In: Wavelet Applications in Signal and Image Processing II,
A.F. Laine and M.A. Unser, eds., SPIE Proceedings Series
Volume 2303, SPIE, Bellingham, WA (1994), pp. 15-24.

Asymptotic Singular Value Decay of Time-Frequency Localization Operators

Christopher Heil

School of Mathematics
Georgia Institute of Technology
Atlanta, Georgia 30332-0160
and
The MITRE Corporation
Bedford, Massachusetts 01730

Jayakumar Ramanathan

Department of Mathematics
Eastern Michigan University
Ypsilanti, Michigan 48197

Pankaj Topiwala

The MITRE Corporation
Bedford, Massachusetts 01730

ABSTRACT

The Weyl correspondence is a convenient way to define a broad class of time-frequency localization operators. Given a region Ω in the time-frequency plane $\mathbf{R}^2$ and given an appropriate μ, the Weyl correspondence can be used to construct an operator $L(\Omega, \mu)$ which essentially localizes the time-frequency content of a signal on Ω. Different choices of μ provide different interpretations of localization. Empirically, each such localization operator has the following singular value structure: there are several singular values close to 1, followed by a sharp plunge in values, with a final asymptotic decay to zero. The exact quantification of these qualitative observations is known only for a few specific choices of Ω and μ. In this paper we announce a general result which bounds the asymptotic decay rate of the singular values of any $L(\Omega, \mu)$ in terms of integrals of $|\chi_\Omega * \tilde{\mu}|^2$ and $|(\chi_\Omega * \tilde{\mu})^\wedge|^2$ outside squares of increasing radius, where $\tilde{\mu}(a, b) = \mu(-a, -b)$. More generally, this result applies to all operators $L(\sigma, \mu)$ allowing window functions σ in place of the characteristic functions χ_Ω. We discuss the motivation and implications of this result. We also sketch the philosophy of proof, which involves the construction of an approximating operator through the technology of Gabor frames—overcomplete systems which allow basis-like expansions and Plancherel-like formulas, but which are not bases and are not orthogonal systems.

Keywords: Cohen's class, frames, Gabor systems, singular values, time-frequency localization, Weyl correspondence, Wigner distribution

1. INTRODUCTION

Because of the uncertainty principle, there is no single, ideal time-frequency localization methodology. The Weyl correspondence provides one convenient way to define a broad class of operators which "essentially" localize signals on a domain Ω in the time-frequency plane $\mathbf{R}^2$. Starting with a region $\Omega \subset \mathbf{R}^2$ and an appropriate function μ, the Weyl correspondence can be used to define an operator $L(\Omega, \mu)$ which maps the space $L^2(\mathbf{R})$, consisting of all square-integrable functions on $\mathbf{R}$, into itself. Details of the Weyl correspondence are given in Section 2. Briefly, the process required to construct $L(\Omega, \mu)$ begins with the joint time-frequency distribution $W(f, g)$ known as the *Wigner distribution*. Then $W(f, g)$ is convolved with μ to obtain $W_\mu(f, g)$, a *joint time-frequency distribution in the Cohen class*. Finally

$L(\Omega, \mu)$ is constructed from $W_\mu(f, g)$. Cohen refers to the Fourier transform $\hat{\mu}$ of μ as the *kernel* of the time-frequency distribution $W_\mu(f, g)$.

Each $L(\Omega, \mu)$ "localizes" functions on Ω. With Ω fixed, each choice of μ gives a different interpretation of localization. If $L(\Omega, \mu)$ is self-adjoint then its maximal eigenvalue represents the highest degree of concentration possible on Ω (with respect to that μ). The corresponding eigenfunction is that square-integrable function whose time-frequency content is most concentrated on Ω. If $L(\Omega, \mu)$ is not self-adjoint then the notion of concentration can be formulated in terms of the singular values and singular functions of $L(\Omega, \mu)$.

The problem we address in this note was first posed by Flandrin:[7] determine the singular values and singular functions of the localization operator $L(\Omega, \mu)$.

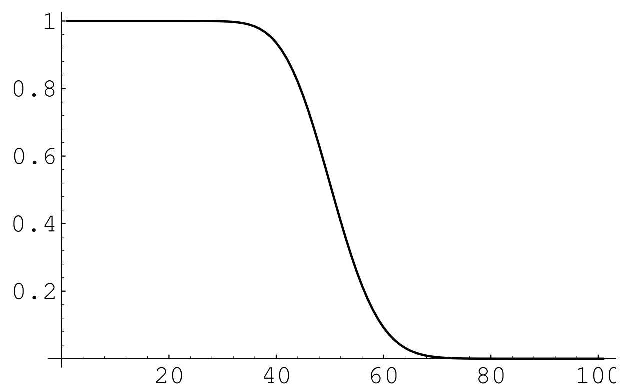

FIGURE 1. Plot of the eigenvalues of $L(D, \mu_1)$ for D the disk of area 50 and $\mu_1(a, b) = e^{-(a^2+b^2)}$.

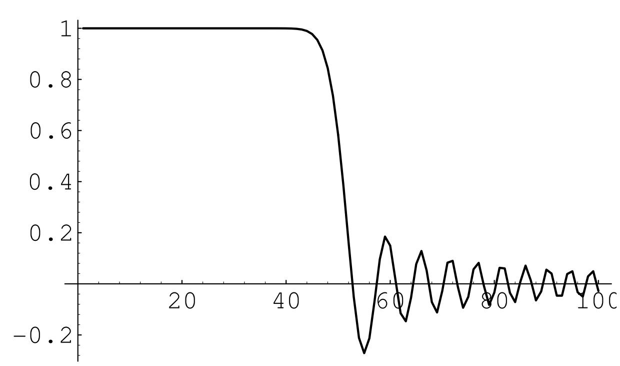

FIGURE 2. Plot of the eigenvalues of $L(D, \mu_2)$ for D the disk of area 50 and $\mu_2 = \delta$.

Figures 1 and 2 display features typical of the singular values of $L(\Omega, \mu)$. In both plots the region is $\Omega = D$, the disk of area 50 centered at the origin. Figures 1 and 2 differ in their choice of μ. However, for both of these particular choices μ_1 and μ_2, the operators $L(D, \mu_1)$ and $L(D, \mu_2)$ are self-adjoint. Therefore Figures 1 and 2 plot the eigenvalues of $L(D, \mu_1)$ and $L(D, \mu_2)$, respectively (the singular values of a self-adjoint operator are simply the absolute values of the eigenvalues).

For Figure 1, we have selected μ_1 to be the two-dimensional Gaussian function, $\mu_1(a,b) = e^{-(a^2+b^2)}$. The corresponding Cohen class time-frequency distribution $W_{\mu_1}(f,g)$ is called the *spectrogram*. The eigenvalues and eigenfunctions of the localization operator $L(\Omega, \mu_1)$ were first determined by Daubechies for circular or elliptical domains Ω.[3] Actually, Daubechies formulated this operator in terms of coherent states, a point of view beautifully exposited in her recent monograph.[4] A previous paper[16] reinterpreted this operator in the light of the Weyl correspondence, and extended the analysis of its eigenvalues and eigenfunctions to any bounded, measurable domain Ω.

For Figure 2, $\mu_2 = \delta$, the point mass at the origin. The corresponding Cohen class time-frequency distribution $W_{\mu_2}(f,g)$ is simply the Wigner distribution $W(f,g)$. The eigenvalue and eigenfunction structure of $L(\Omega, \mu_2)$ has been analyzed for all bounded, measurable domains Ω with piecewise once-differentiable boundaries.[17]

Before proceeding, we must remark on the obvious similarity of Figure 1 to plots arising in the classic energy concentration problem solved by Landau, Pollack, and Slepian.[11,12,19] The problem there was as follows: given a timespan $[-T,T]$ and frequency band $[-S,S]$, determine that function f which is bandlimited to $[-S,S]$ and which concentrates the greatest possible energy into the timespan $[-T,T]$. The elegant solution is that f is the eigenfunction corresponding to the maximal eigenvalue of the positive operator BAB, where $Ah = h \cdot \chi_{[-T,T]}$ (timepass filtering) and $Bh = (\hat{h} \cdot \chi_{[-S,S]})^\vee$ (bandpass filtering). A plot of the eigenvalues of BAB is remarkably similar to Figure 1. Flandrin[7] has shown that the operator BAB can be realized in terms of the Weyl correspondence: $BAB = L^*L$ where $L = L(\Omega, \mu)$ with Ω the rectangle $[-T,T] \times [-S,S]$ and $\mu(a,b) = e^{\pi iab}$. The Cohen class time-frequency distribution $W_\mu(f,g)$ corresponding to this μ is known as the *Rihaczek distribution*.

Figures 1 and 2 are typical of those $L(\Omega, \mu)$ which are self-adjoint. The singular value structure of non-self-adjoint $L(\Omega, \mu)$ is similar, although naturally the singular values are always positive. Empirically, each singular value plot has the following behavior: several singular values cluster near 1 at the beginning, followed by a sharp plunge in the values, with a final asymptotic decay to zero. We therefore desire answers to the following questions, in terms of computable properties of Ω and μ:

(1) How many singular values lie near 1?

(2) What is the width of the plunge region?

(3) What is the asymptotic decay rate?

(4) What are the singular functions?

With answers to these questions in hand we could, for example, obtain an implementable time-frequency filter by projecting signals onto the finite-dimensional subspace of $L^2(\mathbf{R})$ spanned by the principal singular functions of $L(\Omega, \mu)$. A procedure along these lines has been adopted by Hlawatsch, Kozek and Krattenthaler.[10]

The answers to questions 1–4 are known only for the special cases already described above: (a) the spectrogram case, where $\mu(a,b) = e^{-(a^2+b^2)}$, with results known for any bounded, measurable domain Ω; (b) the Wigner case, where μ is the point mass δ, with results known for any bounded, measurable domain Ω with piecewise differentiable boundary; and (c) the Rihaczek case, where $\mu(a,b) = e^{\pi iab}$, with results known if Ω is a rectangle $[-T,T] \times [-S,S]$.

In this note, we announce a new, general result which applies to arbitrary $L(\Omega, \mu)$. The price is that we obtain information on only one of the four questions: question 3, the asymptotic decay rate of the singular values. We are hopeful that our technique will yield insights into the other questions. We have already shown that that this technique leads to other interesting results; in particular, we have an improvement of the *Calderón–Vaillancourt Theorem*, which gives sufficient conditions for the operator $L(\sigma, \mu)$ to be bounded, where $L(\sigma, \mu)$ is the generalization of $L(\Omega, \mu)$ allowing a window function σ to be used in place of the cutoff function χ_Ω. Full details of the proof of the asymptotic decay rate theorem and the statement and proof of the Calderón–Vaillancourt improvement will appear in a forthcoming journal article.[9]

Stated in terms of the general operator $L(\sigma, \mu)$ allowing window functions σ, our asymptotic decay rate result is as follows.

Theorem 1. *Let σ be a square-integrable function on $\mathbf{R}^2$, and let μ be such that $\hat{\mu}$ is a bounded function on $\mathbf{R}^2$. Set*

$$F(n) \;=\; \iint_{\mathbf{R}^2 \setminus B_n} |(\sigma * \tilde{\mu})(p,q)|^2 \, dp\, dq \;+\; \iint_{\mathbf{R}^2 \setminus B_n} |(\sigma * \tilde{\mu})^\wedge(p,q)|^2 \, dp\, dq \;+\; e^{-\frac{\pi}{8} n^2},$$

where $\tilde{\mu}(a,b) = \mu(-a,-b)$ and B_n is the box $B_n = [-n/4, n/4] \times [-n/4, n/4]$. Then there is a constant C so that the singular values s_n of $L(\sigma, \mu)$ satisfy

$$s_n \;\le\; C \left(\frac{F\big((n/8)^{1/2}\big)}{n} \right)^{1/2}.$$

Thus, knowledge of the decay of $\sigma * \tilde{\mu}$ and $(\sigma * \tilde{\mu})^\wedge$ leads to knowledge of the decay of the singular values of $L(\sigma, \mu)$.

Example 1. Consider the spectrogram case: $\mu(a,b) = e^{-(a^2+b^2)}$. If $\sigma = \chi_\Omega$ where Ω is any bounded, measurable domain, then both $(\sigma * \tilde{\mu})(p,q)$ and $(\sigma * \tilde{\mu})^\wedge(p,q)$ have quadratic exponential decay in both p and q.[9] Theorem 1 then yields quadratic exponential decay of the singular values of $L(\sigma, \mu)$. This improves known results.[16]

Example 2. Consider next the Wigner case: $\mu = \delta$. If Ω is bounded with piecewise differentiable boundary, then Theorem 1 leads to a decay rate for the singular values of only $\mathcal{O}(n^{-3/4})$.[9] This result is sharp if Ω is an annulus.[17]

Example 3. Finally, consider the Rihaczek case: $\mu(a,b) = e^{\pi i a b}$. With Ω the rectangle $[-T,T] \times [-S,S]$, Theorem 1 leads only to a decay rate for the singular values of $\mathcal{O}(n^{-3/4})$.[9] Yet it is known that the decay rate is far quicker, at least $\mathcal{O}(n^{-2n})$.[11]

Thus Theorem 1 is sharp in some cases but not in others.

Full details of the proof of Theorem 1 will appear elsewhere.[9] Here we will be content to briefly sketch the philosophy of the proof and to point out its novelty: we bring to bear on this time-frequency problem a time-frequency technique whose roots reach back to Duffin and Schaeffer,[5] but which has seen limited application. That technique is the use of *Gabor frames*. In general, frames provide an alternative to bases. Whether they are practical or convenient depends on the application. In time-frequency analysis, the *Balian–Low theorem*[1] imposes a strict limitation on the sorts of functions available to generate *Gabor bases*, i.e., bases constructed from a single function g by time-frequency shifts. The Balian–Low Theorem implies that if g is continuous and has even moderate decay at infinity then $\{e^{2\pi ibnx}g(x+am)\}_{m,n\in\mathbf{Z}}$ cannot be an orthonormal basis for $L^2(\mathbf{R})$, or even merely an unconditional basis. However, by relaxing the requirements of orthogonality and unique decompositions, we can obtain collections $\{e^{2\pi ibnx}g(x+am)\}$ with well-behaved g which do form frames. These prove sufficient to our need.

The remainder of this paper provides a sketch of how frames play a role in proving Theorem 1. First, Section 2 collects the necessary machinery on the Weyl correspondence. Section 3 introduces frames, especially those formed via time-frequency shifts of a single g. Finally, Section 4 indicates the connection—how Gabor frames can be used with the Weyl correspondence.

2. THE WEYL CORRESPONDENCE

We use the standard energy norm for signals f in $L^2(\mathbf{R})$, the space of all square-integrable functions on the real line $\mathbf{R}$. This is defined by $\|f\| = (\int |f(t)|^2\, dt)^{1/2}$. The inner product of two signals f and g is $\langle f, g\rangle = \int f(t)\, \overline{g(t)}\, dt$, the overbar denoting complex conjugation.

The *time-frequency shift* of a signal f by (a, b) is the function $\rho(a, b)f$ defined by

$$\rho(a,b)f(t) \;=\; e^{\pi iab}\, e^{2\pi ibt}\, f(t+a).$$

The *cross ambiguity function* $A(f, g)$ of two signals f and g is the function defined on the time-frequency plane $\mathbf{R}^2$ by

$$\begin{aligned} A(f,g)(p,q) \;=\; \langle \rho(p,q)f, g\rangle \;&=\; \int e^{\pi ipq}\, e^{2\pi iqt}\, f(t+p)\, \overline{g(t)}\, dt \\ &=\; \int e^{2\pi iqs}\, f(s+p/2)\, \overline{g(s-p/2)}\, ds. \end{aligned}$$

Thus, the ambiguity function is a time-frequency cross-correlation between the two signals. If $f,\, g \in L^2(\mathbf{R})$ then $A(f, g)$ is a bounded, square-integrable function on $\mathbf{R}^2$.

The Wigner distribution is the two-dimensional Fourier transform of the cross-ambiguity function. It is a function $W(f, g)$ defined on the time-frequency plane $\mathbf{R}^2$, and can be written

$$W(f,g)(p,q) \;=\; (A(f,g))^{\wedge}(p,q) \;=\; \int e^{-2\pi isp}\, f(q+s/2)\, \overline{g(q-s/2)}\, ds.$$

Therefore, the Wigner distribution can be interpreted as a time-varying spectrum. If $f,\, g \in L^2(\mathbf{R})$ then $W(f, g)$ is a bounded and square-integrable function on $\mathbf{R}^2$.

There are many other possible ways to obtain time-varying spectra. A broad class of these are the time-frequency distributions $W_\mu(f, g)$ in Cohen's class. Cohen's survey paper[2] contains an excellent discussion of the definitions, properties, and uses of these distributions. To obtain these we take a μ satisfying certain conditions, and obtain $W_\mu(f, g)$ by convolving the Wigner distribution $W(f, g)$ with μ:

$$W_\mu(f,g)(p,q) \;=\; (W(f,g)*\mu)(p,q) \;=\; \iint W(f,g)(a,b)\; \mu(p-a, q-b)\, da\, db.$$

A relevant fact is that the kernel $\hat{\mu}$ always has the property that it is a bounded function. However, μ itself need not be realizable as a function. For example, we may take μ to be the point mass δ; this is not a function, but its Fourier transform is the function which is identically 1. Although μ need not be a function, the convolution of $W(f,g)$ with μ will always produce a function $W_\mu(f,g)$ which is square-integrable. For, if we begin on the Fourier transform side, we know that the inverse Fourier transform $\check{\mu}$ is a bounded function. Therefore $A(f,g) \cdot \check{\mu}$ is square-integrable, and hence

$$W_\mu(f,g) \;=\; W(f,g) * \mu \;=\; (A(f,g) \cdot \check{\mu})^\wedge$$

is a well-defined, square-integrable function. Note that if $\mu = \delta$ then $W_\mu(f,g) = W(f,g)$.

The Weyl correspondence uses the Wigner distribution to define a correspondence between functions $\sigma(p,q)$ and operators $L(\sigma,\mu)$ mapping $L^2(\mathbf{R})$ into itself. The operator $L(\sigma,\mu)$ is defined implicitly by the equation

$$\langle L(\sigma,\mu)f, g\rangle \;=\; \langle \sigma, \overline{W_\mu(f,g)}\rangle \;=\; \iint \sigma(p,q)\, W_\mu(f,g)(p,q)\, dp\, dq.$$

A simple extension of a result of Pool[14] shows that if $\sigma \in L^2(\mathbf{R}^2)$ then $L(\sigma,\mu)$ is a *Hilbert–Schmidt* operator on $L^2(\mathbf{R})$. In particular, this means that the singular values s_n of $L(\sigma,\mu)$ must be square-summable: $\sum |s_n|^2 < \infty$.

Actually, each $L(\sigma,\mu)$ can be realized as $L(\tilde{\sigma},\delta)$ for a new symbol $\tilde{\sigma}$ defined by

$$\tilde{\sigma}(a,b) \;=\; (\sigma * \tilde{\mu})(a,b) \;=\; \iint \sigma(p,q)\, \tilde{\mu}(a-p, b-q)\, dp\, dq,$$

where $\tilde{\mu}(a,b) = \mu(-a,-b)$.

3. FRAMES

Let $\{g_{mn}\}_{m,n\in\mathbf{Z}}$ be a collection of functions from $L^2(\mathbf{R})$. If $\{g_{mn}\}$ is a basis for $L^2(\mathbf{R})$ then we know, by definition, that each signal $f \in L^2(\mathbf{R})$ can be written

$$f \;=\; \sum_{m,n} c_{mn}(f)\, g_{mn}, \tag{1}$$

for some *unique* choice of coefficients $\{c_{mn}(f)\}$. We say that the basis is *unconditional* if the order of the summation in (1) is unimportant: every reording of the sum also converges, in which case it must converge to the same value.

An *orthonormal* basis has the further property that the functions are orthonormal: $\langle g_{mn}, g_{m'n'}\rangle = 0$ unless $m = m'$ and $n = n'$, in which case the inner product is 1. In this case, the coefficients $c_{mn}(f)$ are easy to calculate: they are the inner products $c_{mn}(f) = \langle f, g_{mn}\rangle$. Moreover, we have for orthonormal bases a *Plancherel formula*:

$$\sum_{m,n} |\langle f, g_{mn}\rangle|^2 \;=\; \|f\|^2, \qquad \text{all } f \in L^2(\mathbf{R}).$$

An ideal situation for time-frequency analysis would be the existence of a function g which is well-localized in both time and frequency and such that we can obtain an orthonormal basis $\{g_{mn}\}$ by simply letting g_{mn} be a time-frequency shift of g, i.e.,

$$g_{mn}(x) \;=\; \rho(am,bn)g(x) \;=\; e^{\pi iabmn}\, e^{2\pi ibnx}\, g(x+am), \tag{2}$$

where a and b are fixed time and frequency step-sizes, respectively. However, the *Balian–Low Theorem*[1] prevents this. Relaxing the requirement of orthogonality does not help: we cannot construct even an unconditional basis $\{g_{mn}\}$ by the method in (2) unless g has poor localization properties, specifically, g must satisfy $\|xg(x)\| \cdot \|\gamma\hat{g}(\gamma)\| = \infty$.

However, we can choose to further relax the requirements on $\{g_{mn}\}$. In addition to relaxing the orthogonality requirement, we can also relax the requirement for uniqueness of the decompositions in (1). Merely having *computable* decompositions may be sufficient, even if they are not unique. Frames provide a method of achieving this.

Definition. A collection $\{g_{mn}\}$ of functions from $L^2(\mathbf{R})$ is a *frame* if there exist constants A, $B > 0$ such that

$$A\,\|f\|^2 \;\le\; \sum_{m,n} |\langle f, g_{mn}\rangle|^2 \;\le\; B\,\|f\|^2, \qquad \text{all } f \in L^2(\mathbf{R}). \tag{3}$$

The numbers A, B are the *frame bounds.*

Thus, $\{g_{mn}\}$ is a frame if the *pseudo-Plancherel formula* (3) holds. All orthonormal bases are frames. In fact, all unconditional bases which have the additional "norm boundedness" property

$$0 \;<\; \inf_{m,n} \|g_{mn}\| \;\le\; \sup_{m,n} \|g_{mn}\| \;<\; \infty \tag{4}$$

are frames. Conversely, any frame which is a basis is in fact an unconditional basis which satisfies the boundedness criteria (4).

It is a remarkable fact that the pseudo-Plancherel formula (3) alone ensures that frames are practical alternatives to bases. In fact, the pseudo-Plancherel formula implies the existence of decompositions of signals with respect to the frame elements. (For details, we refer to Duffin and Schaeffer's original paper,[5] to Daubechies' monograph,[4] or to the research-tutorial of Heil and Walnut.[8]) If $\{g_{mn}\}$ is a frame, then there will be a collection $\{\tilde{g}_{mn}\}$ which is its *dual frame*: the collection $\{\tilde{g}_{mn}\}$ will also be a frame (with frame bounds $1/B$, $1/A$), and every $f \in L^2(\mathbf{R})$ can be written

$$f \;=\; \sum_{m,n} c_{mn}(f)\, g_{mn} \qquad \text{and} \qquad f \;=\; \sum_{m,n} \tilde{c}_{mn}(f)\, \tilde{g}_{mn} \tag{5}$$

with

$$c_{mn}(f) \;=\; \langle f, \tilde{g}_{mn}\rangle \qquad \text{and} \qquad \tilde{c}_{mn}(f) \;=\; \langle f, g_{mn}\rangle. \tag{6}$$

Importantly, the summations in (5), with coefficients given by (6), converge unconditionally, i.e., they converge regardless of the order of summation. However, the decompositions in (5) need not be unique: there may exist other choices of coefficients which can be used to represent f as a combination of the frame elements. Yet there always is at least one computable choice of coefficents, given by (6). Moreover, this choice satisfies the pseudo-Plancherel formula (3), meaning that the energy of the signal f is equivalent (up to lower and upper constant multiples) to the "energy" of the coefficients $\{c_{mn}(f)\}$ or $\{\tilde{c}_{mn}(f)\}$.

Of special interest are collections $\{g_{mn}\}$ obtained as in (2) by time-frequency shifts of a single g. That is, $g_{mn}(x) = \rho(am, bn)g$ for some fixed choice of a and b. We call such collections *Gabor systems.* If a Gabor system is a frame then we call it simply a *Gabor frame.* (These are also known as *windowed Fourier transform frames* or *Weyl–Heisenberg frames.*) In order to obtain a frame there must be restrictions on g,

a, and b. In particular, if the product ab of the time-frequency step sizes is too large, namely $ab > 1$, then the Gabor system $\{g_{mn}\}$ constructed from any choice of g must be incomplete.[15] Thus, a Gabor system can be a frame only when $ab \le 1$ (this is not sufficient: additional requirements must be imposed on g). A Gabor frame can be a basis only when $ab = 1$.[15] A Gabor frame with $ab < 1$ is necessarily *overcomplete*, in fact, it must contain infinitely many redundant elements. Yet it is the presence of those redundant elements that allows us to construct Gabor frames whose elements have good time-frequency localization.

Example 4. Let $G(x) = e^{-x^2}$, i.e., G is the Gaussian function, the function which has the best possible simultaneous localization in time and frequency, the function which minimizes the quantity $\|xg(x)\| \cdot \|\gamma\hat{g}(\gamma)\|$. Because of the Balian–Low Theorem, the Gabor system $\{G_{mn}\}$ cannot be a basis for any choice of a, b. However, it is easy to show[4] that $\{G_{mn}\}$ does form a Gabor frame when $ab = 1/2$. In fact, Seip and Wallstén[18] have shown that $\{G_{mn}\}$ is a Gabor frame for any choice of a, b satisfying $ab < 1$.

4. FRAMES AND THE WEYL CORRESPONDENCE

The *singular values* $s_n(L)$ of a compact operator L on $L^2(\mathbf{R})$ are obtained from the eigenvalues $\lambda_n(L^*L)$ of the compact, positive, self-adjoint operator L^*L. With the eigenvalues $\lambda_n(L^*L)$ arranged in order of decreasing value (they are all nonnegative), we set:

$$s_n(L) \; = \; \lambda_n(L^*L)^{1/2}.$$

If L is self-adjoint then $s_n(L) = |\lambda_n(L)|$. The operator L is *Hilbert–Schmidt* if the singular values are square-summable. In this case we say that the *Hilbert–Schmidt norm* of L is

$$\|L\|_{HS} \; = \; \left(\sum_{n=1}^{\infty} s_n(L)^2 \right)^{1/2}.$$

The basic tool which allows us to estimate the singular values of $L(\sigma, \mu)$ is the following standard result,[6] which is the extension to singular values of a result first proved by Weyl for eigenvalues of certain self-adjoint operators.[20]

Lemma. *If L_1, L_2 are compact operators then*

$$s_{m+n+1}(L_1 + L_2) \; \le \; s_{m+1}(L_1) \; + \; s_{n+1}(L_2).$$

If one of the operators L_1 or L_2 has finite rank, then that operator has only finitely many nonzero singular values. This fact can be used to obtain an estimate for the "tail" of the singular values of the other operator. In particular, if L is Hilbert–Schmidt and L_N is an approximation of L with finite rank N, then

$$\sum_{n=N+1}^{\infty} s_n(L)^2 \; \le \; \|L - L_N\|_{HS}^2.$$

This tail estimate can be turned into an estimate on the individual singular values:[9] there must be a constant C such that

$$s_{2N}(L) \; \le \; C \, \frac{\|L - L_N\|_{HS}}{\sqrt{N}}.$$

Thus, the problem of estimating the decay rate of the singular values of $L = L(\sigma, \mu)$ reduces to a problem of finding a sequence of finite-rank approximations so that the Hilbert–Schmidt norms $\|L - L_N\|_{HS}$ are computable. Here is where frames prove useful. Let $G(x) = e^{-x^2}$ be the Gaussian function. We know then that the Gabor system $\{G_{mn}\}_{m,n\in\mathbf{Z}}$ forms a frame for $L^2(\mathbf{R})$ if we choose the step sizes a, b so that $ab < 1$. By applying properties of the cross ambiguity function, we can use this frame for $L^2(\mathbf{R})$ to construct a frame for $L^2(\mathbf{R}^2)$: the collection

$$\{A(G_{kl}, G_{mn})\}_{k,l,m,n\in\mathbf{Z}}$$

is a frame for $L^2(\mathbf{R}^2)$. Therefore, there is a dual frame, and it can be shown that it has the form

$$\{A(\tilde{G}_{kl}, \tilde{G}_{mn})\}_{k,l,m,n\in\mathbf{Z}}$$

for some appropriate function $\tilde{G} \in L^2(\mathbf{R})$. It is possible to specifically calculate $\tilde{G}$, but this is not necessary for our purpose—we only need to know that it exists.

Now we construct a finite-rank approximation of $L(\sigma, \mu)$. Set $\tilde{\sigma} = \sigma * \tilde{\mu} \in L^2(\mathbf{R}^2)$. Then the decomposition property of frames tells us that

$$\tilde{\sigma} \;=\; \sum_{k,l,m,n\in\mathbf{Z}} \langle \tilde{\sigma}, A(G_{kl}, G_{mn})\rangle \, A(\tilde{G}_{kl}, \tilde{G}_{mn}),$$

with a pseudo-Plancherel relation between $\|\tilde{\sigma}\|^2$ and $\sum |\langle \tilde{\sigma}, A(G_{kl}, G_{mn})\rangle|^2$. By summing only finitely many terms, we can obtain a function $\tilde{\sigma}_N$ which approximates $\tilde{\sigma}$: define

$$\tilde{\sigma}_N \;=\; \sum_{(k,l,m,n)\in D_N} \langle \tilde{\sigma}, A(G_{kl}, G_{mn})\rangle \, A(\tilde{G}_{kl}, \tilde{G}_{mn}),$$

where $D_N = \{(k,l,m,n) : |k|, |l|, |m|, |n| \leq N\}$. It is then possible to translate this approximation of the *function* $\tilde{\sigma}$ back to an *approximating operator* L_N which is finite-rank. Moreover, the approximation in the Hilbert–Schmidt norm is

$$\|L - L_N\|_{HS} \;=\; \|\tilde{\sigma} - \tilde{\sigma}_N\|,$$

and we know that the square of the latter quantity is equivalent up to upper and lower constant factors (via the pseudo-Plancherel formula) to

$$\sum_{(k,l,m,n)\notin D_N} |\langle \tilde{\sigma}, A(G_{kl}, G_{mn})\rangle|^2.$$

Now the problem becomes one of calculating these factors. We are helped by the fact that the cross ambiguity functions $A(G_{kl}, G_{mn})$ are easy to calculate: they are essentially time and frequency shifted two-dimensional Gaussian functions. Therefore they are each well-localized. After doing the calculations, we complete a proof of Theorem 1.

5. ACKNOWLEDGMENTS

We gratefully acknowledge the following financial support. Heil was partially supported by the MITRE Sponsored Research Program and by National Science Foundation Grants DMS-9007212 and DMS-9401340. Ramanathan was supported by a Faculty Research and Creative Projects Fellowship Award from Eastern Michigan University. Topiwala was supported by the MITRE Sponsored Research Program.

6. REFERENCES

1. J. Benedetto, C. Heil, and D. Walnut, *Remarks on the proof of the Balian–Low theorem*, preprint.

2. L. Cohen, "Time-frequency distributions—a review," *Proceedings IEEE*, vol. 77, pp. 941–981, 1989.

3. I. Daubechies, "Time-frequency localization operators–a geometric phase space approach", *IEEE Transactions on Information Theory*, vol. 24, pp. 605–612, 1988.

4. I. Daubechies, *Ten Lectures on Wavelets*, SIAM Press, Philadelphia, 1993.

5. R.J. Duffin and A.C. Schaeffer, "A class of nonharmonic Fourier series," *Transactions Amer. Math. Soc.*, vol. 72, pp. 341–366, 1952.

6. N. Dunford and J.T. Schwartz, *Linear Operators, Part II*, Wiley, New York, 1988.

7. P. Flandrin, "Maximal signal energy concentration in a time-frequency domain," *ICASSP 88*, pp. 2176–2179, 1988.

8. C. Heil and D. Walnut, "Continuous and discrete wavelet transforms," *SIAM Review*, vol. 31, pp. 628–666, 1989.

9. C. Heil, J. Ramanathan, and P. Topiwala, "Singular values of Weyl operators," preprint.

10. F. Hlawatsch, W. Kozek and W. Krattenthaler, "Time frequency subspaces and their application to time-varying filtering," *ICASSP 90*, pp. 1609–1610, 1990.

11. H.J. Landau, "The eigenvalue behavior of certain convolution equations," *Transactions Amer. Math. Soc.*, vol. 115, pp. 242–256, 1965.

12. H.J. Landau and H.O. Pollack, "Prolate Spheroidal Wavefunctions, Fourier Analysis and Uncertainty, II," *Bell Systems Technical J.*, vol. 40, pp. 65–84, 1961.

13. H.J. Landau and H.O. Pollack, "Prolate Spheroidal Wavefunctions, Fourier Analysis and Uncertainty, III," *Bell Systems Technical J.*, vol. 41, pp. 1295–1336, 1962.

14. J. Pool, "Mathematical aspects of the Weyl correspondence," *J. Mathematical Physics*, vol. 7, pp. 66–76, 1966.

15. J. Ramanathan and T. Steger, "Incompleteness of Sparse Coherent States," preprint.

16. J. Ramanathan and P. Topiwala, "Time-frequency localization and the spectrogram," *Applied and Computational Harmonic Analysis*, vol. 1, pp. 209–215, 1994.

17. J. Ramanathan and P. Topiwala, "Time-frequency localization via the Weyl correspondence," *SIAM J. Mathematical Analysis*, vol. 24, pp. 1378–1383, 1993.

18. K. Seip and R. Wallstén, "Sampling and interpolation in the Bargmann–Fock space," preprint.

19. D. Slepian and H.O. Pollack, "Prolate Spheroidal Wavefunctions, Fourier Analysis and Uncertainty, I," *Bell Systems Technical J.*, vol. 40, pp. 43–63, 1961.

20. H. Weyl, "Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung)," *Mathematische Annalen*, vol. 71, pp. 441-479, 1911.