\|V_g f\|_2 = \|f\|_2 \, \|g\|_2 \,.$$

In particular, if $\|g\|_2 = 1$ then

$$\|f\|_2 = \|V_g f\|_2 \quad \text{for all } f \in L^2(\mathbb{R}^d) \,. \tag{3.16}$$

Thus, in this case the STFT is an isometry from $L^2(\mathbb{R}^d)$ into $L^2(\mathbb{R}^{2d})$.

It follows from (3.16) that f is completely determined by $V_g f$. Furthermore, the implication $\langle f, M_\omega T_x g\rangle = 0$, $\forall x, \omega \in \mathbb{R}^d \Rightarrow f = 0$ is equivalent to saying that for each fixed $g \in L^2(\mathbb{R}^d)$ the set $\{M_\omega T_x g : x, \omega \in \mathbb{R}^d\}$ spans a dense subspace of $L^2(\mathbb{R}^d)$. This still leaves open the question of how f can be recovered from $V_g f$. We will show that the orthogonality relations imply a remarkable inversion formula.

For the formulation of such an inversion formula, we need a brief explanation of vector-valued or operator-valued integrals. In this book vector- or operator-valued integrals are always understood in a weak sense (unless specifically stated otherwise). If g is a function on $\mathbb{R}^d$ that takes values in a Banach space B, that is, $g(x) \in B$ for all $x \in \mathbb{R}^d$, then $f = \int_{\mathbb{R}^d} g(x)\, dx$ means that

$$\langle f, h\rangle = \int_{\mathbb{R}^d} \langle g(x), h\rangle \, dx$$

for all $h \in B^*$. If the mapping $\ell(h) \longmapsto \int_{\mathbb{R}} \langle g(x), h\rangle \, dx$ is a bounded (conjugate-)linear functional on B^*, then ℓ defines a unique element $f \in B^{**}$. Although $g(x) \in B$, in general we can only say that the vector-valued integral is in the bidual B^{**}. This technical difficulty need not worry us, because we will work mostly with reflexive Banach spaces, that is, $B^{**} = B$.

The most important vector-valued integrals in time-frequency analysis are superpositions of time-frequency shifts of the form

$$f = \iint_{\mathbb{R}^{2d}} F(x, \omega) M_\omega T_x g \, dx \, d\omega \,. \tag{3.17}$$

For example, if $F \in L^2(\mathbb{R}^{2d})$, then the conjugate-linear functional

$$\ell(h) = \iint_{\mathbb{R}^{2d}} F(x, \omega) \overline{\langle h, M_\omega T_x g\rangle} \, dx \, d\omega \tag{3.18}$$

is a bounded functional on $L^2(\mathbb{R}^d)$. To see this, apply the Cauchy-Schwartz inequality to (3.18) and use Corollary 3.2.2:

$$|\ell(h)| \leq \|F\|_2 \, \|V_g h\|_2 = \|F\|_2 \, \|g\|_2 \, \|h\|_2 \,. \tag{3.19}$$

This means that ℓ defines a unique function $f = \int_{\mathbb{R}^{2d}} F(x,\omega) M_\omega T_x g \, dx \, d\omega \in L^2(\mathbb{R}^d)$ with norm $\|f\|_2 \leq \|F\|_2 \|g\|_2$ and satisfying $\ell(h) = \langle f, h \rangle$.

We are now ready to state a precise version of the inversion formula for the STFT.

Corollary 3.2.3 (Inversion formula for the STFT). *Suppose that $g, \gamma \in L^2(\mathbb{R}^d)$ and $\langle g, \gamma \rangle \neq 0$. Then for all $f \in L^2(\mathbb{R}^d)$*

$$f = \frac{1}{\langle \gamma, g \rangle} \iint_{\mathbb{R}^{2d}} V_g f(x,\omega) \, M_\omega T_x \gamma \, d\omega \, dx \,. \tag{3.20}$$

Proof. Since $V_g f \in L^2(\mathbb{R}^{2d})$ by Corollary 3.2.2, the vector-valued integral

$$\tilde{f} = \frac{1}{\langle \gamma, g \rangle} \iint_{\mathbb{R}^{2d}} V_g f(x,\omega) \, M_\omega T_x \gamma \, dx \, d\omega$$

is a well-defined function in $L^2(\mathbb{R}^d)$. Further, using the orthogonality relations, we see that

$$\begin{aligned} \langle \tilde{f}, h \rangle &= \frac{1}{\langle \gamma, g \rangle} \iint_{\mathbb{R}^{2d}} V_g f(x,\omega) \, \overline{\langle h, M_\omega T_x \gamma \rangle} \, dx \, d\omega \\ &= \frac{1}{\langle \gamma, g \rangle} \langle V_g f, V_\gamma h \rangle = \langle f, h \rangle \,. \end{aligned}$$

Thus $\tilde{f} = f$, and the inversion formula is proved. ∎

REMARKS: 1. The inversion formula (3.20) shows that f can be expressed as a continuous superposition of time-frequency shifts with the STFT as weight function. In this sense, (3.20) is similar to the inversion formula for the Fourier transform, that is, $f(x) = \int \hat{f}(\omega) e^{2\pi i x \cdot \omega} \, d\omega$. However, in Fourier inversion the elementary functions $e^{2\pi i x \cdot \omega}$ are not in $L^2(\mathbb{R}^d)$, whereas in Corollary 3.2.3 the elementary functions $M_\omega T_x \gamma$ are particularly nice functions in $L^2(\mathbb{R}^d)$.

2. The vector-valued integrals of the form (3.17) are closely related to the STFT itself. Let A_g be the linear operator defined by $A_g F = \iint_{\mathbb{R}^{2d}} F(x,\omega) M_\omega T_x g \, dx \, d\omega$. By (3.19) A_g is a bounded operator from $L^2(\mathbb{R}^{2d})$ onto $L^2(\mathbb{R}^d)$. Moreover, A_g is exactly the adjoint operator of the STFT V_g

(viewed as an operator from $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^{2d})$). This claim follows from the computation

$$\begin{aligned}\langle A_g F, h\rangle &= \iint_{\mathbb{R}^{2d}} F(x,\omega)\langle M_\omega T_x g, h\rangle \, dx\, d\omega \\ &= \langle F, V_g h\rangle = \langle V_g^* F, h\rangle ,\end{aligned}$$

where $h \in L^2(\mathbb{R}^d)$ and $F \in L^2(\mathbb{R}^{2d})$. Thus indeed $V_g^* = A_g$. The inversion formula (3.20) thus reads as

$$\frac{1}{\langle \gamma, g\rangle} V_\gamma^* V_g = \mathrm{I}\,. \tag{3.21}$$

This point of view will be important in Chapter 11, where we extend the theory of the STFT to Banach spaces and norms other than the L^2-norm.

The inversion formula reveals what kind of time-frequency analysis is possible despite the uncertainty principle. Suppose that γ is a "nice" function with essential support $T \subseteq \mathbb{R}^d$ and essential spectrum (=supp $\hat{\gamma}$) $\Omega \subseteq \mathbb{R}^d$, in the sense of Definition 2.3.1 for small ϵ_T and ϵ_Ω. Then $M_\omega T_x \gamma$ is essentially supported on $x + T$, and its essential spectrum is $\omega + \Omega$. Thus $M_\omega T_x \gamma$ occupies the cell $(x + T) \times (\omega + \Omega)$ in the time-frequency plane, and the size of $V_g f(x,\omega)$ measures the contribution of this time-frequency atom in the decomposition of f. The uncertainty principle of Donoho and Stark (Theorem 2.3.1) limits the possible time-frequency concentration to $|T|\,|\Omega| \geq 1 - \delta$.

At least in a qualitative sense, we can now understand the role of the window for the properties of the STFT. Good time resolution requires a window with small support, that is, small $|T|$, but this comes at the price of a poor frequency resolution since $|\Omega| > \frac{1-\delta}{|T|}$ becomes large. In the same way, good frequency resolution by means of a band-limited window implies poor resolution in time. These brief remarks demonstrate the influence of the window on the time-frequency properties of the STFT. In practice one will choose a window such that both g and $\hat{g}$ decay rapidly. For example, a Schwartz function or a C^∞-function with compact support is suitable. However, a characteristic function is not! Indeed, if $g = \chi_{[0,1]^d}$, then $V_g f$ provides an accurate picture of the temporal behavior of f since $V_g f(x, 0) = \int_{x+[0,1]^d} f(t)\, dt$ is the average value of f in a neighborhood of x. On the other hand, since $\widehat{\chi_{[0,1]^d}}(\omega) = \prod_{j=1}^d \frac{1-e^{-2\pi i\omega_j}}{2\pi i\omega_j}$ decays slowly and is not even in $L^1(\mathbb{R}^d)$, the STFT $V_g f(x,\omega) = V_{\hat{g}} \hat{f}(\omega, -x) e^{2\pi i x\cdot\omega}$ provides a completely inadequate frequency resolution. In particular, one signal processing application known commonly as "signal segmentation" amounts to using a STFT whose window is a characteristic function. The lack of frequency resolution then presents itself as a severe problem in this approach.

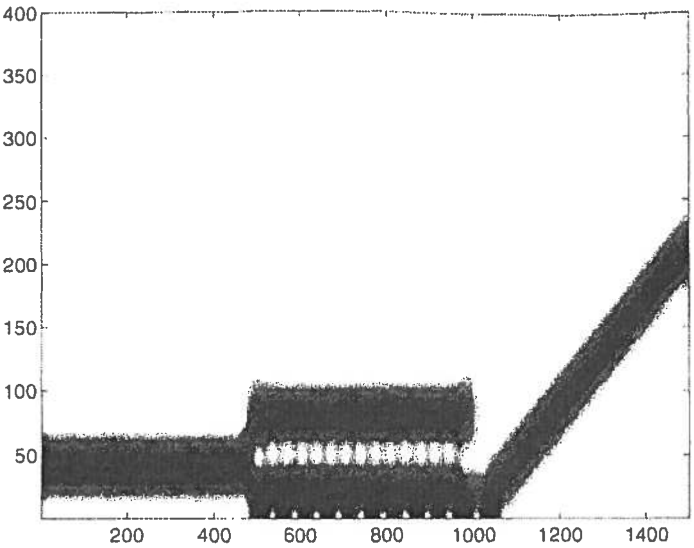

Figure 3.3: The short-time Fourier transform of a multi-component signal with respect to a "short window."

Figures 3.3 and 3.4 illustrate the dependence of the STFT of the window. Both plots show the STFT of the multi-component signal of Figure 2.1. In Figure 3.3 the window is a bump function with small support and large spectrum. Consequently, the STFT possesses good time resolution, which is clearly visible in the neat separation of the three components of the signal. On the other hand, the wide frequency bands and the interference patterns in the second segment show the poor frequency resolution of this STFT. By contrast, the window in Figure 3.4 is a bump function with large support and narrow spectrum. In this case, the STFT provides a clean resolution of the frequencies in each segment of the signal. However, the time resolution is only mediocre, the transitions between the segments are fuzzy, and the chirp is badly localized.

Schematically the time-frequency analysis of a signal consists of three distinct steps:

A. *Analysis:* Given a signal (or image) f, its STFT $V_g f$ with respect to a suitable window is computed, and is interpreted as a joint time-frequency information for f.

B. *Processing:* $V_g f(x,\omega)$ is transformed into some new function $F(x,\omega)$. A typical processing step often contains a truncation of $V_g f$ to a

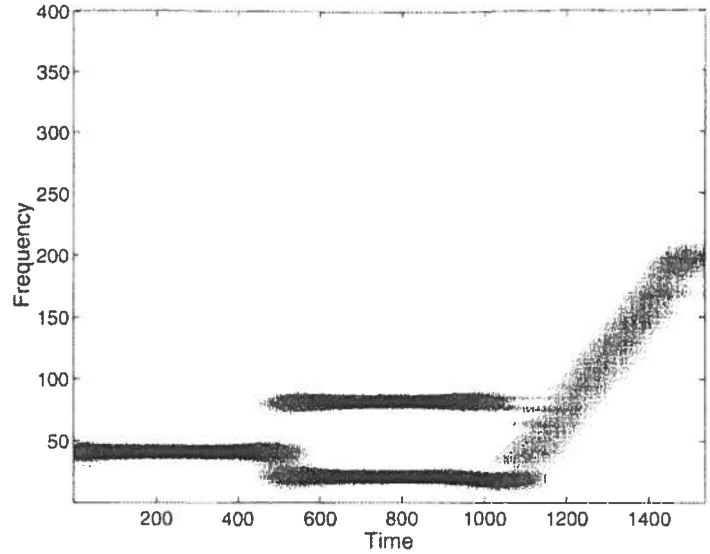

Figure 3.4: The short-time Fourier transform of a multi-component signal with respect to a "long window."

region where something interesting seems to happen or where $|V_g f|$ is above a given threshold. In an application-oriented language, one speaks of feature extraction, separation of signal components, and signal compression.

C. *Synthesis:* The processed signal (or image) is then reconstructed by using the modified inversion formula

$$\tilde{f} = \iint_{\mathbb{R}^{2d}} F(x,\omega) M_\omega T_x \gamma \, dx \, d\omega \tag{3.22}$$

with respect to a suitable synthesis window γ. We remark explicitly that distinct windows may be used for the analysis and synthesis.

Theoretical physics uses a different language to describe the inversion formula (3.20) [1, 177, 207]. The time-frequency shifts $T_x M_\omega g$ of a fixed window are called generalized coherent states, and the inversion formula is interpreted as a decomposition of a quantum-mechanical state f into coherent states. The coherent states in the strict sense are the time-frequency shifts of Gaussian functions. In this case, (3.20) amounts to a decomposition into states of minimal uncertainty. It is customary to write the inversion formula as a superposition of rank one operators. Let $\mathcal{H}$ be a Hilbert space, and let $u \otimes \bar{v}$ denote the rank one operator defined by $(u \otimes \bar{v})(h) = \langle h, v \rangle u$ for $u, v, h \in \mathcal{H}$. Then (3.20) is the following continuous resolution of the identity operator

$$I = \frac{1}{\langle \gamma, g \rangle} \iint_{\mathbb{R}^{2d}} M_\omega T_x \gamma \otimes \overline{M_\omega T_x g} \, dx \, d\omega .$$

Next we prove a strong version of the inversion formula. Its formulation resembles the definition of the Fourier transform of an L^2-function by an approximation procedure (see the discussion of Plancherel's theorem in Chapter 1.1). For the approximation we consider a nested sequence of compact sets $K_n \subseteq \mathbb{R}^{2d}$ that exhaust $\mathbb{R}^{2d}$. This means that $\bigcup_{n \geq 1} K_n = \mathbb{R}^{2d}$ and $K_n \subseteq \operatorname{int} K_{n+1}$. Then any compact set is contained in some K_n. The cubes $[-n, n]^{2d}$ or the balls $\bar{B}(0, n) = \{x \in \mathbb{R}^{2d} : |x| \leq n\}$ are common choices for K_n.

Theorem 3.2.4. *Fix $g, \gamma \in L^2(\mathbb{R}^d)$ and let $K_n \subseteq \mathbb{R}^{2d}$ for $n \geq 1$ be a nested exhausting sequence of compact sets. Define f_n to be*

$$f_n = \frac{1}{\langle \gamma, g \rangle} \iint_{K_n} V_g f(x, \omega) M_\omega T_x \gamma \, dx \, d\omega .$$

Then $\lim_{n \to \infty} \|f - f_n\|_2 = 0$.

Proof. Using the Cauchy–Schwartz inequality and Corollary 3.2.2, we estimate for $h \in L^2(\mathbb{R}^d)$ that

$$\begin{aligned} |\langle f_n, h \rangle| &= \frac{1}{|\langle \gamma, g \rangle|} \left| \iint_{K_n} V_g f(x, \omega) \overline{V_\gamma h(x, \omega)} \, dx \, d\omega \right| \\ &\leq \frac{1}{|\langle \gamma, g \rangle|} \|V_g f\|_2 \, \|V_\gamma h\|_2 \\ &= \frac{1}{|\langle \gamma, g \rangle|} \|g\|_2 \, \|f\|_2 \, \|\gamma\|_2 \, \|h\|_2 . \end{aligned}$$

Therefore for each n, f_n is a well-defined element of $L^2(\mathbb{R}^d)$, and furthermore, $\|f_n\|_2 \leq |\langle \gamma, g \rangle|^{-1} \|g\|_2 \, \|\gamma\|_2 \, \|f\|_2$ by Corollary 3.2.2. Next, we estimate similarly that

$$\begin{aligned} |\langle f - f_n, h \rangle| &= \frac{1}{|\langle \gamma, g \rangle|} \left| \left(\iint_{\mathbb{R}^{2d}} - \iint_{K_n} \right) V_g f(x, \omega) \overline{V_\gamma h(x, \omega)} \, dx \, d\omega \right| \\ &= \frac{1}{|\langle \gamma, g \rangle|} \left| \iint_{K_n^c} V_g f(x, \omega) \overline{V_\gamma h(x, \omega)} \, dx \, d\omega \right| \\ &\leq \frac{1}{|\langle \gamma, g \rangle|} \|V_\gamma h\|_2 \left(\iint_{K_n^c} |V_g f(x, \omega)|^2 \, dx \, d\omega \right)^{1/2} \\ &= \frac{1}{|\langle \gamma, g \rangle|} \|\gamma\|_2 \|h\|_2 \left(\iint_{K_n^c} |V_g f(x, \omega)|^2 \, dx \, d\omega \right)^{1/2} . \end{aligned}$$

Since this is true for all $h \in L^2(\mathbb{R}^d)$, we therefore have

$$\|f - f_n\|_2 = \sup_{\|h\|_2 \le 1} |\langle f - f_n, h\rangle| \le \frac{1}{|\langle \gamma, g\rangle|} \|\gamma\|_2 \left(\iint_{K_n^c} |V_g f(x,\omega)|^2 \, dx \, d\omega \right)^{1/2} .$$

Since $V_g f \in L^2(\mathbb{R}^{2d})$, and K_n is exhausting, the right-hand side becomes arbitrarily small as n increases. ∎

Benedetto, Heil, and Walnut [16, 142] use approximate units for their version of the inversion formula. *Let $\{u_n\}$ be an approximate unit in $L^1 \cap \mathcal{F}L^1(\mathbb{R}^d)$ and let $g, \gamma \in L^1 \cap L^\infty(\mathbb{R}^d)$. Given $f \in L^p(\mathbb{R}^d)$, $1 \le p < \infty$, define*

$$f_n(t) = \frac{1}{\langle \gamma, g\rangle} \iint_{\mathbb{R}^{2d}} V_g f(x,\omega) M_\omega T_x \gamma(t) \hat{u}_n(\omega) \, d\omega \, dx .$$

Then $\lim_{n\to\infty} \|f - f_n\|_p = 0$.

Under additional assumptions and with proper mathematical care, the inversion formulas of this section can be extended to other function spaces. For instance, in Chapter 11.2 we will extend Corollary 3.2.3 to tempered distributions.

3.3 Lieb's Uncertainty Principle

In the discussion of the inversion formula for the STFT we saw how the time-frequency resolution of the STFT depends on the choice of the window function g. In particular, the time-frequency resolution of $V_g f$ is limited by the size of the essential supports of g and $\hat{g}$. The classical uncertainty principle (Theorem 2.2.1) for g thus implies an uncertainty principle for $V_g f$. By contrast, in this section we present uncertainty principles that apply directly to the STFT. They are the first manifestation of the following generic principle:

> *A function cannot be concentrated on small sets in the time-frequency plane, no matter which time-frequency representation is used.*

Here is an easy version of such an uncertainty principle. It is analogous to the uncertainty principle of Donoho and Stark for the pair $(f, \hat{f})$ of Chapter 2.3.

Proposition 3.3.1 (Weak uncertainty principle for the STFT). *Suppose that* $\|f\|_2 = \|g\|_2 = 1$ *and that* $U \subseteq \mathbb{R}^{2d}$ *and* $\epsilon \geq 0$ *are such that*

$$\iint_U |V_g f(x,\omega)|^2 \, dx \, d\omega \geq 1 - \epsilon \, .$$

Then $|U| \geq 1 - \epsilon$.

Proof. The Cauchy–Schwartz inequality implies that

$$|V_g f(x,\omega)| = |\langle f, M_\omega T_x g\rangle| \leq \|f\|_2 \|g\|_2 = 1 \quad \text{for all } (x,\omega) \in \mathbb{R}^{2d} \, . \tag{3.23}$$

Therefore,

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

■

A much deeper and stronger inequality for the STFT was proved by E. Lieb [190].

Theorem 3.3.2. *If* $f, g \in L^2(\mathbb{R}^d)$ *and* $2 \leq p < \infty$, *then*

$$\iint_{\mathbb{R}^{2d}} |V_g f(x,\omega)|^p \, dx \, d\omega \leq \left(\tfrac{2}{p}\right)^d \left(\|f\|_2 \, \|g\|_2\right)^p \, . \tag{3.24}$$

Proof. Let p' be the conjugate index defined by $\frac{1}{p} + \frac{1}{p'} = 1$. Since $2 \leq p < \infty$, we have $1 < p' \leq 2$. We note that $f \cdot T_x \bar{g} \in L^1(\mathbb{R}^d)$ by the Cauchy–Schwartz inequality. Moreover, since $V_g f(x,\omega) = (f \cdot T_x \bar{g})\hat{}\,(\omega) \in L^2(\mathbb{R}^{2d})$ by Corollary 3.2.2, we conclude from Fubini's theorem (Appendix A.13) that $(f \cdot T_x \bar{g})\hat{} \in L^2(\mathbb{R}^d)$ for almost all $x \in \mathbb{R}^d$. Thus $f \cdot T_x \bar{g} \in L^1 \cap L^2(\mathbb{R}^d)$ for almost all x, which implies $f \cdot T_x \bar{g} \in L^{p'}(\mathbb{R}^d)$ for a.a. x. The Hausdorff–Young inequality (1.4) implies

$$\begin{aligned}
\left(\int_{\mathbb{R}^d} |V_g f(x,\omega)|^p \, d\omega\right)^{1/p} &= \left(\int_{\mathbb{R}^d} |(f \cdot T_x \bar{g})\hat{}\,(\omega)|^p \, d\omega\right)^{1/p} \\
&\leq A_{p'}^d \left(\int_{\mathbb{R}^d} |(f \cdot T_x \bar{g})(y)|^{p'} \, dy\right)^{1/p'} \\
&= A_{p'}^d \left(\int_{\mathbb{R}^d} |f(y)|^{p'} \, |\bar{g}(y - x)|^{p'} \, dy\right)^{1/p'} \\
&= A_{p'}^d \left(|f|^{p'} * |g^*|^{p'}(x)\right)^{1/p'} ,
\end{aligned}$$

where $A_{p'} = \left((p')^{1/p'} p^{-1/p}\right)^{1/2}$ and $g^*(x) = \overline{g(-x)}$. Hence

$$\begin{aligned} \|V_g f\|_p &= \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |(f \cdot T_x \bar{g})\hat{}\,(\omega)|^p \, d\omega\right) dx\right)^{1/p} \\ &\leq A_{p'}^d \left(\int_{\mathbb{R}^d} \left(|f|^{p'} * |g^*|^{p'}(x)\right)^{p/p'} dx\right)^{1/p} \qquad (3.25) \\ &= A_{p'}^d \, \| \, |f|^{p'} * |g^*|^{p'} \, \|_{p/p'}^{1/p'} . \end{aligned}$$

Now we apply Young's inequality (Theorem 1.2.1) to the functions $|f|^{p'}$ and $|g^*|^{p'}$, which are elements of $L^{2/p'}(\mathbb{R}^d)$, with the triple (p, q, r) being replaced by (s, s, t), where $s = \frac{2}{p'} \geq 1$, $t = \frac{p}{p'}$ (note that $\frac{1}{s} + \frac{1}{s} = 1 + \frac{1}{t}$). We obtain

$$\| \, |f|^{p'} * |g^*|^{p'} \, \|_t \leq A_s^{2d} A_{t'}^d \, \| \, |f|^{p'} \, \|_s \, \| \, |g^*|^{p'} \, \|_s .$$

However, $\| \, |f|^{p'} \, \|_s = (\int |f(x)|^{p' \cdot \frac{2}{p'}})^{p'/2} = \|f\|_2^{p'}$ and $\| \, |g^*|^{p'} \, \|_s = \|g\|_2^{p'}$. Inserting these into (3.25) yields

$$\begin{aligned} \|V_g f\|_p &\leq A_{p'}^d \left(A_s^{2d} A_{t'}^d \|f\|_2^{p'} \|g\|_2^{p'}\right)^{1/p'} \\ &= A_{p'}^d \, A_s^{2d/p'} \, A_{t'}^{d/p'} \, \|f\|_2 \, \|g\|_2 . \end{aligned}$$

Finally, we leave as an exercise the verification that $A_{p'}^d \, A_s^{2d/p'} \, A_{t'}^{d/p'} = \left(\frac{2}{p}\right)^d$.
■

REMARKS: 1. A careful analysis of the minimizing functions in the sharp version of the inequalities of Young and Hausdorff–Young shows that equality in Lieb's uncertainty principle is obtained if and only if f and g are time-frequency shifts of Gaussians [190].

2. Lieb's paper contains a number of other important inequalities about the STFT. Here we only mention the counterpart of Theorem 3.3.2 for the case that $1 \leq p \leq 2$. *If* $f, g \in L^2(\mathbb{R}^d)$ *and* $1 \leq p \leq 2$, *then*

$$\iint_{\mathbb{R}^{2d}} |V_g f(x, \omega)|^p \, dx \, d\omega \geq \left(\frac{2}{p}\right)^d \left(\|f\|_2 \, \|g\|_2\right)^p . \qquad (3.26)$$

Equality holds if and only if $p > 1$ and f, g are certain Gaussians. The proof is similar to the one for $p \geq 2$, but is technically more involved.

3. Lieb's uncertainty principle carries over verbatim to general locally compact abelian groups [124].

Next we show that Lieb's uncertainty principle improves Proposition 3.3.1 and yields a sharper estimate for the essential support of $V_g f$ (which was apparently not previously observed).

Theorem 3.3.3. *Suppose that* $\|f\|_2 = \|g\|_2 = 1$. *If* $U \subseteq \mathbb{R}^{2d}$ *and* $\epsilon \geq 0$ *are such that*

$$\iint_U |V_g f(x,\omega)|^2 \, dx \, d\omega \geq 1 - \epsilon \,,$$

then

$$|U| \geq (1-\epsilon)^{\frac{p}{p-2}} \left(\frac{p}{2}\right)^{\frac{2d}{p-2}} \qquad \text{for all } p > 2 \,.$$

In particular,

$$|U| \geq \sup_{p>2} (1-\epsilon)^{\frac{p}{p-2}} \left(\frac{p}{2}\right)^{\frac{2d}{p-2}} \geq (1-\epsilon)^2 2^d \,.$$

Proof. We first apply Hölder's inequality with exponents $q = \frac{p}{2}$ and $q' = \frac{p}{p-2}$, and then we use Lieb's inequality in the second step:

$$\begin{aligned} 1 - \epsilon &\leq \iint_U |V_g f(x,\omega)|^2 \, dx \, d\omega \\ &\leq \left(\iint_{\mathbb{R}^{2d}} |V_g f(x,\omega)|^{2\cdot\frac{p}{2}} \, dx \, d\omega \right)^{2/p} \left(\iint_{\mathbb{R}^{2d}} \chi_U(x,\omega)^{q'} \, dx \, d\omega \right)^{\frac{p-2}{p}} \\ &\leq \left(\frac{2}{p}\right)^{\frac{2d}{p}} \left(\|f\|_2 \|g\|_2 \right)^2 |U|^{\frac{p-2}{p}} \,. \end{aligned}$$

Thus for all $p > 2$

$$|U| \geq (1-\epsilon)^{\frac{p}{p-2}} \left(\frac{p}{2}\right)^{\frac{2d}{p-2}} \,.$$

For $p = 4$ this becomes $|U| \geq (1-\epsilon)^2 2^d$. ■

Note that taking $\epsilon = 0$ in Theorem 3.3.3 yields the following lower bound for the support of $V_g f$:

$$|\operatorname{supp} V_g f| \geq \lim_{p \to 2+} \left(\frac{p}{2}\right)^{\frac{2d}{p-2}} = e^d \,. \qquad (3.27)$$

In analogy to Theorem 2.3.3 for the pair $(f, \hat{f})$, it can be shown that if $|\operatorname{supp} V_g f| < \infty$, then either $f = 0$ or $g = 0$. See [159, 173, 254].

3.4 The Bargmann Transform

Since Gaussian functions minimize the uncertainty principle (Theorem 2.2.1), it is of special interest to study the STFT with respect to a Gaussian window. In the light of our previous discussion, this STFT will provide the optimal resolution of signals in the time-frequency plane. In quantum mechanics and in quantum optics, these states of minimal uncertainty, that is, time-frequency shifts of a Gaussian, play an important role. They are a widely used tool usually referred to by the name coherent states, see [177, 207].

Let $\varphi(x) = 2^{d/4}e^{-\pi x^2}$ be the Gaussian on $\mathbb{R}^d$, normalized such that $\|\varphi\|_2 = 1$. Then

$$\begin{aligned} V_\varphi f(x,\omega) &= 2^{d/4}\int_{\mathbb{R}^d} f(t)\, e^{-\pi(t-x)^2}\, e^{-2\pi i\omega\cdot t}\, dt \\ &= 2^{d/4}\int_{\mathbb{R}^d} f(t)\, e^{-\pi t^2}\, e^{2\pi x\cdot t}\, e^{-\pi x^2} e^{-2\pi i\omega\cdot t}\, dt \qquad (3.28) \\ &= 2^{d/4}\, e^{-\pi i x\cdot\omega}\, e^{-\frac{\pi}{2}(x^2+\omega^2)} \int_{\mathbb{R}^d} f(t) e^{-\pi t^2}\, e^{2\pi t\cdot(x-i\omega)} e^{-\frac{\pi}{2}(x-i\omega)^2}\, dt\,. \end{aligned}$$

Let us convert $(x,\omega) \in \mathbb{R}^{2d}$ into a complex vector $z = x + i\omega \in \mathbb{C}^d$. We will keep the notation consistent with $\mathbb{R}^d$, that is, we will write $z^2 = (x+i\omega)\cdot(x+i\omega)$ and $|z|^2 = z\cdot\bar{z} = (x+i\omega)\cdot(x-i\omega) = x^2+\omega^2$. Further, dz denotes the Lebesgue measure on $\mathbb{C}^d$. Then, comparing to (3.28) the following definition is quite natural [10].

Definition 3.4.1. The *Bargmann transform* of a function f on $\mathbb{R}^d$ is the function Bf on $\mathbb{C}^d$ defined by

$$Bf(z) = 2^{d/4}\int_{\mathbb{R}^d} f(t) e^{2\pi t\cdot z - \pi t^2 - \frac{\pi}{2}z^2}\, dt\,. \qquad (3.29)$$

The (Bargmann–)*Fock space* $\mathcal{F}^2(\mathbb{C}^d)$ is the Hilbert space of all entire functions F on $\mathbb{C}^d$ for which the norm

$$\|F\|_{\mathcal{F}}^2 = \int_{\mathbb{C}^d} |F(z)|^2 e^{-\pi|z|^2}\, dz$$

is finite. The inner product on $\mathcal{F}^2(\mathbb{C}^d)$ is

$$\langle F, G\rangle_{\mathcal{F}} = \int_{\mathbb{C}^d} F(z)\overline{G(z)}\, e^{-\pi|z|^2}\, dz\,.$$

By means of the little calculation in (3.28), the previous results of this chapter can be written as follows.

Proposition 3.4.1. *(a) If f is a function on $\mathbb{R}^d$ that has polynomial growth, then its Bargmann transform Bf is an entire function on $\mathbb{C}^d$. If we write $z = x + i\omega$, then*

$$V_\varphi f(x, -\omega) = e^{\pi i x \cdot \omega} Bf(z) e^{-\pi |z|^2/2} . \tag{3.30}$$

(b) If $f \in L^2(\mathbb{R}^d)$, then

$$\|f\|_2 = \left(\int_{\mathbb{C}^d} |Bf(z)|^2 e^{-\pi |z|^2} \, dz \right)^{1/2} = \|Bf\|_{\mathcal{F}} .$$

Thus B is an isometry from $L^2(\mathbb{R}^d)$ into $\mathcal{F}^2(\mathbb{C}^d)$.

Proof. If $|f(t)| = \mathcal{O}(|t|^N)$, then the integral (3.29) converges absolutely for every $z \in \mathbb{C}^d$ and uniformly over compact sets in $\mathbb{C}^d$. Therefore one can differentiate under the integral and Bf is an entire function.

Equation (3.30) is just (3.28) rewritten in new notation. Statement (b) follows from Corollary 3.2.2. ∎

Our next goal is to show that B is a unitary mapping from $L^2(\mathbb{R}^d)$ onto $\mathcal{F}^2(\mathbb{C}^d)$. In light of Proposition 3.4.1 we need only prove that the range of B is dense in $\mathcal{F}^2(\mathbb{C}^d)$. This requires a more detailed study of entire functions in several variables. This should not deter the reader; all we need are power series expansions, and thanks to multi-index notation there will not even be a visible difference between the theory in one or in several complex variables.

Theorem 3.4.2. *(a) The collection of all monomials of the form*

$$e_\alpha(z) = \left(\frac{\pi^{|\alpha|}}{\alpha!} \right)^{1/2} z^\alpha = \prod_{j=1}^{d} \left(\frac{\pi^{\alpha_j}}{\alpha_j!} \right)^{1/2} z_j^{\alpha_j}$$

for $\alpha = (\alpha_1, \dots, \alpha_d)$ with $\alpha_j \geq 0$, forms an orthonormal basis for $\mathcal{F}^2(\mathbb{C}^d)$.
(b) $\mathcal{F}^2(\mathbb{C}^d)$ is a reproducing kernel Hilbert space, that is,

$$|F(z)| \leq \|F\|_{\mathcal{F}} \, e^{\pi |z|^2/2} \quad \text{for all } z \in \mathbb{C}^d .$$

The reproducing kernel is $K_w(z) = e^{\pi \bar{w} \cdot z}$; this means that $F(w) = \langle F, K_w \rangle$.

Proof. (a) Write each variable in polar coordinates; that is, let $z_j = r_j e^{i\theta_j}$. Let us first compute the inner product of z^α with z^β restricted to the (poly)disc $P_R = \{z \in \mathbb{C}^d : |z_j| \leq R\}$:

$$\int_{P_R} z^\alpha \, \overline{z^\beta} \, e^{-\pi |z|^2} \, dz = \prod_{j=1}^{d} \int_{|z_j| \leq R} z_j^{\alpha_j} \bar{z}_j^{\beta_j} e^{-\pi |z_j|^2} \, dz_j$$

$$= \prod_{j=1}^{d} \int_0^R \int_0^{2\pi} r_j^{\alpha_j + \beta_j + 1} e^{i(\alpha_j - \beta_j)\theta_j} e^{-\pi r_j^2} \, dr_j \, d\theta_j .$$

If $\alpha \neq \beta$, then this integral equals zero for all $R > 0$, and therefore

$$\langle z^\alpha, z^\beta \rangle_{\mathcal{F}} = \lim_{R\to\infty} \int_{P_R} z^\alpha \overline{z^\beta} e^{-\pi|z|^2}\, dz = 0\,.$$

On the other hand, if $\alpha = \beta$ then

$$\int_{P_R} |z^\alpha|^2 e^{-\pi|z|^2}\, dz = \prod_{j=1}^{d} \left(2\pi \int_0^R r_j^{2\alpha_j+1} e^{-\pi r_j^2} dr_j \right) = \mu_{\alpha,R}\,.$$

For $R = \infty$, by making the change of variables $s = \pi r^2$, we can continue as follows:

$$\mu_{\alpha,\infty} = \prod_{j=1}^{d} \left(\int_0^\infty \left(\frac{s}{\pi}\right)^{\alpha_j} e^{-s} ds \right) = \prod_{j=1}^{d} \frac{\alpha_j!}{\pi^{\alpha_j}} = \frac{\alpha!}{\pi^{|\alpha|}}\,.$$

Consequently $\{\mu_{\alpha,R}^{-1/2} z^\alpha : \alpha \geq 0\}$ is an orthonormal system in $L^2(P_R, e^{-\pi|z|^2} dz)$. In particular, $\{e_\alpha, \alpha \geq 0\}$ is an orthonormal system in $\mathcal{F}^2(\mathbb{C}^d)$.

To prove completeness of $\{e_\alpha\}$ in $\mathcal{F}^2$, we start from the power series expansion of F in $\mathcal{F}^2$, which has the form

$$F(z) = \sum_{\alpha \geq 0} c_\alpha z^\alpha\,.$$

Suppose that $\langle F, e_\beta \rangle = 0$ for all $\beta \geq 0$. Then

$$\langle F, e_\beta \rangle_{\mathcal{F}} = \lim_{R\to\infty} \left(\frac{\pi^{|\beta|}}{\beta!}\right)^{1/2} \int_{P_R} \Big(\sum_\alpha c_\alpha z^\alpha\Big) \overline{z^\beta} e^{-\pi|z|^2}\, dz\,.$$

Since the power series in the integral converges uniformly on compact sets, we can interchange the integration and summation to obtain

$$\int_{P_R} F(z)\, \overline{z^\beta}\, e^{-\pi|z|^2}\, dz = \sum_\alpha c_\alpha \int_{P_R} z^\alpha \overline{z^\beta}\, e^{-\pi|z|^2}\, dz = c_\beta \mu_{\beta,R}\,.$$

Thus $\langle F, e_\beta \rangle_{\mathcal{F}} = (\pi^{|\beta|}/\beta!)^{1/2} \lim_{R\to\infty} \mu_{\beta,R}\, c_\beta = 0$. This implies $c_\beta = 0$ for all β and thus $F \equiv 0$. Since we have already shown that $\{e_\alpha\}$ is an orthonormal system, it follows that $\{e_\alpha\}$ is an orthonormal basis for $\mathcal{F}^2(\mathbb{C}^d)$.

(b) Since $F(z) = \sum_{\alpha \geq 0} \langle F, e_\alpha \rangle_{\mathcal{F}}\, e_\alpha(z)$, we obtain by the Cauchy–Schwarz inequality that

$$|F(z)| \leq \Big(\sum_{\alpha\geq 0} |\langle F, e_\alpha \rangle_{\mathcal{F}}|^2\Big)^{1/2} \Big(\sum_{\alpha \geq 0} \frac{\pi^\alpha}{\alpha!} |z^\alpha|^2\Big)^{1/2} = \|F\|_{\mathcal{F}} \cdot e^{\pi|z|^2/2}\,.$$

Thus point evaluations are continuous linear functionals on the Hilbert space $\mathcal{F}^2$. It follows that for each $w \in \mathbb{C}^d$ there is a function $K_w \in \mathcal{F}^2$ such that

$$F(w) = \langle F, K_w \rangle . \tag{3.31}$$

Expanding K_w with respect to the orthonormal basis e_α and applying (3.31), we obtain K_w explicitly as

$$\begin{aligned} K_w(z) &= \sum_\alpha \langle K_w, e_\alpha \rangle_{\mathcal{F}} e_\alpha(z) \\ &= \sum_\alpha \overline{e_\alpha(w)} e_\alpha(z) \\ &= \sum_\alpha \frac{\pi^{|\alpha|}}{\alpha!} \overline{w^\alpha} z^\alpha = e^{\pi \bar{w} \cdot z} , \end{aligned}$$

so we are done. ∎

Equipped with the reproducing kernel K_w for $\mathcal{F}^2(\mathbb{C}^d)$, we can now prove that the Bargmann transform maps onto $\mathcal{F}^2$.

Theorem 3.4.3. *The Bargmann transform is a unitary operator from $L^2(\mathbb{R}^d)$ onto $\mathcal{F}^2(\mathbb{C}^d)$.*

Proof. We have already seen in Proposition 3.4.1 that B is an isometry. Thus its range is a closed subspace of $\mathcal{F}^2(\mathbb{C}^d)$. Therefore, if we show that $B(L^2(\mathbb{R}^d))$ is dense in $\mathcal{F}^2(\mathbb{C}^d)$, then it follows that $B(L^2(\mathbb{R}^d)) = \mathcal{F}^2(\mathbb{C}^d)$, so B is surjective and the proof is complete.

We start by rewriting Lemma 1.5.2 in terms of the Bargmann transform. Using Lemma 1.5.2 with $a = 1$ and taking the normalization of φ into account, we compute

$$\begin{aligned} V_\varphi(T_u M_{-\eta}\varphi)(x, -\omega) &= \langle T_u M_{-\eta}\varphi, M_{-\omega} T_x \varphi \rangle \\ &= e^{2\pi i x \cdot \omega} \langle T_u M_{-\eta}\varphi, T_x M_{-\omega}\varphi \rangle \\ &= e^{\pi i (x-u)\cdot(-\eta-\omega)} e^{2\pi i x \cdot \omega} e^{-\pi[(x-u)^2 + (\eta-\omega)^2]/2} . \end{aligned}$$

On the other hand, writing $z = x + i\omega$ and $w = u + i\eta$, we obtain after some bookkeeping in the exponents and after applying Proposition 3.4.1 that

$$\begin{aligned} B(T_u M_{-\eta}\varphi)(z) &= e^{-\pi i x \cdot \omega} e^{\pi |z|^2/2} V_\varphi(T_u M_{-\eta}\varphi)(x, -\omega) \\ &= e^{\pi i u \cdot \eta} e^{-\pi(u^2+\eta^2)/2} e^{\pi(x\cdot u + \eta\cdot\omega + i(\omega\cdot u - x\cdot\eta))} \\ &= e^{\pi i u \cdot \eta} e^{-\pi |w|^2/2} e^{\pi \bar{w} \cdot z} . \end{aligned}$$

We can rewrite this in short as

$$B(T_u M_{-\eta}\varphi)(z) = e^{\pi i u\cdot\eta} e^{-\pi|w|^2/2} K_w(z)\,. \tag{3.32}$$

This shows that the reproducing kernel of $\mathcal{F}^2(\mathbb{C}^d)$ is in the range of B. Now suppose that for some $F \in \mathcal{F}^2(\mathbb{C}^d)$ we have $\langle F, Bf\rangle_{\mathcal{F}} = 0$ for all $f \in L^2(\mathbb{R}^d)$. In particular, using (3.32) we have for all $w \in \mathbb{C}^d$ that

$$\begin{aligned} 0 &= \langle F, B(T_u M_{-\eta}\varphi)\rangle \\ &= e^{-\pi i u\cdot\eta}\, e^{-\pi|w|^2/2}\,\langle F, K_w\rangle \\ &= e^{-\pi i u\cdot\eta}\, e^{-\pi|w|^2/2}\, F(w)\,. \end{aligned}$$

Therefore $F \equiv 0$, and consequently the range of B is dense in $\mathcal{F}^2(\mathbb{C}^d)$. ∎

Since the Bargmann transform is a unitary operator, the pre-image of the orthonormal basis $\{e_\alpha\}$ consisting of the functions $H_\alpha = B^{-1}e_\alpha \in L^2(\mathbb{R}^d)$ is an orthonormal basis for $L^2(\mathbb{R}^d)$. The functions H_α are called the *Hermite functions*. Even without an explicit formula for H_α we can derive their most important property for Fourier analysis.

Proposition 3.4.4. *The Hermite functions are eigenfunctions of the Fourier transform; specifically, for all* $\alpha \geq 0$ *we have*

$$\widehat{H_\alpha} = (-i)^{|\alpha|} H_\alpha\,.$$

Proof. Combine the fundamental identity (3.10) of time-frequency analysis with Proposition 3.4.1. Writing $z = x + i\omega \in \mathbb{C}^d$, we have

$$\begin{aligned} V_\varphi H_\alpha(x, -\omega) &= e^{\pi i x\cdot\omega} BH_\alpha(z)\, e^{-\pi|z|^2/2} \\ &= e^{\pi i x\cdot\omega}\, e^{-\pi|z|^2/2}\, e_\alpha\,. \end{aligned}$$

On the other hand, using (3.10) and the property $\varphi = \hat{\varphi}$ of the Gaussian, we obtain

$$\begin{aligned} V_\varphi \widehat{H_\alpha}(x, -\omega) &= V_{\hat{\varphi}} \widehat{H_\alpha}(x, -\omega) \\ &= e^{2\pi i x\cdot\omega}\, V_\varphi H_\alpha(\omega, x) \\ &= e^{2\pi i x\cdot\omega} e^{-\pi i x\cdot\omega} BH_\alpha(\omega - ix)\, e^{-\pi|z|^2/2} \\ &= e^{\pi i x\cdot\omega}\, e^{-\pi|z|^2/2} \Big(\frac{\pi^{|\alpha|}}{\alpha!}\Big)^{1/2} (\omega - ix)^\alpha \\ &= e^{\pi i x\cdot\omega}\, e^{-\pi|z|^2/2} \Big(\frac{\pi^{|\alpha|}}{\alpha!}\Big)^{1/2} (-iz)^\alpha \\ &= (-i)^{|\alpha|}\, e^{\pi i x\cdot\omega}\, e^{-\pi|z|^2/2} e_\alpha(z)\,. \end{aligned}$$

Since V_φ is one-to-one, it follows that $\widehat{H_\alpha} = (-i)^{|\alpha|} H_\alpha$. ■

As a consequence we obtain Plancherel's theorem.

Corollary 3.4.5. *If $f \in L^2(\mathbb{R}^d)$ is a finite linear combination of Hermite functions, then $\|f\|_2 = \|\hat{f}\|_2$.*

Proof. Since the collection of Hermite functions forms an orthonormal basis for $L^2(\mathbb{R}^d)$, the L^2-norm of a finite linear combination $f = \sum_\alpha c_\alpha H_\alpha$ is given by $\|f\|_2 = \|c\|_2$. Since $\hat{f} = \sum_\alpha c_\alpha (-i)^{|\alpha|} H_\alpha$, its norm is $\|\hat{f}\|_2 = \|((-i)^{|\alpha|} c_\alpha)\|_2 = \|c\|_2 = \|f\|_2$.

This isometry extends to all of $L^2(\mathbb{R}^d)$, because the finite linear combinations of Hermite functions are dense in $L^2(\mathbb{R}^d)$. ■

The attentive reader will observe that the preceding argument is not an independent proof of Plancherel's theorem. The argument is circular because we have used Plancherel's theorem to show that the Bargmann transform is an isometry. The insight from Corollary 3.4.5 is that the unitarity of the Fourier transform and that of the Bargmann transform are equivalent. Each statement can be derived from the other.

The results of this section show once again how special Gaussians are in time-frequency analysis. The STFT with respect to a Gaussian window is, except for a weighting factor, an entire function. Consequently, for the investigation of V_φ the entire arsenal of complex analysis is at our disposal. This explains why, for many questions about the STFT, the results for the case of Gaussian windows are much more precise than for other windows.