The Wavelet Transform, Time-Frequency Localization and Signal Analysis

INGRID DAUBECHIES, MEMBER, IEEE

Abstract —**Two different procedures are studied by which a frequency analysis of a time-dependent signal can be effected, locally in time. The first procedure is the short-time or windowed Fourier transform, the second is the "wavelet transform," in which high frequency components are studied with sharper time resolution than low frequency components. The similarities and the differences between these two methods are discussed. For both schemes a detailed study is made of the reconstruction method and its stability, as a function of the chosen time-frequency density. Finally the notion of "time-frequency localization" is made precise, within this framework, by two localization theorems.**

I. Introduction

A. The Windowed Fourier Transform and Coherent States

IN SIGNAL ANALYSIS one often encounters the so-called short-time Fourier transform, or windowed Fourier transform. This consists of multiplying the signal $f(t)$ with a usually compactly supported window function g, centered around 0, and of computing the Fourier coefficients of the product gf. These coefficients give an indication of the frequency content of the signal f in a neighborhood of $t=0$. This procedure is then repeated with translated versions of the window function (i.e., $g(t)$ is replaced by $g(t\pm t_0)$, $g(t\pm 2t_0),\cdots$, where t_0 is a suitably chosen time translation step). This results in a collection of Fourier coefficients

$$c_{mn}(f)=\int_{-\infty}^{\infty} dt\, e^{im\omega_0 t} g(t-nt_0) f(t) \qquad (m,n\in Z). \quad (1.1)$$

Similar coefficients also occur in a transform first proposed by Gabor [1] for data transmission. The original proposal used a Gaussian function g, and parameters ω_0, t_0 such that $\omega_0\cdot t_0=2\pi$. A Gaussian window function is, of course, not compactly supported, but it has many other qualities. One of these is that it is the function which is optimally concentrated in both time and frequency, and therefore well-suited for an analysis in which both time and frequency localization are important. Gabor's original proposal, with $\omega_0\cdot t_0=2\pi$, leads to unstable reconstruction (see [2], [3]; we shall come back to this in Section II-C-1). The Gabor functions have been used in many different settings in signal analysis, either in discrete lattices (with $\omega_0\cdot t_0<2\pi$ for stable reconstruction) or in the continuous form described next. In many of these applications their usefulness stems from their time-frequency localization properties (see e.g., [4]).

Whatever the choice for g (Gaussian, supported on an interval, etc.), it is interesting to know to which extent the coefficients $c_{mn}(f)$ of (1.1) define the function f. This is one of the main issues of this paper.

The coefficients $c_{mn}(f)$ in (1.1) can also be viewed as inner products of the signal f to be analyzed with a discrete lattice of coherent states. Let us clarify this statement. By "coherent states" we understand here the family of square integrable functions $g^{(p,q)}$, generated from a single $L^2(\mathbb{R})$-function g, by phase space translations (p,q). A "discrete lattice" of coherent states is a discrete subset of the whole family obtained by restricting the labels (p,q) to a regular rectangular lattice in phase space. "Phase space," a term we borrow here from physics, stands for the two-dimensional time-frequency space, considered as one geometric whole. More precisely, $g^{(p,q)}(x)$ is obtained from $g(x)$ by a translation of x by q, and by a similar translation of the Fourier transform $\hat{g}$ by p, i.e.,

$$g^{(p,q)}(x)=e^{ipx}g(x-q). \qquad (1.2)$$

Families of coherent states, as defined by (1.2), are used in many different areas of theoretical physics. Their name stems from their use in quantum optics (where the Gaussian choice for g is favored, $g(x)=\pi^{-1/4}\exp(-x^2/2)$) [5], [6], but they have since spilled over to many other fields. See e.g., [7] for a review, including many generalizations of the original concept. In quantum mechanics, they are particularly useful in semiclassical arguments because they make it possible to study quantum phenomena in a phase-space setting. If the original function g is centered around (0,0) in phase space, i.e., if the mean value of position, $\int dx\, x|g(x)|^2$, and of momentum, $\int dk\, k|\hat{g}(k)|^2$, are both zero, then the state $g^{(p,q)}$ will be centered around (p,q), i.e.,

$$\int dx\, x\left|g^{(p,q)}(x)\right|^2=q$$

Manuscript received November 9, 1987; revised December 7, 1988. This work was presented in part at the International Conference on Mathematical Physics, Marseille, France, July 1986.

The author was with the Mathematics Department, University of Michigan, Ann Arbor, MI 48109-1003. She is now with AT&T Bell Laboratories, 600 Mountain Ave., Murray Hill, NJ 07974.

IEEE Log Number 9036002.

0018-9448/90/0900-0961$01.00 ©1990 IEEE

$$\int dk\, k\left|\left(g^{(p,q)}\right)^{\wedge}(k)\right|^2 = p.$$

Here $\hat{g}$ denotes the Fourier transform of g, $\hat{g}(k) = (2\pi)^{-1/2}\int dx\, e^{ikx} g(x)$. The inner products $(g^{(p,q)}, \phi)$ will therefore "measure" the phase space content of the function ϕ around the phase space point (p,q). See e.g., [8, Section V-A.] and [9] for two different methods of using coherent states in a study of semiclassical approximations to quantum mechanics.

A very important property of the coherent states is the so-called "resolution of the identity." It has been discovered and rediscovered many times (see the earlier papers in [7]). A discussion of its relevance for signal analysis can be found in [10]. For the sake of completeness, we give its (easy) proof in Appendix A. The "resolution of the identity" says that the map Φ from $L^2(\mathbb{R})$ into $L^2(\mathbb{R}^2)$, defined by $\Phi f(p,q) = \langle g^{(p,q)}, f\rangle$, is an isometry (up to a constant factor), i.e.,

$$\int dp \int dq \left|\langle g^{(p,q)}, f\rangle\right|^2 = 2\pi \|g\|^2 \|f\|^2. \tag{1.3a}$$

Here, as in the remainder of this paper, $\langle f, g\rangle$ stands for the L^2-inner product of the functions f and g,

$$\langle f, g\rangle = \int dx\, \overline{f(x)} g(x)$$

(note that this inner product is antilinear in the first argument and linear in the second argument, following the physicists' convention), while $\|f\|$ stands for the L^2-norm of f,

$$\|f\|^2 = \langle f, f\rangle = \int dx\, |f(x)|^2.$$

Formula (1.3a) implies that

$$\frac{1}{2\pi}\int dp \int dq\, \langle g^{(p,q)}, f\rangle g^{(p,q)} = f. \tag{1.3b}$$

This means that a function f can be recovered completely, and easily, from the phase space "projections" $\langle g^{(p,q)}, f\rangle$. Note that, since (1.3) holds for any g, one can use this freedom to choose g optimally for the application at hand. This freedom of choice was exploited in e.g., [8, Section V-A]; it will also be important to us here.

If, instead of letting (p,q) roam over all of phase space, in a continuous fashion, we rather restrict ourselves to a discrete sublattice of phase space, then we revert to (1.1). That is, if we choose $p_0, q_0 > 0$, and we define

$$\begin{aligned} g_{mn}(x) &= g^{(mp_0, nq_0)}(x) \\ &= e^{imp_0x} g(x - nq_0) \end{aligned}$$

then clearly (with x interpreted as "time," $p_0 = \nu_0$, $q_0 = t_0$)

$$c_{mn}(f) = \langle g_{mn}, f\rangle = \langle g^{(mp_0, nq_0)}, f\rangle.$$

This shows that the short-time Fourier transform can indeed be viewed as the computation of inner products with a discrete lattice of coherent states.

B. Phase Space in Signal Analysis

The appearance of a phase space concept, such as the coherent states, in signal analysis, is not altogether surprising. As pointed out by N. G. de Bruijn [11], it is an entirely natural concept in music. Let us consider a time-dependent signal, which is a piece of music played (for the sake of simplicity), by a single instrument. If we disregard problems with high harmonics, with the "attack" of the notes, etc. (this is a simple example!), the musical score corresponding to the piece can be considered as a satisfactory representation of the time-dependent acoustic signal. A musical score indicates which notes have to be played at consecutive time steps. Thus, it gives a frequency analysis, locally in time, and is much closer to a short-time Fourier transform or a coherent states analysis, than to e.g., the Fourier transform, in which all track of time-dependence is lost, or at least not explicitly recognizable. The notation of a musical score, indicating time horizontally, and frequency vertically, is really a phase space notation. Phase space is thus seen to be a natural concept in signal analysis. This is also illustrated by the successful use, in signal analysis, of that other phase space concept, the Wigner distribution, as in [12] or [13].

C. Frames

As long as the continuously labelled coherent states $g^{(p,q)}$ are used, we know, from (1.3), that knowing a signal f is equivalent to knowing the inner products $(g^{(p,q)}, f)$. This need not be automatically true if one restricts the labels (p,q) to the discrete sublattice (mp_0, nq_0), $m,n \in \mathbb{Z}$; some conditions on g, p_0, q_0 will be required. Let us define the map T

$$\begin{aligned} &T: L^2(\mathbb{R}) \to l^2(\mathbb{Z}^2) \\ &(Tf)_{m,n} = \langle g_{mn}, f\rangle = c_{mn}(f). \end{aligned} \tag{1.4}$$

This map is the discrete analog of the "continuous" map Φ in Section I-A. For all the cases of interest to us, the map T will be bounded. To have complete characterization of a function f by its coefficients $c_{mn}(f)$ we shall require that T be one-to-one. If the characterization of f by means of the $c_{mn}(f)$ is to be of any use for practical purposes, one needs more than this, however. It is important that the reconstruction of f from the coefficients $c_{mn}(f)$ (which is possible, in principle, if T is one-to-one) be numerically stable: if the sequences $c_{mn}(f), c_{mn}(g)$ are "close" for two given functions f and g, then we want this to mean that f and g are "close" as well. More concretely, we require that

$$A\|f\|^2 \le \sum_{m,n} |\langle g_{mn}, f\rangle|^2 \le B\|f\|^2 \tag{1.5}$$

with $A > 0$, $B < \infty$, and A, B independent of f. This can be rewritten, in operator language, as

$$A\mathbb{1} \le T^*T \le B\mathbb{1}. \tag{1.6}$$

Here the inequality $T_1 \le T_2$, where T_1, T_2 are symmetric operators on $L^2(\mathbb{R})$, stands for $\langle f, T_1 f\rangle \le \langle f, T_2 f\rangle$ for all

$f \in L^2(\mathbb{R})$. A set of vectors $\{\phi_j; j \in J\}$ in a Hilbert space $\mathscr{H}$ for which the sums $\sum_{j \in J} |\langle \phi_j, f\rangle|^2$ yield upper and lower bounds for the norms $\|f\|^2$, as in (1.5), is also called a *frame*. The concept "frame" was introduced by Duffin and Schaeffer [14] in the context of nonharmonic Fourier analysis; see also [15]. We shall thus require that the $\{g_{mn}; m,n \in \mathbb{Z}\}$ constitute a frame; we shall give the name *frame bounds* to the constants A, B.

In a previous paper [16] particular functions g were constructed, for given $p_0, q_0 > 0$ (with $p_0 \cdot q_0 < 2\pi$) such that the corresponding g_{mn} constitute a frame. For this special construction, the frame bounds A, B are known explicitly in terms of g, and explicit inversion formulas for T can be given. In a particular case of this construction one even finds that the inequalities in (1.5), (1.6) becomes equalities, i.e., $A = B$. Whenever $A = B$ the frame is called a *tight frame*. The inversion $c_{mn}(f) \to f$ then becomes trivial; since $T^*T = A\mathbb{1}$, one has

$$f = A^{-1}T^*Tf = A^{-1}\sum_{m,n} g_{mn}\langle g_{mn}, f\rangle$$

where the sum converges strongly, i.e.,

$$\left\| f - A^{-1} \sum_{\substack{m,n \\ |m|,|n| \le K}} g_{mn}\langle g_{mn}, f\rangle \right\| \xrightarrow[K \to \infty]{} 0.$$

Remark: Note that frames, even tight frames, are not bases in general, in spite of the resemblance between the reconstruction formula for a tight frame and the standard expansion with respect to an orthonormal basis. In general, a frame contains "too many" vectors. An example in the finite-dimensional space $\mathbb{C}^2$ is given by $u_1 = e_1$, $u_2 = -1/2e_1 + \sqrt{3}/2e_2$, $u_3 = -1/2e_1 - \sqrt{3}/2e_2$, where $e_1 = (1,0)$ and $e_2 = (0,1)$ constitute the standard basis for $\mathbb{C}^2$. One easily checks that, for all $v \in \mathbb{C}^2$,

$$\sum_{l=1}^{3} |\langle u_l, v\rangle|^2 = \frac{3}{2}\|v\|^2$$

so that the $\{u_l; l = 1,2,3\}$ constitute a tight frame, with inversion formula

$$v = \frac{2}{3}\sum_{l=1}^{3} u_l\langle u_l, v\rangle.$$

The $\{u_l; l = 1,2,3\}$ do not constitute a basis because they are not linearly independent. In the infinite-dimensional frames we shall consider in this paper any finite number of vectors will be linearly independent in general, but there will still be "too many" vectors in the sense that any of them lies in the closed linear span of all the others. If the vectors constituting a tight frame are normalized, then the frame constant $A = B$ indicates the "rate of redundancy" of the frame; if $A = B = 1$ then the frame is automatically an orthonormal basis.

While the constructions in [16] lead to satisfactory and easy inversion formulas for T, they have the drawback that the function g cannot be chosen freely, but has to be of the very particular type constructed in [16]. For some applications however, the window function g might be imposed *a priori*, and not be of this particular type. In that case it is of interest to find ranges for the parameters p_0, q_0 such that the g_{mn}, associated with the triplet g, p_0, q_0, constitute a frame. As we shall see next, *snug frames*, which are close to tight frames, i.e., for which the ratio B/A is close to 1, are particularly interesting, because they lead to easy inversion formulas with rapid convergence properties. It is therefore important to have good estimates for the frame bounds A and B. One of the questions we shall address in this paper is therefore the following: given g,

1) find a range R such that for $(p_0, q_0) \in R$ the associated g_{mn} are a frame, and
2) for $(p_0, q_0) \in R$, compute estimates for the frame bounds A, B.

Once good estimates for the frame bounds A, B are obtained, one can construct a dual function $\tilde{g}$ (depending on p_0, q_0 as well as on g; see Section II-A next) which leads to the easy reconstruction formula

$$f(x) = \sum_{m,n} \langle g_{mn}, f\rangle e^{imp_0x}\tilde{g}(x - nq_0).$$

The function $\tilde{g}$ is essentially a multiple of g, with correction terms of the order of $B/A - 1$; the closer B/A is to 1, the faster the series for $\tilde{g}$ converges. Obtaining good frame bounds is important even if different reconstruction procedures are considered, or if only characterization of f by means of the $c_{mn}(f) = \langle g_{mn}, f\rangle$ and not full reconstruction (after e.g., transmission) is the main goal, since *any* stable reconstruction or characterization procedure can exist only if the g_{mn} constitute a frame.

All of these concern expansions with respect to coherent states constructed according to (1.2). We shall also discuss a different type of expansion, the so-called "wavelet expansion" [17], [18], which corresponds to coherent states of a different type.

D. *Wavelets—A Different Kind of Coherent States*

The coherent states $g^{(p,q)}$ all have the same envelope function g, which is translated by the amount q, and "filled in" with oscillations with frequency p (see Fig. 1a). They typically have all the same width in time and in frequency. "Wavelets" are similar to the $g^{(p,q)}$ in that they also constitute a family of functions, derived from one single function, and indexed by two labels, one for position and one for frequency. More explicitly,

$$h^{(a,b)}(x) = |a|^{-1/2}h\left(\frac{x-b}{a}\right)$$

where h is a square integrable function such that

$$C_h = \int dy\,|y|^{-1}|\hat{h}(y)|^2 < \infty \tag{1.7}$$

and where $a, b \in \mathbb{R}$, $a \neq 0$. Note that if h has some decay at infinity, then (1.7) is equivalent to the requirement $\int dx\, h(x) = 0$. The parameter b in the $h^{(a,b)}$ gives the position of the wavelet, while the dilation parameter a

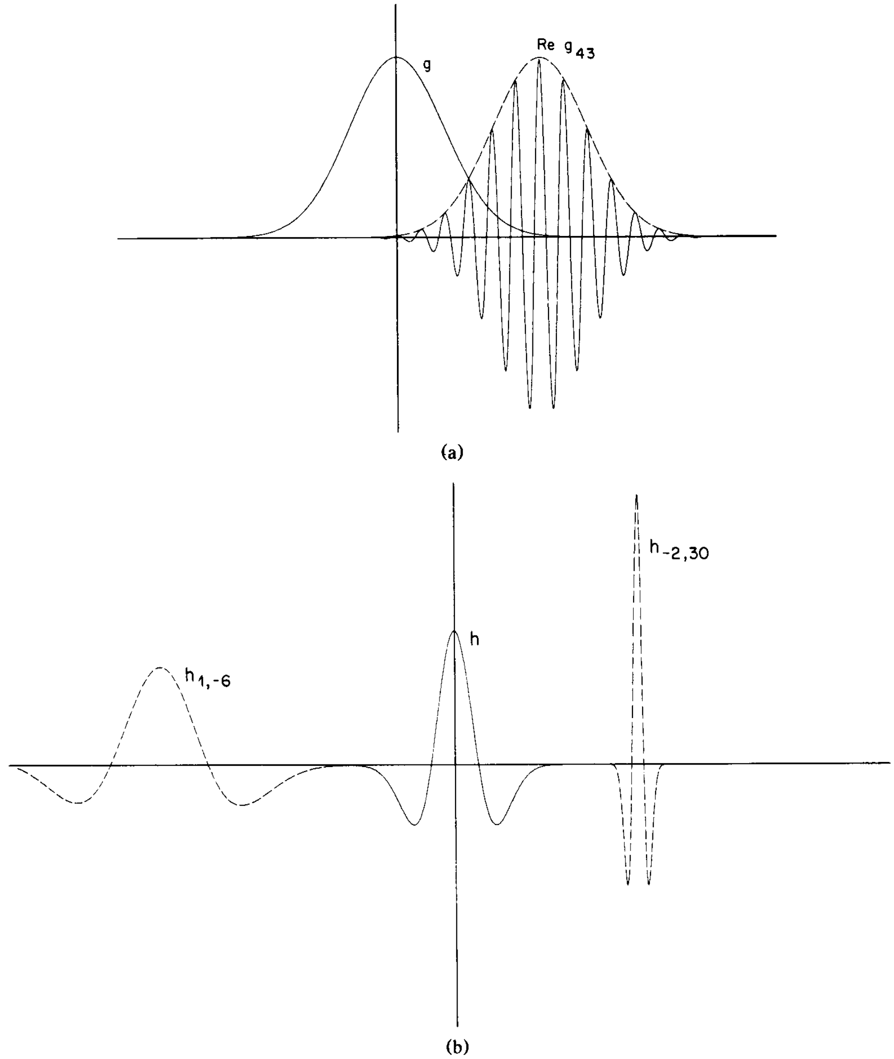

Fig. 1. (a) Typical choice for the window function $g = g_{00}$, and typical g_{mn}. In this case $g(x) = \pi^{-1/4}\exp(-x^2/2)$, $p_0 = \pi$, $q_0 = 1$; figure shows $\operatorname{Re} g_{43}(x) = \pi^{-1/4}\cos(4\pi x)\exp[-(x-3)^2/2]$. (b) Typical choice for basic wavelet $h = h_{00}$, and a few typical h_{mn}. In this case $h(x) = 2/\sqrt{3}\,\pi^{-1/4}(1-x^2)\exp(-x^2/2)$, $a_0 = 2$, $b_0 = 1$.

governs its frequency. For $|a| \ll 1$, the wavelet $h^{(a,b)}$ is a very highly concentrated "shrunken" version of h, with frequency content mostly in the high frequency range. Conversely, for $|a| \gg 1$, the wavelet $h^{(a,b)}$ is very much spread out and has mostly low frequencies (see Fig. 1(b)). As a result of this construction, wavelets will be a better tool than the "canonical" coherent states $g^{(p,q)}$ in situations where better time-resolution at high frequencies than at low frequencies is desirable.

There exists a resolution of the identity for wavelets as well as for the canonical coherent sates. One finds, for all $f_1, f_2 \in L^2(\mathbb{R})$,

$$\int \frac{da}{a^2} \int ab \langle f_1, h^{(a,b)}\rangle \langle h^{(a,b)}, f_2\rangle = C_h \langle f_1, f_2\rangle.$$

See Appendix A. Again, this implies that a function f can be recovered easily from the inner products $\langle h^{(a,b)}, f\rangle$, since

$$f = C_h^{-1} \int \frac{da}{a^2} \int db\, \langle h^{(a,b)}, f\rangle h^{(a,b)}. \tag{1.8}$$

The similarities between the $g^{(p,q)}$ associated with the short-time Fourier transform and the wavelets $h^{(a,b)}$ are no accident: both families are special cases of "coherent states associated with a Lie-group," in the first case the Weyl–Heisenberg group, in the second case the "$ax+b$"-group or affine group. Wavelets are therefore also called "affine coherent states." They were first introduced, under this name, in [19]. They are of great interest for their own sake, and have led to interesting applications in quantum mechanics (see, e.g., [20], [21]). They are also related to the use of dilation and translation techniques in harmonic analysis, which have turned out to be very powerful tools. It is likely that these techniques, and affine coherent states, will find interesting applications in quantum mechanics problems (see e.g., [22]). We shall not go into these aspects here. Note that the "resolutions of the identity" (1.3) and (1.8) can be used as starting points for the construction of time-frequency localization or filter operators [23], [24].

E. Discrete Lattices of Wavelets—The Wavelet Transform

In analogy to the lattice of coherent states associated with the short-time Fourier transform, which can be viewed as a discrete subset of the continuously labeled $g^{(p,q)}$, we shall also consider discrete lattices of wavelets. We choose to discretize the dilation parameter a by taking powers of a fixed dilation step $a_0>1$, $a=a_0^m$ with $m\in\mathbb{Z}$. For different values of m the wavelets will be more or less concentrated, and we adapt the discretized translation steps to the width of the wavelet by choosing $b=nb_0a_0^m$, with $n\in\mathbb{Z}$. This leads to

$$h_{mn}(x)=h^{(a_0^m,nb_0a_0^m)}(x)=a_0^{-m/2}h(a_0^{-m}x-nb_0). \quad (1.9)$$

As we shall see, it is also important for the discrete case to have $\int dx\, h(x)=0$.

The "discrete wavelet transform" is the analog of the map defined by (1.4) for the Weyl–Heisenberg case.[1]

Again one can investigate, as in the Weyl–Heisenberg case, whether the h_{mn} and the associated wavelet transform lead to a "discrete approximation" of the resolution of the identity (1.8), i.e., whether a family of wavelets h_{mn} constitutes a frame. As was shown in [16], it is possible, given a_0, b_0, to construct explicit functions h such that the associated h_{mn} constitute a frame, or even, in particular cases, a tight frame. These functions h are, as in the Weyl–Heisenberg case, of a very particular type. Typically their Fourier transform $\hat{h}$ has a compact support; since $\hat{h}$ may be chosen in C^∞, the function h may have arbitrarily fast decay, or

$$|h(x)|\le C_k(1+|x|)^{-k} \quad (1.10)$$

with $C_k<\infty$ for all $k\in\mathbb{N}$. In practice, however, the constants C_k turn out to be fairly large for functions h of the type constructed in [16]. This means that while h satisfies (1.10), its graph is very spread out (see [16]); typically the distance between the maximum x_0 of h and the point x defined by $\sup\{|h(y)|;\, y\ge x\}=10^{-2}|h(x_0)|$ will be an order of magnitude larger than b_0, the translation step parameter. For practical purposes, it is desirable to use functions h that are more concentrated than this, such as e.g., $h(x)=(1-x^2)e^{-x^2/2}$. We shall therefore address the same questions as the Weyl–Heisenberg case, i.e., for given h:

1) find a range R for the parameters such that for $(a_0,b_0)\in R$ the associated h_{mn} constitute a frame; and
2) for $(a_0,b_0)\in R$, compute estimates on the frame bounds A, B.

[1] To distinguish the g_{mn} from the wavelets h_{mn}, we shall call the g_{mn} "Weyl–Heisenberg" coherent states, after the group with which they are associated (see Section I-D).

The wavelet transform can be used, like the short-time Fourier transform, for signal analysis purposes. As in the short-time Fourier transform the two integer indices, m and n, control respectively, the frequency range and the time translation steps. There are however some significant differences between the two transforms. Some of these differences may well make the less widely used wavelet transform a better tool for the analysis of some types of signals (e.g., acoustic signals, such as music or speech) than the short-time Fourier transform. Let us digress a little on a qualitative discussion of these differences.

F. Qualitative Comparison of the Short-Time Fourier Transform and the Wavelet Transform

To illustrate this comparison we give graphs of typical g_{mn} and h_{mn} in Fig. 1, and of the associated phase space lattices in Fig. 2. More specifically, we represent each g_{mn} or h_{mn} by the point in phase space around which that function is mostly concentrated. In the Weyl–Heisenberg case, assuming that $\int dx\,|g(x)|^2=1$ and $\int dx\,x|g(x)|^2=0=\int dk\,k|\hat{g}(k)|^2$, the lattice points are given by

$$(nq_0, mp_0)=\left(\int dx\,x|g_{mn}(x)|^2, \int dk\,k|(g_{mn})^\wedge(k)|^2\right).$$

In the affine case we again associate to every h_{mn} the space localization point $\int dx\,x|h_{mn}(x)|^2=nb_0a_0^m$ (assuming that $\int dx\,|h(x)|^2=1$ and $\int dx\,x|h(x)|^2=0$). Since the function $|\hat{h}|$ and consequently all the $|(h_{mn})^\wedge|$ is even in many applications, the choice $\int dk\,k|(h_{mn})^\wedge(k)|^2$ is not appropriate for the frequency localization, since this integral is zero. This is due to the fact that the $(h_{mn})^\wedge$ have *two* peaks, one for positive and one for negative frequencies. We therefore represent the frequency content of h_{mn} by *two* points, namely

$$\int_0^\infty dk\,k|(h_{mn})^\wedge(k)|^2 \quad\text{and}\quad \int_{-\infty}^0 dk\,k|(h_{mn})^\wedge(k)|^2.$$

The two lattice points corresponding to the positive and negative frequency localizations of h_{mn} are thus

$$(nb_0a_0^m, a_0^{-m}\omega_\pm) = \left(\int dx\,x|h_{mn}(x)|^2, \int_{0\le\pm k<\infty} dk\,k|(h_{mn})^\wedge(k)|^2\right)$$

where $\omega_\pm=\int_{0\le\pm k<\infty}dk\,k|\hat{h}(k)|^2$. In Fig. 1(a) a few typical Weyl–Heisenberg coherent states were given, in Fig. 1(b) some typical wavelets. Fig. 1 shows one very basic difference between the two approaches; while the size of the support of the g_{mn} is fixed, the support of the h_{mn} is essentially proportional to a_0^m. As a result, high frequency h_{mn}, which correspond to $m\ll 0$, are very much concentrated. This means, of course, that the time-translation step (if x is interpreted as "time") has to be smaller for high-frequency h_{mn}, as shown by the phase space lattice in Fig. 2(b). It also means, however, that the wavelet transform will be able to "zoom in" on singularities, using more and more concentrated h_{mn} corresponding to higher and higher frequencies.

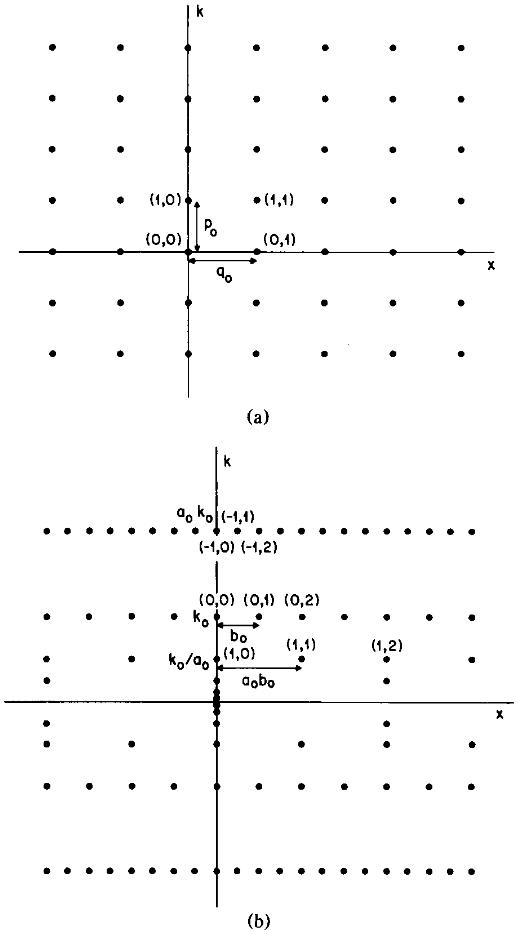

Fig. 2. (a) Phase space lattice corresponding to short-time Fourier transform (see text). (b) Phase space lattice corresponding to wavelet transform (see text). Constant k_0 is given by $k_0 = \int_0^\infty dk\, k^{-1}|\hat{h}(k)|^2$; we have assumed $\hat{h}$ to be even, and we have chosen $a_0 = 2$.

Let us illustrate this by the following simple example, taken from a grossly simplified problem in the synthesis of music. Typically one needs to be able to handle relatively low frequencies corresponding to the lowest notes and very high frequencies corresponding to high harmonics. Suppose one wants to be able to represent tones with frequency of the order of $2\pi/T$. Suppose also that one wants to be able to render faithfully the "attack" of notes. This "attack" consists of very high harmonics at the start of the note which die out very quickly, typically in a time $t_0 \ll T$. We have represented one component of such an "attack" very schematically in Fig. 3. Intuitively, the function $f(t)$ in Fig. 3 seems to correspond to a signal with "frequency" $2\pi 3/t_0$ during the time interval $[0,t_0]$, while its amplitude on $[t_0,T]$ is zero. Let us compare the performances of discrete Weyl–Heisenberg coherent states and of wavelets for this problem. In the first case, the support of g, and hence that of all the g_{mn}, needs to have a width of at least T. The high frequency $6\pi/t_0$ corresponds to a value for m of approximately $6\pi t_0^{-1}p_0^{-1}$. In practice however, since $T \gg t_0$, much higher values of m will be needed to reproduce, by means of the g_{mn} sketched in Fig. 1(a) (and which all have width T), a function f which is nonzero only in the interval $[0,t_0]$. This is not the case if wavelets are used. The high frequency wavelets have very small support, so that the above problem (having to bring in much higher frequencies than intuitively needed) does not occur. Moreover, even for the high frequencies corresponding to f, which correspond in the wavelet transform to very negative values of m and a very small time translation step (see Fig. 2(b)), only a few of the many time-steps necessary to cover $[0,T]$ would be needed, namely only those corresponding to $[0,t_0]$. This is what is meant by the "zooming in" property of the wavelets. For this kind of problem, wavelets thus provide a more efficient way (needing fewer coefficients) of representing the signal.

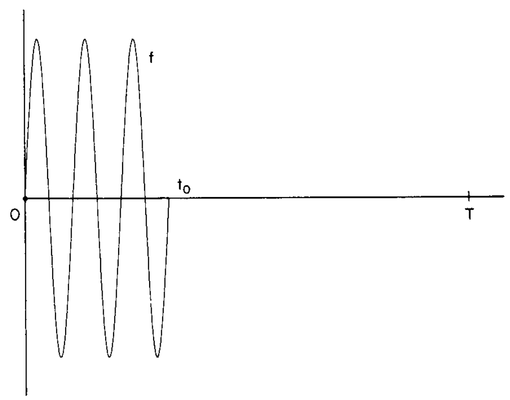

Fig. 3. One component of attack of note (see text). We take, as model,

$$f(t) = \begin{cases} \sin(6\pi t/t_0), & 0 \le t \le t_0 \\ 0 \quad t \le 0, & \text{or } t \ge t_0. \end{cases}$$

Lowest frequency of interest is $2\pi T^{-1}$; typically $t_0 \ll T$.

This example is so much simplified that it is rather unrealistic. The "zooming in" faculty of the wavelets, illustrated by this example, does however play an important role in more realistic applications; it makes the wavelets a useful tool in the areas of signal analysis such as seismic analysis [25] and music synthesis [26]. This same property also makes the wavelets a choice tool for the detection of singularities [27], [28], which is of great interest to the analysis of vision [29], [38], [40], and for the study of fractals [67].

As a final remark we note that the wavelet transform, unlike the short-time Fourier transform, treats frequency in a logarithmic way (as clearly shown by Fig. 2), which is similar to our acoustic perception. This corresponds to constant Q as opposed to constant bandwidth filters. This is another argument for the use of wavelets for the analysis and/or synthesis of acoustic signals.

G. Short Historical Review of the Wavelet Transform—Orthonormal Bases of Wavelets

As was already pointed out at the end of Section I-E, the use of functions $h^{(a,b)}$ generated from a single initial function h by means of dilations and translations is not new in either mathematics or physics, although their use is not so widespread as the Weyl–Heisenberg coherent states $g^{(p,q)}$ described in Section I-A. The use of the $h^{(a,b)}$ for signal analysis purposes, analogous to the use of the short-time Fourier transform, is much more recent, however. As a tool for signal analysis, the wavelet transform was first proposed by the geophysicist J. Morlet in view of applications for the analysis of seismic data [17], [25]. J. Morlet's original name for the wavelets was "wavelets of constant shape", to contrast them with the analyzing functions in the short-time Fourier transform, which do not have a constant shape (see Fig. 1). The numerical success of J. Morlet's method prompted A. Grossmann to make a more detailed study of the wavelet transform; this resulted in the recognition that the wavelets $h^{(a,b)}$ corresponded to a square integrable representation of the $ax+b$-group. The resolution of the identity (1.8) and associated interpolation techniques were then proposed (and implemented) for reconstruction schemes. This work was presented in a series of papers [17], [18], [30], in which the original longer name was shortened to "wavelet." These papers were concerned with the map associating to a function $f \in L^2(\mathbb{R})$ the function $\Phi f(a,b) = \langle h^{(a,b)}, f \rangle$, where a, b ran continuously over $\mathbb{R}^* \times \mathbb{R}$. (Here $\mathbb{R}^* = \mathbb{R} \backslash \{0\}$.) For closer comparison with the numerical situation, it is necessary to limit oneself to the discrete sublattices described in Section I-F. Again it was A. Grossmann who first realized the importance of the "frame" concept in this connection. It was around this time that I first became involved in the subject. Around the same time Y. Meyer, having learned about A. Grossmann's and J. Morlet's results, pointed out to them that there was a connection between their signal analysis methods and existing, powerful techniques in the mathematical study of singular integral operators. All this resulted in our first construction of a special type of frames [16]. It also was the start of a cross-fertilization between the signal analysis applications and the purely mathematical aspects of techniques based on dilations and translations. The wavelet frames constructed in [16] are based on functions h with compactly supported and C^∞ Fourier transform. A similar construction can be made for Weyl–Heisenberg coherent states [16]. In that case, due to the Balian–Low theorem (see Section II-C-1), there is a tradeoff between redundancy of the frame and smoothness of the frame functions: if one requires the g_{mn} to constitute a basis (i.e., no redundancy), then either $xg(x)$ or $k\hat{g}(k)$ is not square integrable. For the Weyl–Heisenberg case one is thus forced to consider frames rather than bases if the phase space localization properties of g are important. It seemed natural to assume that the same would be true for the affine case, and [16] was written under that implicit assumption (although it is never explicitly stated). Shortly after [16] was written, however, Y. Meyer constructed an orthonormal basis, for $L^2(\mathbb{R})$, of wavelets h_{mn}, based on a function h with compactly supported and C^∞ Fourier transform $\hat{h}$ [31]. This quite amazing basis turned out to be an unconditional basis for all L^p-spaces, $1 < p < \infty$, all the Sobolev spaces, etc, [31]. A basis $\{e_l\}_{l \in J}$ in a Banach space E is called "unconditional" if the convergence of $\sum_l \lambda_l e_l$ depends only on the $|\lambda_l|$, and is therefore independent of, e.g., the order in which the terms are summed. For a normalized basis $\{e_l\}_{l \in J}$ ($\|e_l\| = 1$ for all l) in a Hilbert space $\mathscr{H}$, this is equivalent to requiring that there exist $A > 0$, $B < \infty$ so that, for all $\nu \in \mathscr{H}$,

$$A\|\nu\|^2 \leq \sum_{l \in J} |\langle e_l, \nu \rangle|^2 \leq B\|\nu\|^2.$$

In a Hilbert space a normalized basis is thus an unconditional basis if and only if it is also a frame. Such bases are also called Riesz bases. Every orthonormal basis is automatically a Riesz basis; if $(\psi_n)_{n \in \mathbb{N}}$ is an orthonormal basis, then $e_n = (1+n^2)^{-1/2}(n\psi_1 + \psi_n)$ is an example of a basis of normalized vectors that is *not* a Riesz basis. The Meyer basis was generalized to more than one dimension by P. G. Lemarié and Y. Meyer [32]. Even though the function h in Y. Meyer's basis has fast decay in the mathematical sense (it decays faster than any inverse polynomial), the constants involved are very large, so that it is not very well localized in practice. This limits its usefulness for signal analysis. If one is willing to accept functions h with less regularity (C^k instead of C^∞) and to forego the compact support of $\hat{h}$, then bases with better localization can be used. G. Battle [33] and P. G. Lemarié [34] independently, using very different approaches, constructed orthonormal bases with exponential decay. Ph. Tchamitchian [35] constructed bases of compactly supported wavelets, which are however nonorthogonal.

A major breakthrough in the understanding of orthonormal wavelet bases came with the concept of multiresolution analysis, developed by S. Mallat and Y. Meyer [36], [37]. A multiresolution analysis consists in breaking up $L^2(\mathbb{R})$ into a ladder of spaces V_j,

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots$$

where $V_{-m} \to L^2(\mathbb{R})$ for $m \to \infty$. The orthonormal projections onto the V_m correspond to approximations with "resolution" 2^m; a typical (but too simple to be practical) example is

$$V_j = \{f \in L^2(\mathbb{R});$$
$$f \text{ is constant on every interval } [l2^j, (l+1)2^j),\ l \in \mathbb{Z}\}.$$

The V_j are all obtained from V_0 by the appropriate dilation,

$$f(x) \in V_j \Leftrightarrow f(2^j x) \in V_0$$

and V_0 is generated by the integer translates $\phi_{0n}(x) = \phi(x-n)$ of one single function ϕ. (In this example, $\phi(x) = 1$ for $0 \leq x < 1$, $\phi(x) = 0$ otherwise). Then there exists a function ψ such that the $\psi_{jn}(x)=2^{-j/2}\psi(2^{-j}x-n)$, j fixed, are an orthonormal basis of the orthogonal complement W_j of V_j in V_{j-1}. In other words, the inner products $\{\langle\psi_{jn},f\rangle;\ n\in\mathbb{Z}\}$ contain exactly all the information that is present in the approximation to f at level $j-1$ (with resolution 2^{j-1}) but which is lacking in the next coarser level j (resolution 2^j). Because of the "ladder property" of the V_j it follows that the whole collection $\{\psi_{jn};\ j,n\in\mathbb{Z}\}$ is an orthonormal basis of wavelets for $L^2(\mathbb{R})$. Readers who would like to learn more about multiresolution analysis should consult [36], [37], or [38], where explicit recipes are given for the construction of wavelet bases, together with examples. Wavelet decompositions using multiresolution analysis have been implemented in vision analysis [38], [40]. A short review is also included in [39]. Note that there is a connection between orthonormal wavelet bases and quadrature mirror filters, used in subband coding [39], [40].

Incidentally, the discovery of multiresolution analysis is another instance in which signal analysis applications provided the first intuition leading to the mathematical construction. Existing, more rudimentary techniques in vision analysis inspired S. Mallat to view orthonormal wavelet bases as a more refined tool for multiresolution approximations, which led to [36], [37]. S. Mallat has implemented these ideas, using the Battle–Lemarié bases, into an algorithm for decomposition and reconstruction of images [29], [38]. Finally, in February 1987, using Mallat's algorithm as inspiration, I succeeded in constructing orthonormal bases of wavelets with compact support, which are discussed elsewhere [39].

These different families of orthonormal wavelet bases have created quite a stir among mathematicians. Apart from applications to signal analysis, they should be useful in physics also. A first application, to quantum field theory, can be found in [41]. All these orthonormal bases are, however, rather spread out numerically, if one wants h to be reasonably regular ($h\in C^k$, with k large enough). If one is willing to give up the requirement of a basis and to settle for a frame (giving up the linear independence—see the remark in Section I-C), then much better localized C^∞ h can be chosen. This makes frames more interesting than orthonormal bases for certain wavelet applications in signal analysis. Another reason is that, as will be shown later, for a given desired reconstruction precision, frames allow one to calculate the wavelet coefficients with less precision than would be needed in the orthonormal case; the number of coefficients calculated is, of course, higher. This may be useful in some numerical applications. The present paper addresses "frame" questions. We shall formulate criteria under which the h_{mn} constitute a frame, and then derive some properties of these frames. Since similar techniques can be used for Weyl–Heisenberg coherent states, we address the same questions for that case as well. The basic results of this paper were reported, in an abridged version, at two mathematics conferences, in March 1986 [42] and July 1986 [43], respectively.

H. Organization of the Paper

In Section II we study "frame questions." We start with some generalities concerning frames in Hilbert spaces. We review the construction of [16], leading to tight frames. For more general choices of the initial function g or h, we answer the questions 1) and 2) asked previously (sufficient condition for frame, estimates for the frame constants A and B), both for the Weyl–Heisenberg and for the wavelet case. We discuss whether certain ranges for p_0,q_0 or a_0,b_0 are *a priori* excluded, and we give inequalities linking A, B, g, or h and p_0,q_0 or a_0,b_0. We give many examples, and numerical tables of frame constants for these examples. For particular choices of g or h, our results can be translated into estimates on entire functions [44]. A special case of this kind of estimate was already given in [45]. Since numerically one is also interested in convergence in other topologies than only L^2, we address the question of convergence in other spaces as well. To improve the readability of the paper, we have relegated many of the technical proofs to appendices.

In Section III we show how to use these frames for phase space localization. Concretely this means the following. Suppose that a signal f is mostly concentrated in $[-T,T]$ in time, and that its Fourier transform $\hat{f}$ is concentrated mostly between the frequencies Ω_1 and Ω_2. This means that in phase space, f is mostly concentrated on $[-T,T]\times([-\Omega_2,-\Omega_1]\cup[\Omega_1,\Omega_2])$. One would then expect that only those phase space lattice points, in Fig. 2, lying within this box (plus those lying very close to it) would suffice to approximately reconstruct f. It turns out that this intuition is essentially right. Note that such a "box" contains only a finite number of points. We also show how the "over-sampling" inherent to working with frames permits, for fixed desired precision on the reconstruction of f, to compute the coefficients $c_{mn}(f)$ with less precision than would be needed in the orthonormal case.

I. Some Remarks

A first remark we want to make is that while we shall stick, throughout the paper, to one dimension (i.e., a two-dimensional phase space), it is possible to generalize the results to more than one dimension. For the Weyl–Heisenberg case this is trivially true. For the wavelet case two possibilities exist. In the first case dilations and translations are used, independently in the n dimensions, and then again the extension is trivial. In the second case one uses n-dimensional translations, but only one dilation parameter, which acts simultaneously on all dimensions. In this case several (a finite number) h_j have to be introduced. This construction is then similar to the generalization in [32] of Y. Meyer's basis to n dimensions: in that case 2^n-1 different functions h_j are used.

A second remark is that we have restricted ourselves here to regularly spaced lattices, in the Weyl–Heisenberg (WH) case as well as the wavelet case. While it is possible to relax this regularity somewhat (we can also handle the

WH case if one of the two variables, p or q, is not regularly spaced, as long as the other is; likewise, in the wavelet case, we can handle dilations that are more irregular than the geometric sequence a_0^m), the methods presented here are essentially unable to cope with distributions of phase space points that would have the same density (and this probably suffices to define frames, as in nonharmonic Fourier analysis [15]), but which would not be given by lattices. Using different methods, one can indeed show that less regular phase space point distributions with a sufficiently high density do indeed lead to frames. For very special choices of g or h for which the "frame questions" can be formulated as properties of entire function spaces, that was proved in [45], [47]. Using the full power of the underlying group structure, Feichtinger and Gröchenig [48] proved the same result for much more general functions g or h (essentially, they only require that g or h is reasonably "nice"), without recourse to entire function spaces. The results in [46]–[48] are, however, more qualitative than quantitative in that they establish that certain discrete families of coherent states constitute frames without caring too much about the values of the frame constants. Moreover, the phase space densities of the discrete families considered in [46]–[48] seems to be considerably higher than in our examples.

II. Frames and Frame Bounds

A. Generalities Concerning Frames

We start by reviewing some general properties of frames [14], [15].

Suppose that the (ϕ_l; $l \in J$) constitute a frame in the Hilbert space $\mathscr{H}$, i.e., there exist $A > 0$, $B < \infty$ such that, for all $f \in \mathscr{H}$,

$$A\|f\|^2 \leq \sum_{l \in J} |\langle \phi_l, f\rangle|^2 \leq B\|f\|^2. \tag{2.1.1}$$

Define T: $\mathscr{H} \to l^2(J)$ by $(Tf)_l = \langle \phi_l, f\rangle$. Here $l^2(J)$ stands for the space of square summable complex sequences indexed by J. The operator T is clearly bounded, $\|Tf\| \leq B^{1/2}\|f\|$. We shall call T the "frame operator" associated with the frame $(\phi_l)_{l \in J}$. Its adjoint operator T^* maps $l^2(J)$ onto $\mathscr{H}$; it is defined by $T^*c = \sum_{l \in J} c_l \phi_l$, where $c = (c_l)_{l \in J} \in l^2(J)$. The frame inequalities (2.1.1) can be rewritten as

$$A\mathbb{1} \leq T^*T \leq B\mathbb{1}.$$

Since the symmetric operator T^*T is bounded below by a strictly positive constant, it is invertible, with a bounded inverse. This inverse satisfies

$$B^{-1}\mathbb{1} \leq (T^*T)^{-1} \leq A^{-1}\mathbb{1}. \tag{2.1.2}$$

Define $\tilde{\phi}_l = (T^*T)^{-1}\phi_l$. Then the family $(\tilde{\phi}_l)_{l \in J}$ constitutes another frame. More precisely, we have

Proposition 2.1:

1) The family $(\tilde{\phi}_l)_{l \in J}$, with $\tilde{\phi}_l = (T^*T)^{-1}\phi_l$, constitutes a frame with bounds B^{-1} and A^{-1}.

2) The associated frame operator $\tilde{T}$ is given by $\tilde{T} = T(T^*T)^{-1}$ and satisfies

$$\begin{aligned}\tilde{T}^*\tilde{T} &= (T^*T)^{-1}\\ \tilde{T}^*T &= \mathbb{1} = T^*\tilde{T}\end{aligned} \tag{2.1.3}$$

where $\tilde{T}T^* = T\tilde{T}^*$ is the orthogonal projection operator, in $l^2(J)$, onto the range of T.

Proof:

1) For any $f \in \mathscr{H}$, we have

$$\langle \tilde{\phi}_l, f\rangle = \left\langle (T^*T)^{-1}\phi_l, f\right\rangle = \left\langle \phi_l, (T^*T)^{-1}f\right\rangle.$$

Hence

$$\begin{aligned}\sum_l |\langle \tilde{\phi}_l, f\rangle|^2 &= \|T(T^*T)^{-1}f\|^2\\ &= \left\langle (T^*T)^{-1}f, T^*T(T^*T)^{-1}f\right\rangle\\ &= \left\langle f, (T^*T)^{-1}f\right\rangle.\end{aligned}$$

It then follows from (2.1.2) that the $(\tilde{\phi}_l)_{l \in J}$ constitute a frame, with frame bounds B^{-1} and A^{-1}.

2) By the definition of the $\tilde{\phi}_l$ we have

$$\begin{aligned}(\tilde{T}f)_l &= \langle \tilde{\phi}_l, f\rangle = \langle (T^*T)^{-1}\phi_l, f\rangle\\ &= \langle \phi_l, (T^*T)^{-1}f\rangle = \left(T(T^*T)^{-1}f\right)_l,\end{aligned}$$

hence $\tilde{T} = T(T^*T)^{-1}$. It follows that

$$\tilde{T}^*\tilde{T} = (T^*T)^{-1}T^*T(T^*T)^{-1} = (T^*T)^{-1}$$

and

$$\begin{aligned}\tilde{T}^*T &= (T^*T)^{-1}T^*T = \mathbb{1},\\ T^*\tilde{T} &= T^*T(T^*T)^{-1} = \mathbb{1}.\end{aligned}$$

3) Finally $\tilde{T}T^* = T(T^*T)^{-1}T^* = T\tilde{T}^*$; it remains to prove that this is the orthogonal projection operator P onto the range of T. Since $T^*c = 0$ for any c orthogonal to the range of T, it suffices to prove that $T(T^*T)^{-1}T^*c = c$ for c in the range of T. If c is in the range of T, $c = Tf$, then

$$T(T^*T)^{-1}T^*c = T(T^*T)^{-1}T^*Tf = Tf = c.$$

This proves $\tilde{T}T^* = P = T\tilde{T}^*$. □

The operation $\{\phi_l;\ l \in J\} \to \{\tilde{\phi}_l;\ l \in J\}$ defines, in a sense, a duality operation. The same procedure, applied to the frame $\{\tilde{\phi}_l;\ l \in J\}$, gives the original frame $\{\phi_l;\ l \in J\}$ back again. We shall therefore call $\{\tilde{\phi}_l;\ l \in J\}$ the *dual frame* of $\{\phi_l;\ l \in J\}$. The duality $\phi_l \leftrightarrow \tilde{\phi}_l$ is also expressed by (2.1.3); for any $f, g \in \mathscr{H}$ we have

$$\begin{aligned}\langle f, g\rangle &= \sum_{l \in J} \langle f, \tilde{\phi}_l\rangle\langle \phi_l, g\rangle\\ &= \sum_{l \in J} \langle f, \phi_l\rangle\langle \tilde{\phi}_l, g\rangle.\end{aligned}$$

In what follows we shall so often encounter T^*T (rather than T alone) that it makes sense to introduce a new notation for this operator. We define

$$\mathbb{T} = T^*T, \qquad \tilde{\mathbb{T}} = \tilde{T}^*\tilde{T} = \mathbb{T}^{-1}.$$

In particular

$$\mathbb{T} = \sum_{l \in J} \phi_l \langle \phi_l, \cdot \rangle$$
$$A\mathbb{1} \le \mathbb{T} \le B\mathbb{1}$$
$$\tilde{\phi}_l = \mathbb{T}^{-1}\phi_l.$$

Proposition 2.1 gives us an inversion formula for T. If the elements $f \in \mathscr{H}$ are characterized by means of the inner products $(\langle f, \phi_l \rangle;\ l \in J)$, then f can be reconstructed from these by means of

$$f = \sum_l \tilde{\phi}_l \langle \phi_l, f \rangle. \tag{2.1.4}$$

In (2.1.4) the vectors $\tilde{\phi}_l$ are defined by $\tilde{\phi}_l = \mathbb{T}^{-1}\phi_l$. Note that, in general, if the frame is redundant (i.e., contains "more" vectors than a basis would), there exist other vectors in $\mathscr{H}$ that could equally well play the role of the $\tilde{\phi}_l$ and lead to a reconstruction formula. This is due to the fact that the ϕ_l are not linearly independent in the general case. This phenomenon can be illustrated with the two-dimensional example of Section I-C. Take $\mathscr{H} = \mathbb{C}^2$, and define

$$\phi_1 = e_1, \qquad \phi_2 = -\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2, \qquad \phi_3 = -\frac{1}{2}e_1 - \frac{\sqrt{3}}{2}e_2$$

where $e_1 = (1,0)$, $e_2 = (0,1)$ constitute the standard orthonormal basis in $\mathbb{C}^2$. One can check very easily that for all $u \in \mathscr{H}$,

$$\sum_{l=1}^{3} |\langle \phi_l, u \rangle|^2 = \frac{3}{2}\|u\|^2$$

so that $\mathbb{T} = 3/2\mathbb{1}$, hence $\tilde{\phi}_l = \mathbb{T}^{-1}\phi_l = 2/3\phi_l$, and (2.1.4) becomes

$$u = \frac{2}{3}\sum_{l=1}^{3} \phi_l \langle \phi_l, u \rangle.$$

Since $\Sigma_{l=1}^3 \phi_l = 0$ in this example, it is clear, however, that for any choice of $a \in \mathscr{H}$, an equally valid reconstruction formula is given by

$$u = \frac{2}{3}\sum_{l=1}^{3} (\phi_l + a)\langle \phi_l, u \rangle.$$

The choice $a = 0$, corresponding to (2.1.4), is the "minimal solution" in the following sense. The image, in $\mathbb{C}^3$, of $\mathbb{C}^2$ under the frame operator T is the two-dimensional subspace with equation $x_1 + x_2 + x_3 = 0$; we denote this subspace by *RanT*. Vectors in $\mathbb{C}^3$ orthogonal to *RanT* are all of the type $c = \lambda(1,1,1)$. When the components of such a vector are substituted for the $\langle \phi_l, u \rangle$ in (2.1.4), then the "reconstruction" leads to 0, since

$$\sum_{l=1}^{3} \tilde{\phi}_l c_l = \frac{2}{3}\lambda \sum_{l=1}^{3} \phi_l = 0.$$

This is no longer the case if one of the other "equally valid" choices is used, where $\tilde{\phi}_l$ is replaced by $2/3(\phi_l + a)$. Of all the possible "quasi-inverses" for T, $\tilde{T}^*$ is therefore the only one that automatically projects sequences in $(RanT)^{\perp}$ to zero while effecting the reconstruction.

A similar phenomenon occurs with the infinite-dimensional frames we shall consider. While no finite number of them will be linearly dependent, there generally do exist convergent linear combinations, involving infinitely many ϕ_l, which sum to zero. This again causes "reconstruction formulas" to be nonunique, but again the $\tilde{\phi}_l = \mathbb{T}^{-1}\phi_l$ will offer the optimal solution, in the sense that they are the *only* choice leading to

$$\sum_l c_l \tilde{\phi}_l = 0, \qquad \text{for all } c = (c_l)_{l \in J} \perp RanR.$$

This is easy to check by the following argument. Write $\tilde{\phi}_l = \mathbb{T}^{-1}\phi_l + u_l$, and suppose that (2.1.4) holds. It follows that $\Sigma u_l \langle \phi_l, f \rangle = 0$, for all $f \in \mathscr{H}$, or $\Sigma u_l c_l = 0$, for all $c \in RanT$. The optimality condition implies $\Sigma u_l c_l = 0$ for all $c \perp RanT$. It follows that $\Sigma u_l c_l = 0$ for all $c \in l^2(J)$, hence $u_l = 0$ for all l.

To apply (2.1.4) it is, of course, necessary to construct the $\tilde{\phi}_l$ first. Writing $\mathbb{T}$ as $2(A+B)^{-1}[\mathbb{1} - (2(A+B)^{-1}\mathbb{T})]$, one has the following series for the $\tilde{\phi}_l$,

$$\tilde{\phi}_l = \frac{2}{A+B}\sum_{k=0}^{\infty}\left(\mathbb{1} - \frac{2\mathbb{T}}{A+B}\right)^k \phi_l. \tag{2.1.5}$$

Since

$$\frac{A-B}{A+B}\mathbb{1} \le \mathbb{1} - \frac{2\mathbb{T}}{A+B} \le \frac{B-A}{B+A}\mathbb{1}$$

or

$$\left\|\mathbb{1} - \frac{2\mathbb{T}}{A+B}\right\| \le \frac{B/A - 1}{B/A + 1} < 1 \tag{2.1.6}$$

the series in (2.1.5) always converges. The closer B/A is to 1, the faster the convergence, of course.

Remark: For the reconstruction formula (2.1.4) for f from its coefficients $\langle \phi_l, f \rangle$, we have used (2.1.3), namely that $\tilde{T}^*T = \mathbb{1}$. One can also interpret (2.1.3) differently. From $T^*\tilde{T} = \mathbb{1}$ it follows that, for all $f \in \mathscr{H}$,

$$f = \sum_l \langle \tilde{\phi}_l, f \rangle \phi_l.$$

This tells us that any function f can be expanded in the ϕ_l, and gives us a recipe for calculating appropriate coefficients. This is the point of view taken in so-called "atomic decompositions." It is thus clear that frames and atomic decompositions are dual notions [48].

All the previous formulas become much simpler if the frame is *tight*, i.e., if $B = A$. In that case

$$\mathbb{T} = A\mathbb{1}, \qquad \tilde{\mathbb{T}} = A^{-1}\mathbb{1}$$
$$\tilde{\phi}_l = A^{-1}\phi_l$$
$$f = A^{-1}\sum_{l \in J} \phi_l \langle \phi_l, f \rangle.$$

If the frame constant for a tight frame is equal to one, $A = 1$, then the reconstruction formula looks exactly like the decomposition of a function with respect to an orthonormal basis,

$$f = \sum_{l \in J} \phi_l \langle \phi_l, f \rangle.$$

In fact, if the vectors in a tight frame with frame constant 1 are all normalized, then the frame constitutes an orthonormal basis, as can easily be seen by the following argument:

$$\begin{aligned} \|\phi_k\|^2 &= \sum_{l \in J} |\langle \phi_l, \phi_k \rangle|^2 \\ &= \|\phi_k\|^4 + \sum_{l \notin J} |\langle \phi_l, \phi_k \rangle|^2 \end{aligned}$$

which implies $\langle \phi_l, \phi_k \rangle = 0$ for $l \neq k$ if $\|\phi_k\| = 1$. As we explained in the introduction, there are circumstances in which one prefers to work with nontight frames. In that case the $\tilde{\phi}_l$ have to be constructed explicitly from the ϕ_l. Formula (2.1.5) then gives an algorithm for the computation of the $\tilde{\phi}_l$; as (2.1.6) shows, it pays to have a ratio B/A as close to 1 as possible. Frames that are almost tight (i.e., B/A close to 1) will be called *snug frames*. The "snugness" can be measured by $S = [B/A - 1]^{-1}$. In the examples below we shall encounter values of $S \geq 100$, corresponding to B/A of the order of 1.01 or even smaller, for quite realistic frames (i.e., reasonably large values of p_0, q_0 or a_0, b_0—see Introduction).

Note also that, while in principle all the different $\tilde{\phi}_l$, $l \in J$, have to be calculated separately, in practice simplifications may occur. Let us show what happens for the discrete Weyl–Heisenberg coherent state frames and for the wavelet frames defined in the introduction. We take therefore $\mathscr{H} = L^2(\mathbb{R})$, $J = \mathbb{Z}^2$.

In the Weyl–Heisenberg case, we can rewrite $\mathbb{T}$ as

$$\mathbb{T} = \sum_{m,n \in \mathbb{Z}} W(mp_0, nq_0) g \langle W(mp_0, nq_0) g, \cdot \rangle$$

where the operator $W(p,q)$ is defined by $[W(p,q)g](x) = e^{ipx} g(x-q)$. One easily checks that $W(p,q)W(p',q') = e^{-ip'q} W(p+p', q+q')$. Using this multiplication formula for the operators W, one easily finds that

$$W(mp_0, nq_0)\mathbb{T} = \mathbb{T} W(mp_0, nq_0).$$

Hence

$$W(mp_0, nq_0)\mathbb{T}^{-1} = \mathbb{T}^{-1} W(mp_0, nq_0).$$

This implies

$$\begin{aligned} (g_{mn})^{\sim} &= \mathbb{T}^{-1} g_{mn} \\ &= \mathbb{T}^{-1} W(mp_0, nq_0) g \\ &= W(mp_0, nq_0) \mathbb{T}^{-1} g, \end{aligned}$$

or $(g_{mn})^{\sim}(x) = e^{imp_0x}(g_{00})^{\sim}(x - nq_0) = \tilde{g}_{mn}(x)$, where $\tilde{g} = (g_{00})^{\sim} = \mathbb{T}^{-1} g$. We have thus only one function to compute, i.e., $\tilde{g}_{00} = \mathbb{T}^{-1} g$, instead of a double infinity of different $(g_{mn})^{\sim} (m, n \in \mathbb{Z})$. If moreover B/A is close to 1, then the rapidly converging series (2.1.5) can be used for this computation.

In the affine case, we find

$$\mathbb{T} = \sum_{m,n \in \mathbb{Z}} U(a_0^m, a_0^m nb_0) h \langle U(a_0^m, a_0^m nb_0) h, \cdot \rangle$$

where $[U(a,b)f](x) = |a|^{-1/2} f(x - b/a)$. Again we can use the composition law

$$U(a', b') U(a'', b'') = U(a'a'', b' + a'b'')$$

to show that

$$U(a_0^m, 0)\mathbb{T} = \mathbb{T} U(a_0^m, 0).$$

Note that we have no translation in these U-operators; it turns out that $\mathbb{T}$ does *not* commute with $U(a_0^m, nb_0 a_0^m)$ if $n \neq 0$. Since $U(a_0^m, a_0^m nb_0) = U(a_0^m, 0) U(1, nb_0)$, we find

$$(h_{mn})^{\sim} = U(a_0^m, 0) \mathbb{T}^{-1} U(1, nb_0) h$$

or

$$(h_{mn})^{\sim}(x) = a_0^{-m/2} (h_{0n})^{\sim} (a_0^{-m} x).$$

The simplification is less drastic than in the WH case, since only one of the two indices is eliminated, and one still has to compute the infinite number of $(h_{0n})^{\sim} = \mathbb{T}^{-1} h_{0n}$. In practice, however, only a finite number are needed, and for the computation of these it is again a big help if B/A is close to 1.

Note that if one stops at the first term ($k = 0$) in (2.1.5), one obtains the approximate inversion formula

$$f_{\text{approx}} = \frac{2}{A+B} \sum_{m,n} \phi_{mn} \langle \phi_{mn}, f \rangle. \tag{2.1.7}$$

If B/A is close enough to one, this is a fairly good approximation. In the first calculations with wavelets, the formula used for analysis and reconstruction was very similar to (2.1.7) [25], without theoretical basis. It turns out that in those frames B/A was indeed very close to 1, so that f_{approx} is a good approximation to f.

B. Tight Frames

1) Explicitly Constructed Tight Frames: In this subsection we review shortly, for the sake of completeness, the explicit construction of tight frames given in [16]. We have drawn inspiration from [31] to make the construction more elegant. For examples, graphs and a more detailed discussion we refer to [16].

In both cases an auxiliary function ν of the following type is used

$$\nu: \mathbb{R} \to \mathbb{R}$$

$$\nu(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1, & \text{if } x \geq 1. \end{cases}$$

If ν is chosen to be a C^k (or C^∞) function, then the resulting frames will consist of C^k (C^∞) functions.

a) The Weyl–Heisenberg case: Choose $p_0, q_0 > 0$, with $p_0 \cdot q_0 < 2\pi$. (In Section II-C we shall have more to say about this restriction.) We shall assume here that $p_0 \cdot q_0 \geq \pi$. (See [16, Appendix] for indications how the construction can be adapted if $p_0 \cdot q_0 < \pi$). Define $\lambda = 2\pi/p_0 - q_0$; one has $0 < \lambda \leq \pi/p_0$. The function g is then constructed as follows:

$$g(x)=q_0^{-1/2}\begin{cases}0, & x\le -\pi/p_0\\ \sin\left[\frac{\pi}{2}\nu\left(\lambda^{-1}\left(\frac{\pi}{p_0}+x\right)\right)\right], & -\pi/p_0\le x\le -\frac{\pi}{p_0}+\lambda\\ 1, & -\frac{\pi}{p_0}+\lambda\le x\le\frac{\pi}{p_0}-\lambda\\ \cos\left[\frac{\pi}{2}\nu\left(\lambda^{-1}\left(x-\frac{\pi}{p_0}+\lambda\right)\right)\right], & \frac{\pi}{p_0}-\lambda\le x\le \pi/p_0\\ 0, & x\ge\pi/p_0.\end{cases}$$

The overall factor $q_0^{-1/2}$ normalizes g; one easily checks that $\int dx|g(x)|^2=1$. The following two properties of g are crucial.

$$\begin{aligned}&1)\quad |\text{support } g|=2\pi/p_0,\\ &2)\quad \sum_{k\in\mathbb{Z}}|g(x-kq_0)|^2=1/q_0.\end{aligned}\tag{2.2.1}$$

These two properties together ensure that the g_{mn}, $g_{mn}(x)=e^{imp_0x}g(x-nq_0)$, constitute a tight frame. By Plancherel's theorem we have indeed for all $f\in L^2(\mathbb{R})$,

$$\begin{aligned}&\sum_{m,n\in\mathbb{Z}}|\langle g_{mn},f\rangle|^2\\ &=\sum_{m,n\in\mathbb{Z}}\left|\int_{|x-nq_0|\le\pi/p_0}dx\,e^{imp_0x}g(x-nq_0)^*f(x)\right|^2\\ &=\sum_{n\in\mathbb{Z}}\frac{2\pi}{p_0}\int dx|g(x-nq_0)|^2|f(x)|^2\\ &=\frac{2\pi}{p_0q_0}\int dx|f(x)|^2\end{aligned}$$

where we have used the notation $g(y)^*$ to denote the complex conjugate of $g(y)$. Note that the same calculation can be made whenever $|\text{support } g|\le 2\pi/p_0$, even if (2.2.1) is not satisfied. In that case the operator $\mathbb{T}$ of Section II-A reduces to multiplication by the periodic function $G(x)=2\pi p_0^{-1}\Sigma_{n\in\mathbb{Z}}|g(x-nq_0)|^2$ (see also [50]).

One finds thus $\tilde{g}_{mn}(x)=G(x)^{-1}g_{mn}(x)$, and the inversion formula becomes

$$f=G^{-1}\sum_{m,n}g_{mn}\langle g_{mn},f\rangle.$$

Remark: If the function ν is chosen to be C^∞, then g is a compactly supported C^∞-function. While the uncertainty principle can of course not be violated, this nevertheless allows fairly good localization of g in both time and frequency. The coefficients $\langle g_{mn},f\rangle$ do therefore correspond to a reasonably accurate time-frequency localized picture of the signal f. Moreover, there exists a constant $A=2\pi/p_0\cdot q_0$ such that

$$\int dx|f(x)|^2=A^{-1}\sum_{m,n}|\langle g_{mn},f\rangle|^2.$$

The role of this constant A is crucial. If one restricts oneself to $A=1$, then this forces the g_{mn} to be an orthonormal basis, which leads to functions g that are badly localized in either time or frequency. This is a consequence of the Balian–Low theorem [51], [52] (Theorem 2.3 next), rediscovered in [53]. As soon, however, as one allows $A\neq 1$, the picture changes drastically, and much better time-frequency localization is attainable.

b) The wavelet case: In this case there are no *a priori* restrictions on the choice of the parameters a_0,b_0, other than $a_0\neq 0$ or 1 and $b_0\neq 0$. We may choose, without loss of generality, $a_0>1$, and $b_0>0$. Define $l=2\pi/[b_0(a_0^2-1)]$. The tight frame of wavelets will be based on the function h with Fourier transform $\hat{h}$ constructed as follows [16]:

$$\hat{h}(y)=(\log a_0)^{-1/2}\begin{cases}0, & y\le l\\ \sin\left[\frac{\pi}{2}\nu\left(\frac{y-l}{l(a_0-1)}\right)\right], & l\le y\le a_0l\\ \cos\left[\frac{\pi}{2}\nu\left(\frac{y-a_0l}{la_0(a_0-1)}\right)\right], & a_0l\le y\le a_0^2l\\ 0, & y\ge a_0^2l.\end{cases}\tag{2.2.2}$$

The function h itself is given by the inverse Fourier transform,

$$h(x)=\frac{1}{\sqrt{2\pi}}\int dy\, e^{-ixy}\hat{h}(y).$$

The normalization of h has been chosen so that

$$\int dy|y|^{-1}|\hat{h}(y)|^2=1.$$

The following two properties of h are again crucial:

$$\begin{aligned}&1)\quad |\text{support}\,\hat{h}|=(a_0^2-1)l=2\pi/b_0,\\ &2)\quad \sum_{k\in\mathbb{Z}}\left|\hat{h}(a_0^k y)\right|^2=(\log a_0)^{-1}\chi_{[0,\infty)}(y),\end{aligned}\tag{2.2.3}$$

where $\chi_{[0,\infty)}$ is the indicator function of the right half line, $\chi_{[0,\infty)}(y)=1$ if $y\geq 0$, $\chi_{[0,\infty)}(y)=0$, otherwise. These properties together enable us to construct a tight frame based on h. By Parseval's theorem, we have, for all $f\in L^2(\mathbb{R})$,

$$\begin{aligned}\sum_{m,n\in\mathbb{Z}}|\langle h_{mn},f\rangle|^2 &=\sum_{m,n\in\mathbb{Z}}a_0^m\left|\int dy\, e^{inb_0a_0^m y}\hat{h}(a_0^m y)^*\hat{f}(y)\right|^2\\ &=\sum_{m\in\mathbb{Z}}\frac{2\pi}{b_0}\int dy\left|\hat{h}(a_0^m y)\right|^2|\hat{f}(y)|^2\\ &=\frac{2\pi}{b_0\log a_0}\int_0^\infty dy\,|\hat{f}(y)|^2.\end{aligned}\tag{2.2.4}$$

If we define $h^+=h$, $h^-=h^*$ (i.e., $(h^-)^\wedge(k)=\hat{h}(-k)$), then this calculation shows that the $(h_{mn}^{\pm};\ m,n\in\mathbb{Z})$ constitute a tight frame. Specifically,

$$\sum_{\epsilon=+\text{ or }-}\ \sum_{m,n\in\mathbb{Z}}|\langle h_{mn}^{\epsilon},f\rangle|^2=\frac{2\pi}{b_0\log a_0}\|f\|^2.\tag{2.2.5}$$

Similarly the $(h_{mn}^{(\lambda)};\ m,n\in\mathbb{Z},\ \lambda=1\text{ or }2)$, with $h^{(1)}=\operatorname{Re}h$, $h^{(2)}=\operatorname{Im}h$, constitute a tight frame. Strictly speaking, the frames $\{h_{mn}^{\epsilon};\ \epsilon=+\text{ or }-,\ m,n\in\mathbb{Z}\}$ or $\{h_{mn}^{(\lambda)};\ \lambda=1\text{ or }2,\ m,n\in\mathbb{Z}\}$ are not quite of the type described in the introduction, since they are not obtained by dilating and translating one single function. As shown by the computation leading to (2.2.4), the operator $\mathbb{T}=T^*T$ handles the positive and negative frequency domains independently. Since the Fourier transform $\hat{h}$ of h, defined by (2.2.2), is entirely supported on the positive half line, the use of a second function, which will handle the negative frequencies, is therefore unavoidable. In other examples (see e.g., Section II-C-2) the function h will be chosen such that $\hat{h}$ is supported on both the negative and the positive half lines, and one function suffices.

Let us return to the construction (2.2.2). For a special set of signals f to be analyzed, one may restrict oneself, even if support $\hat{h}\subset[0,\infty)$, to only one basic function h, and the associated wavelets h_{mn}. This happens if one knows *a priori*, as is often the case, that the signal f is real. Since then $\hat{f}(-y)=\hat{f}(y)^*$, one finds

$$\sum_{m,n\in\mathbb{Z}}|\langle h_{mn},f\rangle|^2=\frac{\pi}{b_0\log a_0}\|f\|^2.$$

As for the Weyl–Heisenberg case, a calculation similar to (2.2.4) can be carried out whenever the support width of $\hat{h}$ is smaller than $2\pi/b_0$, even if (2.2.3) is not satisfied. The operator $\mathbb{T}$ becomes then

$$(\mathbb{T}f)^\wedge(y)=\hat{H}(y)\hat{f}(y)$$

with

$$\hat{H}(y)=\frac{2\pi}{b_0}\sum_m\left[|h(a_0^m y)|^2+|h(-a_0^m y)|^2\right]$$

or, equivalently,

$$\begin{aligned}(\mathbb{T}f)(x)&=\frac{1}{\sqrt{2\pi}}(H*f)(x)\\ &=\frac{1}{\sqrt{2\pi}}\int dx'\,H(x-x')f(x').\end{aligned}$$

Hence

$$\left[(h_{mn}^{\pm})^\sim\right]^\wedge(y)=\left[\hat{H}(y)\right]^{-1}(h_{mn}^{\pm})^\wedge(y).$$

In the construction of the orthonormal basis, in [31], the procedure starts in the same way. The function ν is chosen to be C^∞, and has one additional property,

$$\nu(x)+\nu(1-x)=1,\qquad\text{for all }x\in\mathbb{R}.\tag{2.2.6}$$

The "doubling" (using superscripts $\pm$ to be able to cover, in Fourier space, the whole real line instead of only the half line) is avoided in [31] by incorporating also the mirror image of h into the basis function ψ. Explicitly,

$$\hat{\psi}(y)=\left(\frac{b_0\log a_0}{2\pi}\right)^{1/2}e^{ib_0y/2}\left[\hat{h}(y)+\hat{h}(-y)\right].$$

Due to the extra property (2.2.6) of ν, one easily checks that $\|\psi\|^2=\int dx\,|\psi(x)|^2=1$. Note that the size of the support of $\hat{\psi}$ is no longer equal to $2\pi/b_0$, so that the straightforward calculation made previously no longer works. In fact, the set of functions $\psi_{mn}(x)=a_0^{-m/2}\psi(a_0^{-m}x-nb_0)$, $m,n\in\mathbb{Z}$, turns out to be an orthonormal basis of $L^2(\mathbb{R})$ as the result of some miraculous cancellations, for which the property (2.2.6) of the function ν turns out to be quite crucial [31]. These cancellations occur only when a_0 takes on a very special value, namely when

$$a_0=\frac{k+1}{k},\qquad\text{with }k\in\mathbb{N}.$$

The case $k=1$, or $a_0=2$, corresponds to Y. Meyer's construction in [31]. With $b_0=1$ (as in [31]), Y. Meyer's basic wavelet is thus given by

$$\hat{\psi}(y)=\frac{1}{\sqrt{2\pi}}e^{iy/2}[\omega(y)+\omega(-y)]\tag{2.2.7}$$

with

$$\omega(y)=\begin{cases}0, & y\le \dfrac{2\pi}{3}\\ \sin\left[\dfrac{\pi}{2}\nu\left(\dfrac{3y}{2\pi}-1\right)\right], & \dfrac{2\pi}{3}\le y\le\dfrac{4\pi}{3}\\ \cos\left[\dfrac{\pi}{2}\nu\left(\dfrac{3y}{4\pi}-1\right)\right], & \dfrac{4\pi}{3}\le y\le\dfrac{8\pi}{3}\\ 0, & y\ge\dfrac{8\pi}{3}\end{cases}$$

where ν is a C^∞ function from $\mathbb{R}$ to $[0,1]$ such that $\nu(y)=0$ for $y\le 0$, $\nu(y)=1$ for $y\ge 1$ and $\nu(y)+\nu(1-y)=1$ for all y. As pointed out in the introduction, the concept of multiscale analysis allows one to understand more deeply why this construction works, so that the "miraculous cancellations" just mentioned become less so.

2) Relations Between the Frame Parameters and the Frame Bounds: In the explicitly constructed tight frames, the frame constant A is given by $A=2\pi/p_0q_0$, for the Weyl–Heisenberg case, and by $A=\pi/b_0\log a_0$, for the affine or wavelet case. This is no coincidence. We show in this subsection that these values for A are imposed by the normalizations chosen for g,h, and are independent of the details of the construction, i.e., they are generally true for *all* tight frames. More generally, we prove inequalities for the frame bounds A,B for *all* Weyl–Heisenberg frames or wavelet frames, tight or not.

a) The Weyl–Heisenberg case: Let us assume that the $(\phi_{mn};\ m,n\in\mathbb{Z})$ are an arbitrary frame of discrete Weyl-Heisenberg coherent states, with frame constants A,B, and lattice spacings p_0,q_0, i.e.,

$$\phi_{mn}(x)=e^{imp_0x}\phi(x-nq_0),$$
$$A\|f\|^2\le\sum_{m,n}|\langle\phi_{mn},f\rangle|^2\le B\|f\|^2.$$

We shall see next that a frame is possible only if $p_0\cdot q_0\le 2\pi$. Let us therefore restrict ourselves to this case. Then there exists, for the same values p_0,q_0, a tight frame $\{g_{mn};\ m,n\in\mathbb{Z}\}$ with $\|g\|=1$. (For $p_0\cdot q_0<2\pi$, we take g as constructed in Section II-B-1a); for $p_0\cdot q_0=2\pi$, take $g(x)=q_0^{-1/2}$ if $|x|\le q_0/2$, $g(x)=0$, otherwise). One easily checks that

$$\langle g_{mn},\phi\rangle=e^{-imnp_0q_0}\langle g,\phi_{-m-n}\rangle. \tag{2.2.8}$$

It follows that

$$A\|g\|^2\le\sum_{m,n}|\langle\phi_{mn},g\rangle|^2$$
$$=\sum_{m,n}|\langle g_{mn},\phi\rangle|^2=\frac{2\pi}{p_0q_0}\|\phi\|^2.$$

Similarly $B\|g\|^2\ge 2\pi/p_0q_0\|\phi\|^2$. Since $\|g\|=1$, we find

$$A\le\frac{2\pi}{p_0q_0}\|\phi\|^2\le B. \tag{2.2.9}$$

This is true for any Weyl–Heisenberg-frame with lattice spacings p_0,q_0. In particular, if the ϕ_{mn} constitute a tight frame (i.e., $A=B$), then necessarily

$$A=\frac{2\pi}{p_0q_0}\|\phi\|^2.$$

a) The wavelet case: In this case also we shall derive inequalities similar to (2.2.9). Since there is no equivalent to (2.2.8) for wavelet frames, the derivation will be slightly more complicated. Let us assume, again, that $(\phi_{mn};\ m,n\in\mathbb{Z})$ is an arbitrary frame of wavelets, with lattice spacings determined by a_0,b_0, and with frame constants A,B, i.e.,

$$\phi_{mn}(x)=a_0^{-m/2}\phi(a_0^{-m}x-nb_0)$$
$$A\|f\|^2\le\sum_{m,n}|\langle\phi_{mn},f\rangle|^2\le B\|f\|^2. \tag{2.2.10}$$

Take now any positive operator C which is trace-class, i.e., which is of the form

$$C=\sum_l c_lu_l\langle u_l,\cdot\rangle$$

where the u_l are orthonormal, $c_l>0$ and $\operatorname{tr}C=\Sigma_lc_l<\infty$. Then (2.2.10) implies

$$\sum_l c_lA\|u_l\|^2\le\sum_l\sum_{m,n}c_l|\langle u_l,\phi_{mn}\rangle|^2\le\sum_l c_lB\|u_l\|^2$$

or

$$A\mathit{Tr}C\le\sum_{m,n}\langle\phi_{mn},C\phi_{mn}\rangle\le B\mathit{Tr}C. \tag{2.2.11}$$

We shall apply this to the following operator,

$$C=\int\int\frac{da\,db}{a^2}h^{(a,b)}\langle h^{(a,b)},\cdot\rangle t(a,b) \tag{2.2.12}$$

constructed with the help of the affine coherent states introduced in Section I-D. The function $h\in L^2(\mathbb{R})$ should satisfy $\int dy\,|y|^{-1}|\hat{h}(y)|^2<\infty$ (see Section I-D). Here $t(a,b)$ is a positive function in $L^1(\mathbb{R}^*\times\mathbb{R};\ a^{-2}da\,db)$. The operator (2.2.12) is positive and trace-class, with

$$\operatorname{tr}C=\int\frac{da\,db}{a^2}t(a,b)$$

where we have assumed $\|h\|=1$. On the other hand,

$$\begin{aligned}\langle\phi_{mn},C\phi_{mn}\rangle&=\langle\phi,U(a_0^{-m},a_0^{-m}nb_0)^{-1}CU(a_0^{-m},a_0^{-m}nb_0)\phi\rangle\\&=\int\frac{da\,db}{a^2}\left|\langle\phi,U(a_0^{-m},a_0^{-m}nb_0)^{-1}U(a,b)h\rangle\right|^2t(a,b)\\&=\int\frac{da\,db}{a^2}|\langle\phi,U(a,b)h\rangle|^2t(a_0^ma,a_0^m(b+nb_0))\end{aligned} \tag{2.2.13}$$

where we have used the composition law of the U-operators. We now restrict ourselves to functions t of the form

$$t(a,b)=\chi_{[1,a_0[}(|a|)\cdot t_1(b/|a|)$$

where $\chi_{[1,a_0)}$ is the indicator function of the half open interval $[1,a_0)$, i.e., $\chi_{[1,a_0)}u = 1$ if $1 \le u < a_0$, 0 otherwise. Since $\Sigma_m \chi_{[1,a_0)}(a_0^m|a|) = 1$, (2.2.13) then leads to

$$\sum_{m,n} \langle \phi_{mn}, C\phi_{mn}\rangle = \int \frac{da\,db}{a^2} |\langle \phi, U(a,b)h\rangle|^2 \sum_{n\in\mathbb{Z}} t_1\left(\frac{b+nb_0}{|a|}\right). \quad (2.2.14)$$

This sum over n can be estimated by the following lemma.

Lemma 2.2: Let f be a positive, continuous, bounded function on $\mathbb{R}$, with $f(x) \to 0$ as $|x| \to \infty$. Assume that f has a finite number of local maxima, at x_j, $j = 1, \cdots, N$. Define

$$\Delta_j = \sup_{\delta \in [0,1]} \int_{x_j-\delta}^{x_j-\delta+1} dx f(x).$$

Then

$$\int_{-\infty}^{\infty} dx f(x) - \sum_{j=1}^{N} \Delta_j \le \sum_{n\in\mathbb{Z}} f(n) \le \int_{-\infty}^{\infty} dx f(x) + \sum_{j=1}^{N} f(x_j).$$

For the sake of completeness we provide a proof for the case $N = 1$; the case for general N can be proved analogously, though the inequalities can be sharpened if some of the x_j are within a distance 1 of each other.

Proof: Let n_0 be the largest integer not exceeding x_1, the point where f reaches its maximum. Since f is increasing on $(-\infty, x_1]$ and decreasing on $[x_1, \infty)$, and since $n_0 \le x_1 < n_0 + 1$:

$$\sum_{n=-\infty}^{n_0-1} f(n) \le \sum_{n=-\infty}^{n_0-1} \int_n^{n+1} dx f(x) = \int_{-\infty}^{n_0} dx f(x),$$

$$\sum_{n=n_0+1}^{\infty} f(n) \le \sum_{n=n_0+1}^{\infty} \int_{n-1}^{n} dx f(x) = \int_{n_0}^{\infty} dx f(x).$$

Hence

$$\sum_{n=-\infty}^{\infty} f(n) \le \int_{-\infty}^{\infty} dx f(x) + f(n_0) \le \int_{-\infty}^{\infty} dx f(x) + f(x_1).$$

On the other hand,

$$\sum_{n=-\infty}^{n_0} f(n) \ge \sum_{n=-\infty}^{n_0} \int_{n-1}^{n} dx f(x) = \int_{-\infty}^{n_0} dx f(x),$$

$$\sum_{n=n_0+1}^{\infty} f(n) \ge \sum_{n=n_0+1}^{\infty} \int_n^{n+1} dx f(x) = \int_{n_0+1}^{\infty} dx f(x).$$

Hence

$$\sum_{n=-\infty}^{\infty} f(n) \ge \int_{-\infty}^{\infty} dx f(x) - \int_{n_0}^{n_0+1} dx f(x) \ge \int_{-\infty}^{\infty} dx f(x) - \Delta_1. \quad \blacksquare$$

In particular, if f in Lemma 2.2 has only one local maximum, at $x = x_1$, then

$$\int_{-\infty}^{\infty} dx f(x) - f(x_1) \le \sum_{n=-\infty}^{\infty} f(n) \le \int_{-\infty}^{\infty} dx f(x) + f(x_1).$$

Let us apply this to (2.2.14). Choose F to be any positive, continuous L^1 function on $\mathbb{R}$, tending to zero at infinity, with one local maximum, at $x = 0$. Choose $\lambda > 0$, and define

$$t_1(x) = F(\lambda x).$$

Applying Lemma 2.2 to (2.2.14) then leads to

$$\left|\sum_{m,n} \langle \phi_{mn}, C\phi_{mn}\rangle - \frac{1}{\lambda b_0}\left[\int dx F(x)\right] \cdot \int \frac{da\,db}{|a|} |\langle \phi, U(a,b)h\rangle|^2 \right| \le F(0) \int \frac{da\,db}{a^2} |\langle \phi, U(a,b)h\rangle|^2$$

or

$$\left|\sum_{m,n} \langle \phi_{mn}, C\phi_{mn}\rangle - \frac{1}{\lambda b_0}\left[\int dx F(x)\right] 2\pi \|h\|^2 \cdot \left[\int dy\, |y|^{-1} |\hat{\phi}(y)|^2\right]\right| \le F(0) \cdot 2\pi \|\phi\|^2 \cdot \left[\int dy\, |y|^{-1} |\hat{h}(y)|^2\right]. \quad (2.2.15)$$

On the other hand,

$$TrC = 2\int_1^{a_0} \frac{da}{a^2} \int db F\left(\lambda \frac{b}{a}\right) = \frac{2}{\lambda} \log a_0 \cdot \int dx F(x). \quad (2.2.16)$$

Inserting (2.2.15) and (2.2.16) into (2.2.11), and taking the limit for $\lambda \to 0$ leads to

$$A \le \frac{\pi}{b_0 \log a_0} \int dy\, |y|^{-1} |\hat{\phi}(y)|^2 \le B. \quad (2.2.17)$$

These inequalities hold for *any* frame of wavelets ϕ_{mn} Incidentally, (2.2.17) shows that the basic function ϕ for a frame of wavelets must satisfy the same "admissibility condition," i.e., $\int dy |y|^{-1} |\hat{\phi}(y)|^2 < \infty$, as the functions from which continuously labeled affine coherent states are constructed (see Section I-D). In particular, if the ϕ_{mn} constitute a tight frame of wavelets, then

$$A = \frac{\pi}{b_0 \log a_0} \int dy\, |y|^{-1} |\hat{\phi}(y)|^2.$$

In the explicitly constructed tight frames in Section II-B-1b, two functions (either $h^{\pm}$, with $h^+ = h$, $h^- = h^*$, or $h^{(\lambda)}$, $\lambda = 1,2$, with $h^{(1)} = \mathrm{Re}\, h$, $h^{(2)} = \mathrm{Im}\, h$) were needed to construct tight frames of wavelets (i.e., the $h^{\pm}_{mn}$, or the $h^{(\lambda)}_{mn}$, $\lambda = 1,2$). For such frames the arguments lead to the frame constant

$$A = \frac{2\pi}{b_0 \log a_0} \int dy\, |y|^{-1} |\hat{h}(y)|^2.$$

This agrees with the value $A = 2\pi / b_0 \log a_0$ obtained in Section II-B-1b (see (2.2.5)), since we chose the normalization of h such that $\int dy\,|y|^{-1}|\hat{h}(y)|^2 = 1$.

C. General (Not Necessarily Tight) Frames in $L^2(\mathbb{R})$—Ranges for the Lattice Spacings—Frame Bounds

As we already explained in the introduction, it may be necessary in some applications to resort to nontight frames. This can be the case if the basic function g or h is imposed *a priori* (because of its adaptation to the problem at hand), or in the case of wavelets, when the explicit examples of functions h leading to tight frames are too spread out.

In this section we treat the following questions:

1) Is there a range of parameters that is excluded *a priori* (i.e., independently of the choice of g or h)?
2) Given g (or h), determine a range R for the parameters p_0, q_0 (resp. a_0, b_0) such that if $(p_0, q_0) \in R$ (resp. $(a_0, b_0) \in R$), then the associated g_{mn} (resp. h_{mn}) constitute a frame.
3) For g, p_0, q_0 (or h, a_0, b_0) chosen as in 2), compute estimates for the frame bounds A and B.

In order to interrupt the flow of the exposition as little as possible, we relegate all the technical proofs for this section to Appendix C.

1) The Weyl–Heisenberg Case:

a) Critical value for the product $p_0 \cdot q_0$: In the Weyl–Heisenberg case there exists a critical value, 2π, for the product $p_0 \cdot q_0$. This is already illustrated by the construction in Section II-A-1a), which only works if $p_0 \cdot q_0 < 2\pi$. The following theorem states that at the critical value $p_0 \cdot q_0 = 2\pi$, only functions g that are either not very smooth or do not decay very fast can give rise to a frame.

Theorem 2.3 (Balian–Low–Coifman–Semmes): Choose $g \in L^2(\mathbb{R})$, $p_0 > 0$. If the g_{mn} associated with $g, p_0, q_0 = 2\pi / p_0$, constitute a frame, then either $xg \notin L^2$ or $g' \notin L^2$.

Remark: This theorem was first published by R. Balian [51] for the case where the g_{mn} are an orthonormal basis. In the 1985 Festschrift for the 60th birthday of the physicist G. Chew, F. Low also discusses this problem [52]. He gives (independently) essentially the same proof as in [51]. The proof presented in [51], [52] contains a technical gap that was filled by R. Coifman and S. Semmes. The proof extends easily from the basis case to the frame case. We give here the proof as completed by R. Coifman and S. Semmes.

The proof of Theorem 2.3 uses the Zak transform U_Z. This transform maps $L^2(\mathbb{R})$ unitarily onto $L^2([0,1]^2)$; it was first systematically studied by J. Zak [54], in connection with solid state physics. Some of its properties were known long before Zak's work, however. In [55] the same transform is called the Weil–Brezin map, and it is claimed that the same transform was already known to Gauss. It was also used by Gel'fand (see, e.g., Ch. XIII in [56]). J. Zak seems, however, to have been the first to recognize it as the versatile tool it is and to have studied it systematically. It has many very interesting properties; its applications range from solid state physics to the derivation, in [57], of new relationships between Jacobi's theta functions. Before embarking on the proof of Theorem 2.3, we briefly review the definition and some of the properties of U_Z. The Zak transform U_Z is defined by

$$(U_Z f)(t,s) = \lambda^{1/2} \sum_{l \in \mathbb{Z}} e^{2\pi i t l} f(\lambda(s-l)) \tag{2.3.1}$$

where the parameter $\lambda > 0$ can be adjusted to the problem at hand. For the proof of Theorem 2.3 we shall take $\lambda = q_0$.

Strictly speaking, the definition (2.3.1) does only make sense for the subspace of the L^2-functions for which the series converges. It is, however, easy to extend (2.3.1) from those functions for which it is well-defined, to all of $L^2(\mathbb{R})$. One way of doing this is to observe that the images under (2.3.1) of the orthonormal basis e_{mn} of $L^2(\mathbb{R})$,

$$e_{mn}(x) = \lambda^{-1/2} e^{2\pi i m x/\lambda} \chi_{[0,\lambda)}(x - n\lambda) (m, n \in \mathbb{Z})$$

are well-defined, and constitute again an orthonormal basis of $L^2([0,1]^2)$,

$$(U_Z e_{mn})(t,s) = e^{-2\pi i t n} e^{2\pi i m s}.$$

It follows that (2.3.1) defines a unitary map from $L^2(\mathbb{R})$ to $L^2([0,1]^2)$. On the other hand, (2.3.1) can also be extended to values of (t,s) outside $[0,1]^2$. For $f \in L^2(\mathbb{R})$, the resulting function is in $L^2_{loc}(\mathbb{R}^2)$, and satisfies

$$\begin{aligned} (U_Z f)(t+1,s) &= (U_Z f)(t,s) \\ (U_Z f)(t,s+1) &= e^{2\pi i t}(U_Z f)(t,s) \end{aligned} \tag{2.3.2}$$

almost everywhere (a.e.) with respect to Lebesgue measure on $\mathbb{R}^2$. A rather remarkable consequence of (2.3.2) is the following. Suppose that f is such that $U_Z f$ is continuous (on $\mathbb{R}^2$). Then $U_Z f$ must necessarily have a zero in $[0,1]^2$. The proof of this fact, first pointed out in [58], is quite simple. The presentation given here is borrowed from [59]. If $U_Z f$ had no zero in $[0,1]^2$, then $\log U_Z f$ would be a univalued, continuous function on $[0,1]^2$ extending to a continuous function in $\mathbb{R}^2$ by (2.3.2). On the other hand, (2.3.2) implies that

$$\begin{aligned} (\log U_Z f)(t+1,s) &= (\log U_Z f)(t,s) + 2\pi i k \\ (\log U_Z f)(t,s+1) &= (\log U_Z f)(t,s) + 2\pi i t + 2\pi i l \end{aligned} \tag{2.3.3}$$

where k, l are integers, independent of t, s because of the continuity of $\log U_Z f$. But (2.3.3) leads immediately to the following contradiction:

$$\begin{aligned} 0 &= (\log U_Z f)(0,0) - (\log U_Z f)(1,0) \\ &\quad + (\log U_Z f)(1,0) - (\log U_Z f)(1,1) \\ &\quad + (\log U_Z f)(1,1) - (\log U_Z f)(0,1) \\ &\quad + (\log U_Z f)(0,1) - (\log U_Z f)(0,0) \\ &= -2\pi i k - 2\pi i - 2\pi i l + 2\pi i k + 2\pi i l = -2\pi i \neq 0. \end{aligned}$$

This proves that we started from a false premise, i.e., that $U_Z f$ has a zero in $[0,1]^2$. A similar argument proves Theorem 2.3. The above is only one of the many properties of the Zak transform. For more of these properties, and interesting applications of the Zak transform to signal analysis, we urge the reader to consult [60].

The images, under the Zak transform, of functions g_{mn} (with $p_0 \cdot q_0 = 2\pi$), are remarkably simple. One easily checks that, for the choice $\lambda = q_0$,

$$(U_Z g_{mn})(t,s) = e^{2\pi i m s} e^{-2\pi i t n} (U_Z g)(t,s). \quad (2.3.4)$$

An immediate consequence of (2.3.4) is

$$\begin{aligned}\sum_{m,n} |\langle g_{mn}, f\rangle|^2 &= \sum_{m,n} |\langle U_Z g_{mn}, U_Z f\rangle|^2 \\ &= \sum_{m,n} \left| \int_0^1 dt \int_0^1 ds\, e^{-2\pi i m s} e^{2\pi i t n} \right. \\ &\qquad \left. \cdot (U_Z g)^*(t,s)(U_Z f)(t,s) \right|^2 \\ &= \int_0^1 dt \int_0^1 ds\, |(U_Z g)(t,s)|^2 |(U_Z f)(t,s)|^2.\end{aligned}$$

It follows that the g_{mn} constitute a frame, with frame bounds A, B, if and only if, for $(t,s) \in [0,1]^2$ a.e. (and hence, by (2.3.2), for $(t,s) \in \mathbb{R}^2$ a.e.)

$$A^{1/2} \le |(U_Z g)(t,s)| \le B^{1/2}.$$

This is another crucial ingredient of the proof of Theorem 2.3, to which we now turn.

Proof of Theorem 2.3: Suppose that the g_{mn} constitute a frame, and assume also that xg, $g' \in L^2(\mathbb{R})$. We want to show that this leads to a contradiction. Define $G(t,s) = (U_Z g)(t,s)$. Since the g_{mn} constitute a frame, we have

$$a \le |G(t,s)| \le b \quad (2.3.5)$$

for some $a > 0$, $b < \infty$, and for $(t,s) \in \mathbb{R}^2$, a.e. For compactly supported f one finds

$$\begin{aligned}&[U_Z(xf)](t,s) \\ &\quad = q_0^{3/2} s (U_Z f)(t,s) - q_0^{3/2} \sum_l e^{2\pi i t l} l f(q_0(s-l)) \\ &\quad = q_0^{3/2} s (U_Z f)(t,s) - q_0^{3/2} (2\pi i)^{-1} \partial_t (U_Z f)(t,s).\end{aligned}$$

This shows that $xg \in L^2$ implies $\partial_t G = \partial_t(U_Z g) \in L^2_{loc}(\mathbb{R}^2)$. Similarly $g' \in L^2$ implies $\partial_s G \in L^2_{loc}(\mathbb{R}^2)$.

If the square integrability of the partial derivatives of G implied that G was continuous (which it does not, since we are in more than one dimension), then the proof would be finished. By the argument given before the proof, $\inf |G|$ would then be zero, which is in contradiction with (2.3.5). This is essentially the argument of Balian in [51], where the implicit assumption that G is continuous seems to be made. Lemma 2.4, due to R. Coifman and S. Semmes, which we state below, shows how the boundary conditions (2.3.2), together with the bounds (2.3.5), lead to a contradiction, without assuming continuity for G. The main idea is to use an averaged version of G. This averaged function is automatically continuous, and, if the averaging is done on a small enough scale, close enough to G so that the properties inherited from (2.3.5) and (2.3.2) still lead to a contradiction. This then proves Theorem 2.3. □

Lemma 2.4: Assume that G is a bounded function on $\mathbb{R}^2$ that is locally square integrable and which satisfies

$$G(t+1,s) = G(t,s)$$
$$G(t,s+1) = e^{2\pi i t} G(t,s).$$

If both $\partial_t G$ and $\partial_s G$ are locally square integrable, then $\operatorname{ess\,inf}_{0 \le t,s \le 1} |G(t,s)| = 0$.

Here the "essential infimum" of a measurable function f is defined by

$$\operatorname{ess\,inf} f = \inf\{\lambda; |\{x; f(x) \le \lambda\}| > 0\}$$

where $|V|$ denotes the Lebesgue measure of the set V. This definition avoids values taken by f on sets of measure zero. For instance, for $f(x) = 1$ for $x \ne 0$, $f(0) = 0$, one has $\inf f = 0$, but $\operatorname{ess\,inf} f = 1$.

Proof of Lemma 2.4: This proof is rather technical; it is given in Appendix B. □

Note added in proof: After this paper was written, G. Battle produced a very elegant new proof of Theorem 2.3 which avoids the use of the Zak transform [69]. Battle's paper only treats the case where the g_{mn} are an orthonormal basis (as did the original papers by Balian and Low). His argument was extended to frames by A. J. E. M. Janssen and I. Daubechies [70].

For the critical value $p_0 \cdot q_0 = 2\pi$ we see thus that not much regularity and/or decay can be expected from a function g leading to a frame. This is in marked contrast with the case $p_0 \cdot q_0 < 2\pi$. In that case, as shown in Section II-B-1a, there even exist C^∞-functions g with compact support such that the associated g_{mn} constitute a tight frame.

This critical value $p_0 \cdot q_0 = 2\pi$, has a physical meaning. As shown by Fig. 1(a), $(p_0 \cdot q_0)^{-1}$ is the density, in phase space, of the discrete lattice of functions g_{mn}. The density $(2\pi)^{-1}$ is nothing but the Nyquist density; it is well-known in information theory that time-frequency densities at least as high as the Nyquist density are needed for a full transmission of information. It is therefore not surprising to encounter this same critical value here. One encounters the same argument in quantum physics, often used in semiclassical approximations, where a complete set of independent states (i.e., a set of linearly independent functions in $L^2(\mathbb{R})$ whose linear combinations span a dense subspace in $L^2(\mathbb{R})$) heuristically corresponds to a density $(2\pi)^{-1}$ in phase space. In other words, every state "occupies" a "cell" of area 2π in phase space.

For $p_0 \cdot q_0 > 2\pi$, the intuition from physics or information theory suggests that the associated phase space lattice is "too loose," i.e., that the g_{mn} cannot span the whole Hilbert space. More precisely, we expect that for *any* $g \in L^2(\mathbb{R})$, there exists at least one $f \in L^2(\mathbb{R})$, $f \neq 0$, such that $\langle g_{mn}, f\rangle = 0$ for all $m,n \in \mathbb{Z}$. This is indeed the case.

If $p_0 \cdot q_0 / 2\pi > 1$ is rational, then the following argument, again using the Zak transform, shows how to construct a function f.

By a dilation argument we can restrict ourselves to the case $p_0 = 2\pi$, $q_0 = K/L$, with $K, L \in \mathbb{N}$, $K > L > 0$. Let F, G, G_{mn} be the Zak transforms of respectively f, g, g_{mn}, as defined by (2.3.1), where we take $\lambda = 1$. Then

$$G_{mn}(t,s) = e^{2\pi i m s} G\left(t, s - n\frac{K}{L}\right).$$

For $n = kL + l$, with $l, k \in \mathbb{Z}$, $0 \le l < L$, this reduces to

$$G_{mn}(t,s) = e^{2\pi i m s} e^{-2\pi i k K t} G\left(t, s - \frac{l}{L}K\right)$$

where we have used (2.3.2). The function F will be orthogonal in $L^2([0,1]^2)$ to each G_{mn} if, for all $s \in [0,1]$ and for all $k, l \in \mathbb{Z}$ with $0 \le l < L$,

$$\int_0^1 dt\, \overline{F(t,s)} e^{-2\pi i k K t} G\left(t, s - \frac{l}{L}K\right) = 0. \quad (2.3.6)$$

We can rewrite (2.3.6) as

$$\int_0^{1/K} dt\, e^{-2\pi i k K t} \sum_{m=0}^{K-1} G\left(t + \frac{m}{K}, s - \frac{l}{L}K\right) \overline{F\left(t + \frac{m}{K}, s\right)} = 0. \quad (2.3.7)$$

We are thus led to the linear system of equations

$$\sum_{m=0}^{K-1} A_{lm}(t,s)\phi_m(t,s) = 0 \qquad 0 \le l < L \quad (2.3.8)$$

where

$$A_{lm}(t,s) = G\left(t + \frac{m}{K}, s - \frac{l}{L}K\right) \qquad (2.3.9)$$

$$\phi_m(t,s) = \overline{F\left(t + \frac{m}{K}, s\right)}. \qquad (2.3.10)$$

Since the system (2.3.8) has L equations for $K > L$ unknowns, it always has a nonzero solution, for every pair $(t,s) \in [0, 1/K) \times [0,1]$. The $\phi_m(t,s)$ solving (2.3.8) can, moreover, be chosen in $L^2([0,1/K) \times [0,1])$. One way of doing this is to choose a fixed $u \in \mathbb{C}^K$, and define $\phi_m(t,s) = \lim_{\tau \to \infty} u_m(t,s;\tau)$, where

$$u(t,s;\tau) = \exp\left[-\tau A^*(t,s)A(t,s)\right]u.$$

Clearly $\|u(t,s;\tau)\|^2 \le \|u\|^2$; hence $\sum_{m=1}^K |\phi_m(t,s)|^2 \le \|u\|^2$. On the other hand, since all the A_{lm} are in $L^2([0,1/K) \times [0,1])$, the $u(t,s;\tau)$ are clearly measurable in t,s; as pointwise limits of measurable functions, the ϕ_m are measurable too. Putting the ϕ_m together according to (2.3.10) then defines a function F in $L^2([0,1]^2)$ which is orthogonal to all the G_{mn}. Note that while F may be zero a.e. for some choices of u, it is impossible that the functions $F(u_k)$ associated with K linearly independent vectors $u_1, \cdots, u_K$ in $\mathbb{C}^K$ all be zero a.e., since this would mean that $A^*(t,s)A(t,s) > 0$ a.e. in t,s, which contradicts rank $[A(t,s)] < K$. For appropriately chosen $u \in \mathbb{C}^K$, the previous construction leads thus to a nontrivial function $F \in L^2([0,1]^2)$ orthogonal to all the G_{mn}.

This argument does not work if $p_0 \cdot q_0 / 2\pi$ is irrational. It is nevertheless still true that the g_{mn} do not span $L^2(\mathbb{R})$, whatever g in $L^2(\mathbb{R})$ is chosen, even for irrational values of $p_0 \cdot q_0 / 2\pi$. The only proof that I know of this fact uses von Neumann algebras; it was pointed out to me by R. Howe and T. Steger. The proof consists in the computation of the coupling constant of the von Neumann algebra spanned by the Weyl operators $(W(mp_0, nq_0);\ m,n \in \mathbb{Z})$. The coupling constant for this von Neumann algebra was computed explicitly by M. Rieffel [49]; for $p_0 \cdot q_0 > 2\pi$ it is larger than 1, which implies that the von Neumann algebra has no cyclic element. This means that for any $g \in L^2(\mathbb{R})$, the closed linear span of the g_{mn} is a proper subspace of $L^2(\mathbb{R})$, which was the desired result. Unfortunately, this proof does not seem very illuminating from the signal analyst's point of view.

Note added in proof: Recently H. Landau [71] found a different, intuitively much more appealing argument to prove that the g_{mn} cannot constitute a frame if $p_0 \cdot q_0 > 2\pi$. His proof works for all g which are reasonably "nice" (decaying in both time and frequency).

This concludes what we have to say about the critical value $(p_0 \cdot q_0 = 2\pi)$ in the Weyl–Heisenberg case. We shall always assume $p_0 \cdot q_0 \le 2\pi$ in what follows.

b) Ranges for the parameters p_0, q_0, and estimates for the frame bounds: In the preceding subsection we have excluded parameters p_0, q_0 for which $p_0 \cdot q_0 > 2\pi$. Even if $p_0 \cdot q_0 \le 2\pi$, however, we do not automatically have a frame for arbitrary functions g. When, for example, $g(x) = 1$ for $0 \le x \le 1$, and $g(x) = 0$, otherwise, the g_{mn} cannot constitute a frame if $q_0 > 1$, even if $p_0 \cdot q_0 < 2\pi$. Indeed, for $q_0 > 1$, one finds that $\langle g_{mn}, f\rangle = 0$, for all $m,n \in \mathbb{Z}$ if the support of $f \subset [1, q_0]$, independently of the choice of p_0. This is thus a case where an inappropriate choice of q_0 excludes the possibility of a frame, for all values of p_0.

The theorem below gives sufficient conditions on g and q_0 under which this cannot happen, i.e., there always exist some $p_0 > 0$ (in fact, a whole interval) leading to a frame.

Theorem 2.5: If

1) $$m(g;q_0) = \operatorname*{ess\,inf}_{x \in [0,q_0]} \sum_n |g(x - nq_0)|^2 > 0 \qquad (2.3.11)$$

2) $$M(g;q_0) = \operatorname*{ess\,sup}_{x \in [0,q_0]} \sum_n |g(x - nq_0)|^2 < \infty \qquad (2.3.12)$$

and

3) $$\sup_{s \in \mathbb{R}} \left[(1+s^2)^{(1+\epsilon)/2} \beta(s)\right] = C_\epsilon < \infty \qquad \text{for some } \epsilon > 0$$

where

$$\beta(s) = \sup_{x \in [0,q_0]} \sum_{n \in \mathbb{Z}} |g(x - nq_0)|\,|g(x + s - nq_0)| \qquad (2.3.13)$$

then there exists a $P_0^c > 0$ such that

$\forall p_0 \in (0, P_0^c)$: the g_{mn} associated with g, p_0, q_0 are a frame

$\forall \delta > 0 : \exists p_0$ in $[P_0^c, P_0^c + \delta]$ such that the g_{mn} associated to g, p_0, q_0 are not a frame.

Proof: see Appendix C. □

The conditions (2.3.11)–(2.3.13) may seem rather technical. They are, in fact, extremely reasonable. Condition (2.3.11) specifies that the collection of g and its translates should not have any "gaps." This already excludes the example given at the start of this subsection. The conditions (2.3.12), (2.3.13) are satisfied if g has sufficient decay at ∞, in particular, if $|g(x)| \le C[1 + x^2]^{-3/2}$.

Note that both (2.3.11) and (2.3.12) are necessary conditions. If (2.3.11) is not satisfied, then for every $\epsilon > 0$ one can find a nonzero f in $L^2(\mathbb{R})$ such that

$$\sum_{m,n} |\langle g_{mn}, f \rangle|^2 \le \epsilon \|f\|^2$$

which means there exists no nonzero lower frame bound A for the g_{mn}. Similarly there exists no finite upper frame bound B for the g_{mn} if (2.3.12) is not satisfied.

Remarks

1) At the end of the preceding subsection we showed that the g_{mn} can constitute a frame only if $p_0 \cdot q_0 \le 2\pi$. Hence necessarily $P_0^c \le 2\pi / q_0$.
2) The set $\{p_0$; the g_{mn} associated to g, p_0, q_0 constitute a frame$\}$ with g and q_0 fixed need not be connected. It is possible that this set contains values of p_0 larger than P_0^c. An example is given by the following construction. Let ϕ be a C^∞ function with support $[0, 1/3]$ such that $|\phi|$ has no zeros in $(0, 1/3)$. Define g by

$$\hat{g}(y) = \begin{cases} 0, & y \le 0 \\ \phi(y), & 0 \le y \le 1/3 \\ \phi(y - 1/3), & 1/3 \le y \le 2/3 \\ \phi(y - 2/3), & 2/3 \le y \le 1 \\ 0, & y \ge 1. \end{cases}$$

Take $q_0 = 2\pi$. Then, since support $\hat{g} = [0, 1]$,

$$\sum_{m,n} |\langle g_{mn}, f \rangle|^2 = \int dy \sum_m |\hat{g}(y - mp_0)|^2 |\hat{f}(y)|^2.$$

This implies that the g_{mn} constitute a frame if and only if $\inf \sum_m |\hat{g}(y - mp_0)|^2 > 0$. Consequently $P_0^c = 1/3$, while the set of all p_0 leading to a frame is $(0, 1/3) \cup (1/3, 2/3)$.

For reasonably smooth g with sufficient decay at ∞, the constants $m(g; q_0), M(g; q_0)$ and the function $\beta(s)$ can easily be computed numerically. These constants are useful in estimations of the frame bounds A, B, as the following theorem shows.

Theorem 2.6: Assume that (2.3.11), (2.3.12), (2.3.13) are satisfied. Define

$$p_0^c = \inf \left\{ p_0 \,\middle|\, \sum_{k=1}^{\infty} \left[\beta\left(\frac{2\pi}{p_0} k\right) \beta\left(-\frac{2\pi}{p_0} k\right) \right]^{1/2} \ge \frac{1}{2} m(g; q_0) \right\}.$$

Then $p_0^c \le P_0^c$, and for $0 < p_0 < p_0^c$, the following estimates for the frame bounds A, B holds

$$A \ge \frac{2\pi}{p_0} \left(m(g; q_0) - 2 \sum_{k=1}^{\infty} \left[\beta\left(\frac{2\pi}{p_0} k\right) \beta\left(-\frac{2\pi}{p_0} k\right) \right]^{1/2} \right)$$

$$B \le \frac{2\pi}{p_0} \left(M(g; q_0) + 2 \sum_{k=1}^{\infty} \left[\beta\left(\frac{2\pi}{p_0} k\right) \beta\left(-\frac{2\pi}{p_0} k\right) \right]^{1/2} \right).$$

Proof: see Appendix C. □

These bounds for A and B can be improved by the following observation. If the $g_{mn}(x) = e^{imp_0x} g(x - nq_0)$ constitute a frame, then so do the

$$(g_{mn})^\wedge(\zeta) = e^{-inq_0\zeta} \hat{g}(\zeta - mp_0),$$

where $\hat{g}$ denotes the Fourier transform of g. It follows that

$$A \ge \max \left(\frac{2\pi}{p_0} \left\{ m(g; q_0) - 2 \sum_{k=1}^{\infty} \left[\beta\left(\frac{2\pi}{p_0} k\right) \beta\left(-\frac{2\pi}{p_0} k\right) \right]^{1/2} \right\}, \frac{2\pi}{q_0} \left\{ m(\hat{g}; p_0) - 2 \sum_{k=1}^{\infty} \left[\hat{\beta}\left(\frac{2\pi}{q_0} k\right) \hat{\beta}\left(-\frac{2\pi}{q_0} k\right) \right]^{1/2} \right\} \right) \tag{2.3.14}$$

$$B \le \min \left(\frac{2\pi}{p_0} \left\{ M(g; q_0) + 2 \sum_{k=1}^{\infty} \left[\beta\left(\frac{2\pi}{p_0} k\right) \beta\left(-\frac{2\pi}{p_0} k\right) \right]^{1/2} \right\} \frac{2\pi}{q_0} \left\{ M(\hat{g}; p_0) + 2 \sum_{k=1}^{\infty} \left[\hat{\beta}\left(\frac{2\pi}{q_0} k\right) \hat{\beta}\left(-\frac{2\pi}{q_0} k\right) \right]^{1/2} \right\} \right). \tag{2.3.15}$$

Here $M(\hat{g}; p_0)$, $m(\hat{g}; p_0)$ and $\hat{\beta}(s)$ are the obvious extensions of the quantities in (2.3.11), (2.3.12), and (2.3.13). For instance,

$$\hat{\beta}(s) = \sup_{\xi \in [0, p_0]} \sum_{m \in \mathbb{Z}} |\hat{g}(\xi + mp_0)| \, |\hat{g}(\xi + s + mp_0)|.$$

Similarly, one has the following better lower bound for P_0^c,

$$P_0^c \ge p_0^c = \inf \left\{ p_0 \,\middle|\, 2 \sum_{k=1}^{\infty} \left[\beta\left(\frac{2\pi}{p_0} k\right) \beta\left(-\frac{2\pi}{p_0} k\right) \right]^{1/2} \ge m(g, q_0) \text{ and } 2 \sum_{k=1}^{\infty} \left[\hat{\beta}\left(\frac{2\pi}{q_0} k\right) \hat{\beta}\left(-\frac{2\pi}{q_0} k\right) \right]^{1/2} \ge m(\hat{g}; p_0) \right\}. \tag{2.3.16}$$

C) Examples

i) The Gaussian case

In this case

$$g(x) = \pi^{-1/4} e^{-x^2/2}. \tag{2.3.17}$$

This is the basic function for the so-called "canonical coherent states" in physics [5]. It is also the basic function chosen by Gabor [1] in the definition of his expansion. In the notations used in this paper, Gabor's approach amounts to writing an expansion with respect to the g_{mn} associated to g, p_0, q_0, where $p_0 \cdot q_0 = 2\pi$. This choice seems very natural from the point of view of information theory, since it corresponds exactly to the Nyquist density. However, since both xg and g' are square integrable in this case, Theorem 2.3 tells us that the g_{mn} cannot possibly constitute a frame. In fact, the Zak transform $U_Z g$ of g (see Section II-C-1a) can be constructed explicitly in this case; it is one of Jacobi's theta-functions, and it

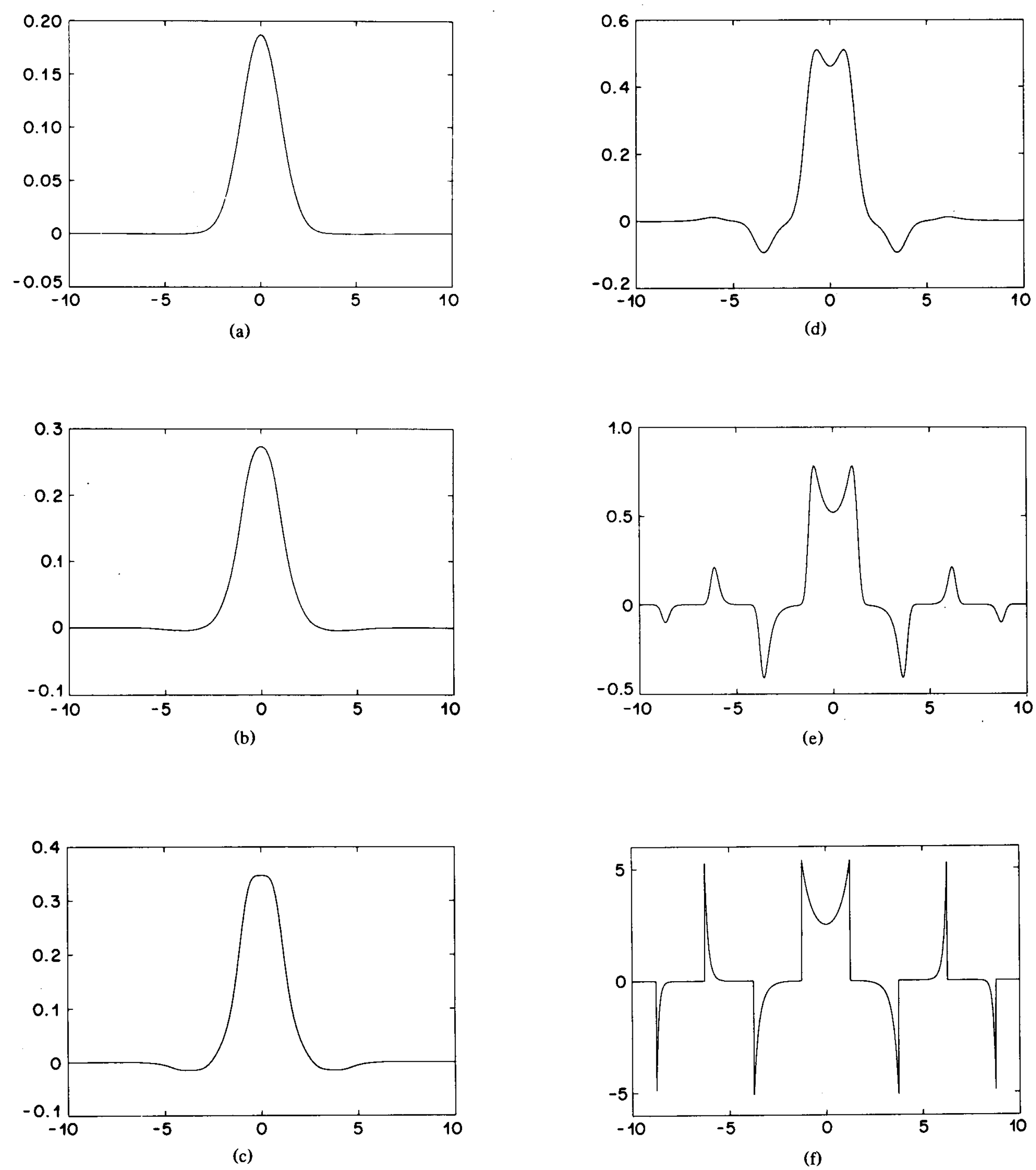

Fig. 4. Function $\tilde{g}$ for different values of $p_0 \cdot q_0$. We have taken $g(x) = \pi^{-1/4} \exp(-x^2/2)$, and $p_0 = q_0$ in every case. (a) $p_0 \cdot q_0 = \pi/2$. (b) $p_0 \cdot q_0 = 3\pi/4$. (c) $p_0 \cdot q_0 = \pi$. (d) $p_0 \cdot q_0 = 3\pi/2$. (e) $p_0 \cdot q_0 = 1.9\pi$. (f) $p_0 \cdot q_0 = 2\pi$. In this last case $\mathbb{T}$ does not have bounded inverse, and $\tilde{g} \notin L^2$. This function $\tilde{g}$ was first computed by Bastiaans [2].

has a zero in $(1/2, 1/2)$ [2], [3], [60], [63]. The operator $\mathbb{T}$ of Section II-A is unitarily equivalent with multiplication by $|U_Z g|^2$ on $L^2([0,1]^2)$, and is therefore one-to-one, but it has an unbounded inverse, i.e.,

$$\inf_{f \in L^2(\mathbb{R})} \|f\|^{-2} \sum_{m,n} |\langle h_{mn}, f\rangle|^2 = 0.$$

One can still write an expansion of type

$$f = \sum_{m,n} \tilde{g}_{mn} \langle g_{mn}, f\rangle$$

but this expansion has, in general, bad convergence properties because $\tilde{g} \notin L^2(\mathbb{R})$ (since $U_Z \tilde{g} = U_Z(\mathbb{T}^{-1} g) = |U_Z g|^{-2} U_Z g$ has a pole at $(1/2, 1/2)$, and is therefore not in $L^2([0,1]^2)$). It is this fact that makes the expansion associated to the Gabor wave functions unstable [2], [3], [60]. The function $\tilde{g}$ was explicitly constructed in [2]; it turns out to be discontinuous as well as nonsquare-integrable (see Fig. 4(f)). Explicit computation of $\tilde{g}$ (via the Zak transform), for $p_0 = q_0 = \sqrt{2\pi}$, leads to [2]

$$\tilde{g}(x) = C e^{x^2/2} \sum_{n+1/2 \geq x/\sqrt{2\pi}} (-1)^n e^{-\pi(n+1/2)^2}$$

where C is a normalization constant.

For g Gaussian, as in (2.3.17), the fact that the g_{mn} do not span $L^2(\mathbb{R})$ if $p_0 \cdot q_0 > 2\pi$, has been known for quite a while [61], [62]. The proofs in [61], [62] use entire function techniques. These techniques do not work, however, for non-Gaussian g.

Our first table of numerical results, for Gaussian g, lists p_0^c, for different values of q_0 (see Table I). It turns out that p_0^c is always very close to $2\pi/q_0$; for this reason we have tabulated $(2\pi/q_0 \cdot p_0^c) - 1$, rather than p_0^c itself. The difference $(2\pi/q_0 \cdot p_0^c) - 1$ is largest for $q_0 = \sqrt{2\pi}$, where it is about 4×10^{-3}.

TABLE I*

q_0	$(2\pi/q_0 \cdot p_0^c) - 1$
1.0	6×10^{-8}
1.5	3×10^{-7}
2.0	2×10^{-4}
2.5	4×10^{-3}
3.0	4×10^{-4}
3.5	2×10^{-5}
4.0	8×10^{-7}

*The deviation, for $g(x) = \pi^{-1/4} \exp(-x^2/2)$, of the estimated value p_0^c from the optimal value $2\pi/q_0$ (see text).

Note that $P_0^c \leq 2\pi/q_0$. The numerical results show therefore that in this case the estimate p_0^c is remarkably close to the true critical value P_0^c. In fact they suggest the conjecture that $P_0^c = 2\pi/q_0$, for all $q_0 > 0$, in this case. I believe that this conjecture holds for all positive functions g with positive Fourier transform.

Next we list the estimates (2.3.14) and (2.3.15) on the frame bounds for a few values of q_0, $p_0 < p_0^c$. We also tabulate the corresponding B/A, and $r = (B/A - 1)/(B/A + 1)$. This parameter measures how snug the frame is, i.e., how close it is to a tight frame; as explained in Section II-A, this parameter is essential if one wants to apply the inversion formula as indicated in Section II-A. We have grouped together different values of q_0, p_0 corresponding to the same value of $q_0 \cdot p_0$. If $q_0 \cdot p_0 = 2\pi/k$, with k integer, then a different method, based on the Zak transform, enables us to write explicit, exact expressions for the frame bounds A and B. With U_Z defined as in Section II-C-1a (see (2.3.1), with $\lambda = q_0$), one finds indeed that, for $m \in \mathbb{Z}$, $m = m'k + r, m'$, $r \in \mathbb{Z}$, $0 \leq r < k$,

$$\begin{aligned}(U_Z g_{mn})(t,s) &= (U_Z g_{m'k+rn})(t,s) \\ &= e^{2\pi i r n/k} e^{-2\pi i t n} e^{2\pi i m' s} (U_Z g^r)(t,s)\end{aligned}$$

where we have used $p_0 = 2\pi/q_0$, and where

$$g^r(x) = e^{2\pi i r x q_0/k} g(x).$$

Hence

$$(U_Z g^r)(t,s) = (U_Z g)\left(t - \frac{r}{k} s\right).$$

This is entirely similar to computations made in Section II-C-1a; see also [45]. Consequently, as in Section II-C-1a (see also [16], [45])

$$\begin{aligned}&\sum_{m,n} |\langle g_{mn}, f\rangle|^2 \\ &= \int_0^1 dt \int_0^1 ds \left[\sum_{r=0}^{k-1} \left|(U_Z g)\left(t - \frac{r}{k}, s\right)\right|^2\right] |(U_Z f)(t,s)|^2\end{aligned}$$

which implies

$$A = \inf \left[\sum_{r=0}^{k-1} \left|(U_Z g)\left(t - \frac{r}{k}, s\right)\right|^2\right]$$

$$B = \sup \left[\sum_{r=0}^{k-1} \left|(U_Z g)\left(t - \frac{r}{k}, s\right)\right|^2\right].$$

To find A and B, in the case where $2\pi/(p_0 \cdot q_0)$ is an integer, it suffices therefore to compute $U_Z g$ and these two extrema.

In Table II next we list the estimates, given by (2.3.14), (2.3.15) for A, B, B/A and r, for a few values of q_0, in the cases $p_0 \cdot q_0 = \pi/2$, $3\pi/4$, π, $3\pi/2$, and 1.9π. In the cases $p_0 \cdot q_0 = \pi/2$ and π we also list the exact values for A, B, denoted by $A_{\text{exact}}, B_{\text{exact}}$. For values of p_0 close to the critical value $2\pi/q_0$ (see the case $q_0 \cdot p_0 = 1.9\pi$ in Table I), the ratio B/A becomes very large, as was to be expected. The convergence of the formula for $\tilde{g}$, measured by r, will be very slow. For $q_0 = 2.0$, $p_0 = \pi/q_0 = \pi/2$, the ratio r is already of the order of 0.2, while $q_0 = 1.5$, $p_0 = \pi/3$, $q_0 p_0 = \pi/2$ leads to $r = 0.025$. In the latter case a few terms will suffice to obtain an accuracy of 10^{-6} is the computation of $\tilde{g}$. This shows that very good frames can be obtained with lattice spacings that are not very small.

In the two cases where the exact values of A, B can be computed by other means ($q_0 \cdot p_0 = \pi/2$ and $q_0 \cdot p_0 = \pi$), the estimates for A and B, as calculated from (2.3.14) and

TABLE II
VALUES FOR THE FRAME BOUNDS A, B, THEIR RATIO B/A, AND THE CONVERGENCE FACTOR $r=(B/A-1)/(B/A+1)$, FOR THE CASE $g(x)=\pi^{-1/4}\exp(-x^2/2)$, FOR DIFFERENT VALUES OF q_0, p_0*

q_0	Aa	A_{exact}	$q_0\cdot p_0=\pi/2$ B	B_{exact}	B/A	$r=(B-A)/(B+A)$
0.5	1.203	1.221	7.091	7.091	5.896	0.710
1.0	3.853	3.854	4.147	4.147	1.076	0.037
1.5	3.899	3.899	4.101	4.101	1.052	0.025
2.0	3.322	3.322	4.679	4.679	1.408	0.170
2.5	2.365	2.365	5.664	5.664	2.395	0.411
3.0	1.427	1.427	6.772	6.772	4.745	0.652

q_0	A	B	$q_0\cdot p_0=3\pi/4$ B/A	$r=(B-A)/(B+A)$
1.0	1.769	3.573	2.019	0.338
1.5	2.500	2.833	1.133	0.062
2.0	2.210	3.124	1.414	0.172
2.5	1.577	3.776	2.395	0.411
3.0	0.951	4.515	4.745	0.652

q_0	A	A_{exact}	B	$q_0\cdot p_0=\pi$ B_{exact}	B/A	$r=(B-A)/(B+A)$
1.0	0.601	0.601	3.546	3.546	5.901	0.710
1.5	1.519	1.540	2.482	2.482	1.635	0.241
2.0	1.575	1.600	2.425	2.425	1.539	0.212
2.5	1.172	1.178	2.843	2.843	2.426	0.416
3.0	0.713	0.713	3.387	3.387	4.752	0.652

q_0	A	B	$q_0\cdot p_0=3\pi/2$ B/A	$r=(B-A)/(B+A)$
1.0	0.027	3.545	130.583	0.985
1.5	0.342	2.422	7.082	0.753
2.0	0.582	2.089	3.592	0.564
2.5	0.554	2.123	3.834	0.586
3.0	0.393	2.340	5.953	0.712
3.5	0.224	2.656	11.880	0.845
4.0	0.105	3.014	28.618	0.932

q_0	A	B	$q_0\cdot p_0=1.9\pi$ B/A	$r=(B-A)/(B+A)$
1.5	0.031	2.921	92.935	0.979
2.0	0.082	2.074	25.412	0.924
2.5	0.092	2.021	22.004	0.913
3.0	0.081	2.077	25.668	0.925
3.5	0.055	2.218	40.455	0.952
4.0	0.031	2.432	79.558	0.975

*Where possible ($q_0\cdot p_0=\pi/2$, $q_0\cdot p_0=\pi$) the estimates for A, B are compared with the exact values (computed via the Zak transform; see text).

(2.3.15), turn out to be remarkably close to the true values, giving deviations of at most a few percent on A, and less than 0.1% on B. In the next example we shall find similar orders of magnitude for the deviation of our estimates for A, B with respect to the exact values. The formulas (2.3.14) and (2.3.15) seem thus to give quite good results, despite the rather brutal estimating methods used.

For every one of the values of the product $q_0\cdot p_0$ in Table II we have also computed $\tilde{g}$ by means of the inversion formula (2.1.5). These functions $\tilde{g}$ are plotted in Fig. 4, for the choice $q_0=p_0$.

For $p_0=q_0=(\pi/2)^{1/2}$, we know, from Table II, that the frame is snug ($r\simeq 0.02$ in this case), i.e., that $\mathbb{T}$ is very close to a multiple of the unit operator. Consequently $\tilde{g}=\mathbb{T}^{-1}g$ is very close to a (scaled) Gaussian. For increasing $p_0=q_0$, several things happen to $\tilde{g}$. The decrease of both frames bounds A and B, which reflects the decrease in the "oversampling ratio" $2\pi(p_0q_0)^{-1}$, causes the amplitude of $\tilde{g}$ to increase. Moreover, the frame becomes less and less snug (for $p_0=q_0=(1.9\pi)^{1/2}$, one even has $r>0.9$), causing $\tilde{g}$ to deviate more and more from a Gaussian. In all these examples (Figs. 4(a)–4(e)), the function $\tilde{g}$ remains square integrable, however. One can even show that it remains C^∞, with fast decay. This is no longer the case if $p_0=q_0=(2\pi)^{1/2}$. In this limiting case, where the g_{mn} no longer constitute a frame, one can still construct $\tilde{g}=\mathbb{T}^{-1}g$, but, as previously shown, $\tilde{g}$ is no longer in L^2. This singular $\tilde{g}$ was first plotted by Bastiaans [2]; we have replotted it here in Fig. 4(f). One clearly sees how the regular $\tilde{g}$, for lower values of $p_0\cdot q_0$, approaches the singular limiting function as $p_0=q_0$ increases towards $(2\pi)^{1/2}$.

ii) The exponential case: We take $g(x)=e^{-|x|}$. In this case $m(g;q_0)$ and $M(g;q_0)$ can be calculated explicitly. One finds

$$m(g;q_0)=(\sinh q_0)^{-1}$$

$$M(g;q_0)=\coth q_0.$$

The function β cannot be written in closed analytic form. In Table III we list p_0^c for a few values of q_0; we also list again $2\pi/(q_0\cdot p_0^c)-1$, which is much less close to zero in this case. In Table IV we list A, B, B/A and $r=(B/A-1)/(B/A+1)$ for several values of q_0, p_0. Again we have grouped together those pairs q_0, p_0 with the same value of $p_0\cdot q_0$; in the cases $p_0\cdot q_0=\pi/2$, $p_0\cdot q_0=\pi/4$ we compare the estimates with the exact values.

TABLE III
THE ESTIMATED VALUES p_0^c, AND THEIR DEVIATION FROM $2\pi/q_0$, FOR $g(x)=e^{-|x|}$

q_0	p_0^c	$2\pi/(q_0\cdot p_0^c)-1$
0.5	5.32	1.36
1.0	2.99	1.10
1.5	2.52	0.66
2.0	2.21	0.42
2.5	1.92	0.30
3.0	1.68	0.25
3.5	1.48	0.22
4.0	1.33	0.17

In the two cases, in Table IV, where the exact values of A, B can be computed via the Zak transform (for $q_0\cdot p_0=\pi/2$ and $q_0\cdot p_0=\pi/4$), we see again that the estimates for A and B given by (2.3.14) and (2.3.15) are remarkably close to the exact values. The error on B is negligible, and the error on A does not exceed a few percent. Note also that frames based on the exponential used here are much less "snug" than Gaussian frames (compare e.g., the value of r for $q_0=1$, $p_0=\pi/2$, which is 0.399 in the present case, but 0.037 for a Gaussian).

TABLE IV
VALUES FOR THE FRAME BOUNDS A, B, THEIR RATIO B/A, AND THE CONVERGENCE FACTOR $r = (B/A - 1)/(B/A + 1)$, FOR THE CASE $g(x) = \exp(-|x|)$, FOR DIFFERENT VALUES OF q_0, p_0*

$q_0 \cdot p_0 = \pi/2$

q_0	A	A_{exact}	B	B_{exact}	B/A	$r = (B-A)/(B+A)$
1.0	2.600	2.724	6.056	6.056	2.330	0.399
1.5	2.665	2.692	6.781	6.781	2.544	0.436
2.0	2.179	2.190	8.326	8.326	3.821	0.585
2.5	1.648	1.657	10.140	10.140	6.152	0.720
3.0	1.197	1.206	12.060	12.060	10.074	0.819

$q_0 \cdot p_0 = 3\pi/8$

q_0	A	B	B/AQ	$r = (B-A)/(B+A)$
0.5	2.014	8.873	4.405	0.630
1.0	4.203	7.339	1.746	0.272
1.5	3.724	8.872	2.382	0.409
2.0	2.938	11.068	3.767	0.580
2.5	2.204	13.514	6.133	0.720
3.0	1.597	16.080	10.068	0.819

$q_0 \cdot p_0 = \pi/4$

q_0	A	A_{exact}	B	B_{exact}	B/A	$r = (B-A)/(B+A)$
1.0	6.757	6.766	10.554	10.554	1.562	0.219
1.5	5.634	5.645	13.259	13.259	2.353	0.404
2.0	4.412	4.426	16.597	16.597	3.762	0.580
2.5	3.306	3.322	20.271	20.271	6.132	0.720
3.0	2.396	2.413	24.119	24.119	10.068	0.819

*Where possible ($q_0 \cdot p_0 = \pi/2$, $q_0 \cdot p_0 = \pi/4$) the estimates for A, B are compared with the exact values (computed via the Zak transform; see text).

2) The Wavelet Case:

a) Ranges for the parameters a_0, b_0—Estimates for the frame bounds: In Section II-B-1b) we construct tight frames for arbitrary choices of $a_0 > 1$, $b_0 > 0$. This shows that there exists no absolute, *a priori* limitation on a_0, b_0—values leading to frames, unlike the Weyl–Heisenberg case, where $p_0 \cdot q_0 \le 2\pi$ is a necessary condition (see Section II-C-1a)). This freedom in the choice of a_0, b_0 is deceptive, however, because of the behavior of wavelet frames under dilations. If the h_{mn}, based on h, with parameters a_0, b_0, constitute a frame, then so do the $h_{\gamma;mn}$, based on $h_\gamma(x) = \gamma^{1/2} h(\gamma x)$, with frame parameters $a_0, \gamma^{-1} b_0$. This explains, at least partially, why a frame can be constructed for *any* pair a_0, b_0. To eliminate this dilational freedom, let us restrict our attention, in the present discussion, to frames such that $\|h\| = 1$ and $\int dk\, |k|^{-1} |\hat{h}(k)|^2 = 1$. Under this restriction, one might hope again that there exists a critical curve $b_0^c(a_0)$ separating the "frameable" pairs from the "nonframeable," with the orthogonal bases corresponding to the curve itself. This was the situation for the Weyl–Heisenberg case. It turns out however that this picture is not true in the wavelet case. At the end of this subsection, in Theorem 2.10, we establish the following counterexample. We take Y. Meyer's basic wavelet ψ, and look at the ψ_{mn,b_0}, a family of wavelets generated from ψ with $a_0 = 2$, b_0 arbitrary. For $b_0 = 1$, these wavelets constitute an orthonormal basis [31]. If there existed a nice critical curve $b_0^c(a_0)$ separating frameable and nonframeable values, then we would expect that the $\psi_{mn;b_0}$ would not be a frame for $b_0 > 1$ ("not enough" vectors), and might be a frame consisting of nonindependent vectors for $b_0 < 1$ ("too many" vectors). It turns out, however (see Theorem 2.10), that there exists $\epsilon > 0$ such that, for all values of b_0 in $(1-\epsilon,\ 1+\epsilon)$, the associated $\psi_{mn;b_0}$ constitute a basis for $L^2(\mathbb{R})$. This baffling fact shows that the concept of "phase space density," so well-suited for Weyl–Heisenberg frames, is not well adapted to the wavelet situation.

For the wavelet case this is all we have to say in answer to question 1), as formulated at the start of Section II-C. The following theorem addresses question 2), i.e., the determination for a given function h, of a range R such that the h_{mn} are a frame for all choices $(a_0, b_0) \in R$. The formulation of this theorem is very similar to Theorem 2.5, and so is its proof.

Theorem 2.7: If

$$1)\quad m(h;a_0) = \operatorname*{ess\,inf}_{|x| \in [1, a_0]} \sum_{m \in \mathbb{Z}} \left|\hat{h}(a_0^m x)\right|^2 > 0 \tag{2.3.18}$$

$$2)\quad M(h;a_0) = \operatorname*{ess\,sup}_{|x| \in [1, a_0]} \sum_{m \in \mathbb{Z}} \left|\hat{h}(a_0^m x)\right|^2 < \infty \tag{2.3.19}$$

and

$$3)\quad \sup_{s \in \mathbb{R}} \left[(1+s^2)^{(1+\epsilon)/2} \beta(s) \right] = C_\epsilon < \infty \quad \text{for some } \epsilon > 0 \tag{2.3.20}$$

where

$$\beta(s) = \sup_{|x| \in [1, a_0]} \sum_{m \in \mathbb{Z}} \left|\hat{h}(a_0^m x)\right| \left|\hat{h}(a_0^m x + s)\right|. \tag{2.3.21}$$

Then there exists a $B_0^c > 0$ such that

$\forall b_0 \in (0, B_0^c)$: the h_{mn} associated to h, a_0, b_0 constitute a frame,

$\forall \delta > 0$:$\exists b_0$ in $[B_0^c, B_0^c + \delta]$ such that the g_{mn} associated to h, a_0, b_0 are not a frame.

Proof: see Appendix C. □

Remarks:

1) The conditions (2.3.18), (2.3.19) are again necessary conditions. If (2.3.18) is not satisfied, then $\inf_{f \in L^2} \|f\|^{-2} \sum_{m,n} |\langle h_{mn}, f \rangle|^2 = 0$, which excludes the existence of a nonzero lower frame bound. Similarly (2.3.19) is necessary for the existence of a finite upper frame bound.
2) The range R of "good" parameters, i.e., the set of (a_0, b_0) such that the h_{mn}, associated to h, a_0, b_0, constitute a frame, need not be connected. It is possible to construct functions h such that, for fixed b_0, there exist $a_{0,1} < a_{0,2} < a_{0,3}$, for which the h_{mn} associated with $h, a_{0,j}; b_0$ constitute a frame if $j = 1$ or 3, but do not if $j = 2$. (The construction is similar to the one given for the Weyl–Heisenberg case. See Remark 2, following Theorem 2.5.)

3) Theorem 2.7 is only useful for choices of h for which the support of $\hat{h}$ contains negative as well as positive frequencies. In some cases one prefers to work with functions h with support $\hat{h} \subset \mathbb{R}_+$. Functions with this property are also called "analytic signals," because they extend to functions analytic on a half-plane. See e.g., [68]. The frame to be used then consists of $\{h^{\pm}_{mn}; m,n \in \mathbb{Z},\ h^+ = h,\ h^- = h^*\}$ or of $\{h^{(\lambda)}_{mn};\ m,n \in \mathbb{Z},\ \lambda = 1,2\ h^{(1)} = \sqrt{2}\ \mathrm{re}\, h,\ h^{(2)} = \sqrt{2}\ \mathrm{im}\, h\}$. For these frames the conclusions of Theorem 2.7 hold under very similar conditions. The only changes to be made concern the definitions of $m(h;a_0)$, $M(h;a_0)$ and $\beta(s)$. In each of these definitions, the condition $|x| \in [1,a_0]$ should be replaced by $x \in [1,a_0]$.

As in the Weyl–Heisenberg case (Theorem 2.5), the conditions (2.3.18)–(2.3.20) may seem very technical. They are however very easy to check on a computer. Good estimates for $m(h;a_0)$, $M(h;a_0)$ and $\beta(s)$ lead again to useful inequalities for the frame bounds A and B.

Theorem 2.8: Under the same conditions as in Theorem 2.7, the following lower bound for B_0^c holds

$$B_0^c \geq b_0^c = \inf\left\{ b_0 \left| 2\sum_{k=1}^{\infty}\left[\beta\left(-\frac{2\pi}{b_0}k\right)\beta\left(\frac{2\pi}{b_0}k\right)\right]^{1/2} \right. \geq m(h;a_0) \right\}. \tag{2.3.22}$$

For $0 < b_0 < b_0^c$, the following estimates for the frame bounds A, B hold

$$A \geq \frac{2\pi}{b_0}\left\{ m(h;a_0) - 2\sum_{k=1}^{\infty}\left[\beta\left(\frac{2\pi}{b_0}k\right)\beta\left(-\frac{2\pi}{b_0}k\right)\right]^{1/2}\right\}, \tag{2.3.23}$$

$$B \leq \frac{2\pi}{b_0}\left\{ M(h;a_0) + 2\sum_{k=1}^{\infty}\left[\beta\left(\frac{2\pi}{b_0}k\right)\beta\left(-\frac{2\pi}{b_0}k\right)\right]^{1/2}\right\}. \tag{2.3.24}$$

Proof: see Appendix C. □

Remark: The proof in Appendix C applies for the case where both $\mathbb{R}_+ \cap \mathrm{support}\,\hat{h}$ and $\mathbb{R}_- \cap \mathrm{support}\,\hat{h}$ have nonzero measure. If this is not the case, e.g., if support $\hat{h} \subset \mathbb{R}_+$, and the frame considered is either $\{h^{\pm}_{mn} | m,n \in \mathbb{Z},\ h^+ = h,\ h^- = h^*\}$ or $\{h^{(\lambda)}_{mn} | m,(n \in \mathbb{Z},\ \lambda = 1,2,\ h^{(1)} = \sqrt{2}\ \mathrm{re}\, h,\ h^{(2)} = \sqrt{2}\ \mathrm{im}\, h\}$ see Section II-B-1b)), then the definitions of $m(h;a_0), M(h;a_0), \beta(s)$ have to be slightly changed (the restriction $|x| \in [1,a_0]$ is replaced by $x \in [1,a_0]$—see Remark 3 following Theorem 2.7), and the same formulas (2.3.23), (2.3.24) apply.

In most practical examples the dilation parameter a_0 is equal to 2. In this case the estimates (2.3.23) and (2.3.24) can be sharpened. The following corollary is due to Ph. Tchamitchian.

Theorem 2.9: Choose $a_0 = 2$. Under the same conditions as in Theorem 2.7, the following estimates for the frame bounds A, B hold,

$$A \geq \frac{2\pi}{b_0}\left\{ m(h;2) - 2\sum_{l=0}^{\infty}\left[\beta_1\left(\frac{2\pi}{b_0}(2l+1)\right)\cdot\beta_1\left(-\frac{2\pi}{b_0}(2l+1)\right)\right]^{1/2}\right\} \tag{2.3.25}$$

$$B \leq \frac{2\pi}{b_0}\left\{ M(h;2) + 2\sum_{l=0}^{\infty}\left[\beta_1\left(\frac{2\pi}{b_0}(2l+1)\right)\cdot\beta_1\left(-\frac{2\pi}{b_0}(2l+1)\right)\right]^{1/2}\right\} \tag{2.3.26}$$

where

$$\beta_1(s) = \sup_x \sum_{m \in \mathbb{Z}}\left| \sum_{m' \geq 0} \hat{h}(2^{m+m'}x)\hat{h}^*[2m'(2^m x + s)]\right|.$$

Proof: See Appendix C. □

Note that the estimates in Theorem 2.9 use some of the phase information contained in $\hat{h}$, which is completely lost in the estimates in Theorem 2.8. It is therefore to be expected that the estimates (2.3.25), (2.3.26) are a significant improvement (2.3.23), (2.3.24) for complex $\hat{h}$. This is illustrated most dramatically by applying both pairs of estimates to the basic wavelet in Y. Meyer's basis (defined at the end of Section II-B-1b)). For $h = \psi$, with ψ defined by (2.2.7), one finds $\sum_m |\hat{\psi}(2^m y)|^2 = 1$, $\beta(2\pi) = \beta(-2\pi) = 1/2$, $\beta(4\pi) = \beta(-4\pi) = 1/2$, and $\beta(k2\pi) = 0$ if $|k| \geq 3$, hence $\sum_{k=1}^{\infty}[\beta(2\pi k)\beta(-2\pi k)]^{1/2} = 1$. Applying Theorem 2.8 we therefore find $A \geq -1$, $B \leq 3$.

This means that if we had only the bounds (2.3.23), (2.3.24) to go by, we wouldn't be able to recognize that the ψ_{mn} constitute a frame. Since, for $a_0 = 2$ and $b_0 = 1$, the ψ_{mn} do in fact constitute an orthonormal base, this is a rather poor performance. Calculating β_1 we find that

$$\beta_1(2\pi k) = 0, \qquad \text{for all odd } k.$$

The estimates in Theorem 2.9 thus lead to $A \geq 1$, $B \leq 1$, or equivalently to the optimal estimate $A = B = 1$. (Note that this automatically proves that the ψ_{mn} constitute an orthonormal basis, since $\|\psi_{mn}\| = \|\psi\| = 1$. We showed in Section II-A that a tight frame with frame constant 1, consisting of normalized vectors, necessarily is an orthonormal basis.) Since the phase factor in the definition of Y. Meyer's basic wavelet (2.2.7) is essential for the ψ_{mn} to constitute an orthonormal basis (see e.g., [31]), it is natural that the phase-dependent estimates (2.3.25), (2.3.26) perform much better than the phase-independent estimates (2.3.23), (2.3.24) in this case. Using Theorem 2.9 one can prove the result we just announced, i.e., that the wavelets ψ_{mn}, associated with Y. Meyer's ψ, and with $a_0 = 2$, b_0 arbitrary, still constitute a basis for $b_0 \neq 1$, $|b_0 - 1| < \epsilon$, for some $\epsilon > 0$. The following Theorem was proved in collaboration with Ph. Tchamitchian.

Theorem 2.10: Let ψ be the function defined by (2.2.7). Then there exists $\epsilon > 0$ such that the $\psi_{mn;b_0}$,

$$\psi_{mn;b_0}(x) = 2^{-m/2}\psi(2^{-m}x - nb_0), \qquad m,n \in \mathbb{Z}$$

constitute a basis for $L^2(\mathbb{R})$, for any choice $b_0 \in (1-\epsilon, 1+\epsilon)$.

Proof: See Appendix C. □

This result is quite surprising: it shows, as pointed out at the start of Section II-C-2, that it is not always safe to apply "phase space density intuition" to families of wavelets.

Remark: It follows from the proof that ϵ can be estimated explicitly, from computations of the frame bounds A, B, as given by (2.3.25), (2.3.26) on the one hand, and of

$$\sum_{\substack{|m|<1 \\ m \in \mathbb{Z}}} |\langle \psi_{mn;b_0}, \psi \rangle|$$

on the other hand. This estimate for ϵ depends, of course, on the choice for the function ν (see (2.2.7)). For the (non C^∞) choice $\nu(x) = x$ if $0 \le x \le 1$, $\nu(x) = 0$ otherwise, one finds $\epsilon > 0.02$. The ψ_{mn,b_0} remain a frame for $b_0 \le 1.08$ in this case.

In many practical examples $\hat{h}$ decays very fast, and is real. In those cases the differences between estimates using β (i.e., (2.3.23), (2.3.24)) or β_1, (i.e., (2.3.25), (2.3.26)) are very small, and much less dramatic then for Y. Meyer's wavelet ψ. In fact, if $\hat{h}$ is positive (e.g., the Mexican hat function, the 8th derivative of the Gaussian, the modulated Gaussian, $\cdots$: see the next subsection) the estimates using β perform better than those using β_1, as can easily be checked directly by the formulas (2.3.23), (2.3.24), and (2.3.25), (2.3.26).

As already mentioned above, most practical applications use $a_0 = 2$. This choice makes numerical computations much easier since it means that the translation steps $b_0 \cdot 2^m$ (see Fig. 1(b)), for two different frequency levels $m_1 > m_2$, are multiples of each other. It also makes the different frequency levels correspond to "octaves." On the other hand one likes to use functions h with fairly concentrated Fourier transforms $\hat{h}$, corresponding to a good frequency resolution. This means that $\sum_m |\hat{h}(2^m x)|^2$ is bound to have rather large oscillations; since then $m(h;2)$ is much smaller than $M(h;2)$, this leads to high, and therefore unpleasant values of B/A. This can be avoided by the use of several functions h^j, chosen so that the minima of $\sum_m |\hat{h}^j(2^m x)|^2$ are compensated by the maxima of $\sum_m |\hat{h}^{j'}(2^m x)|^2$, for some $j' \neq j$. This is made explicit by the following corollary of Theorem 2.7.

Corollary 2.11: Let $h^0, \cdots, h^{N-1}$ be N functions satisfying the conditions (2.3.18), (2.3.19), and (2.3.20). Define

$$m(h^0, \cdots, h^{N-1}; a_0) = \operatorname*{ess\,inf}_{|x| \in [1, a_0]} \sum_{j=0}^{N-1} \sum_m |\hat{h}^j(a_0^m x)|^2 \tag{2.3.27}$$

$$M(h^0, \cdots, h^{N-1}; a_0) = \operatorname*{ess\,sup}_{|x| \in [1, a_0]} \sum_{j=0}^{N-1} \sum_m |\hat{h}^j(a_0^m x)|^2 \tag{2.3.28}$$

$$\beta^j(s) = \sup_{|x| \in [1, a_0]} \sum_{m \in \mathbb{Z}} |\hat{h}^j(a_0^m x)| \, |\hat{h}^j(a_0^m x + s)|. \tag{2.3.29}$$

Choose b_0 such that

$$2 \sum_{j=0}^{N-1} \sum_{k=1}^{\infty} \left[\beta^j\left(\frac{2\pi}{b_0}k\right) \beta^j\left(-\frac{2\pi}{b_0}k\right) \right]^{1/2} < m(h^0, \cdots, h^{N-1}; a_0).$$

Then the $\{h^j_{mn}; m,n \in \mathbb{Z},\ j = 0, \cdots, N-1\}$ constitute a frame. The following estimates for the frame bounds A and B hold,

$$A \ge \frac{2\pi}{b_0} \left\{ m(h^0, \cdots, h^{N-1}; a_0) - 2 \sum_{j=0}^{N-1} \sum_{k=1}^{\infty} \cdot \left[\beta^j\left(\frac{2\pi}{b_0}k\right) \beta^j\left(-\frac{2\pi}{b_0}k\right) \right]^{1/2} \right\} \tag{2.3.30}$$

$$B \le \frac{2\pi}{p_0} \left\{ M(h^0, \cdots, h^{N-1}; a_0) + 2 \sum_{j=0}^{N-1} \sum_{k=1}^{\infty} \cdot \left[\beta^j\left(\frac{2\pi}{p_0}k\right) \beta^j\left(-\frac{2\pi}{p_0}k\right) \right]^{1/2} \right\}. \tag{2.3.31}$$

Proof: The proof is a simple variant on the proof of Theorem 2.8. □

In the special case where $a_0 = 2$, one can, of course, replace β^j in (2.3.30), (2.3.31) by β^j_1, with β^j_1 defined as in Theorem 2.8, for $j = 0, \cdots, N-1$. Then the sums over $k \neq 0$ also have to be replaced by sums over only odd k.

The number N of functions used is called the number of "voices" per octave [28]. In numerical calculations $N = 4$ seems to be a satisfactory choice.

In practice one often chooses the $h^0, \cdots, h^{N-1}$ to be dilated versions of one function h, i.e.,

$$h^j(x) = 2^{-j/N} h(2^{-j/N}x), \qquad j = 0, \cdots, N-1. \tag{2.3.32}$$

The phase space lattice corresponding to the $\{h^j_{mn};\ m,n \in \mathbb{Z},\ j = 0, \cdots, N-1\}$ is the superposition of N lattices of the type depicted in Fig. 1(b), shifted with respect to each other in frequency. Fig. 5 shows such a lattice, for the case $N = 4$, and for $a_0 = 2$.

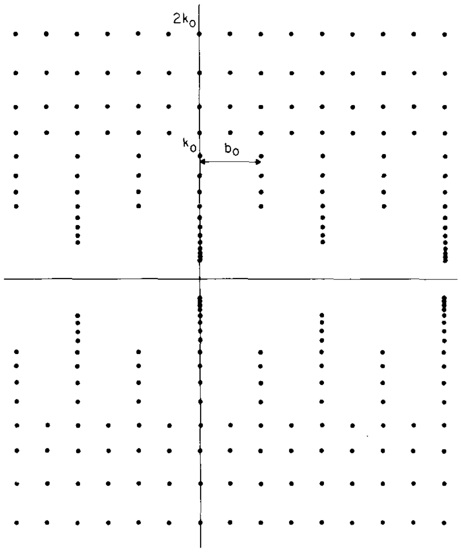

Fig. 5. Phase space lattice corresponding to wavelet frame with $a_0 = 2$, $b_0 = 1$, and 4 voices per octave (see text). The different voice wavelets are obtained by dilating one single function, $h^j(x) = 2^{-j/4}h(2^{-j/4}x)$, $j = 0,1,2,3$. As in Fig. 1(b), $k_0 = \int_0^\infty dk\,|k|\,|\hat{h}(k)|^2$, and $\hat{h}$ is assumed to be even.

TABLE V-A
FRAME BOUNDS FOR WAVELET FRAMES*

	$N=1$				$N=2$		
b_0/π	A	B	B/A	b_0/π	A	B	B/A
0.25	4.038	8.409	2.082	0.25	11.950	16.294	1.364
0.50	1.838	4.386	2.387	0.50	5.711	8.411	1.473
0.75	1.412	3.007	2.634	0.75	3.629	5.785	1.594
1.00	0.585	2.527	4.323	1.00	2.410	4.651	1.930
1.25	0.337	2.152	6.380	1.25	1.709	3.939	2.305
				1.50	0.999	3.708	3.713
				1.75	0.556	3.479	6.259
				2.00	0.202	3.328	16.473

	$N=3$				$N=4$		
b_0/π	A	B	B/A	b_0/π	A	B	B/A
0.25	20.035	24.331	1.214	0.25	27.986	32.500	1.161
0.50	9.655	12.528	1.298	0.50	13.598	16.645	1.224
0.75	6.185	8.603	1.391	0.75	8.687	11.475	1.321
1.00	4.230	6.861	1.622	1.00	6.042	9.080	1.503
1.25	3.050	5.823	1.909	1.25	4.431	7.666	1.730
1.50	2.033	5.361	2.637	1.50	3.076	7.005	2.278
1.75	1.313	5.025	3.828	1.75	2.031	6.610	3.255
2.00	0.736	4.810	6.539	2.00	1.261	6.300	4.998

*Based on the function

$$h(x) = \begin{cases} \pi^{-1/2} \sin x & |x| \le \pi \\ 0, & \text{otherwise} \end{cases}.$$

The dilation parameter $a_0 = 2$ in all cases; N is the number of "voices" (see text).

Remark: Note that the h^j, constructed by dilations from one function h as in (2.3.32), do not have the same L^2-normalization, $\|h^j\|_{L^2} = 2^{-j/2N}\|h\|_{L^2}$. This change in normalization compensates for the fact that the phase space lattice in Fig. 5 is "denser" than the corresponding lattice (see e.g., Fig. 1(b)) would be for the same basic wavelet h, but with $a_0 = 2^{1/N}$, and with only one function (namely h) per dilation step (instead of N, as in Fig. 5).

b) Examples: We illustrate the different bounds with several examples. In practice, it is by far preferable to use $a_0 = 2$, rather than noninteger values. We have therefore, in all our examples, restricted ourselves to $a_0 = 2$, and introduced several voices (see Corollary 2.11). The different voice-functions h^j are always obtained, in these examples, by dilation of one given function h (see (2.3.32)).

i) One cycle of the sine-function: In this case we take

$$h(x) = \begin{cases} \dfrac{1}{\sqrt{\pi}} \sin x, & \text{if } |x| \le \pi \\ 0, & \text{otherwise} \end{cases}.$$

The Fourier transform $\hat{h}$ of h is given by

$$\hat{h}(y) = \frac{\sqrt{2}}{\pi} i \frac{\sin y\pi}{1-y^2}.$$

Since this is an oscillating function, one suspects that the formulas analogous to (2.3.30), (2.3.31), but using β_1 rather than β, will perform better than the β-formulas. This turns out to be true. Table V-A lists the estimates for the frame bounds A, B and their ratio, for a few values of b_0, and of N, the number of voices.

ii) The Mexican hat: The Mexican-hat function is the second derivative of the Gaussian (up to a sign),

$$h(x) = \frac{2}{\sqrt{3}} \pi^{-1/4}(1-x^2)e^{-x^2/2}$$

$$\hat{h}(y) = \frac{2}{\sqrt{3}} \pi^{-1/4} y^2 e^{-y^2/2}.$$

The graph of h looks a bit like a transverse section of a Mexican hat (see Fig. 2(b)), whence the name. Since $\hat{h}$ is positive, the formulas using β rather than β_1 are more effective here. The same will be true in our next examples. Table V-B lists the estimates for the frame bounds A, B and their ratio, for a few values of the translation parameter b_0, and of N, the number of voices.

iii) The eighth derivative of the Gaussian: Functions like the Mexican hat and higher order derivatives of the Gaussian are useful in applications of the wavelet transform to edge detection (see, e.g., [27]). Table V-C lists the estimates for the frame bounds A, B for wavelets based on the eighth derivative of the Gaussian,

$$h(x) = \left(\frac{2^{15}7!}{15!}\right)^{1/2} \pi^{-1/4}(x^8 - 28x^6 + 210x^4 - 405x^2 + 90)e^{-x^2/2}$$

$$\hat{h}(y) = \left(\frac{2^{15}7!}{15!}\right)^{1/2} \pi^{-1/4} y^8 e^{-y^2/2}.$$

This is a typical example of a function where, with $a_0 = 2$ fixed, the introduction of several voices is necessary. For $N = 1$ (i.e., only one voice, corresponding to the phase

TABLE V-B
FRAME BOUNDS FOR WAVELET FRAMES*

$N=1$				$N=2$			
b_0	A	B	B/A	b_0	A	B	B/A
0.25	13.091	14.183	1.083	0.25	27.273	27.278	1.0002
0.50	6.546	7.092	1.083	0.50	13.637	13.639	1.0002
0.75	4.364	4.728	1.083	0.75	9.091	9.093	1.0002
1.00	3.223	3.596	1.116	1.00	6.768	6.870	1.015
1.25	2.001	3.454	1.726	1.25	4.834	6.077	1.257
1.50	0.325	4.221	12.986	1.50	2.609	6.483	2.485
				1.75	0.517	7.276	14.061

$N=3$				$N=4$			
b_0	A	B	B/A	b_0	A	B	B/A
0.25	40.914	40.914	1.0000	0.25	54.552	54.552	1.0000
0.50	20.457	20.457	1.0000	0.50	27.276	27.276	1.0000
0.75	13.638	13.638	1.0000	0.75	18.184	18.184	1.0000
1.00	10.178	10.279	1.010	1.00	13.586	13.690	1.007
1.25	7.530	8.835	1.173	1.25	10.205	11.616	1.138
1.50	4.629	9.009	1.947	1.50	6.594	11.590	1.758
1.75	1.747	9.942	5.691	1.75	2.928	12.659	4.324

*Based on the Mexican hat function

$$h(x)=2/\sqrt{3}\,\pi^{-1/4}(1-x^2)e^{-x^2/2}.$$

The dilation parameter $a_0=2$ in all cases; N is the number of voices (see text).

TABLE V-C
FRAME BOUNDS FOR WAVELET FRAMES*

$N=2$			
b_0	A	B	B/A
0.125	25.515	26.569	1.041
0.250	12.758	13.285	1.041
0.375	8.505	8.856	1.041
0.500	6.379	6.642	1.041
0.625	5.101	5.316	1.042
0.750	3.995	4.686	1.173
0.875	1.669	5.772	3.459
$N=3$			
b_0	A	B	B/A
0.125	39.054	39.073	1.0005
0.250	19.527	19.536	1.0005
0.375	13.018	13.024	1.0005
0.500	9.764	9.768	1.0005
0.625	7.808	7.817	1.001
0.750	6.251	6.770	1.083
0.875	3.563	7.598	2.132
1.000	0.163	9.603	58.929
$N=4$			
b_0	A	B	B/A
0.125	52.085	52.085	1.0000
0.250	26.042	26.042	1.0000
0.375	17.362	17.362	1.0000
0.500	13.022	13.022	1.0000
0.625	10.414	10.420	1.0005
0.750	8.420	8.941	1.062
0.875	5.313	9.568	1.801
1.000	1.127	11.894	10.550

*Based on the 8th derivative of the Gaussian,

$$h(x)=\left(\frac{2^{15}7!}{15!}\right)^{1/2}\cdot\pi^{-1/4}(x^8-28x^6+210x^4-405x^2+90)e^{-x^2/2}.$$

The dilation parameter $a_0=2$ in all cases; N is the number of voices.

space lattice in Fig. 1(b)) one finds that $M(\hat{h};2)/m(\hat{h};2)$ is equal to 3.385. This means that the ratio B/A, which is bounded below by $M(\hat{h};2)/m(\hat{h};2)$, is pretty large, with $N=1$, even for very small values of b_0. As soon as more voices are introduced, the ratio M/m becomes much smaller, and snug frames can be constructed, for appropriate choices of b_0. We have restricted ourselves, in Table Vc, to the choices $N=2,3,4$, excluding $N=1$.

iv) The modulated Gaussian: In this case we take

$$h(x)=\pi^{-1/4}\left(e^{-ikx}-e^{-k^2/2}\right)e^{-x^2/2}$$
$$\hat{h}(y)=\pi^{-1/4}\left(e^{-(y-k)^2/2}-e^{-k^2/2}e^{-y^2/2}\right)$$

where $k=\pi(2/\ln 2)^{1/2}$.

The subtraction term in the definition of $h,\hat{h}$ ensures that $\hat{h}(0)=0$; for the value of k chosen here, this term is negligible in practice. The value of k has been fixed so that the ratio between the highest and the second highest local maxima of $\operatorname{Re} h$ is approximately $1/2$. This wavelet is exactly the wavelet used by J. Morlet in his numerical computations [25], [26]. If only real signals f are decomposed and reconstructed, by means of the $\langle h_{mn},f\rangle$, then the complex wavelet h consists really of two wavelets, $\operatorname{Re} h$ and $\operatorname{Im} h$. In this case the frame bounds for real signals can be rewritten as

$$A=\frac{2\pi}{b_0}\left\{\frac{1}{2}\min_x\sum_{m\in\mathbb{Z}}\left[|\hat{h}(a_0^m x)|^2+|\hat{h}(-a_0^m x)|^2\right]-R\right\}$$

$$B=\frac{2\pi}{p_b}\left\{\frac{1}{2}\max_x\sum_{m\in\mathbb{Z}}\left[|\hat{h}(a_0^m x)|^2+|\hat{h}(-a_0^m x)|^2\right]+R\right\}$$

with

$$R=2\sum_{\epsilon=+,-}\sum_{l=1}^{\infty}\left[\beta_\epsilon\left(\frac{2\pi}{b_0}l\right)\beta_\epsilon\left(-\frac{2\pi}{b_0}l\right)\right]^{1/2}$$

and

$$\beta_\epsilon(s)=\frac{1}{4}\sup_x\sum_m|\hat{h}(a_0^m x)+\epsilon\hat{h}(-a_0^m x)|\cdot|\hat{h}(a_0^m x+s)+\epsilon\hat{h}(-a_0^m x-s)|.$$

Note that $\hat{h}$ is almost completely concentrated on the positive frequency half line; neglecting terms with negative arguments for $\hat{h}$ in the previous formulas leads back to the frame bounds for support $\hat{h}\subset\mathbb{R}_+$. In the first reconstructions with this wavelet, before even the connection with continuously labelled affine coherent states was made (see Sections I-E, I-F), a formula similar to (2.1.7) was used. This reconstruction formula turned out to be extremely precise. Our calculations of frame bounds show why. For $N=4$, $b_0=1$, for instance, which are choices that do correspond to values used in practice (in fact, [25] uses even higher values of N, and smaller values of b_0), we find $B/A=1.0008$. Hence $r=(B-A)/(B+A)\simeq 0.04\%$, which explains why (2.1.7) is such a good approximation to the true reconstruction formula.

As in the previous example, the ratio M/m is rather large for $N=1$, and we have computed the frame bounds only for $N=2$, 3, and 4. They are tabulated in Table V-D.

TABLE V-D
FRAME BOUNDS FOR WAVELET FRAMES*

b_0	A	B	B/A
		$N=2$	
0.5	6.019	7.820	1.299
1.0	3.009	3.910	1.230
1.5	1.944	2.669	1.373
2.0	1.173	2.287	1.950
2.5	0.486	2.282	4.693
		$N=3$	
0.5	10.295	10.467	1.017
1.0	5.147	5.234	1.017
1.5	3.366	3.555	1.056
2.0	2.188	3.002	1.372
2.5	1.175	2.977	2.534
3.0	0.320	3.141	9.824
		$N=4$	
0.5	13.837	13.846	1.0006
1.0	6.918	6.923	1.0008
1.5	4.540	4.688	1.032
2.0	3.013	3.910	1.297
2.5	1.708	3.829	2.242
3.0	0.597	4.017	6.732

*Based on the modulated Gaussian,

$$h(x)=\pi^{-1/4}(e^{-ikx}-e^{-k^2/2})e^{-x^2/2},$$

with $k=\pi(2/\ln 2)^{1/2}$.

The dilation constant $a_0=2$ in all cases; N is the number of voices.

Remark: The tables show that extremely snug frames can be obtained for quite reasonable phase space lattices (i.e., N not too large, b_0 not too small). Note, however, that even when B/A s not very close to 1, the frame may still be useful. For $B/A \simeq 1.5$, e.g., the convergence factor $r=(B-A)/(B+A)$ is still of the order of 0.2; while this is insufficient to permit the use of the approximation formula (2.1.7), it does mean that only a few iterations will suffice for the computation of the $(h_{mn})^{\sim}$ up to e.g., 10^{-3}, leading, again, to a very accurate reconstruction formula.

D. Frames in Other Spaces than $L^2(\mathbb{R})$

The results in this section are more technical and specialized than those in the preceding sections. The reader who is mostly interested in L^2-results can safely omit reading this section and go to Section III, where we again discuss L^2-estimates.

1) Motivation: Why Other Spaces than $L^2(\mathbb{R})$?: So far, we have restricted ourselves to studying frames in $L^2(\mathbb{R})$. The preceding section was mainly concerned with the formulation of conditions under which the short-time Fourier expansions (Weyl–Heisenberg case) or the wavelet expansions (affine case) would converge with respect to the L^2-norm. As explained in the introduction, both types of expansions are used in the analysis, e.g., of time-dependent signals. For such signals $f(t)$, the square of the L^2-norm, $\int_{-\infty}^{\infty}dt|f(t)|^2$, is a natural quantity, often called the "energy" of the signal in electrical engineering literature. It is therefore customary to require L^2-convergence for series representations of these signals.

This does not mean, however, that convergence in other topologies is not important. For example, convergence in $L^p(\mathbb{R})$-spaces,

$$L^p(\mathbb{R})=\left\{f;\|f\|_p=\left[\int dx|f(x)|^p\right]^{1/p}<\infty\right\}$$

with $p\neq 2$, where the coefficients are weighted in a way different from $L^2(\mathbb{R})$, would certainly indicate that the series is more "robust" than if it converged in $L^2(\mathbb{R})$ alone. On the other hand, if the signals themselves have some regularity (consider, e.g., the case where not only f but also its first k derivatives are in $L^2(\mathbb{R})$), then one would wish that the partial sums of the series representing f share that regularity, and that the convergence respects this (in the previous example, this would mean L^2-convergence also for the first k derivatives). It would therefore be interesting to have convergence in the Sobolev spaces

$$\mathscr{H}_s(\mathbb{R})=\left\{f;\|f\|^2_{\mathscr{H}_s}=\int dy(1+y^2)^s|\hat{f}(y)|^2<\infty\right\}$$

for at least some $s>0$.

In the two subsections following this one, we shall show that for suitably chosen g or h, and appropriate parameters p_0, p_0 or a_0, b_0, the resulting frames are frames in $\mathscr{H}_s(\mathbb{R})$, at least for some strip $|s|<s_0$. In the remainder of this subsection we give some examples illustrating "what can go wrong." The examples given here all pertain to the wavelet case. I want to thank Y. Meyer and Ph. Tchamitchian for having pointed them out to me.

If the h_{mn} ($h_{mn}(x)=a_0^{-m/2}h(a_0^{-m}x-nb_0)$) constitute a frame in $L^2(\mathbb{R})$, then, as shown in Section II-A, there exists a dual frame $\{(h_{mn})^{\sim};\ m,n\in\mathbb{Z}\}$ such that, for all f in $L^2(\mathbb{R})$,

$$f=\sum_{m,n}(h_{mn})^{\sim}\langle h_{mn},f\rangle \tag{2.4.1}$$

where the series converges in $L^2(\mathbb{R})$. If the frame is not tight, then the $(h_{mn})^{\sim}$ are not multiples of the h_{mn}, and in general they will not be generated by dilations and translations of a single function (see Section II-A).

There are several ways in which (2.4.1) may fail to extend to $L^p(\mathbb{R})$, with $p\neq 2$, or to $\mathscr{H}_s(\mathbb{R})$, with $s\neq 0$. One possibility is that the coefficients $\langle h_{mn},f\rangle$ are not well defined for all f in the space under consideration, i.e., that h does not belong to the dual of this space. This is not really a problem: it is enough to impose some regularity and/or decay conditions on h to avoid this. Something much more pernicious may happen, however. It is possible that even though h is a "nice" function, the elements $(h_{mn})^{\sim}$ of the dual frame are not, so that (2.4.1) fails. The examples we shall give here illustrate this.

The first example is due to P.-G. Lemarié. It was communicated to me by Y. Meyer. Let ψ be Y. Meyer's wavelet, as As already mentioned, the $\psi_{mn}(x)=$

$2^{-m/2}\psi(2^{-m}x-n)$ constitute an unconditional basis for many function spaces, including all the $L^p(\mathbb{R})$, $1<p<\infty$. The wavelet coefficients $\langle\psi_{mn},f\rangle$ can be used to characterize these spaces and their duals [31]. In particular, L^p-spaces are characterized as follows: for any function f one has the equivalence

$$f\in L^p(\mathbb{R})\Leftrightarrow\left[\sum_{m,n}2^{-m}|\langle\psi_{mn},f\rangle|^2\chi_{mn}\right]^{1/2}\in L^p(\mathbb{R}) \tag{2.4.2}$$

where χ_{mn} is the indicator function of the interval $I_{mn}=[2^m n,2^m(n+1))$. The basic wavelet in Lemarié's construction is the following perturbation of ψ:

$$h(x)=\psi(x)+\sqrt{2}\,r\psi(2x).$$

This function h is clearly in C^∞, and has fast decay at ∞. We use the same dilation and translation parameters as for Y. Meyer's basis, i.e., $h_{mn}(x)=2^{-m/2}h(2^{-m}x-n)$. Then $h_{mn}=\psi_{mn}-r\psi_{m-1\,2n}=(1-rU)\psi_{mn}$, where U is the partial isometry defined by $U\psi_{mn}=\psi_{m-1\,2n}$. For $|r|<1$, the operator $(1-rU)$ is one-to-one and onto, which means that the h_{mn} still constitute a basis for $L^2(\mathbb{R})$ if $|r|<1$. The dual frame $(h_{mn})^\sim$ is given by

$$\begin{aligned}(h_{mn})^\sim&=[(1-rU)(1-r^*U^*)]^{-1}h_{mn}\\&=(1-r^*U^*)^{-1}\psi_{mn}\\&=\sum_{k=0}^{\infty}r^{*k}U^{*k}\psi_{mn}\end{aligned} \tag{2.4.3}$$

where $U^*\psi_{m\,2n+1}=0$, and $U^*\psi_{m\,2n}=\psi_{m+1\,n}$. In particular

$$(h_{00})^\wedge=\sum_{j=0}^{\infty}r^{*k}\psi_{-k0}.$$

It turns out that if $|r|>1/\sqrt{2}$, this function does not belong to $L^p(\mathbb{R})$ for large p, as shown by the following argument. If we apply (2.4.2) to $(h_{00})^\sim$, we find

$$\begin{aligned}(h_{00})^\sim\in L^p(\mathbb{R})&\Leftrightarrow\left[\sum_{m=0}^{\infty}(2|r|^2)^m\chi_{-m0}\right]^{1/2}\in L^p(\mathbb{R})\\&\Leftrightarrow\sum_{m=0}^{\infty}\int_{2^{-(m+1)}}^{2^{-m}}dx\left[\sum_{j=0}^{m}(2|r|^2)^j\right]^{p/2}<\infty\\&\Leftrightarrow\sum_{m=0}^{\infty}2^{-(m+1)}\left|\frac{(2|r|^2)^{m+1}-1}{2|r|^2-1}\right|^{p/2}<\infty.\end{aligned}$$

If $|r|>1/\sqrt{2}$, this shows that $(h_{00})^\sim\notin L^p(\mathbb{R})$ for all $p>2\ln 2/[\ln 2+2\ln|r|]$. This implies that, even though the h_{mn} themselves are C^∞ functions with fast decay, and constitute a frame for $L^2(\mathbb{R})$ (with $A=(1-|r|^2)^{1/2}$, $B=(1+|r|^2)^{1/2}$), the frame expansion (2.4.1) does not extend to all $L^p(\mathbb{R})$-spaces, $1<p<\infty$, if $|r|>1/\sqrt{2}$.

Note that, in the example, the $(h_{m0})^\sim$ are the only functions in the dual frame causing problems. For $n\neq 0$, only a finite number of terms in the series (2.4.3) contribute, and $(h_{mn})^\sim$ is still infinitely differentiable and decaying rapidly. This might lead one to believe that the problems are caused by the fact that the $(h_{mn})^\sim$ are not given, in general, by dilations and translations of a single function $\tilde{h}$. It is true, also, that this phenomenon ($h\in\mathscr{S}$, and some $(h_{mn})^\sim\notin L^p$) does not occur for the tight frames constructed in Section II-B-2b). For these frames, results analogous to those for Y. Meyer's basis hold, and the expansions

$$f=A^{-1}\sum_{m,n}h_{mn}\langle h_{mn},f\rangle$$

are valid, and converge, in particular, for all L^p-spaces ($1<p<\infty$) and all $\mathscr{H}_s(\mathbb{R})$. Nevertheless, there exist (nontight) frames for which the $(h_{mn})^\sim$ happen to be dilations and translations of a single function $\tilde{h}$, and where, as in our first example, h is "nice" and $\tilde{h}$ is not, causing (2.4.1) to fail in at least some function spaces. Our second example illustrates this; it is a special case of a construction made by Ph. Tchamitchian [35].

In [35], Ph. Tchamitchian constructs functions σ,τ such that the dyadic (i.e., $a_0=2$, $b_0=1$) lattices of wavelets generated by σ and τ constitute biorthogonal bases for $L^2(\mathbb{R})$, i.e.,

$$\sum_{m,n}\sigma_{mn}\langle\tau_{mn},\cdot\rangle=\mathbb{1} \tag{2.4.4}$$

where the sum converges strongly in $L^2(\mathbb{R})$, and with $\sigma_{mn}(x)=2^{-m/2}\sigma(2^{-m}x-n)$, $\tau_{mn}(x)=2^{-m/2}\tau(2^{-m}x-n)$. The details of this construction, together with proofs and applications, are given in [35b]. It is possible to choose both σ and τ with compact support. One such example is the following (see [35b]):

$$\hat{\sigma}(y)=y^{-1}e^{iy/2}\left(\frac{1}{2}-\frac{9}{16}\cos\frac{y}{2}+\frac{1}{16}\cos\frac{3y}{2}\right)$$

or

$$\sigma(x)=\frac{1}{4}\sqrt{2\pi}\begin{cases}0, & x\le -1\\ 1/8, & -1\le x\le 0\\ -1, & 0\le x\le 1/2\\ 1, & 1/2\le x\le 1\\ -1/8, & 1\le x\le 2\\ 0, & x\ge 2\end{cases}$$

and

$$\hat{\tau}(y)=ye^{iy/2}\prod_{j=1}^{\infty}P\left[\cos^2\left(2^{-j}\frac{y}{4}\right)\right]$$

where $P(u)=3u^2-2u^3$. Since $P(\cos y)=1+O(y^4)$, the infinite product in the definition of $\hat{\tau}$ is well defined. One finds that $\hat{\tau}$ is entire and of exponential type, implying that τ has compact support. Moreover, since $P(u)\le u^2$ on $[0,1]$,

$$\begin{aligned}|\hat{\tau}(y)|&\le|y|\prod_{j=1}^{\infty}\cos^4[2^{-j}y/4]\\&=|y|\left[\frac{\sin y/4}{y/4}\right]^4\le C(1+|y|^2)^{-3/2}.\end{aligned} \tag{2.4.5}$$

Using this decay of $\hat{\tau}$, the compactness of support τ, and $\int dx\, \tau(x) = \hat{\tau}(0) = 0$, one can easily prove that $\Sigma_{m,n}|\langle \tau, \tau_{mn}\rangle| < \infty$. The decay properties of $\hat{\sigma}$ are rather weak, but one can use $\int dx\, \sigma(x) = 0$, the compactness of support σ and the fact that σ is piecewise constant in an easy proof of $\Sigma_{m,n}|\langle \sigma, \sigma_{mn}\rangle| < \infty$. This implies that, for all sequences $(c_{mn})_{m,n\in\mathbb{Z}}$ in $l^2(\mathbb{Z})$,

$$\left\| \sum_{m,n} c_{mn}\sigma_{mn} \right\|^2 \leq \left(\sum_{m,n} |\langle \sigma, \sigma_{mn}\rangle| \right) \cdot \left(\sum_{m,n} |c_{mn}|^2 \right) \leq C \sum_{m,n} |c_{mn}|^2$$

$$\left\| \sum_{m,n} c_{mn}\tau_{mn} \right\|^2 \leq \left(\sum_{m,n} |\langle \tau, \tau_{mn}\rangle| \right) \cdot \left(\sum_{m,n} |c_{mn}|^2 \right) \leq C \sum_{m,n} |c_{mn}|^2.$$

Together with (2.4.4) (for the proof of which we refer to [35b]), this implies that the σ_{mn} and the τ_{mn} both constitute frames. Moreover (see [35]), both the σ_{mn} and the τ_{mn} constitute bases in $L^2(\mathbb{R})$.

It is now easy to construct the example previously announced. Take $h(x) = \tau(x)$. Then h is of compact support, and h belong to $\mathscr{H}_{5/2-\epsilon}(\mathbb{R})$, for all $\epsilon > 0$, because of (2.4.5). Since the $h_{mn}(x) = 2^{-m/2}h(2^{-m}x - n)$ constitute a frame, we can construct the dual frame, $(h_{mn})^{\sim}$. Since the h_{mn} are linearly independent, however, and because of (2.4.4), we find $(h_{mn})^{\sim} = \sigma_{mn}$. Since $\sigma \in \mathscr{H}_s(\mathbb{R})$ only if $s < 1/2 - \epsilon$, we see that *all* the functions in the dual frame are much less "nice" than those in the original frame; in particular, they are discontinuous, while h itself, together with its first derivative, is continuous.

2) Weyl–Heisenberg Frames in Sobolev Spaces: The examples in the preceding subsection show that one must be wary when trying to generalize frames, and the associated expansions, to other function spaces than $L^2(\mathbb{R})$. One can however apply the techniques used in Section II-C to extend the notion of frame to a "strip" of Sobolev spaces $\mathscr{H}_s$, with $|s| < s_0$.

Proposition 2.11: Define (as in Section II-C-1)

$$m(\hat{g}; p_0) = \inf_{x\in\mathbb{R}} \sum_{m\in\mathbb{Z}} |\hat{g}(x + mp_0)|^2,$$

$$M(\hat{g}; p_0) = \sup_{m\in\mathbb{Z}} \sum_{m\in\mathbb{Z}} |\hat{g}(x + mp_0)|^2.$$

Assume that $m(\hat{g}; p_0) > 0$, $M(\hat{g}; p_0) < \infty$. Define, for $s \geq 0$,

$$\beta_s^{\pm}(y) = \sup_x (1+x^2)^{\mp s/2}\left[1 + (x+y)^2\right]^{\pm s/2} \cdot \sum_{m\in\mathbb{Z}} |\hat{g}(x + mp_0)|\,|\hat{g}(x + mp_0 + y)|.$$

If

$$\sum_{k\neq 0} \left[\beta_s^{+}\left(\frac{2\pi}{q_0}k\right) \beta_s^{-}\left(-\frac{2\pi}{q_0}k\right) \right]^{1/2} < m(\hat{g}; p_0) \quad (2.4.6)$$

then the operator $\mathbb{T}$,

$$\mathbb{T} = \sum_{m,n} g_{mn}\langle g_{mn}, \cdot\rangle$$

is bounded, with a bounded inverse, on both $\mathscr{H}_s$ and $\mathscr{H}_{-s}$. In particular, for all $f_1 \in \mathscr{H}_s$, $f_2 \in \mathscr{H}_{-s}$,

$$A_s\|f_1\|_s \leq \|\mathbb{T}f_1\|_s \leq B_s\|f_1\|_s \quad (2.4.7)$$

$$A_s\|f_2\|_{-s} \leq \|\mathbb{T}f_2\|_{-s} \leq B_s\|f_2\|_{-s} \quad (2.4.8)$$

where

$$A_s = \frac{2\pi}{q_0}\left\{ m(\hat{g}; p_0) - \sum_{k\neq 0}\left[\beta_s^{+}\left(\frac{2\pi}{q_0}k\right) \beta_s^{-}\left(-\frac{2\pi}{q_0}k\right) \right]^{1/2} \right\} \quad (2.4.9)$$

and

$$B_s = \frac{2\pi}{q_0}\left\{ M(\hat{g}; p_0) + \sum_{k\neq 0}\left[\beta_s^{+}\left(\frac{2\pi}{q_0}k\right) \beta_s^{-}\left(-\frac{2\pi}{q_0}k\right) \right]^{1/2} \right\}. \quad (2.4.10)$$

Proof: We shall show that, for all $f \in \mathscr{H}_s$,

$$A_s\|f\|_s\|f\|_{-s} \leq \langle f, \mathbb{T}f\rangle \leq B_s\|f\|_s\|f\|_{-s}. \quad (2.4.11)$$

Here, as before, we use the notation $\langle\ ,\ \rangle$ for the duality extending the L^2-inner product,

$$\langle f_2, f_1\rangle = \int dx\, \overline{f_2(x)} f_1(x).$$

With respect to $\langle\ ,\ \rangle$, the spaces $\mathscr{H}_s$ and $\mathscr{H}_{-s}$ are each other's dual. On the other hand $\mathbb{T}$ is symmetric with respect to $\langle\ ,\ \rangle$. Since $\mathscr{H}_s$ is dense in $\mathscr{H}_{-s}$, it follows therefore from the upper bound in (2.4.11) that $\mathbb{T}$ is bounded on both $\mathscr{H}_s$ and $\mathscr{H}_{-s}$, with norm smaller than B_s. Similarly the lower bound in (2.4.11) implies that, for all $f_1 \in \mathscr{H}_s$, $f_2 \in \mathscr{H}_{-s}$, $\|\mathbb{T}f_1\|_s \geq A_s\|f_1\|_s$, and $\|\mathbb{T}f_2\|_{-s} \geq A_s\|f_2\|_{-s}$. Again using the symmetry of $\mathbb{T}$ with respect to $\langle\ ,\ \rangle$ one easily checks that this implies that $\mathbb{T}$ is invertible, with a bounded inverse with norm smaller than A_s^{-1}, both on $\mathscr{H}_s$ and $\mathscr{H}_{-s}$.

It remains to prove (2.4.11). This is done along the same lines as in Section II-C.

$$\begin{aligned}\langle f, \mathbb{T}f\rangle &= \frac{2\pi}{q_0}\sum_{m,k}\int dx\, \hat{g}(x+mp_0)\,\hat{g}\left(x + mp_0 + \frac{2\pi}{q_0}k\right)^* \cdot \hat{f}(x)^*\hat{f}\left(x + \frac{2\pi}{q_0}k\right) \\ &= \frac{2\pi}{q_0}\int dx\left[(1+x^2)^{s/2}|\hat{f}(x)|\right] \cdot \left[(1+x^2)^{-s/2}|\hat{f}(x)|\right]\sum_m |\hat{g}(x+mp_0)|^2 + r,\end{aligned}$$

with

$$|r| \leq \frac{2\pi}{q_0} \sum_{k \neq 0} \left\{ \int dx\,(1+x^2)^s |\hat{f}(x)|^2 \sum_m (1+x^2)^{-s/2} \cdot \left[1+\left(x+\frac{2\pi}{q_0}k\right)^2\right]^{s/2} |\hat{g}(x+mp_0)| \left|\hat{g}\left(x+mp_0+\frac{2\pi}{q_0}k\right)\right| \right\}^{1/2}$$

$$\cdot \left\{ \int dx\,(1+x^2)^{-s} |\hat{f}(x)|^2 \sum_m (1+x^2)^{s/2} \cdot \left[1+\left(x-\frac{2\pi}{q_0}k\right)^2\right]^{-s/2} |\hat{g}(x+mp_0)| \left|\hat{g}\left(x+mp_0-\frac{2\pi}{q_0}k\right)\right| \right\}^{1/2}$$

$$\leq \frac{2\pi}{q_0}\|f\|_s \|f\|_{-s} \sum_{k\neq 0}\left[\beta_s^+\left(\frac{2\pi}{q_0}k\right)\beta_s^-\left(\frac{2\pi}{q_0}k\right)\right]^{1/2}.$$

The bounds (2.4.11) then follow immediately. □

Remarks:

1) If the conditions of Proposition 2.11 are satisfied, then, since $\mathbb{T}$ is invertible, and has bounded inverse on $\mathscr{H}_s$, we have

$$\tilde{g}_{mn} = \mathbb{T}^{-1} g_{mn} \in \mathscr{H}_s$$

and, for all $f \in \mathscr{H}_s$,

$$f = \sum_{m,n} \tilde{g}_{mn} \langle g_{mn}, f\rangle$$

where the series converges in $\mathscr{H}_s$: the frame expansion holds on $\mathscr{H}_s$.

2) If $\mathbb{T}$ is a bounded and invertible operator on both $\mathscr{H}_s$ and $\mathscr{H}_{-s}$, then, by standard (highly nontrivial) interpolation theorems (see e.g., Section IX-4 in [64]), it is automatically bounded and invertible on any $\mathscr{H}_{s'}$ with $|s'| \leq s$. Proposition 2.11 gives thus a sufficient condition for the frame expansion formula to hold on a "strip" of Sobolev spaces. To obtain a wide "strip," for fixed p_0, it is clear from (2.4.6) that we have to choose q_0 sufficiently small. Typically, we would expect the critical value $q_0^c(s)$ (the value of q_0 for which equality occurs in (2.4.6)) to be a decreasing function of s. Note that one needs to impose a decay condition on $\beta^{\pm}(s)$, similar to (2.3.13), to ensure the existence of $q_0^c(s)$.

3) Since $\sum_{m\in\mathbb{Z}} |\hat{g}(x+mp_0)|\,|\hat{g}(x+mp_0+y)|$ is periodic in x, with period p_0, and since $[1+(x+y)^2]/(1+x^2)$ has its maximum, resp. minimum, at $x = -y/2 \pm \operatorname{sgn}(y)\sqrt{1+y^2/4}$, it suffices, for numerical computation of $\beta_s^{\pm}(y)$, to take the supremum over values of x in an interval of length p_0 around $-y/2 \pm \operatorname{sgn}(y)\sqrt{1+y^2/4}$.

Example: For $g(x) = \pi^{-1/4}\exp(-x^2/2)$, Table VI-A gives the values of A_s and B_s for a few values of s, for $p_0 = \pi/2$, $q_0 = 1$. Table VI-B shows how $q_0^c(s)$ changes with s, for fixed $p_0 = \pi/2$.

TABLE VI-A
FRAME BOUNDS*

s	A_s	B_s	B_s/A_s
0.0	3.853	4.147	1.076
1.0	3.852	4.148	1.077
2.0	3.849	4.151	1.079
3.0	3.836	4.164	1.086
4.0	3.787	4.213	1.112
5.0	3.600	4.400	1.222
6.0	2.865	5.135	1.793

*In the Sobolev spaces $\mathscr{H}_s$ for Weyl–Heisenberg frames based on the Gaussian $g(x) = \pi^{1/4}\exp(-x^2/2)$, with $p_0 = \pi/2$, $q_0 = 1.0$, for changing values of s. For $s = 7$, our estimate for A_s becomes negative: the frame breaks down.

TABLE VI-B*

s	$q_0^c(s)$
0.0	3.99
1.0	2.28
2.0	1.72
3.0	1.42
4.0	1.24
5.0	1.13
6.0	1.06
7.0	0.99

*The critical value $q_0^c(s)$ for the translation parameter q_0, as a function of the index s of the Sobolev space in which the frame is consdered. Values of q_0 smaller than $q_0^c(s)$ lead to a frame in $\mathscr{H}_s$ and $\mathscr{H}_{-s}$. The basic function here is $g(x) = \pi^{1/4}\exp(-x^2/2)$; $p_0 = \pi/2$ is fixed.

3) Wavelet Frames in Sobolev Spaces: Since the techniques used for proving frame bounds in $L^2(\mathbb{R})$ were essentially the same for the wavelet case as for the Weyl–Heisenberg case, it is not surprising that we can prove the following proposition.

Proposition 2.12: Define (as in Theorem 2.7)

$$m(h;a_0) = \inf_x \sum_{m\in\mathbb{Z}} \left|\hat{h}(a_0^m x)\right|^2,$$

$$M(h;a_0) = \sup_x \sum_{m\in\mathbb{Z}} \left|\hat{h}(a_0^m x)\right|^2.$$

Assume that $m(h;a_0) > 0$, $M(h;a_0) < \infty$. Define, for $s \geq 0$,

$$\beta_s^{\pm}(y) = \sup_x (1+x^2)^{\mp s/2} \sum_{m\in\mathbb{Z}} \left|\hat{h}(a_0^m x)\right| \left|\hat{h}(a_0^m x + y)\right| \cdot \left[1+(x+a_0^{-m}y)^2\right]^{\pm s/2}.$$

If

$$\sum_{k\neq 0}\left[\beta_s^+\left(\frac{2\pi}{b_0}k\right)\beta_s^-\left(-\frac{2\pi}{b_0}k\right)\right]^{1/2} < m(h;a_0) \quad (2.4.12)$$

then the operator $\mathbb{T}$,

$$\mathbb{T} = \sum_{m,n} h_{mn}\langle h_{mn}, \cdot\rangle$$

is bounded, with a bounded inverse, on both $\mathscr{H}_s$ and

$\mathscr{H}_{-s}$. In particular, for all $f_1 \in \mathscr{H}_s$, $f_2 \in \mathscr{H}_{-s}$,

$$A_s\|f_1\|_s \le \|\mathbb{T}f_1\|_s \le B_s\|f_1\|_s \tag{2.4.13}$$

$$A_s\|f_2\|_{-s} \le \|\mathbb{T}f_2\|_{-s} \le B_s\|f_2\|_{-s} \tag{2.4.14}$$

where

$$A_s = \frac{2\pi}{b_0}\left\{m(h;a_0) - \sum_{k\neq 0}\left[\beta_s^+\left(\frac{2\pi}{b_0}k\right)\beta_s^-\left(-\frac{2\pi}{b_0}k\right)\right]^{1/2}\right\}$$

$$B_s = \frac{2\pi}{p_0}\left\{M(h;a_0) - \sum_{k\neq 0}\left[\beta_s^+\left(\frac{2\pi}{b_0}k\right)\beta_s^-\left(-\frac{2\pi}{b_0}k\right)\right]^{1/2}\right\}.$$

Proof: The proof is entirely analogous to the proof of Proposition 2.11. □

Remarks:

1) If $\mathbb{T}$ is bounded, with a bounded inverse, on $\mathscr{H}_s$, then

$$(h_{mn})^{\sim} = \mathbb{T}^{-1}h_{mn} \in \mathscr{H}_s$$

and, for all $f \in \mathscr{H}_s$,

$$f = \sum_{m,n}(h_{mn})^{\sim}\langle h_{mn}, f\rangle \tag{2.4.15}$$

where the series converges in $\mathscr{H}_s$. This means that the phenomenon illustrated by the examples in Section II-D-1 cannot happen, at least in $\mathscr{H}_s$, if h satisfies the conditions in Proposition 2.12.

2) By interpolation (see e.g., Section IX-4 in [64]), one finds that if (2.4.13), (2.4.14) hold in $\mathscr{H}_s$, $\mathscr{H}_{-s}$, respectively, then $\mathbb{T}$ is bounded, with a bounded inverse, on all $\mathscr{H}_{s'}$ with $|s'| \le s$. Under the conditions in Proposition 2.12, the frame expansion (2.4.15) is therefore valid in all $\mathscr{H}_{s'}$ with $|s'| \le s$.

3) Typically (as in the Weyl–Heisenberg case), we expect that, for fixed a_0, the critical value $b_0^c(s)$ (i.e., the value of b_0 for which equality holds in (2.4.12)) decreases with s. It is a remarkable fact that for Y. Meyer's basis, the wavelet expansion is valid in $\mathscr{H}_s$, for all $s \in \mathbb{R}$, for $a_0 = 2$, and for *fixed* $b_0 = 1$. Something similar is true for the tight frames constructed in Section II-B-2b). However, for general functions h, leading to nontight frames, we rather expect $b_0^c(s)$ to decrease with s.

4) For a given function h, the strip of Sobolev spaces for which the h_{mn} constitute a frame, i.e., the possible values of s for which (2.4.13), (2.4.14) holds, is constrained by the behavior of $\hat{h}$ around zero. If $\hat{h}(x) = O(x^a)$ for $x \to 0$, then clearly (see the definition of β_s^+)

$$\beta_s^+(y) \ge C\left[\inf_{|u|\le a_0^{-M}}|\hat{h}(y+u)|\,|y+u|^s\right]\sum_{m=-\infty}^{-M} a_0^{m(\alpha-s)}.$$

This diverges for $s \ge \alpha$.

TABLE VII
FRAME BOUNDS*

	$b_0 = 1.0$				$b_0 = 1.25$		
s	A_s	B_s	B_s/A_s	s	A_s	B_s	B_s/A_s
0.00	3.223	3.596	1.116	0.00	2.001	3.454	1.726
0.25	3.223	3.596	1.116	0.25	1.995	3.460	1.734
0.50	3.222	3.597	1.116	0.50	1.985	3.470	1.748
0.75	3.221	3.597	1.117	0.75	1.971	3.484	1.768
1.00	3.220	3.599	1.118	1.00	1.953	3.502	1.793
1.25	3.218	3.600	1.119	1.25	1.926	3.529	1.832
1.50	3.216	3.602	1.120	1.50	1.877	3.578	1.906
1.75	3.214	3.605	1.122	1.75	1.794	3.661	2.041

In the Sobolev spaces $\mathscr{H}_s$ for wavelet frames based on the Mexican hat function $h(x) = 23^{-1/2}\pi^{-1/4}(1-x^2)\exp(-x^2/2)$, with $a_0 = 2.0$, $b_0 = 1.0$ and 1.25 for changing values of s.

Example: Table VII gives A_s and B_s, for a few values of s, for the Mexican hat function $h(x) = 2/\sqrt{3}\pi^{-1/4}(1-x^2)e^{-x^2/2}$, for $a_0 = 2.0$, and for $b_0 = 1.0$ and 1.25.

III. Phase Space Localization

Let us recall the intuition, already mentioned in the introduction, which leads us to expect the phase space localization results we shall give here.

Both the Weyl–Heisenberg coherent states and the wavelets can be used to give a representation in the time-frequency plane of time-dependent signals, provided the basic functions g or h and the parameters p_0, q_0 or a_0, b_0 are suitably chosen (see Section II). A graphical picture of these representations is given in Fig. 1. Ideally, one would like these representations to be reasonably "sharp." If e.g., the pair (m_0, n_0) corresponds to the time t_0 and the frequency ω_0 (see Fig. 1), then we would wish that the frequency content, in a band around ω_0, of the signal f, during a time interval around t_0, is essentially mirrored by the coefficients $\langle g_{mn}, f\rangle$ or $\langle h_{mn}, f\rangle$ with m close to m_0, n close to n_0. It is intuitively clear that the basic analyzing functions g or h have to be themselves well localized in time and frequency for such "sharpness" of the associated representations to be attainable.

In this section we shall try to make these qualitative statements more precise. We shall do this by giving a sense to the "sharpness" of the time-frequency representation, and by showing how the localization, in time and frequency, of the analyzing functions g or h matters. The Weyl–Heisenberg case shall be discussed in Section III-A, the wavelet case in Section III-B. In both cases we shall see that signals that are essentially limited to a given finite time interval and to a given finite range in frequency can essentially be represented by a finite number of expansion coefficients. (All these qualifications will be made more quantitative next). We shall use this fact to explain, in Section III-C, a phenomenon that was first noticed by J. Morlet in numerical calculations. To reconstruct a signal f with a precision ϵ, it is sufficient, for some frames, to calculate the expansion coefficients with a precision $C\epsilon$, where C turns out to be significantly larger than would be expected for orthonormal bases. We have called this phenomenon the "reduction of calculational noise" (for example, it often reduces round-off errors), and we explain it in Section III-C.

A. The Weyl–Heisenberg Case

Assume that g is normalized, $\int dx\,|g(x)|^2 = 1$, and that

$$\int dx x|g(x)|^2 = 0 = \int dy y|\hat{g}(y)|^2.$$

Then g_{mn} is localized around the phase space point (mp_0, nq_0),

$$\int dx x|g_{mn}(x)|^2 = nq_0, \qquad \int dy y|(g_{mn})^{\wedge}(y)|^2 = mp_0.$$

Suppose, on the other hand, that we restrict ourselves to the analysis and reconstruction of signals that are essentially time-limited to the interval $[-T,T]$ and essentially band-limited (limited in frequency) to the interval $[-\Omega,\Omega]$. We have introduced the word "essentially" in this statement because, as is well-known, no function f can have both a compact support and a Fourier transform with compact support. A more precise statement of the "essential" time- and band-limitedness is

$$\|(1-Q_T)f\| \ll \|f\|$$
$$\|(1-P_\Omega)f\| \ll \|f\|$$

where

$$(Q_T f)(t) = \chi_{[-T,T]}(t) f(t)$$
$$(P_\Omega f)^{\wedge}(w) = \chi_{[-\Omega,\Omega]}(w)\hat{f}(w)$$

and where χ_i denotes the indicator function of the interval I. Effectively, these limitations single out a rectangle of phase space as more important than other regions.

We shall assume, in what follows, that the g_{mn} constitute a frame, with frame bounds A and B and with dual frame $\tilde{g}_{mn}$. The reconstruction formula, valid for all functions in $L^2(\mathbb{R})$, and therefore in particular for the functions f of interest here, is (see Section II-A)

$$f = \sum_{m,n\in\mathbb{Z}} \tilde{g}_{mn}\langle g_{mn}, f\rangle. \tag{3.1}$$

If the g_{mn} are "well localized" in phase space around (mp_0, nq_0), then it is to be expected that $\langle g_{mn}, f\rangle$ will be small if the distance, in phase space, from (mp_0, nq_0) to the rectangle $[-\Omega,\Omega]\times[-T,T]$, is large. In other words, one expects that only those m,n for which (mp_0, nq_0) lies in or close to $[-\Omega,\Omega]\times[-T,T]$ will play a significant role in the reconstruction (3.1) of f. Fig. 6 represents the situation. As in Fig. 1(a), the g_{mn} are represented by their "phase space centers" (mp_0, nq_0).

The signal f is essentially concentrated, in phase space, on the rectangle $[-\Omega,\Omega]\times[-T,T]$. Under suitable conditions on g, it turns out that, for all $\epsilon > 0$, there exist $t_\epsilon, \omega_\epsilon$ such that the partial reconstruction of f using only the (finitely many) m,n associated to the "enlarged rectangle" $[-(\Omega+\omega_\epsilon),(\Omega+\omega_\epsilon)]\times[-(T+t_\epsilon),(T+t_\epsilon)]$ (in

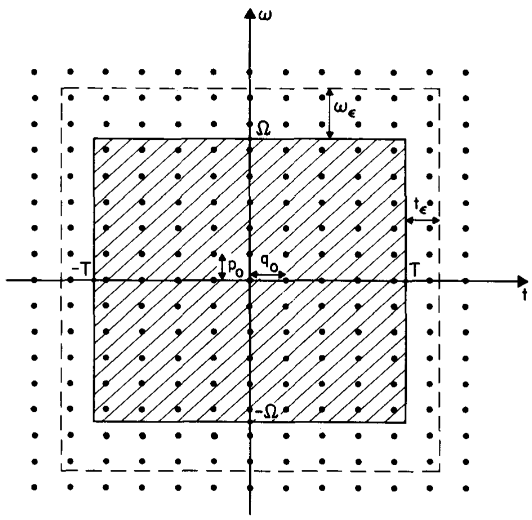

Fig. 6. Rectangular lattice (nq_0, mp_0) indicating localization of g_{mn}, and rectangle in phase space $[-T,T]\times[-\Omega,\Omega]$ on which signal f is mainly concentrated. Coefficients $\langle g_{mn}, f\rangle$ corresponding to (nq_0, mp_0) in enlarged rectangle (in dashed lines) suffice to reconstruct f up to error of order ϵ.

dashed lines in the figure), is equal to f, up to an error of order ϵ. This is essentially the content of the following theorem.

Theorem 3.1: Suppose that the $g_{mn}(x) = e^{imp_0x}g(x - nq_0)$ constitute a frame, with frame bounds A, B. Assume that

$$|g(x)| \le C(1+x^2)^{-\alpha}, \qquad |\hat{g}(y)| \le C(1+y^2)^{-\alpha}$$

for some $C < \infty$, $\alpha > 1/2$. Then, for any $\epsilon > 0$, there exist $t_\epsilon, \omega_\epsilon > 0$ such that, for all $f \in L^2(\mathbb{R})$, and for all T, $\Omega > 0$,

$$\left\| f - \sum_{\substack{|mp_0| \le \Omega + \omega_\epsilon \\ |nq_0| \le T + t_\epsilon}} \tilde{g}_{mn}\langle g_{mn}, f\rangle \right\| \le \sqrt{\frac{B}{A}}\,\big[\|(\mathbb{1} - P_\Omega)f\| + \|(\mathbb{1} - Q_T)f\| + \epsilon\|f\|\big] \tag{3.2}$$

where the $\tilde{g}_{mn}$ constitute the dual frame to the g_{mn}.

Proof: See Appendix D. □

Remarks:

1) In the limit as $\epsilon \to 0$, one finds $t_\epsilon \to \infty$ or $\omega_\epsilon \to \infty$. (With our proof, in Appendix D, both t_ϵ and ω_ϵ tend to ∞ as $\epsilon \to 0$. If g or $\hat{g}$ has compact support, then t_ϵ or ω_ϵ can be kept finite. In all cases at least one of $t_\epsilon, \omega_\epsilon$ must diverge as $\epsilon \to 0$.) This is natural; infinite precision cannot be obtained when using only a finite set of g_{mn}. In particular, for $g(x) = \pi^{-1/4}\exp(-x^2/2)$, one finds $\omega_\epsilon, t_\epsilon \sim_{\epsilon\to 0} C|\log\epsilon|^{1/2}$.
2) Note that the decay conditions on g and $\hat{g}$ exclude all orthonormal bases g_{mn}, by Theorem 2.3.
3) The important fact about Theorem 3.1 is that $t_\epsilon, \omega_\epsilon$ are *independent* of T, Ω: the "enlargement proce-

dure" depends only on ϵ, the desired precision of the finite construction. For fixed ϵ, the number of points $N_\epsilon(T,\Omega)$ used in this finite reconstruction is $4(T+t_\epsilon)(\Omega+\omega_\epsilon)/p_0\cdot q_0$. Hence

$$\lim_{T,\Omega\to\infty}(4T\Omega)^{-1}N_\epsilon(T,\Omega)=(p_0\cdot q_0)^{-1}, \quad (3.3)$$

which is independent of ϵ. This can be compared with an analogous result involving prolate spheroidal wave functions (see [65], [66]; the formula discussed here is explained very clearly in [66]). The prolate spheroidal wave functions are the eigenfunctions of the compact operator $P_\Omega Q_T P_\Omega$ (which also describes localization in phase space, singling out the rectangle $[-\Omega,\Omega]\times[-T,T]$). Signals f that are essentially time-limited to $[-T,T]$ and bandlimited to $[-\Omega,\Omega]$ can be expanded in prolate spheroidal wave functions. Again, if a reconstruction up to a finite error ϵ is desired, one can truncate the expansion to a finite sum. The number, $n_\epsilon(T,\Omega)$, of terms needed is the number of eigenfunctions of $P_\Omega Q_T P_\Omega$ with eigenvalue larger than ϵ. One has [65], [66],

$$n_\epsilon(T,\Omega)=\frac{2T\Omega}{\pi}+C\epsilon\log(T\Omega)+o(\epsilon).$$

In this case

$$\lim_{T,\Omega\to\infty}(4T\Omega)^{-1}n_\epsilon(T,\Omega)=(2\pi)^{-1}. \quad (3.4)$$

This limit is exactly the Nyquist density. In fact, this result was the first rigorous formulation of the intuitive Nyquist density idea [65]. If we compare (3.3) with (3.4), then we see that in our present approach, the density $(p_0\cdot q_0)^{-1}$ is higher than the Nyquist density $(2\pi)^{-1}$. We know indeed (see Section II-C-1a)) that $p_0\cdot q_0\le 2\pi$ for all frames $\{g_{mn};\ m,n\in\mathbb{Z}\}$. For the examples with good phase space localization (measured by e.g., higher moments of $|g|$ and $|\hat{g}|$), we even have, by Theorem 2.3, that $p_0\cdot q_0<2\pi$. The "oversampling" of our phase space density with respect to the optimal Nyquist density, measured by the ratio $2\pi(p_0\cdot q_0)^{-1}>1$, is the price we have to pay for the fact that the frames of interest to us are, in general, not orthonormal. We also gain something with respect to the prolate spheroidal wave functions, however. The construction of the different g_{mn}, obtained from one function g by translations in phase space, is simpler than the construction of the (orthonormal) prolate spheroidal wave functions.

Note that (3.3) can in fact be considered as a *definition* of the phase space density for the frame under consideration. While we have used the term "phase space density" before in heuristic discussions, (3.3) is the first mathematically precise statement justifying this terminology.

Example: In the Table VIII we give the values of $t_\epsilon=\omega_\epsilon$ corresponding to given ϵ for $g(x)=\pi^{-1/4}\exp(-x^2/2)$, in the cases $p_0=q_0=\pi^{1/2}$ and $p_0=q_0=(\pi/2)^{1/2}$. In each case we have used (D.6) in the estimates.

TABLE VIII
THE "ENLARGEMENT PARAMETER" t_ϵ*

ϵ	$t_\epsilon=\omega_\epsilon$
$p_0=q_0=\sqrt{\pi/2}=1.25$	
0.1	2.52
0.05	2.78
0.01	3.31
0.005	3.51
0.001	3.94
0.0005	4.11
0.0001	4.48
$p_0=q_0=\sqrt{\pi}=1.77$	
0.1	2.55
0.05	2.81
0.01	3.33
0.005	3.54
0.001	3.97
0.0005	4.14
0.0001	4.50

*As function of the error ϵ, for the Weyl-Heisenberg frame of Gaussians, respectively for $p_0=q_0=\sqrt{\pi/2}$, $p_0=q_0=\sqrt{\pi}$ (see text)

B. The Wavelet Case

In this case we assume that h is normalized so that $\int dx\,|h(x)|^2=1$, and that $\int dx x|h(x)|^2=0$. Let us assume, for the sake of simplicity, that $|\hat{h}|$ is even (in practice, $\hat{h}$ is either even or odd; even if the frame is constructed by means of a function h with support $\hat{h}\subset\mathbb{R}_+$, then the effective basic wavelets $h^1=\mathrm{Re}\,h$, $h^2=\mathrm{Im}\,h$ satisfy the condition that $|\hat{h}^j|$ is even—see the remarks following Theorem 2.7 and 2.8). Let us also assume that

$$\int_0^\infty dk\,k|\hat{h}(k)|^2=1$$

(this does not imply any loss of generality: it can easily be achieved by dilating h). Then h_{mn} is concentrated around the phase space points $(\pm a_0^{-m}, a_0^m nb_0)$

$$\int dx\,x|h_{mn}(x)|^2=a_0^m nb_0$$

$$\int_0^\infty dk\,k|(h_{mn})^\wedge(k)|^2=a_0^{-m}=-\int_{-\infty}^0 dk\,k|(h_{mn})^\wedge(k)|^2.$$

Again, we suppose we are mainly interested in functions f localized in phase space. In this case, we assume that they are essentially time-limited to $[-T,T]$, and with frequencies $|\omega|$ mainly concentrated in $[\Omega_0,\Omega_1]$, where $0<\Omega_0<\Omega_1<\infty$. The need for a lower bound Ω_0 on the frequencies $|\omega|$, as opposed to the Weyl–Heisenberg case, where we only introduced an upper bound, stems from the logarithmic rather than linear treatment of frequencies by the wavelet transform.

Fig. 7 represents the situation graphically. The dots represent the "phase space localization centers" of the h_{mn}; for the sake of simplicity we have taken $a_0=2$. The

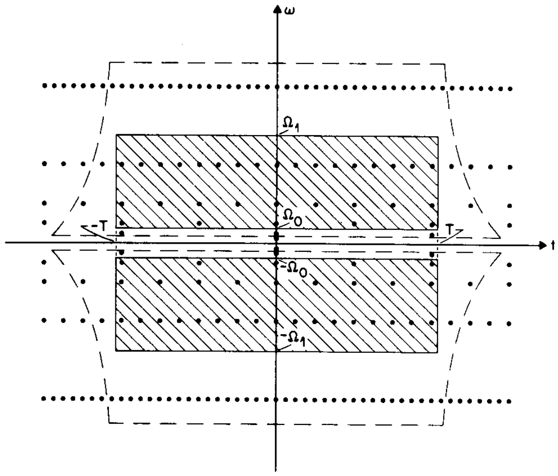

Fig. 7. Lattice $(nb_0a_0^m, \pm a_0^{-m}k_0)$, indicating localization of h_{mn} (see Fig. 1(b)), and two rectangles $[-T,T]\times[\Omega_0,\Omega_1]$, $[-T,T]\times[-\Omega_1, -\Omega_0]$ on which signal f is mainly concentrated. Coefficients $\langle h_{mn}, f\rangle$ corresponding to lattice points within set B_ϵ (in dashed lines) suffice to reconstruct f up to error ϵ.

signals f of interest are essentially concentrated in the set $[-T,T]\times([(-\Omega_1,-\Omega_0]\cup[\Omega_0,\Omega_1])$, marked with full lines in the figure.

We shall assume, of course, that the h_{mn} constitute a frame, with frame bounds A, B and with dual frame $(h_{mn})^{\sim}$. The associated reconstruction formula is then

$$f=\sum_{m,n\in\mathbb{Z}}(h_{mn})^{\sim}\langle h_{mn},f\rangle. \tag{3.5}$$

This formula is valid for all functions in $L^2(\mathbb{R})$. For signals essentially concentrated in $[-T,T]\times([-\Omega_1,-\Omega_0]\cup[\Omega_0,\Omega_1])$ we can again restrict ourselves to a finite subset $B_\epsilon(T,\Omega_0,\Omega_1)$ of indexes, provided the basic wavelet h is itself sufficiently concentrated. Partial reconstruction of such a signal f, using only the pairs $(m,n)\in B_\epsilon$, is then an approximation of f with an error at most equal to ϵ. The set B_ϵ includes all the (m,n) for which $\Omega_0\le a_0^{-m}\le\Omega_1$, and $|a_0^m nb_0|\le T$; it is indicated in dashed lines in Fig. 7. The following theorem gives an exact statement of this result.

Theorem 3.2: Suppose that the

$$h_{mn}(x)=a_0^{-m/2}h(a_0^{-m}x-nb_0)$$

constitute a frame, with frame bounds A, B, and dual frame $(h_{mn})^{\sim}$. Assume that

$$|\hat{h}(y)|\le C|y|^{\beta}(1+y^2)^{-(\alpha+\beta)/2}$$

where $\beta>0$, $\alpha>1$, and that, for some $\gamma>1/2$

$$\int dx(1+x^2)^{\gamma}|h(x)|^2<\infty.$$

Fix $T>0$, $0<\Omega_0<\Omega_1$. Then, for any $\epsilon>0$, there exists a finite subset $B_\epsilon(T,\Omega_1,\Omega_2)$ of $\mathbb{Z}^2$ such that, for all $f\in L^2(\mathbb{R})$,

$$\left\|f-\sum_{(m,n)\in B_\epsilon(T,\Omega_1,\Omega_2)}(h_{mn})^{\sim}\langle h_{mn},f\rangle\right\|$$
$$\le(B/A)^{1/2}\left[\|(\mathbb{1}-Q_T)f\|+\left\|(\mathbb{1}-P_{\Omega_1}+P_{\Omega_0})f\right\|+\epsilon\|f\|\right]. \tag{3.6}$$

Proof: See Appendix D. □

Remarks:

1) The "enlargement procedure," i.e., the transition from the original "box" $[-T,T]\times([-\Omega_1,-\Omega_0]\cup[\Omega_0,\Omega_1])$ to $B_\epsilon(T,\Omega_0,\Omega_1)$, is more complicated here than in the Weyl–Heisenberg case; the parameters m_0, m_1, and t also depend on Ω_0,Ω_1 and not only on ϵ (see the proof in Appendix D).
2) The construction of the set B_ϵ (see also Fig. 7) exhibits the fact that wavelets give a higher resolution at high frequencies than at low frequencies. For fixed m, the "extension" of the box in the time-variable, as defined by (D.7), corresponds to consideration of all the pairs (m,n) with time-localization $|a_0^m nb_0|\le T+a_0^m t$. The "extension" $a_0^m t$ thus depends on m and is small for large negative values of m, which correspond to high frequencies.
3) A crude estimate leads to

$$\#B_\epsilon(T,\Omega_0,\Omega_1)$$
$$\le C\frac{2T}{b_0\ln a_0}\left(\epsilon^{-1/\beta}\Omega_1-\epsilon^{1/\alpha-\kappa}\Omega_0\right)$$
$$+C\frac{\epsilon^{-2}}{b_0\ln a_0}\left(C+(\ln\Omega_1-\ln\Omega_0)^2\right). \tag{3.7}$$

If $\Omega_0\to 0$, T, and $\Omega_1\to\infty$, then

$$(4T\Omega_1)^{-1}[\#B_\epsilon(T,\Omega_0,\Omega_1)]\xrightarrow{\le}\frac{C\epsilon^{-1/\beta}}{2b_0\ln a_0} \tag{3.8}$$

which is not independent of ϵ. While the estimate (3.7) is admittedly crude, finer estimates lead to a similar result. This is in contrast with the analogous result for the Weyl–Heisenberg case (see (3.3)) where the corresponding limit was independent of ϵ, and gave the phase space density of the frame. In fact, it gave a procedure to *define* the phase space density of the frame. Because of the ϵ-dependence of the right-hand side of (3.7), we cannot use the same procedure to define a phase space density corresponding to wavelet frames. This illustrates again (see also the discussion in Section II-C-2a)) that the concept "phase space density" is not well-suited to the wavelet representation.
4) The estimates made in the proof of Theorem 3.2 lead to rather crude bounds. As in the Weyl–Heisenberg case, it is possible to write more complicated, but less coarse estimates. For the choice

$h(x) = 2.3^{-1/2}\pi^{-1/4}(1-x^2)e^{-x^2/2}$, $a_0 = 2.0$, $b_0 = 0.5$, with $\Omega_0 = 10$, $\Omega_1 = 1000$, and $T = 100$, one finds

ϵ	m_0	m_1	t
0.1	−11	−2	4.95
0.05	−12	−2	4.97
0.01	−13	−2	5.75.

C. The Reduction of Calculational Noise

J. Morlet noticed, some time ago, that in numerical wavelet calculations, it often sufficed to calculate the wavelet coefficients to a precision of, say, 10^{-2}, to be able to reconstruct the original signal with a precision of, say, 10^{-3}. This rather surprising fact can be explained as a consequence of both phase space localization, and "oversampling."

Phase space localization is necessary to restrict oneself to a finite number of coefficients. We cannot hope to control an infinite number of coefficients if they can all induce an error of the same order. The role of "oversampling" is the following. Let us go back to the frame operator T (not to be confused with the time-limit T in Sections III-A and III-B) defined in Sections I-C and II-A. We have

$$T: L^2(\mathbb{R}) \to l^2(\mathbb{Z}^2)$$
$$(Tf)_{m,n} = \langle \phi_{mn}, f \rangle$$

where ϕ_{mn} stands for either h_{mn} or g_{mn}. Since the ϕ_{mn} constitute a frame, this operator is bounded and has a bounded inverse on its closed range. The operator T is onto (its range is all of $l^2(\mathbb{Z}^2)$) if and only if the ϕ_{mn} constitute a basis. In general, however, the ϕ_{mn} are not independent, and the range of T ($RanT$) is a proper subspace of $l^2(\mathbb{Z}^2)$. The inversion procedure,

$$f = \sum_{m,n} (\phi_{mn})^{\sim} \langle \phi_{mn}, f \rangle$$

when applied to elements c of $l^2(\mathbb{Z}^2)$ not necessarily in *Ran T*:

$$\sum_{m,n} (\phi_{mn})^{\sim} c_{mn}$$

in fact consists of 1) a projection of $l^2(\mathbb{Z}^2)$ onto *RanT*, 2) the inversion of T on its range (see Section II-A). We shall model the finite precision of numerical calculations by adding random "noise" to the coefficients $\langle \phi_{mn}, f \rangle$, thus leading to modified coefficients $c_{mn}(f)$. The "noise" component of these coefficients "lives" on all of $l^2(\mathbb{Z}^2)$. If we apply the inversion procedure, this component will therefore be reduced in norm by the projection onto *RanT*. This reduction will be the more pronounced the "smaller" *RanT* is, as a subspace of $l^2(\mathbb{Z}^2)$, i.e., the more pronounced the oversampling of redundancy in the frame. The calculations below show how this work in practice.

Let us assume that we are interested in signals f that are essentially localised in the time interval $[-T, T]$, and in the frequency range $[-\Omega, \Omega]$ (in the Weyl–Heisenberg case) or $[-\Omega_1, -\Omega_0] \cup [\Omega_0, \Omega_1]$ (in the wavelet case), i.e.,

$$\|(\mathbb{1} - Q_T)f\| \le \epsilon\|f\|$$
$$\|(\mathbb{1} - P_\Omega)f\| \le \epsilon\|f\| \quad \text{or} \quad \|(\mathbb{1} - P_{\Omega_1} + P_{\Omega_0})f\| \le \epsilon\|f\|.$$

Then, by Theorems 3.1 and 3.2, there exists an "enlarged box" B_ϵ such that

$$\left\| f - \sum_{(m,n)\in B_\epsilon} (\phi_{mn})^{\sim} \langle \phi_{mn}, f \rangle \right\| \le 3(B/A)^{1/2}\epsilon\|f\|$$

where ϕ denotes either g or h. Since B_ϵ is a finite subset of $\mathbb{Z}^2$, we restrict ourselves therefore to the finitely many coefficients $\langle \phi_{mn}, f \rangle$.

In practical calculations, the coefficients $\langle \phi_{mn}, f \rangle$ will be computed with finite precision. Let us take the following model for the errors. Assume that the coefficients to be used in the calculations are given by

$$c_{mn}(f) = \langle \phi_{mn}, f \rangle + \gamma_{mn}$$

where the γ_{mn} are independent identically distributed random variables, with mean zero, and with variance α^2,

$$\mathbb{E}(\gamma_{mn}^2) = \alpha^2.$$

This means that the $\langle \phi_{mn}, f \rangle$ are known with "precision" α. Note that our model is only a first approximation. In general the ϕ_{mn}, and hence the $\langle \phi_{mn}, f \rangle$, are not linearly independent, which means that the round-off errors should not be regarded as independent random variables. With this approximation, we find that the estimated error between f and a partial reconstruction, using only the finitely many coefficients associated to $(m,n) \in B_\epsilon$, and even those only with finite precision (i.e., replace $\langle \phi_{mn}, f \rangle$ by $c_{mn}(f)$), is given by

$$\begin{aligned}
&\mathbb{E}\left(\left\| f - \sum_{(m,n)\in B_\epsilon} c_{mn}(f)(\phi_{mn})^{\sim} \right\|^2\right) \\
&= \mathbb{E}\left(\left\| \left(f - \sum_{(m,n)\in B_\epsilon} (\phi_{mn})^{\sim} \langle \phi_{mn}, f \rangle \right) - \sum_{(m,n)\in B_\epsilon} \gamma_{mn}(\phi_{mn})^{\sim} \right\|^2\right) \\
&\le 9(B/A)\epsilon^2\|f\|^2 + A^{-2}\alpha^2 N_\epsilon
\end{aligned} \tag{3.9}$$

where $N_\epsilon = \#B_\epsilon$, and where we have used $\mathbb{E}(\gamma_{mn}) = 0$, $\mathbb{E}(\gamma_{mn}\gamma_{m'n'}) = \delta_{mm'}\delta_{nn'}\alpha^2$, and $\|(\phi_{mn})^{\sim}\|^2 = \|\mathbb{T}^{-1}\phi_{mn}\|^2 \le A^{-2}\|\phi_{mn}\|^2 = A^{-2}$ (with $\mathbb{T}$ as defined in Section II-A).

The "reduction of calculational noise," observed by J. Morlet, is contained in the second term in (3.9), more particular in the factor $N_\epsilon A^{-2}$. Let us show how.

Assume that we are considering a snug Weyl–Heisenberg frame, g_{mn}. If we assume that B_ϵ is large with respect to the lattice mesh, then (see Section III-A)

$$N_\epsilon = \#B_\epsilon \simeq \frac{4T\Omega}{p_0 \cdot q_0}.$$

On the other hand, if the frame is snug (i.e., $B \simeq A$), we find, by (2.2.9),

$$A \simeq 2\pi(p_0 \cdot q_0)^{-1}$$

(we assume $\|g\| = 1$). Hence

$$N_\epsilon A^{-2} \simeq \pi^{-2} T\Omega(p_0 \cdot q_0). \quad (3.10)$$

If the g_{mn} had constituted an orthonormal basis, then (provided we neglect the loss in phase space localisation due to the use of an orthonormal basis) this factor would have been

$$(N_\epsilon A^{-2})_{\text{orthon. basis}} \simeq \left(\frac{2T\Omega}{\pi}\right). \quad (3.11)$$

The frame gives thus a net gain of $2\pi/p_0 \cdot q_0$ with respect to the orthonormal basis situation.

Something similar happens for wavelets. In this case we don't have such a simple expression for N_ϵ, but we can easily see that the same phenomenon takes place by the following argument. Suppose h, a_0, b_0 are chosen so that the frame is snug (i.e., it is tight for all practical purposes). In particular this means that $m(h; a_0) \simeq M(h; a_0)$ while the β-terms in (2.3.23), (2.3.24) are negligible. Consequently $A \simeq B \simeq 2\pi b_0^{-1} m(h; a_0)$. Consider now the frame with the same h, a_0, but with $b_0' = b_0/2$. This frame will obviously also be snug, with $A' \simeq B' \simeq 2\pi b_0'^{-1} m(h; a_0) = 2A$. On the other hand, there are twice as many points in the graphical representation of this new frame, for every frequency level. Hence $N_\epsilon' = 2N_\epsilon$. Combining these two, we find $N_\epsilon' A'^{-2} = \frac{1}{2} N_\epsilon A^{-2}$, i.e., halving b_0 leads to a gain of 2 in the total error on f, for the same precision on the coefficients.

For the frames used by J. Morlet when he noticed this phenomenon, which were heavily oversampled (e.g., he used up to 15 "voices"—see the end of Section II-C-2a for a definition), a gain factor of 10 or more can be obtained easily. Note, however, that oversampling does not explain completely the observed calculational noise (or quantization noise) reduction. As in vision analysis [38] part of the reduction is a consequence of the fact that, unlike the original signal, the coefficients $c_{mn}(f)$ at every fixed m-level are distributed around zero, with a sharp peak at zero. This makes it possible to drastically reduce the number of quantization steps in the c_{mn}, without significantly altering the quality of the reconstructed signal [38].

IV. Conclusion

We have shown how to characterize functions f by means of a local time-frequency analysis, by computing their inner products with either Weyl–Heisenberg coherent states,

$$c_{mn}(f) = \int dx f(x) g(x - nx_0) e^{im\omega_0 x}$$

or with wavelets,

$$c_{mn}(f) = a_0^{-m/2} \int dx f(x) h(a_0^{-m} x - nb_0).$$

The first approach corresponds to the windowed Fourier transform, the second is the wavelet transform. The wavelet transform handles frequencies in a logarithmic rather than linear way, and seems better adapted to the analysis of acoustic and visual signals than the windowed Fourier transform.

In both cases we have formulated necessary and sufficient conditions for the stable reconstruction of f from the $c_{mn}(f)$. For such a stable reconstruction algorithm to exist, we require that, for some $A > 0$, $B < \infty$, and all $f \in L^2(\mathbb{R})$,

$$A\|f\|^2 \leq \sum_{m,n \in \mathbb{Z}} |c_{mn}(f)|^2 \leq B\|f\|^2. \quad (4.1)$$

We have provided a reconstruction algorithm, which converges at least as fast as a geometric series in $r = B/A - 1$, and we have presented efficient methods for estimating A and B. These methods are illustrated by many examples.

Finally, if g respectively, h is well localized in both time and frequency, and if (4.1) is satisfied, we have shown that the characterization of functions f by their coefficients $c_{mn}(f)$ is truly local in time-frequency. If f is essentially concentrated on a limited region in time-frequency, then only the $c_{mn}(f)$ corresponding to time-frequency points within or close to this region are needed for an approximate reconstruction of f.

Acknowledgment

It is a pleasure for me to thank the many people with whom I have had enjoyable discussions on the subject of this paper, who pointed out relevant references, and who raised interesting questions. First of all, I want to express my gratitude to A. Grossmann, who introduced me to the many properties of coherent states years ago, and more recently stimulated my interest in "frame problems." This paper would not have existed without his continued interest and encouragement. It has been a pleasure to discuss the topics in this paper with him.

The completed proof of the Balian theorem, in Section II-C-1a), is due to R. Coifman and S. Semmes. I want to thank them for letting me include it here. Thanks are also due to R. Howe and T. Steger for solving one of my conjectures and for drawing my attention to reference [49]. The refinements of the frame bounds in Section II-C-2a) are due to Ph. Tchamitchian. Section II-D, on frame bounds and convergence in other spaces than $L^2(\mathbb{R})$, would not have existed without Y. Meyer and Ph. Tchamitchian, who pointed out the problem to me, and who suggested that a modification of the L^2-estimates would lead to useful Sobolev-space estimates.

I would also like to thank P. Deift, H. Feichtinger, P. Jones, H. Landau, P. Lax, R. Littlejohn, M. Sirugue, and D. Thomson for many interesting discussions on one or several subjects of this paper. I am also grateful to C. Tomei and to J. R. Klauder, who suggested the terminology "tight frame" and "snug frame," respectively.

Finally, I would like to thank the referees for making many suggestions that improved the readability of this paper.

Appendix A
Resolutions of the Identity

By explicit computation we have

$$\int dp \int dq \langle f_1, g^{(p,q)}\rangle\langle g^{(p,q)}, f_2\rangle$$
$$= \int dp \int dq \int dx \int dy f_1(x)^* e^{ip(x-y)} g(x-q) g(y-q)^* f_2(y)$$
$$= \int dq \int dx \int dy f_1(x)^* g(x-q) g(y-q)^* f_2(y) 2\pi\delta(x-y)$$
$$= 2\pi \int dq \int dx f_1(x)^* |g(x-q)|^2 f_2(x)$$
$$= 2\pi \int dy \int dx f_1(x)^* f_2(x) |g(y)|^2 = 2\pi \|g\|^2 \langle f_1, f_2\rangle.$$

This proves (1.3a).

In the wavelet case, we work via the Fourier transform,

$$\int \frac{da}{a^2} \int db \langle f_1, h^{(a,b)}\rangle\langle h^{(a,b)}, f_2\rangle$$
$$= \frac{1}{2\pi} \int \frac{da}{a^2} \int db \int dx \int dy \hat{f}_1(x)^* |a| \hat{h}(ax) \hat{h}(ay)^* e^{ib(x-y)} \hat{f}_2(y)$$
$$= \int \frac{da}{a^2} \int dx \int dy \hat{f}_1(x)^* |a| \hat{h}(ax) \hat{h}(ay)^* \hat{f}_2(y) \delta(x-y)$$
$$= \int da \int dx \hat{f}_1(x)^* \hat{f}_2(x) |a|^{-1} |\hat{h}(ax)|^2$$
$$= \int dy \int dx \hat{f}_1(x)^* \hat{f}_2(x) |y|^{-1} |\hat{h}(y)|^2 = C_h \langle f_1, f_2\rangle.$$

Appendix B
Proof of Lemma 2.4

We assume that

$$|G(t,s)| \le b < \infty \tag{B.1}$$

and that G, $\partial_t G$ and $\partial_s G$ are locally square integrable. We show that, together with the "periodicity conditions"

$$G(t+1,s) = G(t,s)$$
$$G(t,s+1) = e^{2\pi i t} G(t,s) \tag{B.2}$$

the extra assumption

$$|G(t,s)| \ge a > 0 \tag{B.3}$$

necessarily leads to a contradiction. To do this we introduce an averaged version G_r of G. More precisely, define, for $r > 0$,

$$G_r(t,s) = \frac{1}{4r^2} \int_{|t'-t| \le r} dt' \int_{|s'-s| \le r} ds' G(t',s').$$

Then G_r is obviously continuous. For $|t-t'| < r$, $|s-s'| < r$ one has

$$|G_r(t,s) - G_r(t',s')| \le br^{-1}(|t-t'| + |s-s'|). \tag{B.4}$$

The function G_r does not satisfy the "periodicity" conditions (B.2), but instead a modified version of them:

$$G_r(t+1,s) = G_r(t,s)$$
$$G_r(t,s+1) = \frac{1}{4r^2} \int_{|t'-t|<r} dt' \int_{|s'-s|<r} e^{2\pi i t'} G(t',s')$$
$$= e^{2\pi i t} G_r(t,s) + \psi(t,s) \tag{B.5}$$

where

$$|\psi(t,s)| \le 2\pi br. \tag{B.6}$$

By choosing r small enough, we can make the deviation of (B.5) from (B.2), as small as desired. The only other ingredient needed is a set of bounds (upper and lower) on $|G_r|$. The upper bound is immediate from (B.1),

$$|G_r(t,s)| \le b.$$

For the lower bound we restrict ourselves to a neighborhood U of $[0,2]^2$. Choose r such that, for all $(t,s) \in U$,

$$\left[\int_{|t'-t| \le 2r} dt' \int_{|s'-s| \le 2r} ds' |\partial_t G(t',s')|^2\right]^{1/2} + \left[\int_{|t'-t| \le 2r} dt' \int_{|s'-s| \le 2r} ds' |\partial_s G(t',s')|^2\right]^{1/2} \le \epsilon$$

the value of ϵ will be fixed later. Then, for any $\alpha > 0$, $\alpha < 1$, and for all $(t,s) \in U$

$$\int_{|t'-t| \le \alpha r} dt' \int_{|s'-s| \le \alpha r} ds' |G(t',s') - G_r(t',s')| \tag{B.7}$$
$$\le \frac{1}{4r^2} \int_{|t'-t| \le \alpha r} dt' \int_{|s'-s| \le \alpha r} ds' \int_{|t''-t'| \le r} dt'' \cdot \int_{|s''-s'| \le r} ds'' \left[\left|\int_{t'}^{t''} dt''' \partial_t G(t''',s'')\right| + \left|\int_{s'}^{s''} ds''' \partial_s G(t',s''')\right|\right]$$
$$\le \frac{1}{4r^2} \int_{|t'''-t| \le (1+\alpha)r} dt''' \int_{|s''-s| \le (1+\alpha)r} ds'' |\partial_t G(t''',s'')| \cdot 8\alpha^2 r^3 + \frac{1}{4r^2} \int_{|t'-t| \le \alpha r} dt' \int_{|s'''-s| \le (1+\alpha)r} ds''' |\partial_s G(t',s''')| 8\alpha r^3$$
$$\le 8\alpha r^2 \left\{\left[\int_{|t'-t| \le 2r} dt' \int_{|s'-s| \le 2r} ds' |\partial_t G(t,s)|^2\right]^{1/2} + \left[\int_{|t'-t| \le 2r} dt' \int_{|s'-s| \le 2r} ds' |\partial_s G(t,s)|^2\right]^{1/2}\right\}$$
$$\le 8\alpha r^2 \epsilon. \tag{B.8}$$

We can however also compute a lower bound on (B.7). For $|t'-t| \le \alpha r$, $|s'-s| \le \alpha r$, we have, from (B.4),

$$|G(t',s') - G_r(t',s')|$$
$$\ge |G(t',s')| - |G_r(t',s') - G_r(t,s)| - |G_r(t,s)|$$
$$\ge a - 2\alpha b - |G_r(t,s)|. \tag{B.9}$$

Together with (B.8) this leads to

$$8\alpha r^2 \epsilon \ge \int_{|t'-t| \le \alpha r} dt' \int_{|s'-s| \le \alpha r} ds' |G(t',s') - G_r(t',s')|$$
$$\ge 4\alpha^2 r^2 (a - 2\alpha b - |G_r(t,s)|)$$

or

$$|G_r(t,s)| \ge a - 2\alpha b - 2\alpha^{-1}\epsilon. \tag{B.10}$$

If we choose $\alpha = a/8b$, $\epsilon = \alpha a/8 = a^2/64b$, then the right hand side of (B.10) reduces to $a/2$, and we conclude

$$a/2 \le |G_r(t,s)| \le b. \tag{B.11}$$

We have now all the necessary ingredients. Since G_r is continuous, and satisfies (B.11) on U, with $a > 0$, $b < \infty$, $\gamma_r = \log G_r$ can be defined as a continuous, univalued function on U. As a consequence of (B.5) we see that, for all $(t,s) \in V = U \cap (U-(1,0)) \cap (U-(0,1))$,

$$\begin{aligned} \gamma_r(t+1,s) &= \gamma_r(t,s) + 2\pi ik \\ \gamma_r(t,s+1) &= \gamma_r(t,s) + 2\pi il + 2\pi it + \phi(t,s) \end{aligned} \tag{B.12}$$

where k,l are integers, constant over all of V, because of the continuity of γ_r, and where

$$\begin{aligned} |\phi(t,s)| &\le -\log\left(1 - \frac{|\psi(t,s)|}{|G_r(t,s)|}\right) \\ &\le 2\frac{|\psi(t,s)|}{|G_r(t,s)|} \end{aligned}$$

provided $|\psi(t,s) \le |G_r(t,s)|/2$. For this it is sufficient (see (B.6)) to impose $r \le a/(4\pi b)$, which we can do without loss of the previous estimates. One finds then

$$|\phi(t,s)| \le 1, \qquad \text{for all } (t,s) \in V.$$

By construction V is a neighborhood of $[0,1]^2$. In particular (B.12) is valid on $[0,1]^2$, which is enough to again lead to a contradiction. We have

$$\begin{aligned} 0 &= \gamma_r(1,1) - \gamma_r(1,0) + \gamma_r(1,0) - \gamma_r(0,0) \\ &\quad + \gamma_r(0,0) - \gamma_r(0,1) + \gamma_r(0,1) - \gamma_r(1,1) \\ &= 2\pi il + 2\pi i + \phi(1,0) + 2\pi ik - 2\pi il - \phi(0,0) - 2\pi ik \\ &= 2\pi i + \phi(1,0) - \phi(0,0) \ne 0 \end{aligned}$$

since

$$|\phi(1,0)|, |\phi(0,0)| \le 1.$$

This contradiction concludes the proof. □

Appendix C
Proof of the Theorems in Section III-B-3

Proof of Theorems 2.5, 2.6: Using the Poisson formula

$$\sum_{l \in \mathbb{Z}} e^{ilax} = \frac{2\pi}{a} \sum_{k \in \mathbb{Z}} \delta\left(x - \frac{2\pi}{a}k\right)$$

we find

$$\begin{aligned} &\sum_{m,n} |\langle g_{mn}, f\rangle|^2 \\ &= \sum_{m,n} \int dx \int dx'\, e^{imp_0(x-x')} \\ &\quad \cdot g(x-nq_0) g(x'-nq_0)^* f(x)^* f(x') \\ &= \frac{2\pi}{p_0} \sum_{n,k} \int dx\, g(x-nq_0) \\ &\quad \cdot g\left(x - nq_0 - \frac{2\pi}{p_0}k\right)^* f(x)^* f\left(x - \frac{2\pi}{p_0}k\right). \end{aligned}$$

Hence, separating the sum over k into the term for $k=0$ and a sum over $k \ne 0$, and applying Cauchy–Schwarz (once on the integral over x, and once on the sum over k),

$$\begin{aligned} &\sum_{m,n} |\langle g_{mn}, f\rangle|^2 \\ &\ge \frac{2\pi}{p_0} \int dx \sum_n |g(x-nq_0)|^2 |f(x)|^2 \\ &\quad - \frac{2\pi}{p_0} \sum_{k \ne 0} \left[\sum_n \int dx\, |\cdot g(x-nq_0)| \right. \\ &\quad \left. \cdot \left| g\left(x - nq_0 - \frac{2\pi}{p_0}k\right)\right| |f(x)|^2 \right]^{1/2} \\ &\quad \cdot \left[\sum_n \int dx\, |g(x-nq_0)| \right. \\ &\quad \left. \cdot \left| g\left(x - nq_0 - \frac{2\pi}{p_0}k\right)\right| \left| f\left(x - \frac{2\pi}{p_0}k\right)\right|^2 \right]^{1/2} \\ &\ge \frac{2\pi}{p_0} \left\{ m(g;q_0) - \sum_{k \ne 0} \left[\beta\left(\frac{2\pi}{p_0}k\right) \beta\left(-\frac{2\pi}{p_0}k\right)\right]^{1/2} \right\} \|f\|^2. \end{aligned} \tag{C.1}$$

The decay condition (2.3.13) on β implies that the sum over k always converges. It also implies that this sum tends to zero for $p_0 \to 0$, so that the coefficient of $\|f\|^2$ in (C.1) is strictly positive for small enough p_0. On the other hand, we also have

$$\begin{aligned} &\sum_{m,n} |\langle g_{mn}, f\rangle|^2 \\ &\le \frac{2\pi}{p_0} \left\{ M(g;q_0) + \sum_{k \ne 0} \left[\beta\left(\frac{2\pi}{p_0}k\right) \beta\left(-\frac{2\pi}{p_0}k\right)\right]^{1/2} \right\} \|f\|^2 \end{aligned} \tag{C.2}$$

for all $p_0 > 0$. Together, the lower and upper bounds (C.1) and (C.2) imply that $P_0^c > 0$, with

$$P_0^c = \inf\{p_0 | \text{the } g_{mn} \text{ associated to } g, p_0, q_0 \text{ do not constitute a frame}\}.$$

This proves Theorem 2.5. The inequalities (C.1) and (C.2) also immediately prove Theorem 2.6. □

The proofs of Theorems 2.7 and 2.8 are very similar.

Proofs of Theorems 2.7 and 2.8: One uses the unitarity of the Fourier transform to write

$$\begin{aligned} \sum_{m,n} |\langle h_{mn}, f\rangle|^2 &= \sum_{m,n} a_0^m \int dx \int dx'\, e^{inb_0 a_0^m (x-x')} \hat{h}(a_0^m x) \\ &\quad \cdot \hat{h}(a_0^m x')^* \hat{f}(x)^* \hat{f}(x'). \end{aligned} \tag{C.3}$$

As in the proof of Theorem 2.5, this can be rewritten with the help of the Poisson formula,

$$\sum_{m,n} |\langle h_{mn}, f\rangle|^2 = \frac{2\pi}{b_0} \sum_{m,k} \int dx \hat{h}(a_0^m x) \hat{h}\left(a_0^m x + \frac{2\pi}{b_0}k\right)^* \cdot \hat{f}(x)^* \hat{f}\left(x + \frac{2\pi}{b_0} k a_0^{-m}\right). \quad \text{(C.4)}$$

The same arguments as in the proofs of Theorems 2.5 and 2.6 then lead to the desired conclusions. □

In the case where support $\hat{h} \subset \mathbb{R}_+$, and the frame $\{h_{mn}^{\pm} | m,n \in \mathbb{Z}\}$ is used (see Section II-C-2a)), one has to be a little more careful. The equivalent of (C.1) is then

$$\sum_{\epsilon=\pm} \sum_{m,n} |\langle h_{mn}^{\epsilon}, f\rangle|^2 = \frac{2\pi}{b_0} \int_0^\infty dx \sum_m |\hat{h}(a_0^m x)|^2 \left[|\hat{f}(x)|^2 + |\hat{f}(-x)|^2\right] + r$$

with

$$|r| \le \frac{2\pi}{b_0} \sum_{k\neq 0} \left\{ \sum_m \left[\int_0^\infty dx |\hat{h}(a_0^m x)| \left|\hat{h}\left(a_0^m x + \frac{2\pi}{b_0}k\right)\right| |\hat{f}(x)|^2\right]^{1/2} \cdot \left[\int_0^\infty dx \left|\hat{h}\left(a_0^m x - \frac{2\pi}{b_0}k\right)\right| |\hat{h}(a_0^m x)| |\hat{f}(x)|^2\right]^{1/2} \right.$$
$$\left. + \sum_m \left[\int_0^\infty dx |\hat{h}(a_0^m x)| \left|\hat{h}\left(a_0^m x + \frac{2\pi}{b_0}k\right)\right| |\hat{f}(-x)|^2\right]^{1/2} \cdot \left[\int_0^\infty dx \left|\hat{h}\left(a_0^m x - \frac{2\pi}{b_0}k\right)\right| |\hat{h}(a_0^m x)| |\hat{f}(-x)|^2\right]^{1/2} \right\}.$$

This leads again to the same bounds (2.3.23), (2.3.24).

Proof of Theorem 2.9: Applying the Poisson formula gives

$$\sum_{m,n} |\langle h_{mn}, f\rangle|^2 = \frac{2\pi}{b_0} \int dx \left[\sum_m |\hat{h}(2^m x)|^2\right] |\hat{f}(x)|^2 + \frac{2\pi}{b_0} \sum_{k\neq 0} \sum_m \int dx \hat{h}(2^m x) \hat{h}\left(2^m x - \frac{2\pi}{b_0}k\right)^* \hat{f}(x)^* \cdot \hat{f}\left(x - \frac{2\pi}{b_0} 2^{-m} k\right). \quad \text{(C.5)}$$

Any $k \in \mathbb{Z}$, $k \neq 0$ can be decomposed, in a unique way, as $k = 2^{m'} k'$, where $m', k' \in \mathbb{Z}$, $m' \ge 0$ and k' odd. Substituting this into (C.5), and defining $m'' = m - m'$, one finds

$$\sum_{m,n} |\langle h_{mn}, f\rangle|^2 = \frac{2\pi}{b_0} \int dx \left[\sum_m |\hat{h}(2^m x)|^2\right] |\hat{f}(x)|^2 + \frac{2\pi}{b_0} \sum_{k\, odd} \sum_{m' \ge 0} \sum_{m'' \in \mathbb{Z}} \int dx \hat{h}(2^{m'+m''} x) \cdot \hat{h}\left[2^{m'}\left(2^{m''}x - \frac{2\pi}{b_0}k\right)\right]^* \hat{f}(x)^* \hat{f}\left(x - \frac{2\pi}{b_0} 2^{-m''} k\right).$$

Applying Cauchy–Schwarz twice, in the second term (once on the sum over m'', and once on the integral over x) leads to the estimates (2.3.25) and (2.3.26). □

Proof of Theorem 2.10: For $b_0 = 1$, the $\psi_{mn;1}$ are Y. Meyer's basis. We start by showing that the $\psi_{m,n;b_0}$ still constitute a frame for b_0 in a neighborhood of 1. For $b_0 = 2$, the $\psi_{mn;2} = \psi_{m\,2n;1}$ do not span $L^2(\mathbb{R})$. We therefore restrict our attention to $b_0 < 2$.

In the computation of β_1, it is sufficient to take the supremum over $x \in [2\pi/3, 4\pi/3]$. For x in this interval, and for $|s| \ge 2\pi b_0^{-1} \ge \pi$, one finds that only the couples $(m, m') \in \{(-2,2), (-1,1), (-1,2), (0,0), (0,1), (1,0)\}$ lead to nonzero contributions to β_1. As the supremum of the sum of a finite number of continuous functions, β_1 is therefore continuous. On the other hand, since support $\psi \subset [-8\pi/3, 8\pi/3]$, one finds that $\beta_1(s) = 0$ if $|s| \ge 16\pi/3$. For $b_0 < 2$, this implies that only the choices $l = 0$, 1 or 2 lead to nonzero contributions in the sums (2.3.25), (2.3.26). This implies that the right hand sides of (2.3.25), (2.3.26) are continuous in b_0. For $b_0 = 1$ these two expressions are equal to 1. By continuity we find therefore that (2.3.25) > 0 and (2.3.26) < ∞ on a neighborhood of $b_0 = 1$.

To show that the $\psi_{mn;b_0}$ constitute a basis, we introduce the operator

$$S(b_0) = \sum_{m,n} \psi_{mn;b_0} \langle \psi_{mn;1}, \cdot\rangle.$$

In the terminology of Section II-A, we can write

$$S(b_0) = T(b_0)^* T(1)$$

where $T(b_0)$, $T(1)$ are the frame operators for the frames $\psi_{mn;b_0}$, $\psi_{mn;1}$ respectively. To prove that the $\psi_{mn;b_0}$ constitute a basis, it is sufficient to prove that $S(b_0)$ is one-to-one and onto, since $S(b_0)\psi_{mn;1} = \psi_{mn;b_0}$. Since $T(1)$ is unitary, and $T(b_0)^*$ is onto (see Proposition 2.1), we only need to prove that $T(b_0)^*$ or $S(b_0)$ is one-to-one. We have

$$\begin{aligned}\|S(b_0)f\|^2 &= \sum_{m,n} \sum_{m',n'} \langle f, \psi_{mn;1}\rangle \langle \psi_{mn;b_0}, \psi_{m'n';b_0}\rangle \langle \psi_{m'n';1}, f\rangle \\ &\ge \|f\|^2 - \sum_{\substack{m,n,m',n' \\ (m,n)\neq(m',n')}} |\langle f, \psi_{mn;1}\rangle| \, |\langle \psi_{m'n';1}, f\rangle| \cdot |\langle \psi_{mn;b_0}, \psi_{m'n';b_0}\rangle| \\ &\ge \|f\|^2 \left[1 - \sup_{m,n} \sum_{\substack{m',n' \\ (m',n')\neq(m,n)}} |\langle \psi_{mn;b_0}, \psi_{m'n';b_0}\rangle|\right] \\ &= \|f\|^2 \left[1 - \sum_{\substack{m,n \\ (m,n)\neq(0,0)}} |\langle \psi_{mn;b_0}, \psi\rangle|\right]. \qquad \text{(C.6)}\end{aligned}$$

Since support $\hat{\psi} = [-8\pi/3, -2\pi/3] \cup [2\pi/3, 8\pi/3]$, and $a_0 = 2$, $\langle \psi_{mn;b_0}, \psi \rangle = 0$ for $|m| > 1$. One checks easily, using

$$|\psi(x)| \le C_N (1 + |x|^2)^{-N}$$

that

$$\sum_{n \in \mathbb{Z}} |\langle \psi_{mn;b_0}, \psi \rangle|$$

converges, and is continuous in b_0, for $m = 0, \pm 1$. Hence the coefficient of $\|f\|^2$ in the right-hand side of (C.6) is continuous in b_0. Since this coefficient is equal to 1 for $b_0 = 1$, it is therefore strictly positive on a neighborhood of $b_0 = 1$. This implies that, on that neighborhood, $S(b_0)$ is one-to-one. This concludes the proof. □

Appendix D
Proofs of the Theorems in Section III

Proof of Theorem 3.1: Since the g_{mn} constitute a frame, with dual frame $\tilde{g}_{mn}$, we have, for all $f_1, f_2 \in L^2(\mathbb{R})$,

$$\langle f_2, f_1 \rangle = \sum_{m,n \in \mathbb{Z}} \langle f_2, \tilde{g}_{mn}, f_1 \rangle.$$

Fix $T, \Omega > 0$. For $t, \omega > 0$ we define

$$B(t,\omega) = \{(m,n) \in \mathbb{Z}^2; |mp_0| \le \Omega + \omega, |nq_0| \le T + t\}.$$

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

$$\begin{aligned}
&\left\| f - \sum_{(m,n) \in B(t,\omega)} \tilde{g}_{mn} \langle g_{mn}, f \rangle \right\| \\
&= \sup_{\|f_1\|=1} \left| \sum_{(m,n) \notin B(t,\omega)} \langle f_1, \tilde{g}_{mn} \rangle \langle g_{mn}, f \rangle \right| \\
&\le \sup_{\|f_1\|=1} \sum_{\substack{|mp_0| > \Omega + \omega \\ n \in \mathbb{Z}}} |\langle f_1, \tilde{g}_{mn} \rangle| \left[|\langle g_{mn}, P_\Omega f \rangle| \right. \\
&\quad \left. + |\langle g_{mn}, (1 - P_\Omega) f \rangle| \right] \\
&\quad + \sup_{\|f_1\|=1} \sum_{\substack{|nq_0| > T + t \\ m \in \mathbb{Z}}} |\langle f_1, \tilde{g}_{mn} \rangle| \left[|\langle g_{mn}, Q_T f \rangle| \right. \\
&\quad \left. + |\langle g_{mn}, (1 - Q_T) f \rangle| \right] \\
&\le (B/A)^{1/2} \left\{ \|(\mathbb{1} - P_\Omega) f\| + \|(\mathbb{1} - Q_T) f\| \right. \\
&\quad + B^{-1/2} \left(\sum_{\substack{|mp_0| > \Omega + \omega \\ n \in \mathbb{Z}}} |\langle g_{mn}, P_\Omega f \rangle|^2 \right)^{1/2} \\
&\quad \left. + B^{-1/2} \left(\sum_{\substack{|nq_0| > T + t \\ m \in \mathbb{Z}}} |\langle g_{mn}, Q_T f \rangle|^2 \right)^{1/2} \right\}
\end{aligned} \tag{D.1}$$

where we have used the fact that the g_{mn} and the $\tilde{g}_{mn}$ both constitute frames, with frame constants A, B and B^{-1}, A^{-1}, respectively. Similar to the proof of Theorem 2.5 we use the Poisson formula to estimate the last two terms in (D.1). This leads to

$$\begin{aligned}
&\sum_{\substack{|nq_0| > T + t \\ m \in \mathbb{Z}}} |\langle g_{mn}, Q_T f \rangle|^2 \\
&= \frac{2\pi}{p_0} \sum_{\substack{|nq_0| T + t \\ l \in \mathbb{Z}}} \int dx g(x - bq_0)^* g\left(x - nq_0 - \frac{2\pi}{p_0} l\right) \\
&\quad \cdot (Q_T f)(x) (Q_T f)\left(x - \frac{2\pi}{p_0} l\right)^* \\
&\le \frac{2\pi}{p_0} \|Q_T f\|^2 \cdot \sum_{l \in \mathbb{Z}} \sup_{\substack{|x| \le T \\ |x - \frac{2\pi}{p_0} l| \le T}} \sum_{|nq_0| > T + t} |g(x - nq_0)| \\
&\quad \cdot \left| g\left(x - nq_0 - \frac{2\pi l}{p_0}\right) \right|.
\end{aligned} \tag{D.2}$$

From the conditions on g, we have

$$\begin{aligned}
(\text{D.2}) \le C^2 &\sum_{|nq_0| > T + t} \sum_{l \in \mathbb{Z}} \sup_{\substack{|x| \le T \\ |x - \frac{2\pi}{p_0} l| \le T}} \left[1 + (x - nq_0)^2 \right]^{-\alpha} \\
&\cdot \left[1 + \left(x - nq_0 - \frac{2\pi}{p_0} l \right)^2 \right]^{-\alpha}.
\end{aligned} \tag{D.3}$$

The contribution for $n > (T + t)/q_0$ is exactly equal to that for $n < -(T + t)/q_0$, so that we may restrict ourselves to negative n, at the price of a factor 2. By redefining $y = x - (2\pi/p_0) l$ if l is positive, we see that we may restrict ourselves to negative l as well. Hence

$$\begin{aligned}
(\text{D.3}) &\le 4C^2 \sum_{nq_0 > T + t} \sum_{l \ge 0} \sup_{\substack{|x| \le T \\ |x - \frac{2\pi}{p_0} l| \le T}} \left[1 + (x + nq_0)^2 \right]^{-\alpha} \\
&\quad \cdot \left[1 + \left(x + nq_0 + \frac{2\pi}{p_0} l \right)^2 \right]^{-\alpha} \\
&\le 4C^2 \sum_{nq_0 > T + t} \sum_{l \ge 0} \left[1 + (nq_0 - T)^2 \right]^{-\alpha} \\
&\quad \cdot \left[1 + \left(nq_0 - T + \frac{2\pi}{p_0} l \right)^2 \right]^{-\alpha}.
\end{aligned} \tag{D.4}$$

However, for any $a, b > 0$, we have

$$\begin{aligned}
&\sum_{l \ge 0} \left[1 + (a + bl)^2 \right]^{-\alpha} \\
&\le (1 + a^2)^{-\alpha} + \frac{1}{b} \int_a^\infty dx\, (1 + x^2)^{-\alpha} \\
&\le \left[1 + 2^{2\alpha - 1} b^{-1} \left(1 + \frac{1}{2\alpha - 1} \right) \right] (1 + a^2)^{-\alpha + 1/2},
\end{aligned}$$

if $\alpha > 1/2$.

Consequently

$$(D.4) \le 4C^2\left[1+2^{2\alpha-1}\frac{\alpha p_0}{\pi(2\alpha-1)}\right] \cdot \sum_{nq_0 > T+t}\left[1+(nq_0-T)^2\right]^{-2\alpha+1/2}.$$

A similar argument shows that, for $t \ge 2q_0$,

$$\sum_{nq_0 > T+t}\left[1+(nq_0-T)^2\right]^{-2\alpha+1/2} \le \frac{1}{q_0}\left[1+\frac{1}{2(2\alpha-1)}\right]2^{2(2\alpha-1)}\left[1+(t-2q_0)^2\right]^{-2\alpha+1}.$$

We find therefore that there exists a constant $\kappa(p_0,q_0)$, independent of T or t, such that

$$(D.2) \le \kappa(p_0,q_0)[1+t^2]^{-2\alpha+1}\|f\|.$$

The term in $P_\Omega f$ is estimated similarly; putting everything together leads to

$$\left\|f - \sum_{(m,n)\in B(t,\omega)} \tilde{g}_{mn}\langle g_{mn},f\rangle\right\| \le (B/A)^{1/2}\Big\{\|(\mathbb{1}-P_\Omega)f\|+\|(\mathbb{1}-Q_T)f\| + B^{-1/2}\Big[\kappa(p_0,q_0)^{1/2}(1+t^2)^{-2\alpha+1/2} + \kappa(q_0,p_0)^{1/2}(1+\omega^2)^{-2\alpha+1/2}\Big]\|f\|\Big\}, \tag{D.5}$$

since $(1+t^2)^{-2\alpha+1}$, $(1+\omega^2)^{-2\alpha+1} \to 0$ for $t,\omega \to 0$, this proves the theorem. □

Remark: The estimates in this proof cause $t_\epsilon, \omega_\epsilon$, when calculated using (D.5), to be much larger for a given ϵ, and e.g., for Gaussian g, than observed in numerical calculations. The intermediate estimate (D.2) leads to much better values of t_ϵ (a similar formula can of course be written for $P_\Omega f$, leading to estimates for ω_ϵ). If we define

$$\zeta^{\pm}(t;y) = \sup_{\substack{\pm x \ge t\\ \pm(x-y)\ge t}} \sum_{n=0}^{\infty} |g(x\pm nq_0)||g(x-y\pm nq_0)|$$

then $\kappa(p_0,q_0)^{1/2}$, in the estimate (D.5), can be replaced by

$$\left(\frac{2\pi}{p_0}\right)^{1/2}\left\{\sum_{l\in\mathbb{Z}}\left[\zeta^+\left(t;\frac{2\pi}{p_0}l\right)+\zeta^-\left(t;\frac{2\pi}{p_0}l\right)\right]\right\}^{1/2}. \tag{D.6}$$

The same thing can be done for $\kappa(q_0,p_0)^{1/2}$.

Proof of Theorem 3.2: We define the set B_ϵ as

$$B_\epsilon(T,\Omega_0,\Omega_1) = \{(m,n);\, m_0 \le m \le m_1,\, |nb_0| \le a_0^{-m}T+t\} \tag{D.7}$$

where m_0, m_1, and t, to be defined next, depend on Ω_0, Ω_1, and ϵ.

One then has (see the proof of Theorem 3.1)

$$\begin{aligned}
&\left\|f - \sum_{(m,n)\in B_\epsilon(T,\Omega_0,\Omega_1)} (h_{mn})^{\sim}\langle h_{mn},f\rangle\right\| \\
&\le \sup_{\|f\|=1}\sum_{\substack{n\in\mathbb{Z}\\ m<m_0 \text{ or } m>m_1}} |\langle f_1,(h_{mn})^{\sim}\rangle| \cdot\Big[|\langle h_{mn},(P_{\Omega_1}-P_{\Omega_0})f\rangle| + |\langle h_{mn},(1-P_{\Omega_1}+P_{\Omega_0})f\rangle|\Big] \\
&\quad + \sup_{\|f_1\|=1}\sum_{\substack{m_0\le m\le m_1\\ |nb_0|\ge a_0^{-m}T+t}} |\langle f_1,(h_{mn})^{\sim}\rangle| \cdot\Big[|\langle h_{mn},Q_Tf\rangle| + |\langle h_{mn},(1-Q_T)f\rangle|\Big] \\
&\le (B/A)^{1/2}\Bigg\{\|(\mathbb{1}-P_{\Omega_1}+P_{\Omega_0})f\| + \|(\mathbb{1}-Q_T)f\| \\
&\quad + B^{-1/2}\Bigg(\sum_{\substack{n\in\mathbb{Z}\\ m<m_0\\ \text{or } m>m_1}} |\langle h_{mn},(P_{\Omega_1}-P_{\Omega_0})f\rangle|^2\Bigg)^{1/2} \\
&\quad + B^{-1/2}\Bigg(\sum_{\substack{m_0\le m\le m_1\\ |nb_0|\ge a_0^{-m}T+t}} |\langle h_{mn},Q_Tf\rangle|^2\Bigg)^{1/2}\Bigg\}.
\end{aligned} \tag{D.8}$$

One has

$$\begin{aligned}
&\sum_{\substack{m>m_1 \text{ or } m<m_0\\ n\in\mathbb{Z}}} |\langle h_{mn},(P_{\Omega_1}-P_{\Omega_0})f\rangle|^2 \\
&\le \frac{2\pi}{b_0}\sum_{\substack{m>m_1 \text{ or } m<m_0\\ l\in\mathbb{Z}}} \int_{\substack{\Omega_0\le|y|\le\Omega_1\\ \Omega_0\le|y-a_0^{-m}\frac{2\pi}{b_0}l|\le\Omega_1}} dy \cdot|\hat{h}(a_0^m y)|\left|\hat{h}\left(a_0^m y-\frac{2\pi}{b_0}l\right)\right||\hat{f}(y)|\left|\hat{f}\left(y-a_0^{-m}\frac{2\pi}{b_0}l\right)\right| \\
&\le \frac{2\pi}{b_0}\|(P_{\Omega_1}-P_{\Omega_0})f\|^2\cdot\sum_{l\in\mathbb{Z}}\phi\left(\frac{2\pi}{b_0}l\right)^{\kappa/2} \cdot\left[\sup_{\Omega_0\le|y|\le\Omega_1}\sum_{\substack{m>m_1\\ \text{or}\\ m<m_0}}\left[1+(a_0^m y)^2\right]^{\kappa}|\hat{h}(a_0^m y)|\right]^2
\end{aligned}$$

where we have used that, for all $x \in \mathbb{R}$,

$$(1+x^2)\left[1+(x-s)^2\right]\phi(s) \geq 1$$

with ϕ defined by

$$\phi(s) = \begin{cases} \left[1+(s/2)^2\right]^{-2}, & \text{if } |s| \leq 2 \\ s^{-2}, & \text{if } |s| \geq 2. \end{cases}$$

If we choose $1 < \kappa < \alpha$, then

$$K_\kappa(b_0) = \sum_{l \in \mathbb{Z}} \phi\left(\frac{2\pi}{b_0} l\right)^{\kappa/2} \text{ converges}$$

and

$$\sup_{\Omega_0 \leq |y| \leq \Omega_1} \sum_{\substack{m > m_1 \\ \text{or} \\ m < m_0}} \left[1+(a_0^m y)^2\right]^\kappa \left|\hat{h}(a_0^m y)\right|^2$$

$$\leq C^2 \sup_{\Omega_0 \leq |y| \leq \Omega_1} \sum_{\substack{m > m_1 \\ \text{or} \\ m < m_0}} (a_0^m |y|)^{2\beta}\left[1+(a_0^m y)^2\right]^{-\alpha-\beta+s\kappa}.$$

Using the same techniques as in the proof of Lemma 2.2, one finds, for $\Omega_0 \leq |y|$,

$$\sum_{m > m_1} (a_0^m |y|)^{2\beta}\left[1+(a_0^m y)^2\right]^{-\alpha-\beta+\kappa}$$

$$\leq \frac{1}{\ln a_0} \int_{a_0^{m_1}} dt\, t^{2\beta-1}(1+t^2)^{-\alpha-\beta+\kappa}$$

$$\leq \frac{1}{\ln a_0} \frac{1}{2(\alpha-\kappa)} \left[a_0^{m_1}\Omega_0\right]^{-2(\alpha-\kappa)}$$

where we have assumed that $m_1 \geq (\ln \beta - \ln(\alpha-\kappa) - \ln \Omega_0)/\ln a_0$. Similarly, for $|y| \leq \Omega_1$,

$$\sum_{m < m_0} (a_0^m |y|)^{2\beta}\left[1+(a_0^m y)^2\right]^{-\alpha-\beta+\kappa}$$

$$\leq \frac{1}{\ln a_0} \int_0^{a_0^{m_0}\Omega_1} dt\, t^{2\beta-1}(1+t^2)^{-\alpha-\beta+\kappa}$$

$$\leq \frac{1}{\ln a_0} \frac{1}{2\beta} \left[a_0^{m_0}\Omega_1\right]^{2\beta}$$

provided $m_0 \leq (\ln \beta - \ln(\alpha-\kappa) - \ln \Omega_0)/\ln a_0$. It is clear from this that, for Ω_0, Ω_1, ϵ given, we can choose m_0, m_1 so that

$$\sum_{\substack{n \in \mathbb{Z} \\ m > m_1 \text{ or } m < m_0}} \left|\langle h_{mn}, (P_{\Omega_1} - P_{\Omega_0}) f\rangle\right|^2 \leq B\epsilon^2/4\|f\|^2. \quad \text{(D.9)}$$

On the other hand,

$$\sum_{b_0|n| > a_0^{-m}T+t} \left|\langle h_{mn}, Q_T f\rangle\right|^2$$

$$\leq \sum_{b_0|n| > a_0^{-m}T+t} \|Q_T h_{mn}\|^2 \cdot \|f\|^2$$

$$\leq \|f\|^2 \int_{|x| \leq T} dx\, a_0^{-m} \sum_{b_0|n| > a_0^{-m}T+t} \left|h(a_0^{-m}x - nb_0)\right|^2$$

$$\leq \|f\|^2 \left\{ \int_{-\infty}^{-t} du \sum_{l=0}^{\infty} \left|h(u - lb_0)\right|^2 \right.$$

$$\left. + \int_t^\infty du \sum_{l=0}^{\infty} \left|h(u + lb_0)\right|^2 \right\}$$

$$\leq \|f\|^2 \sum_{l=0}^{\infty} \left[1+(t+lb_0)^2\right]^{-\gamma} \int_{|u| \geq t} du (1+u^2)^\gamma \left|h(u)\right|^2. \quad \text{(D.10)}$$

Since $\gamma > 1/2$, this converges. It is moreover clear that t can be chosen so that the coefficient of $\|f\|^2$ in (D.10) is smaller than $B\epsilon^2/4(m_1 - m_0 + 1)$, i.e., such that

$$\sum_{\substack{m_0 \leq m \leq m_1 \\ b_0|n| > a_0^{-m}T+t}} \left|\langle h_{mn}, Q_T f\rangle\right|^2 \leq B\epsilon^2/4\|f\|^2. \quad \text{(D.11)}$$

The statement (3.6) now follows directly from (D.8), (D.9) and (D.11). □

REFERENCES

[1] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.* (London), vol. 93, no. III, pp. 429–457, 1946.

[2] M. J. Bastiaans, "Gabor's signal expansion and degrees of freedom of a signal," *Proc. IEEE*, vol. 68, pp. 538–539, 1980.

——, "A sampling theorem for the complex spectrogram and Gabor's expansion of a signal in Gaussian elementary signals," *Optical Eng.*, vol. 20, pp. 594–598, 1981.

[3] A. J. E. M. Janssen, "Gabor representation of generalized functions," *J. Math. Appl.*, vol. 80, pp. 377–394, 1981.

——, "Gabor representation and Wigner distribution of signals," in *Proc. IEEE*, 1984, pp. 41.B.2.1–41.B.2.4.

[4] M. J. Davis and E. J. Heller, *J. Chem. Phys.*, vol. 71, pp. 3383–3395, 1979.

[5] J. R. Klauder and E. Sudarshan, *Fundamentals of Quantum Optics*, New York: Benjamin, 1968.

[6] R. J. Glauber, "The quantum theory of coherence," *Phys. Rev.*, vol. 130, pp. 2529–2539, 1963; "Coherent and incoherent states of the radiation field," *Phys. Rev.*, vol. 131, pp. 2766–2788, 1963.

[7] J. R. Klauder and B.-S. Skagerstam, "Coherent states," Singapore: World Scientific, 1985.

[8] E. Lieb, "Thomas–Fermi theory and related theories of atoms and molecules," *Rev. Mod. Phys.*, vol. 53, pp. 603–641, 1981.

[9] G. A. Hagedorn, "Semiclassical quantum mechanics I: The $\hbar \to 0$ limit for coherent states," *Comm. Math. Phys.*, vol. 71, pp. 77–93, 1980.

[10] C. W. Helstrom, "An expansion of a signal in Gaussian elementary signals," *IEEE Trans. Inform. Theory*, vol. 12, pp. 81–82, 1966.

[11] N. G. de Bruijn, "Uncertainty principles in Fourier analysis," in *Inequalities*, O. Shisha, Ed. New York: Academic Press, pp. 57–71, 1967.

[12] T. A. C. M. Claassen and W. F. G. Mecklenbräuker, "The Wigner distribution—A tool for time-frequency signal analysis," *Philips J. Res.*, vol. 35, pp. 217–250, 276–300, 372–389, 1980.

[13] C. P. Janse and A. J. M. Kaiser, "Time-frequency distributions of loudspeakers: The application of the Wigner distribution," *J. Audio Eng. Soc.*, vol. 31, pp. 198–223, 1983.

[14] R. J. Duffin and A. C. Schaeffer, "A class of nonharmonic Fourier series," *Trans. Am. Math. Soc.*, vol. 72, pp. 341–366, 1952.

[15] R. M. Young, *An Introduction to Nonharmonic Fourier Series*. New York: Academic Press, 1980.

[16] I. Daubechies, A. Grossmann, and Y. Meyer, "Painless nonorthogonal expansions," *J. Math. Phys.*, vol. 27, pp. 1271–1283, 1986.

[17] A. Grossmann and J. Morlet, "Decomposition of Hardy functions ito square integrable wavelets of constant shape," *SIAM J. Math. Anal.*, vol. 15, pp. 723–736, 1984.
P. Goupillaud, A. Grossmann, and J. Morlet, "Cycle-octave and related transforms in seismic signal analysis," *Geoexploration*, vol. 23, p. 85, 1984.

[18] A. Grossmann, J. Morlet, and T. Paul, "Transforms associated to square integrable group representations," *J. Math. Phys.*, vol. 26, pp. 2473–2479, 1985; *Ann. Inst. Henri Poincaré*, vol. 45, 293–309, 1986.

[19] E. W. Aslaksen and J. R. Klauder, "Unitary representations of the affine group," *J. Math. Phys.*, vol. 9, pp. 206–211, 1968; "Continuous representation theory using the affine group," *J. Math. Phys.*, vol. 10, pp. 2267–2275, 1969.

[20] T. Paul, "Functions analytic on the half-plane as quantum mechanical states," *J. Math. Phys.*, vol. 25, pp. 3252–3263, 1984.

[21] ____, "Affine coherent states and the radial Schrödinger equation. I. Radial harmonic oscillator and hydrogen atom," to be published.

[22] C. Fefferman and R. de la Llave, "Relativistic stability of matter," *Rev. Math. Iberoamericana*, vol. 2, pp. 119–213, 1986.

[23] I. Daubechies, "Time-frequency localization operators—A geometric phase space approach," *IEEE Trans. Inform. Theory*, vol. 34, pp. 605–612, 1988.

[24] I. Daubechies and T. Paul, "Time-frequency localization operators —A geometric phase space approach II. The use of dilations," *Inverse Problems*, vol. 4, pp. 661–680, 1988.

[25] J. Morlet, G. Arens, I. Fourgeau, and D. Giard, "Wave propagation and sampling theory," *Geophys.*, vol. 47, pp. 203–236, 1982; "Sampling theory and wave propagation," NATO ASI Series, vol. 1, *Issues in Acoustic Signal/Image Processing and Recognition*, C. H. Chen, Ed. Berlin: Springer, pp. 233–261.

[26] R. Kronland-Martinet, J. Morlet, and A. Grossmann, "Analysis of sound patterns through wavelet transforms," preprint CTP-87/p. 1981, Centre de Physique Théorique, CNRS, Marseille, France, 1987.

[27] A. Grossmann, "Wavelet transforms and edge detection," to be published in *Stochastic Processes in Physics and Engineering*, Ph. Blanchard, L. Streit, and M. Hasewinkel, Eds.

[28] A. Grossmann, M. Holschneider, R. Kronland-Martinet, and J. Morlet, "Detection of abrupt changes in sound signals with the help of wavelet transforms," in *Inverse Problems: An Interdisciplinary Study, Advances in Electronics and Electron Physics*, Supp. 19. New York: Academic, 1987.

[29] S. Mallat, "A theory for multiresolution signal decomposition," Ph.D. thesis, Univ. of Pennsylvania, Depts. of Elect. Eng. and Comput. Sci., 1988.

[30] A. Grossmann and J. Morlet, "Decomposition of functions into wavelets of constant shape and related transforms," in *Mathematics and Physics, Lectures on Recent Results*. Singapore: World Scientific, 1985.

[31] Y. Meyer, "Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs," *Séminaire Bourbaki*, nr. 662, 1985–1986.

[32] P. G. Lemarié and Y. Meyer, "Ondelettes et bases hilbertinennes," *Rev. Math. Iberoamericana*, vol. 2, pp. 1–18, 1986.

[33] G. Battle, "A block spin construction of ondelettes. Part I: Lemarié functions," *Comm. Math. Phys.*, vol. 110, pp. 601–615, 1987.

[34] P. G. Lemarié, "Une nouvelle base d'ondelettes de $L^2(\mathbb{R}^n)$," to be published in *J. de Math. Pures et Appl.*

[35] a) Ph. Tchamitchian, "Calcul symbolique sur les opérateurs de Caldéron-Zygmund et bases inconditionnelles de $L^2(\mathbb{R}^n)$," C. R. Acad. Sc. Paris, vol. 303, série 1, pp. 215–218, 1986; b) Ph. Tchamitchian, "Biorthogonalité et théorie des opérateurs," to be published in *Rev. Math. Iberoamericana*.

[36] S. Mallat, "Multiresolution approximation and wavelets," *Trans. Am. Math. Soc.*, vol. 135, pp. 69–88, 1989.

[37] Y. Meyer, "Ondelettes, function splines, et analyses graduées," Univ. of Torino, 1986.

[38] S. Mallat, "A Theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 31, pp. 674–693, 1989.

[39] I. Daubechies, "Orthonormal basis of compactly supported wavelets," *Comm. Pure Applied Math.*, vol. 41, pp. 909–996, 1988.

[40] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 37, no. 12, pp. 2091–2110, Dec. 1987.

[41] G. Battle and P. Federbush, "Ondelettes and phase cell cluster expansions: A vindication," *Comm. Math. Phys.*, 1987.

[42] I. Daubechies, "Discrete sets of coherent states and their use in signal analysis," in *Proc. 1986 Int. Conf. Diff. Equations Math. Phys.*, Birmingham, AL, Springer Lecture Notes in Mathematics nr. 1285, pp. 73–82.

[43] I. Daubechies and T. Paul, "Wavelets—Some applications," *Proc. Int. Conf. Math. Phys.*, M. Mebkhout and R. Sénéor, Eds. Singapore: World Scientific, 1987.

[44] I. Daubechies, "How to frame Bargmann and Hardy—Frame bounds for Hilbert spaces of entire functions," unpublished memorandom, AT&T Bell Laboratories, 1987.

[45] I. Daubechies and A. Grossmann, "Frames of entire functions in the Bargmann space," *Comm. Pure Appl. Math.*, vol. 41, pp. 151–164, 1988.

[46] R. R. Coifman and R. Rochberg, "Representation theorems for holomorphic and harmonic functions in L^p," *Astérisque*, vol. 77, pp. 11–66, 1980.

[47] S. Janson, J. Peetre, and R. Rochberg, "Hankel forms and the Fock space," Uppsala Univ., Math. report, 1986.

[48] H. G. Feichtinger and K. Gröchenig, "A unified approach to atomic decompositions via integrable group representations," in *Proc. Function Spaces Applications*, Lund, 1986, to appear in *Springer Lect. Notes in Math.* no. 1302, pp. 52–73; "Banach spaces related to integrable group representations and their atomic decomposition I, II, and III," to be published, *J. Funct. Anal.*, vol. 83, 1989.

[49] M. Rieffel, "Von Neumann algebras associated with pairs of lattices in Lie groups," *Math. Ann.*, vol. 257, pp. 403–418, 1981.

[50] G. Kaiser, "A sampling theorem for signals in the joint time-frequency domain," to be published.

[51] R. Balian, "Un principe d'incertitude fort en théorie du signal on en mécanique quantique," C. R. Acad. Sc. Paris, vol. 292, série 2, 1981.

[52] F. Low, "Complete sets of wave-packets," in *A Passion for Physics —Essays in Honor of Geoffrey Chew*," pp. 17–22. Singapore: World Scientific, 1985.

[53] H. Elbrond Jensen, T. Hoholdt, and J. Justesen, "Double series representation of bounded signals," *IEEE Trans. Inform. Theory*, vol. 34, pp. 613–624, 1988.

[54] J. Zak, "Finite translations in solid state physics." *Phys. Rev. Lett.*, vol. 19, pp. 1385–1397, 1967; "Dynamics of electrons in solids in external fields," *Phys. Rev.*, vol. 168, pp. 686–695; "The *kq*-representation in the dynamics of electrons in solids," *Solid State Phys.*, vol. 27, pp. 1–62, 1972.

[55] W. Schempp, "Radar ambiguity functions, the Heinsenberg group, and holomorphic theta series," *Proc. of the Am. Math. Soc.*, vol. 92, 1984.

[56] M. Reed and B. Simon, *Methods and Modern Mathematical Physics. IV Analysis of Operators*. New York: Academic Press, 1978.

[57] M. Boon, J. Zak, and J. Zucker, "Rational von Neumann lattices," *J. Math. Phys.*, vol. 24, pp. 316–323, 1983.

[58] J. Zak, "Lattice operators in crystals for Bravais and reciprocal vectors," *Phys. Rev.*, vol. B 12, pp. 3023–3026, 1975.

[59] A. J. E. M. Janssen, "Bargmann transform, Zak transform, and coherent states," *J. Math. Phys.*, vol. 23, pp. 720–731, 1982.

[60] A. J. E. M. Janssen, "The Zak transform: A signal transform for sampled time-continuous signals," *Philips J. Res.*, vol. 42, 1987.

[61] V. Bargmann, P. Butera, L. Girardello, and J. R. Klauder, "On the completeness of coherent states," *Rep. Math. Phys.*, vol. 2, pp. 221–228, 1971.

[62] A. M. Perelomov, "Note on the completeness of systems of coherent states," *Teor. i. Matem. Fis.*, vol. 6, pp. 213–224, 1971.

[63] H. Bacry, A. Grossmann, and J. Zak, "Proof of the completeness of lattice states in the *kq*-representation," *Phys. Rev.*, vol. B 12, pp. 1118–1120, 1975.

[64] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, II. Fourier Analysis and Self-Adjointness*. New York: Academic Press, 1975.

[65] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty," I, *Bell. Syst. Tech. J.*, vol. 40,

pp. 43–64, 1961; H. J. Landau and H. O. Pollak, II, *Bell. Syst. Techn. J.*, vol. 40, pp. 65–84, 1961; H. J. Landau and H. O. Pollak, III, *Bell. Syst. Techn. J.*, vol. 41, pp. 1295–1336, 1962.

[66] D. Slepian, "On bandwidth," *Proc. IEEE*, vol. 64, pp. 292–300, 1976.

[67] A. Annéodo, G. Grasseau, and M. Holschneider, "Wavelet transform analysis of invariant measures for some dynamical systems," *Phys. Rev. Lett.*, vol. 61, pp. 2281–2287, 1988.

[68] J. Ville, "Théorie et applications de la notion de signal analytique," *Revue Cables et Transmissions*, vol. 1, pp. 61–74, 1948.

[69] G. Battle, "Heisenberg proof of the Balian-low theorem," *Lett. Math. Phys.*, vol. 15, pp. 175–177, 1988.

[70] I. Daubechies and A. J. E. M. Janssen, "Two theorems on lattice expansions," submitted for publication.

[71] H. Landau, "On the density of phase space functions," submitted for publication.