This Chapter follows closely the text "Foundations of Time-Frequency Analysis" by Karlheinz Gröchenig.

5. Time-Frequency Analysis

In this chapter we develop a quantitative theory of local Fourier analysis.

To motivate this consider the Inversion Formula for the F.T.: if $f, \hat{f} \in L^1(\mathbb{R})$, then

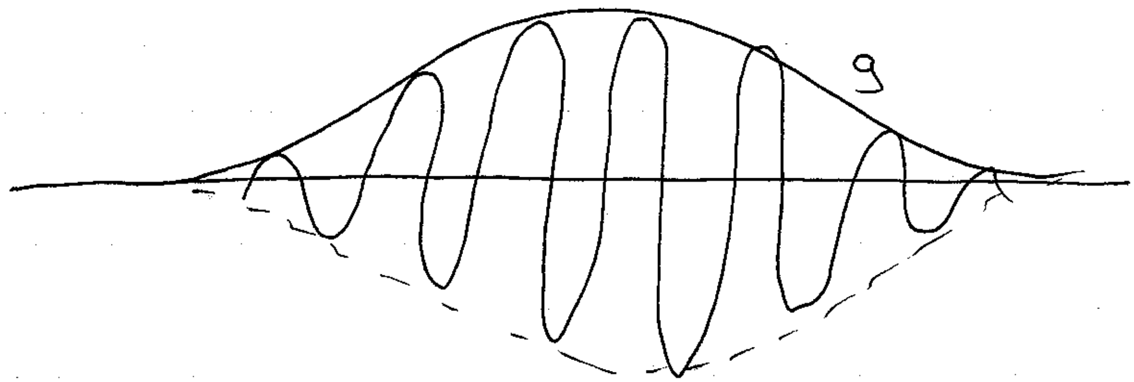
$$f(x) = \int \hat{f}(\xi)\, e^{2\pi i \xi x}\, dx.$$

We represent f as a superposition of the "pure tones" $e^{2\pi i \xi x}$. The scalar $\hat{f}(\xi)$ is the amount of the tone $e^{2\pi i \xi x}$ needed to represent f.

While this description is suggestive of the language of music, it is in fact very unlike the actual production of music by performers with musical instruments. Performers do not play a single tone $e^{2\pi i \xi x}$ at a single amplitude $\hat{f}(\xi)$ that they hold for all eternity. Rather, they play individual notes at individual frequencies at various times & for various lengths of times.

Let us create a (highly) idealized model of a "note." We desire a certain frequency to be heard for a certain amount of time. We can achieve this by "windowing" the tone $e^{2\pi i \xi x}$: Choose a window or envelope function g, & consider

$$M_\xi g(x) = e^{2\pi i \xi x} g(x)$$



Real part of $M_\xi g$

If g is concentrated around the origon, then we can consider this a "note of frequency ξ at time 0."

If the window g is compactly supported, then this note will last for a finite amount of time.

A note at time x, frequency ξ would be obtained by translating the note at time 0 to time x, i.e., it would have the form

$$M_\xi T_x g(t) = e^{2\pi i \xi t} g(t-x)$$

or

$$T_x M_\xi g(t) = e^{2\pi i \xi (t-x)} g(t-x).$$

Obviously notes are not unique: by changing the window we can change their duration and overal "sound." And, of course, real notes produced by real musical instruments are far more complicated, We are ignoring overtones, for example. The "attack" of a note is especially complicated, & is not modelled well here. Nonetheless, this picture has a broad range of applications, & we will pursue it in greater depth in this chapter.

<u>Definition</u>

The compositions $M_\xi T_x$ and $T_x M_\xi$ are called <u>time-frequency shift operators</u>.

If g is a function on $\mathbb{R}$, then $M_\xi T_x g$ & $T_x M_\xi g$ are <u>time-frequency shifts</u> of g.

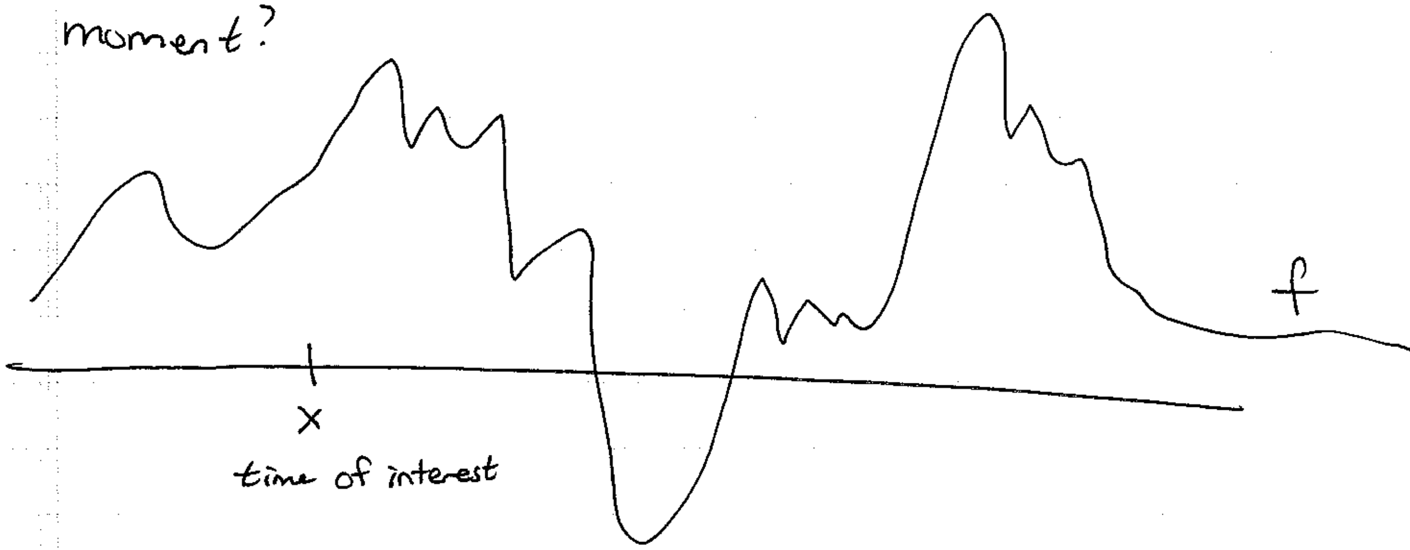
<u>Exercise</u>

$$T_x M_\xi g(t) = e^{-2\pi i \xi x} M_\xi T_x g(t)$$

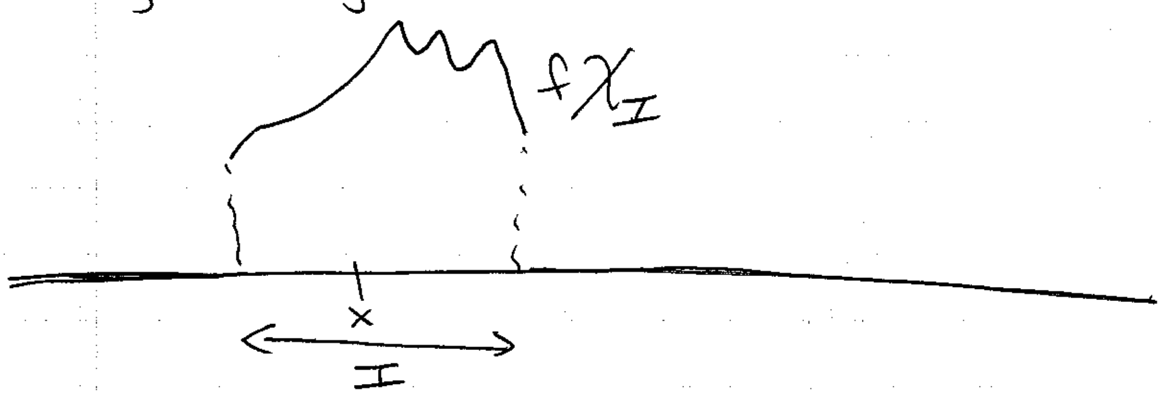
5.1. The STFT

Let us give another motivation for "local frequency analysis."

Given a function f, suppose that we would like to know the frequency content of f over a short time-span — what note is being played at a given moment?



We could simply "isolate" f in a region around a time x by windowing f with a characteristic function:



We might then hope that $(f\chi_I)^\wedge(\xi)$ gives the frequency content of f within an interval I around x. Unfortunately, $(f\chi_I)^\wedge = \hat{f} * \hat{\chi}_I$ is going to be very "blurry" because $\hat{\chi}_I$ is poorly localized in frequency. While f may be smooth, $f\chi_I$ will be discontinuous, & hence $(f\chi_I)^\wedge$ will have poor decay. We can improve this by using a smoother window function g, instead of χ_I, but there is no perfect solution, no way to define the "instantaneous frequency content" of f. Still, if g is centered around the origin, then

$$(fg)^\wedge(\xi) = \int f(x)\, e^{-2\pi i \xi x}\, g(x)\, dx$$

should give us some information on the frequency content of f near time zero. Translating g will then give us frequency content at other times,

i.e.,
$$(f \cdot T_x g)^\wedge(\xi) = \int f(x)\, e^{-2\pi i \xi x}\, g(x)\, dx, \qquad x, \xi \in \mathbb{R},$$
is in some sense the "amount of frequency ξ present in f at time x." Writing this as an inner product, we make the following definition.

Definition

Let f, g be measurable functions on $\mathbb{R}$. Then the Short-Time Fourier Transform (STFT) of f w.r.t. the window g is

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

$$= \int f(t)\, \overline{g(t-x)}\, e^{-2\pi i \xi t}\, dt, \quad (x,\xi) \in \mathbb{R}^2,$$

whenever this integral exists.

Notation

The STFT is also known as the windowed F.T., sliding window F.T., or the continuous Gabor transform.

The domain $\mathbb{R}^2$ of the STFT is called the time-frequency plane or (in physics) phase-space.

The STFT is a time-frequency (or phase space) representation of f. It is an approximation to the ideal of a perfect representation of the ~~[illegible]~~ "instantaneous frequency content" of f. Because of the Uncertainty Principle, no such ideal representation exists. Other closely ~~[illegible]~~ related time-frequency representations are the (cross-) ambiguity function, the (cross-) Wigner distribution, the spectrogram, & others.

<u>Exercise</u>

If $f, g \in L^2(\mathbb{R})$, then $V_g f \in C_b(\mathbb{R})$, with

$$\|V_g f\|_\infty \le \|f\|_2 \, \|g\|_2.$$

In fact, $V_g f$ is uniformly continuous on $\mathbb{R}^2$.

Hint: A Chapter 1 exercise shows that translation & modulation are each strongly continuous families of operators on $L^2(\mathbb{R})$.

Remark

Note that in contrast to the F.T., the STFT is a "local" transform. In particular, if f is perturbed in a small interval $[x_0-\varepsilon, x_0+\varepsilon]$, then the values

$$\hat{f}(\xi) = \int f(t)\, e^{-2\pi i t \xi}\, dt$$

are affected for <u>every</u> frequency ξ. [This is a real problem in medical imaging, for example — a small local motion of the patient during a CAT scan causes a global error in the image].

In contrast, if g is well-localized in time near the origin, then a perturbation of f in $[x_0-\varepsilon, x_0+\varepsilon]$ will have a significant effect on the values of

$$V_g f(x,\xi) = \int f(t)\, \overline{g(t-x)}\, e^{-2\pi i \xi t}\, dt$$

only for x near x_0 (depending on the size of the essential support of g).

Notation

We have used the notation $\langle f, g\rangle$ to denote the action of a linear functional g on an element f of its domain. So, for example, $\langle f, g\rangle$ is defined if $f \in C_0(\mathbb{R})$ & $g \in C_0(\mathbb{R})^* = M_b(\mathbb{R})$, or if $f \in \mathcal{S}(\mathbb{R})$ & $g \in \mathcal{S}'(\mathbb{R})$. In this chapter it will be convenient to allow a symmetric interpretation, where f is the functional acting on g. To achieve this, we simply declare that if $\langle g, f\rangle$ is defined in the sense above (f a linear functional acting on g), then we declare that

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

In time-frequency we usually think of restricting the class that the windows g belong to in order to analyze a larger class of "signals" f.

<u>Exercise</u>: Equivalent forms of the STFT.

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

$$V_g f(x,\xi) = (f \cdot T_x \bar{g})^\wedge(\xi)$$

$$= \langle f, M_\xi T_x g \rangle$$

$$= \langle \hat{f}, T_\xi M_{-x} \hat{g} \rangle$$

$$= e^{-2\pi i x \xi} \langle \hat{f}, M_{-x} T_\xi \hat{g} \rangle$$

$$= e^{-2\pi i x \xi} V_{\hat{g}} \hat{f}(-x, \xi)$$

$$= e^{-2\pi i x \xi} (f * M_{-\xi} \tilde{g})(x)$$

$$= (\hat{f} * M_{-x} \widetilde{\hat{g}})(\xi)$$

$$= e^{\pi i x \xi} \int f(t + \tfrac{x}{2})\, g(t - \tfrac{x}{2})\, e^{-2\pi i t \xi}\, dt.$$

Remark: The final representation in this list relates the STFT to the <u>cross-ambiguity function of f w.r.t. g</u>.

~~Exercise~~

Of the preceding representations of the STFT, one is particularly important.

Fundamental Identity of Time-Frequency Analysis

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

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

Thus, the STFT of $\hat{f}$ is essentially a rotation by $\pi/2$ of the STFT of f. The window g is replaced by the window $\hat{g}$, but in spirit this is not the important issue, e.g., by taking g to be the Gaussian we would have $g = \hat{g}$. Instead, we see that applying the F.T. to f interchanges the roles of time & frequency in the STFT, again demonstrating that the STFT is a <u>time-frequency representation</u> of f.

Exercise/Remark

a. With g fixed, $f \longmapsto V_g f$ is linear in f.

b. With f fixed, $g \longmapsto V_g f$ is antilinear in g.

c. $(f, g) \longmapsto V_g f$ is a sesquilinear form.

Exercise

$$M_{-\eta} T_{-u} M_{\xi} T_{x} = e^{2\pi i u \xi} M_{\xi-\eta} T_{x-u}.$$

The next result shows that the STFT of a time-frequency shift of f is a translation in $\mathbb{R}^2$ of the STFT of f, plus an additional phase factor. Since the phase factor has unit modulus, it ~~is~~ is irrelevant in equations involving only the absolute value of the STFT. But in other equations, it can be extremely important (and often annoying).

Exercise: Covariance Property of the STFT.
If $V_g f$ is defined, then

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

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

In particular,

$$|V_g(T_u M_\eta f)(x,\xi)| = |V_g(M_\eta T_u f)(x,\xi)|$$

$$= |V_g f(x-u, \xi-\eta)|.$$

5.2 The STFT for dual pairs

Exercise

The STFT is defined whenever f, g lie in dual spaces that are closed under translations & modulations. Show that if f, g lie in any of the following dual spaces then

$$V_g f(x,\xi) = \langle f, M_\xi T_x g \rangle, \qquad (x,\xi) \in \mathbb{R}^2,$$

is a continuous function on $\mathbb{R}$. (Note that translation & modulation may be defined in a distributional sense, i.e., by duality.

a. $f \in L^p(\mathbb{R})$, $g \in L^{p'}(\mathbb{R})$, $1 \leq p \leq \infty$, $\frac{1}{p} + \frac{1}{p'} = 1$

b. $f \in M_b(\mathbb{R})$, $g \in C_0(\mathbb{R})$

c. $f \in \mathcal{S}'(\mathbb{R})$, $g \in \mathcal{S}(\mathbb{R})$.

Show further that $V_g f$ is bounded in cases a or b (we will see that $V_g f$ has at most polynomial growth for case c).

Hints: The essential point is to show that the family of time-frequency shift operators $\{M_\xi T_x\}_{(x,\xi)\in\mathbb{R}^2}$ is strongly continuous w.r.t. the topology of

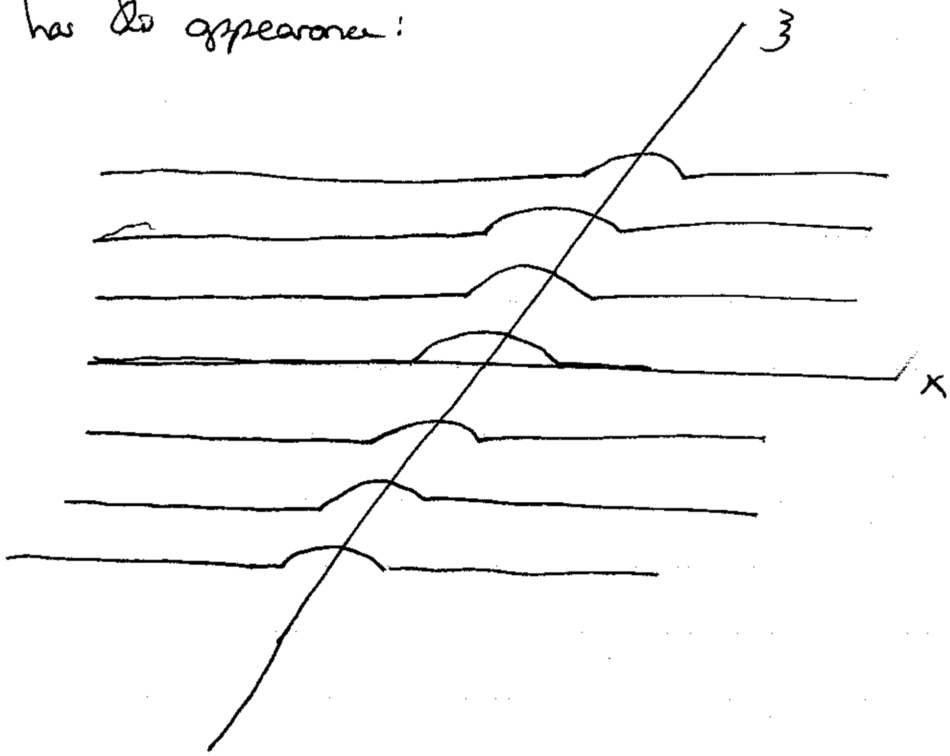
L^p, C_0, or $\mathcal{S}$ (hence weak* continuous on $L^{p'}$, M_b, or $\mathcal{S}'$). Previous results have shown that the families $\{T_x\}_{x \in \mathbb{R}}$ & $\{M_\xi\}_{\xi \in \mathbb{R}}$ are individually strongly continuous.

Example

Let $g \in C_0(\mathbb{R})$ be fixed. The STFT of the δ measure w.r.t. g is

$$
\begin{aligned}
V_g\delta(x,\omega) &= \langle \delta, M_\omega T_x g\rangle \\
&= \overline{\langle M_\omega T_x g, \delta\rangle} \\
&= M_\omega T_x g(0) \\
&= e^{2\pi i 0\omega} g(0-x) \\
&= g(-x).
\end{aligned}
$$

In practice, g is usually a smooth function reasonably well-concentrated around the origin. Hence $V_g\delta$ has the appearance:



An ideal time-frequency representation of δ would show nonzero values only for $x=0$. We cannot have the ideal representation, but we see that the frequency content of δ is locally zero away from time $x=0$. Moreover, for a given time x, every frequency ξ is present with the same amplitude. This matches our intuition of δ being perfectly localized in time at time $x=0$, with frequency content mirroring the fact that $\hat{\delta}=1$.

By making the window more concentrated in time around $x=0$, we can make the graph of $V_g\delta$ more & more concentrated around $x=0$. However, because $V_g\delta$ is constant as a function of ξ for a given x, we cannot see in this example the general phenomenon that the more concentrated in time that the window g becomes, the more "blurry" in frequency the graph of $V_g f$ will become.

<u>Exercise</u>

Write a computer program that will approximately compute a STFT. One method: use the form

$$V_g f(x,\omega) = (f \cdot T_x \bar{g})^\wedge(\xi)$$

and use the FFT to compute the F.T. of discrete functions. Experiment with windows of varying concentration around $x=0$.

Example

If $g \in L^1(\mathbb{R})$ then the STFT of the constant function 1 is defined, & we have

$$V_g 1(x,\xi) = \langle 1, M_\xi T_x g \rangle$$

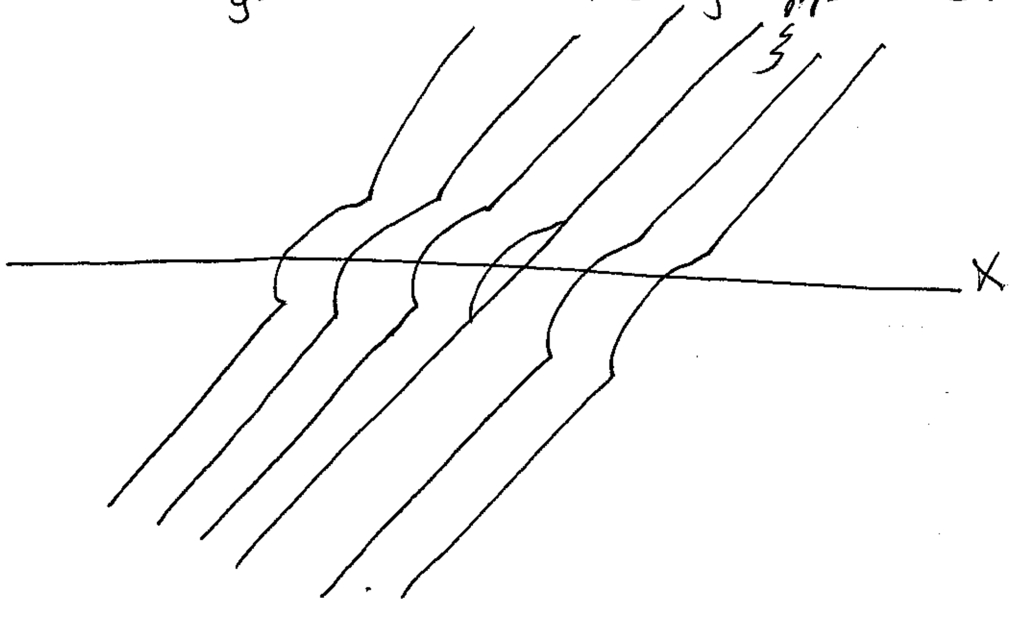
$$= \int 1 \, \overline{g(t-x)} \, e^{-2\pi i t\xi} \, dt$$

$$= \int \overline{g(t)} \, e^{-2\pi i (t+x)\xi} \, dt$$

$$= e^{-2\pi i x\xi} \int \overline{g(t)} \, e^{-2\pi i t\xi} \, dt$$

$$= e^{-2\pi i x\xi} \, \hat{g}(\xi).$$

Ignoring the unit-modulus "phase factor" $e^{-2\pi i x\xi}$ & considering a g s.t. $\hat{g}$ is well-concentrated around $\xi = 0$, we see that $V_g 1$ has the following appearance:



Thus $V_g 1$ has nontrivial frequency content at all times, but the frequency content is localized around $\xi = 0$. In an ideal time-frequency representation it would be perfectly localized around $\xi = 0$, but such an ideal cannot be realized as a general representation applicable to all f, g.

Exercise

~~Remark~~ Note that $1 = \hat{\delta}$, & show that

$$V_g \delta(x, \xi) = e^{-2\pi i x \xi} \; V_{\hat{g}} \hat{\delta}(-\xi, x)$$

This is a special case of the following equivalent forms of the STFT.

The STFT of Tempered Distributions

As remarked before, as long as the window g is taken to lie in the Schwartz class $\mathcal{S}(\mathbb{R})$, the STFT $V_g f$ is a continuous function for all tempered distributions $f \in \mathcal{S}'(\mathbb{R})$. We will show that $V_g f$ actually has polynomial growth at ∞, and consequently $V_g f \in \mathcal{S}'(\mathbb{R}^2)$.

Exercise

Show that if $g \in C^{\infty}(\mathbb{R})$, then for $N \geq 0$ we have

$$(M_\xi T_x g)^{(N)}(t) = \sum_{k=0}^{N} \binom{N}{k} (2\pi i \xi)^k M_\xi T_x g^{(N-k)}(t)$$

Theorem

If $g \in \mathcal{S}(\mathbb{R})$ then for each $f \in \mathcal{S}'(\mathbb{R})$ we have that

$V_g f \in C(\mathbb{R}^2)$ with polynomial growth at ∞. In particular,

$\exists\, C > 0,\ N \geq 0$ s.t.

$$|V_g f(x,\xi)| \leq C\,(1 + |x| + |\xi|)^N, \qquad (x,\xi) \in \mathbb{R}^2.$$

Proof

Fix $f \in \mathcal{S}'(\mathbb{R})$. Since f is a continuous linear functional on $\mathcal{S}(\mathbb{R})$, whose topology is defined by a countable family of seminorms, the "Continuity = Boundedness" theorem for f implies $\exists\, C_1 > 0$ & $N \geq 0$ s.t.

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \quad |\langle f, \varphi \rangle| \leq C_1 \sum_{m=0}^{N} \sum_{n=0}^{N} \| t^m \varphi^{(n)}(t) \|_\infty$$

Applying this to $\varphi = M_\xi T_x g$, we see that

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

$$\leq C_1 \sum_{m=0}^{N} \sum_{n=0}^{N} \| t^m (M_\xi T_x g)^{(n)}(t) \|_\infty$$

$$\leq C_1 \sum_{m=0}^{N} \sum_{n=0}^{N} \sum_{k=0}^{n} \binom{n}{k} |2\pi \xi|^k \| t^m M_\xi T_x g^{(n-k)}(t) \|_\infty$$

(by the exercise)

$$\leq C_2 \sum_{m=0}^{N} \sum_{n=0}^{N} \sum_{k=0}^{N} |\xi|^k \| t^m g^{(n-k)}(t-x) \|_\infty$$

$$= C_2 \sum_{m=0}^{N} \sum_{n=0}^{N} \sum_{k=0}^{N} |\xi|^k \| (t+x)^m g^{(n-k)}(t) \|_\infty$$

$$\leq C_2 \sum_{m=0}^{N} \sum_{n=0}^{N} \sum_{k=0}^{n} |\xi|^k \sum_{j=0}^{m} \binom{m}{j} |x|^{m-j} \| t^j g^{(n-k)}(t) \|_\infty$$

$$\leq C_3 \sum_{m=0}^{N} \sum_{n=0}^{N} |\xi|^m |x|^n = p(|x|, |\xi|).$$

where C_2, C_3 are finite scalars. Hence $V_g f$ is bounded by a polynomial p on $\mathbb{R}^2$ in the variables $|x|, |\xi|$.

Finally, by taking C large enough we will have

$$p(|x|, |\xi|) \leq C(1 + |x| + |\xi|)^N, \quad \text{all } (x,\xi) \in \mathbb{R}^2. \blacksquare$$

<u>Corollary</u>

If $f \in \mathcal{S}'(\mathbb{R})$, $g \in \mathcal{S}(\mathbb{R})$, then $V_g f \in \mathcal{S}'(\mathbb{R}^2)$.

5.3 Orthogonality Relations for the STFT

The Plancherel/Parseval Equality for the F.T. has an analog for the STFT.

Notation

We use the same notation $\langle \cdot , \cdot \rangle$ to denote the inner product on $L^2(\mathbb{R})$ & on $L^2(\mathbb{R}^2)$, the distinction being clear from context.

Theorem: Orthogonality Relations for the STFT.

If $f_1, f_2, g_1, g_2 \in L^2(\mathbb{R})$ then $V_{g_1}f_1, V_{g_2}f_2 \in L^2(\mathbb{R}^2)$, and

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

Proof:

Assume first that $g_1, g_2 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$.

Then for $i=1,2$ we have $f_i \cdot T_x \bar{g}_i \in L^2(\mathbb{R})$ & hence

$(f_i \cdot T_x \bar{g}_i)^\wedge \in L^2(\mathbb{R})$, for every $x \in \mathbb{R}$. Therefore,

we can compute that:

$$\langle V_{g_1}f_1, V_{g_2}f_2\rangle = \iint V_{g_1}f_1(x,\xi)\,\overline{V_{g_2}f_2(x,\xi)}\;d\xi\;dx$$

$$= \int\int (f_1\cdot T_x\bar{g}_1)^\wedge(\xi)\;\overline{(f_2\cdot T_x\bar{g}_2)^\wedge(\xi)}\;d\xi\;dx$$

$$= \int \langle (f_1\cdot T_x\bar{g}_1)^\wedge, (f_2\cdot T_x\bar{g}_2)^\wedge\rangle\;dx$$

$$= \int \langle f_1\cdot T_x\bar{g}_1,\; f_2\cdot T_x\bar{g}_2\rangle\;dx \qquad \text{(Plancherel)}$$

$$= \int\int f_1(t)\,\overline{g_1(t-x)}\;\overline{f_2(t)}\;g_2(t-x)\;dt\;dx.$$

$$= (*)$$

Exercise: Verify that the facts that $f_i, g_i \in L^2(\mathbb{R})$ imply that Fubini's Theorem is applicable. Hence we can continue $(*)$ as

$$(*) = \int\int f_1(t)\,\overline{g_1(t-x)}\;\overline{f_2(t)}\;g_2(t-x)\;dx\;dt$$

$$= \int f_1(t)\,\overline{f_2(t)}\left(\int \overline{g_1(t-x)}\,g_2(t-x)\;dx\right)dt$$

$$= \int f_1(t)\,\overline{f_2(t)}\;\overline{\langle g_1, g_2\rangle}\;dt$$

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

Exercise: Use an extension by density argument to show that the same equality holds for all $f_i, g_i \in L^2(\mathbb{R})$. $\blacksquare$

Corollary

If $g \in L^2(\mathbb{R})$ is fixed then $f \mapsto V_g f$ is a multiple of an isometry of $L^2(\mathbb{R})$ into $L^2(\mathbb{R}^2)$, specifically,

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

Proof:

Apply the theorem with $f_1 = f_2 = f$ & $g_1 = g_2 = g$. ∎

Remark

The range of $f \mapsto V_g f$ is not all of $L^2(\mathbb{R})$ since, for example, $V_g f$ is continuous for all $f, g \in L^2(\mathbb{R})$.

Remark

Since $f \mapsto V_g f$ is a multiple of an isometry (& is an isometry if $\|g\|_2 = 1$), in some sense $V_g f$ contains exactly the same "information" as f, as does the F.T. $\hat{f}$.

However, each of the original function f, the F.T. $\hat{f}$, & the STFT $V_g f$ "display" this information in different ways, each of which may be more "revealing", depending on the context at hand.

There is another equally instructive formulation of the STFT & the orthogonality relations that we present next.

<u>Definiton</u>

The <u>tensor product</u> of $f, g : \mathbb{R} \to \mathbb{C}$ is the function

$f \otimes g : \mathbb{R}^2 \to \mathbb{C}$ defined by

$$(f \otimes g)(x,y) = f(x)\overline{g(y)}.$$

(The complex conjugate is included on g to make certain formulas later more convenient.)

<u>Definition</u>

τ_a will denote the asymmetric change of coordinates function given by

$$\tau_a F(x,t) = F(t, t-x).$$

<u>Definition</u>

$\mathcal{F}_2$ will denote the <u>partial Fourier transform</u> given by

$$\mathcal{F}_2 F(x,\xi) = \int F(x,t)\, e^{-2\pi i \xi t}\, dt,$$

whenever this is defined.

Exercise

The partial F.T. has all the appropriate properties analogous to those of the F.T. on $\mathbb{R}$. In particular, show that $\mathcal{F}_2$ extends to a well-defined isometry of $L^2(\mathbb{R}^2)$ onto itself.

Hints: If $F \in L^2(\mathbb{R})$, then by Fubini's Theorem, we have

$$F_x(t) = F(x,t) \in L^2(\mathbb{R})$$

for a.e. x, & for all such x,

$$\begin{aligned}\int |F_x(t)|^2 \, dt &= \|F_x\|_2^2 \\ &= \|\hat{F}_x\|_2^2 \\ &= \int |\hat{F}_x(\xi)|^2 \, d\xi\end{aligned}$$

And, $\hat{F}_x(\xi) = \mathcal{F}_2 F(x,\xi)$.

In this notation we have the following representation of the STFT.

<u>Exercise</u>

If $f, g \in L^2(\mathbb{R})$, then $V_g f = \mathcal{F}_2 \tau_a (f \otimes \bar{g})$.

Using this we can give another proof of the orthogonality relations.

<u>Second proof of the orthogonality relations</u>

Both τ_a & $\mathcal{F}_2$ are unitary mappings of $L^2(\mathbb{R}^2)$ onto itself, so their composition is as well.

Therefore, given $f_i, g_i \in L^2(\mathbb{R})$, since $f_i \otimes g_i \in L^2(\mathbb{R}^2)$, we have

$$
\begin{aligned}
\langle V_{g_1} f_1, V_{g_2} f_2 \rangle &= \langle \mathcal{F}_2 \tau_a (f_1 \otimes \bar{g}_1), \mathcal{F}_2 \tau_a (f_2 \otimes \bar{g}_2) \rangle \\
&= \langle f_1 \otimes \bar{g}_1, f_2 \otimes \bar{g}_2 \rangle \\
&= \iint f_1(x) \overline{g_1(t)}\, \overline{f_2(x)}\, g_2(t)\, dx\, dt \\
&= \langle f_1, f_2 \rangle \overline{\langle g_1, g_2 \rangle}. \quad \blacksquare
\end{aligned}
$$

Remark

Another benefit of the tensor-product representation is that it indicates how to extend the definition of the STFT beyond dual pairs of windows & functions. For example, we saw before that if $f \in \mathcal{S}'(\mathbb{R})$ & $g \in \mathcal{S}(\mathbb{R})$, then $V_g f(x,\xi) = \langle f, M_\xi T_x g \rangle$ defines the STFT as a continuous function on $\mathbb{R}^2$. Suppose now that μ, ν are any two distributions in $\mathcal{S}'(\mathbb{R})$.

Then we can define their tensor product as follows. For $f, g \in \mathcal{S}(\mathbb{R})$, we declare that

$$\langle f \otimes g, \mu \otimes \nu \rangle = \langle f, \mu \rangle \langle g, \nu \rangle.$$

We then extend this to finite linear combinations:

$$\left\langle \sum_{k=1}^{N} c_k (f_k \otimes g_k), \mu \otimes \nu \right\rangle$$

$$= \sum_{k=1}^{N} c_k \langle f_k \otimes g_k, \mu \otimes \nu \rangle$$

$$= \sum_{k=1}^{N} c_k \langle f_k, \mu \rangle \langle g_k, \nu \rangle.$$

We then show that $\left\{ \sum_{k=1}^{N} c_k (f_k \otimes g_k) : N \in \mathbb{N}, c_k \in \mathbb{C}, f_k, g_k \in \mathcal{S}(\mathbb{R}) \right\}$

is dense in $\mathcal{S}(\mathbb{R}^2)$ & that we can uniquely define $\langle F, \mu \otimes \nu \rangle$ for an arbitrary $F \in \mathcal{S}(\mathbb{R})$ by taking limits, & that $\mu \otimes \nu$ so defined belongs to $\mathcal{S}'(\mathbb{R}^2)$. Once that is done, the STFT of μ w.r.t. ν is defined by

$$V_\nu \mu = \mathcal{F}_2 \mathcal{T}_a (\mu \otimes \bar{\nu}),$$

since both $\mathcal{T}_a$ & $\mathcal{F}_2$ extend by duality to $\mathcal{S}'(\mathbb{R}^2)$.

Thus, even for $f, g \in \mathcal{S}'(\mathbb{R})$, the STFT $V_g f$ is defined. The drawback is that we only have that $V_g f \in \mathcal{S}'(\mathbb{R})$, ie, it is a distribution rather than a function. However, whenever f, g do lie in dual spaces, $V_g f$ will be a function on $\mathbb{R}^2$.

The STFT with windows in $\mathcal{S}(\mathbb{R})$

Yet another benefit of the tensor-product form of the STFT is that it allows us to study the properties of the mapping $f \mapsto V_g f$ on domains other than $L^2(\mathbb{R})$. We illustrate this by considering this mapping on the domain $\mathcal{S}(\mathbb{R})$. If $f \in \mathcal{S}(\mathbb{R})$ then by definition f has rapid decay in the time variable. But also $\hat{f} \in \mathcal{S}(\mathbb{R})$, so $\hat{f}$ has rapid decay in the frequency variable. Hence we expect that the STFT $V_g f$ will have rapid decay in <u>both</u> time & frequency. We will show that this is indeed the case, if the window g is taken to lie in the Schwartz class. Moreover, not only does $V_g f$ have rapid decay, but it is infinitely differentiable – in fact, $V_g f$ lies in $\mathcal{S}(\mathbb{R}^2)$, the Schwartz class for functions on $\mathbb{R}^2$ (discussed in more detail below). That is, we will show that $f \mapsto V_g f$ maps $\mathcal{S}(\mathbb{R})$ into $\mathcal{S}(\mathbb{R}^2)$.

Later, we will see the converse result: if $g \in \mathcal{S}(\mathbb{R})$, then $V_g f \in \mathcal{S}(\mathbb{R}^2)$ implies $f \in \mathcal{S}(\mathbb{R}^2)$. But there is a bigger surprise: rapid decay of $V_g f$ alone, without any assumption of smoothness, is sufficient to imply that $f \in \mathcal{S}(\mathbb{R})$. Thus we will be able to characterize $\mathcal{S}(\mathbb{R})$ by a countable collection of seminorms related to the decay of $V_g f$, without involving any derivatives of $V_g f$.

Definition of $\mathcal{S}(\mathbb{R}^2)$

The definition of the Schwartz class on $\mathbb{R}^2$ is entirely analogous to the definition of $\mathcal{S}(\mathbb{R})$. A function $F \in C^\infty(\mathbb{R}^2)$ belongs to $\mathcal{S}(\mathbb{R}^2)$ if & only if for each $m_1, m_2, n_1, n_2 \geq 0$ we have

$$\rho_{m_1,m_2,n_1,n_2}(F) = \| x^{m_1} y^{m_2} F^{(n_1,n_2)}(x,y) \|_\infty < \infty,$$

where $F^{(n_1,n_2)} = \dfrac{\partial^{n_1}}{\partial x^{n_1}} \dfrac{\partial^{n_2}}{\partial y^{n_2}}$.

$\mathcal{S}(\mathbb{R}^2)$ has properties analogous to $\mathcal{S}(\mathbb{R})$. For example, the F.T. on $\mathbb{R}^2$ maps $\mathcal{S}(\mathbb{R}^2)$ onto itself, as does the partial F.T. $\mathcal{F}_2$.

We also will need the following facts.

Exercise
If $f, g \in \mathcal{S}(\mathbb{R})$, then $(f \otimes g)(x,y) = f(x)\overline{g(y)} \in \mathcal{S}(\mathbb{R}^2)$

Exercise
$\mathcal{S}(\mathbb{R}^2)$ is invariant under the asymmetric change of coordinates map τ_a.

Theorem

Let $g \in \mathcal{S}(\mathbb{R})$ be given.

If $f \in \mathcal{S}(\mathbb{R})$, then $V_g f \in \mathcal{S}(\mathbb{R}^2)$, and for each $N \geq 0$ $\exists\, C_N > 0$ s.t.

$$|V_g f(x,\xi)| \leq C_N \, (1+|x|+|\xi|)^{-N}, \qquad (x,\xi) \in \mathbb{R}^2;$$

i.e.,

$$\| (1+|x|+|\xi|) V_g f(x,\xi) \|_\infty < \infty, \quad N \geq 0.$$

Proof:

Given $f \in \mathcal{S}(\mathbb{R})$, write $V_g f = \mathcal{F}_2 \mathcal{T}_a (f \otimes g)$. Since $f \otimes g \in \mathcal{S}(\mathbb{R}^2)$ and $\mathcal{S}(\mathbb{R}^2)$ is invariant under $\mathcal{F}_2$ & $\mathcal{T}_a$, we conclude that $V_g f \in \mathcal{S}(\mathbb{R}^2)$. Given $N \geq 0$, we therefore have by the Binomial Theorem that

$$\begin{aligned}
&\| (1+|x|+|\xi|)^N V_g f(x,\xi) \|_\infty \\
&\leq \sum_{k=0}^{N} \binom{N}{k} 1^{N-k} \, \| (|x|+|\xi|)^k V_g f(x,\xi) \|_\infty \\
&\leq \sum_{k=0}^{N} \sum_{j=0}^{k} \binom{N}{k} \binom{k}{j} \, \| x^j \xi^{k-j} V_g f(x,\xi) \|_\infty \\
&= \sum_{k=0}^{N} \sum_{j=0}^{N} \binom{N}{k} \binom{k}{j} \rho_{j,k-j,0,0}(V_g f) \\
&= C_N < \infty. \qquad \blacksquare
\end{aligned}$$

Thus, if $f \in \mathcal{S}(\mathbb{R})$, then

$$\tilde{\rho}_N(f) = \| (1+|x|+|\xi|)^N \, V_g f(x,\xi) \|_\infty < \infty.$$

Later we will see that any <u>tempered distribution</u> $f \in \mathcal{S}'(\mathbb{R})$ s.t. $\tilde{\rho}_N(f) < \infty \ \forall N \geq 0$ must actually be a function on $\mathcal{S}(\mathbb{R})$, & furthermore, the collection of seminorms $\{\tilde{\rho}_N\}_{N \geq 0}$ is equivalent to the standard family of seminorms, i.e., they generate the same topology. This is perhaps surprising, since the seminorms $\tilde{\rho}_N$ do explicitly rely on smoothness criteria.

Exercise

Let $f, g \in L^2(\mathbb{R})$ be given. Then given $x \in \mathbb{R}$ we of course have $f \cdot T_x \bar{g} \in L^1(\mathbb{R})$ by Cauchy-Schwarz. Prove that we actually have

$$f \cdot T_x \bar{g} \in L^2(\mathbb{R}) \quad \text{for a.e. } x,$$

and consequently for any $1 \leq p \leq 2$ we have

$$f \cdot T_x \bar{g} \in L^p(\mathbb{R}) \quad \text{for a.e. } x.$$

This exercise will be very useful later.

Hint: $V_g f \in L^2(\mathbb{R})$. Write $V_g f$ in a F.T. formulation, & apply Plancherel & Fubini.

5.4 L^2-Inversion of the STFT.

As a consequence of the orthogonality relations, we have that f is completely determined by its STFT, i.e., $f \mapsto V_g f$ is injective. Rephrasing this gives the following result.

<u>Lemma</u>

Let $g \in L^2(\mathbb{R})$, $g \neq 0$ be given.

a. If $f \in L^2(\mathbb{R})$, then $V_g f = 0$ a.e. $\Leftrightarrow$ $f = 0$ a.e.

b. $\{M_\xi T_x g\}_{(x,\xi)\in\mathbb{R}^2}$ is complete in $L^2(\mathbb{R})$.

<u>Proof:</u>

a. This follows immediately from the fact that

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

b. If $f \in L^2(\mathbb{R})$ is s.t. $\langle f, M_\xi T_x g\rangle = 0$ $\forall (x,\xi) \in \mathbb{R}^2$, then by definition of the STFT we have $V_g f(x,\xi) = 0$ $\forall (x,\xi) \in \mathbb{R}^2$. Hence $f = 0$ by part a, so $\{M_\xi T_x g\}_{(x,\xi)\in\mathbb{R}^2}$ is complete in $L^2(\mathbb{R})$. $\blacksquare$

However, completeness only implies that each $f \in L^2(\mathbb{R})$ can be well-approximated by finite linear combinations of time-frequency shifts of g. This leaves open the question of *how* we can recover f from its STFT $V_g f$. We will address this question in the remaining portion of this section.

Our goal is the following theorem.

<u>Inversion Formula for the STFT</u>

Let $g, \gamma \in L^2(\mathbb{R})$ be s.t. $\langle g, \gamma \rangle \neq 0$. Then

$$\forall f \in L^2(\mathbb{R}), \quad f = \frac{1}{\langle \gamma, g \rangle} \iint V_g f(x,\xi)\, M_\xi T_x \gamma \; d\xi \, dx. \quad (*)$$

However, in order to make sense of this statement, we must first explain the meaning of the "function-valued integral" given by $(*)$.

<u>Motivation</u>

Suppose that $g \in L^2(\mathbb{R})$ & $F \in L^2(\mathbb{R}^2)$ are given.

Formally define

$$f = \iint F(x,\xi)\, M_\xi T_x g \; dx \, d\xi.$$

Then, pretending that this definition makes sense & that we can interchange the order of integration at will, we have for any $h \in L^2(\mathbb{R})$ that

$$\langle f,h\rangle = \left\langle \iint F(x,\xi)\, M_\xi T_x g\, dx\, d\xi,\ h\right\rangle$$

$$= \int \iint F(x,\xi)\, M_\xi T_x g(t)\, dx\, d\xi\, \overline{h(t)}\, dt$$

$$= \iint F(x,\xi) \int M_\xi T_x g(t)\, \overline{h(t)}\, dt\, dx\, d\xi$$

$$= \iint F(x,\xi)\, \langle M_\xi T_x g, h\rangle\, dx\, d\xi.$$

Note that while the preceding equalities are simply fantasy, the final quantity does exist as an ordinary Lebesgue integral. For, $\langle M_\xi T_x g, h\rangle = \overline{V_g h(x,\xi)} \in L^2(\mathbb{R})$ and $F \in L^2(\mathbb{R})$, so

$$\iint F(x,\xi)\, \langle M_\xi T_x g, h\rangle\, dx\, d\xi$$

$$= \iint F(x,\xi)\, \overline{V_g h(x,\xi)}\, dx\, d\xi$$

$$= \langle F, V_g h\rangle$$

is a well-defined scalar. We take this as our definition of the meaning of a function-valued integral.

Definition

If $g \in L^2(\mathbb{R})$ & $F \in L^2(\mathbb{R}^2)$ then the statement

$$f = \iint F(x,\xi)\, M_\xi T_x g \; dx\, d\xi$$

is interpreted weakly, i.e., it means that

$$\forall h \in L^2(\mathbb{R}), \quad \langle f, h\rangle = \iint F(x,\xi) \langle M_\xi T_x g, h\rangle \; dx\, d\xi$$

We can justify this definition more precisely, as follows.

Theorem

Let $g \in L^2(\mathbb{R})$ & $F \in L^2(\mathbb{R}^2)$ be given. Then

$$h \longmapsto \iint F(x,\xi) \langle M_\xi T_x g, h\rangle \; dx\, d\xi, \qquad h \in L^2(\mathbb{R})$$

defines a bounded antilinear functional on $L^2(\mathbb{R})$.

Consequently, $\exists$ unique $f \in L^2(\mathbb{R})$ s.t.

$$(*) \quad \forall h \in L^2(\mathbb{R}), \quad \langle f, h\rangle = \iint F(x,\xi) \langle M_\xi T_x g, h\rangle \; dx\, d\xi.$$

Notation

We write $f = \iint F(x,\xi)\, M_\xi T_x g \; dx\, d\xi$

to denote the unique function in $L^2(\mathbb{R})$ for which $(*)$ holds.

Proof:

Since $V_g h(x,\xi) =$ [redacted] $\langle h, M_\xi T_x g\rangle$ belongs to $L^2(\mathbb{R})$,

$$\left| \iint F(x,\xi)\, \langle M_\xi T_x g, h\rangle \,dx\, d\xi \right|$$

$$= \left|\langle F, V_g h\rangle\right|$$

$$\leq \|F\|_2 \, \|V_g h\|_2$$

$$\leq \|F\|_2 \, \|g\|_2 \, \|h\|_2 .$$

Hence $L(h) = \iint F(x,\xi)\, \langle M_\xi T_x g, h\rangle \,dx\, d\xi$ defines a bounded antilinear functional on $L^2(\mathbb{R})$, with operator norm $\|L\| \leq \|F\|_2 \, \|g\|_2$. Therefore $\exists$ unique $f \in L^2(\mathbb{R})$ s.t. $\langle f, h\rangle = L(h)$ $\forall h \in L^2(\mathbb{R})$. ∎

Now we can prove the Inversion Theorem for the STFT. With the function-valued integral interpreted in a weak sense, the result is simply the orthogonality relations in disguise.

<u>Inversion Formula for the STFT</u>

If $g, \gamma \in L^2(\mathbb{R})$ are s.t. $\langle g, \gamma \rangle \neq 0$, then

$$\forall f \in L^2(\mathbb{R}), \quad f = \frac{1}{\langle g, \gamma\rangle} \iint V_g f(x,\xi)\, M_\xi T_x \gamma \, dx\, d\xi,$$

where the integral is interpreted in a weak sense.

<u>Proof:</u>

Let $f \in L^2(\mathbb{R})$ be given. Then since $V_g f \in L^2(\mathbb{R}^2)$, the preceding discussion implies that there is a unique function $f_0 \in L^2(\mathbb{R})$ s.t. $f_0 = \frac{1}{\langle g,\gamma\rangle} \iint V_g f(x,\xi)\, M_\xi T_x \gamma \, dx\, d\xi$ in a weak sense. By the orthogonality relations, given any $h \in L^2(\mathbb{R})$ we therefore have that

$$\langle f_0, h\rangle = \frac{1}{\langle \gamma, g\rangle} \iint V_g f(x,\xi)\, \overline{\langle M_\xi T_x \gamma, h\rangle}\, dx\, d\xi$$

$$= \frac{1}{\langle \gamma, g\rangle} \iint V_g f(x,\xi)\, \overline{V_\gamma h(x,\xi)}\, dx\, d\xi$$

$$= \frac{1}{\langle \gamma, g\rangle} \langle V_g f, V_\gamma h\rangle$$

$$= \langle f, h\rangle.$$

Since this is true for every $h \in L^2(\mathbb{R})$, we conclude that $f = f_0$. $\blacksquare$

<u>Corollary</u>

If $g \in L^2(\mathbb{R})$, $g \neq 0$ then

$$\forall f \in L^2(\mathbb{R}), \quad f = \frac{1}{\|g\|_2^2} \iint V_g f(x,\xi)\, M_\xi T_x g \; dx\, d\xi$$

Remarks

The mapping $f \longmapsto V_g f$ is the <u>analysis</u> of f.

This process is analogous to listening to music (the function f) and representing it by a musical score (the STFT $V_g f$). The score shows what notes (frequencies) are present in f at what times.

Analogously, $V_g f(x,\xi)$ represents the "amount" of frequency ξ present in f at time x.

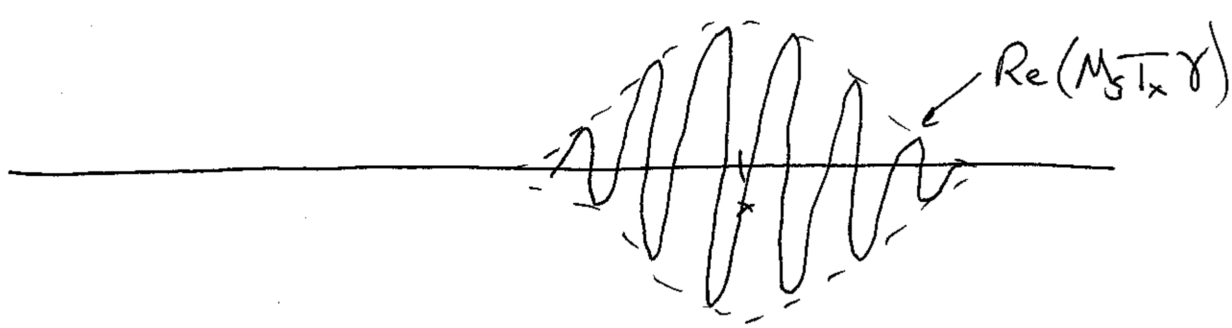
Once we have the musical score, we can reconstruct the music by playing (superimposing) the notes at the appropriate times with the appropriate amplitudes.

The mapping $V_g f \longmapsto f$ is the <u>synthesis</u> of f from its STFT, & is accomplished by the formula

$$f = \frac{1}{\langle \gamma, g\rangle} \iint V_g f(x,\xi)\, M_\xi T_x \gamma \, dx \, d\xi,$$

which instructs us to superimpose the "notes" $M_\xi T_x \gamma$

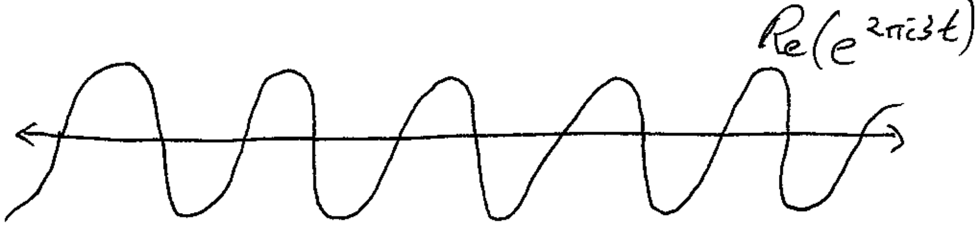
with amplitudes $\frac{1}{\langle \gamma, g \rangle} V_g f(x,\xi)$ to recover f (at least in a weak sense, and we will consider "strong" senses next). Amazingly, any two nonorthogonal window functions g, γ can be used, one for analysis & one for synthesis. If we take the windows g, γ to be well-concentrated around the origin in time & frequency, then the inversion formula is "local" in the sense that each note $M_\xi T_x \gamma$ being superimposed is well-localized at the appropriate time & frequency – it is in fact a note of frequency ξ occurring at time x:



In contrast, the F.T. inversion formula is highly nonlocal:

$$f(t) = \int \hat{f}(\xi)\, e^{2\pi i \xi t}\, dt.$$

The "notes" $e^{2\pi i \xi t}$ with amplitudes $\hat{f}(\xi)$ are superimposed to recover f, but the notes are supported on $\mathbb{R}$, i.e., they have infinite duration:



Consider these two inversion formulas for a compactly supported function f. If g is also compactly supported around the origin, then

$$V_g f(x,\xi) = \int f(t)\, \overline{g(t-x)}\, e^{-2\pi i \xi t}\, dt$$

will be nonzero only for x in & near the support of f, & thus in the inversion formula

$$f = \frac{1}{\langle f, \gamma \rangle} \iint V_g f(x,\xi)\, M_\xi T_x \gamma \, dx\, d\xi,$$

the integral in time will only involve notes $M_\xi T_x \gamma$ for times x near the support of f.

Even if g is not compactly supported but is at least well-concentrated around the origin, the STFT $V_g f$ will decay quickly away from the support of f.

The same will be true on the frequency side: if f is concentrated in frequency (i.e., $\hat{f}$ compactly supported or has good decay) & the window g is s.t. $\hat{g}$ is well-concentrated around the origin, then $V_g f$ will be well-concentrated in the variable ξ, as we can see from the Fundamental Identity $V_g f(x,\xi) = V_{\hat{g}} \hat{f}(\xi, -x)$.

In contrast, if f is compactly supported in time then $\hat{f}$ is not compactly supported. Hence in the inversion formula

$$f(t) = \int \hat{f}(\xi)\, e^{2\pi i \xi t}\, d\xi,$$

for t outside the support of f, all the possible