

Unitary Equivalence: A New Twist on Signal Processing

Richard G. Baraniuk, *Member, IEEE*, and Douglas L. Jones, *Member, IEEE*

Abstract—Unitary similarity transformations furnish a powerful vehicle for generating infinite generic classes of signal analysis and processing tools based on concepts different from time, frequency, and scale. Implementation of these new tools involves simply preprocessing the signal by a unitary transformation, performing standard processing on the transformed signal, and then (in some cases) transforming the resulting output. The resulting unitarily equivalent systems can focus on the critical signal characteristics in large classes of signals and, hence, prove useful for representing and processing signals that are not well matched by current techniques. As specific examples of this procedure, we generalize linear time-invariant systems, orthonormal basis and frame decompositions, and joint time-frequency and time-scale distributions. These applications illustrate the utility of the unitary equivalence concept for uniting seemingly disparate approaches proposed in the literature.

I. INTRODUCTION

THE time and frequency coordinate systems play such a fundamental role in signal analysis and processing that it is virtually impossible to consider the subject of signal processing without them. Aside from their central conceptual function, a vast body of effective, efficient, and robust signal processing algorithms have been developed within these frameworks. The Fourier transform, for example, is precisely the mapping between the time and frequency domains or coordinate systems.

Nonetheless, for many types of signals in important applications, standard tools have proven inadequate. The Fourier transform, to continue the example, does not explicitly indicate how the spectral components of a signal change over time, which is essential in applications such as speech, radar, sonar, biological, and transient signal analysis [1]. Hence, joint representations based simultaneously on both time and frequency have been created, including the narrowband ambiguity function [1], the windowed short-time Fourier transform, the Gabor [2] and Wilson [3] orthonormal bases and frames, and the spectrogram and its generalization, Cohen's class of bilinear time-frequency distributions [1].

Manuscript received April 27, 1993; revised February 13, 1995. This work was completed while R. Baraniuk was with Ecole Normale Supérieure de Lyon, France and was supported by an NSERC-NATO postdoctoral fellowship, the National Science Foundation, grant nos. MIP 90-12747 and MIP 94-57438, the Texas Advanced Technology Program, grant no. TX-ATP 003604-002, and the Joint Services Electronics Program, grant no. N00014-90-J-1270. The associate editor coordinating the review of this paper and approving it for publication was Dr. Boualem Boashash.

R. G. Baraniuk is with the Department of Electrical and Computer Engineering, Rice University, Houston, TX 77251-1892 USA.

D. L. Jones is with the Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801 USA.

IEEE Log Number 9413864.

In turn, however, the constant time-frequency resolution analysis effected by these time-frequency representations has proven unsuitable for the wideband signals appearing in applications such as image processing and wideband Doppler signal processing. For these types of signals, proportional-bandwidth analysis tools based on the concept of "scale" have been developed, including the Mellin transform [4], [5]; the joint time-scale wavelet transform and wavelet orthonormal bases and frames [2]; and bilinear wavelet generalizations such as the Altes-Marinovich distribution [6], [7], the affine Wigner distributions of the Bertrands [8]–[10], the affine class [11]–[13], the time-frequency-scale classes of Cohen [1], [14], and the hyperbolic and power classes of Papandreou *et al.* [15], [16]. Further generalizations of the time-frequency and time-scale techniques well adapted to studying chirping signals have been provided by the chirplet/metaplectic transform framework studied by Berthon [17], Grossman, and Paul [18], Mann and Haykin [19], [20], and Baraniuk and Jones [21], [22], and the polynomial Wigner distribution studied by Boashash and O'Shea [23].

While tools such as the Fourier transform, the short-time Fourier transform, the wavelet transform, and the bilinear time-frequency and time-scale classes are natural for many signals, there still exist large classes of signals (frequency modulated and dispersed signals are two examples) for which neither a time-frequency nor a time-scale analysis is appropriate. These types of signals demand new, better matched analysis and processing tools.

Historically, new signal processing tools have been developed in a piecemeal fashion, with a new tool being created for each new signal class of interest. We do not take this approach here. Rather, in this paper, we present a theory for designing infinite generic classes of signal analysis and processing systems based on alternate coordinate systems. The benefits of the theory are twofold: First, it allows the almost trivial development of many new processing schemes tailored to certain signal characteristics, and second, it allows the use of well-understood, robust, and efficient classical algorithms in the new contexts.

Our approach is based on a special family of "basis changing" operators—the unitary transformations—which convert traditional systems into new systems with different properties [24]. Fig. 1 illustrates the general scheme: An arbitrary conventional processing system P (linear filter, adaptive filter, spectral estimator, detector, time-frequency or time-scale representation, etc.) is cradled between two unitary transformations U and V . As we will see, the unitary transformations



Fig. 1. Prototype unitarily equivalent system: P represents a traditional analysis/processing system, whereas U and V are unitary transformations [24].

change the fundamental coordinate system of the processor P , mapping familiar concepts such as time and frequency to new concepts more natural for the analysis and processing of certain types of signals.¹

As a simple example of the procedure we will develop, consider the recovery of a harmonic signal with slowly time-varying frequencies in a noisy environment, as might arise in the acoustic emissions of rotating machinery in the presence of Doppler shift or in the removal of background noise from a musical solo. Given the signal

$$s(x) = h[m(x)] + n_1(x) \quad (1)$$

with $h(x) = \sum_k a_k e^{j2\pi k f_0 x}$, m a smooth, monotonic function, and n_1 a noise realization, we wish to recover the harmonic component h . Clearly, this signal is instantaneously harmonic and therefore calls for a set of narrow bandwidth filters; however, time-invariant filters would have to include the entire modulation bandwidth. Time-frequency filtering methods [25]–[27] could be applied, but they are usually computationally expensive and, moreover, are not fully characterized in terms of their statistical performance. Demodulation methods are likewise unattractive for this application because they do not preserve the scaling relationship of the harmonics. An alternative solution more suitable for such signals first warps the time axis of s by the function $w = m^{-1}$. Let U represent this warping operator, that is, let

$$(Us)(x) = |\dot{w}(x)|^{1/2} s[w(x)] = |\dot{w}(x)|^{1/2} h(x) + n_2(x) \quad (2)$$

where the envelope $|\dot{w}|^{1/2}$ containing the derivative of w preserves the energy in the signal at the output of U . With respect to the warped time axis, the component h is once again harmonic, and a simple linear time-invariant filter P designed to pass all multiples of the frequency f_0 , such as a comb filter, can be applied to remove most of the noise (under the “slowly varying assumption” on m , the derivative weighting is approximately constant). An inverse warping $V = U^{-1}$ takes the denoised signal back to the normal time domain if desired. Fig. 2 illustrates the results of a numerical simulation of the system $U^{-1}PU$ for a sawtooth wave h and a sinusoidally varying warp m .²

While perhaps somewhat simplistic, this example illustrates the central features of the proposed approach. The key point is that often, the application of a simple transformation or change of basis can turn a difficult, expensive problem into

¹This paper is certainly not the first to propose a processing scheme like that in Fig. 1. Particular unitary transformations have a long history in signal processing and have been used to great advantage in a number of applications, including transform coding, transform domain adaptive filtering, and demodulation. However, the theory developed in this paper both unites these specific approaches and generalizes readily to other applications.

²In the process of mapping n_1 in (1) to n_2 in (2), U changes the statistics of the additive noise. Fortunately, however, if n_1 is white, then n_2 is white as well.

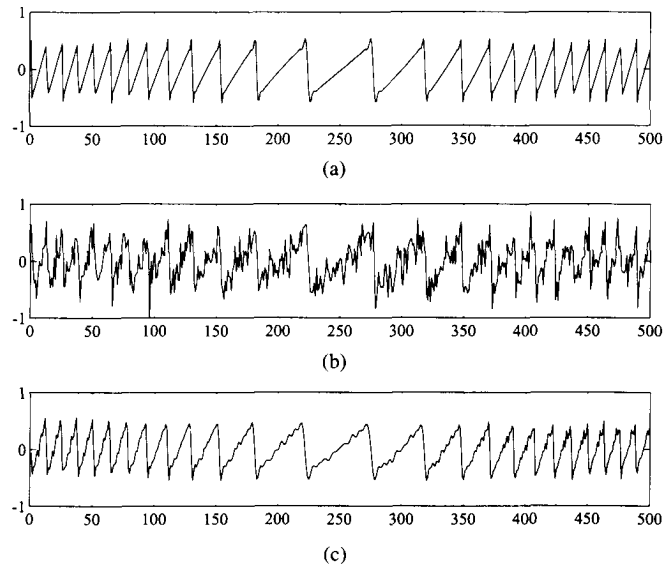


Fig. 2. Retrieval of a slowly time-varying harmonic signal from a noisy environment [see (1) and (2)]. (a) The warped signal $h[m(x)]$, where $h(x)$ is a harmonic sawtooth wave and $m(x) = x + 50.5 \sin(2\pi x/500)$. The signal h could represent a musical note with vibrato or the acoustic emissions from Doppler-shifted rotating machinery. (b) The signal s , obtained by immersing the warped signal from (a) in 3 dB SNR additive white noise. (c) Signal recovered after prewarping the axis of s by the function $w = m^{-1}$, comb filtering to remove noise, and postwarping by the function m .

an easy one that can be solved using standard methods. A first interpretation of the example suggests that the operator U allowed us to apply traditional time-invariant processing techniques to a signal that otherwise would have required much more complicated time-frequency filtering. A second, dual, interpretation suggests that the operator U converted the fixed comb filter—a traditional and well understood signal processing tool—into a new tool well suited to dealing with time-varying frequencies of a certain type. In both interpretations, the operator U provides the key link.

This paper is organized as follows. In the following section, after a brief discussion of operator representations of physical concepts, we introduce the concept of *unitary equivalence*, which is based on the cascade of systems shown in Fig. 1. The bulk of the paper then comprises three sections, each pertaining to a major application of unitary equivalence: linear time-invariant systems in Section III, orthonormal basis and frame decompositions in Section IV, and joint time-frequency and time-scale distributions in Section V. A discussion and conclusion appear in the final section. In order to maintain a manageable scope in this paper, we focus on the power of the unitary equivalence concept for generalizing existing signal processing tools and for uniting seemingly disparate approaches proposed in the literature. We will not address extensively the important question of choosing the basis transformation most appropriate for a given data set.

II. UNITARY EQUIVALENCE

A. Preliminaries on Operators

The foundation of the unitary equivalence principle rests on three classes of operators: parameterized unitary and Hermi-

tian operators representing physical quantities, unitary signal transforms acting as density functions for physical quantities, and unitary coordinate transformations.

Notation: We will operate primarily in the Hilbert space of square-integrable functions $L^2(\mathbb{R})$, which has inner product $\langle g, h \rangle \equiv \int g(x)h^*(x) dx$ for $g, h \in L^2(\mathbb{R})$ and norm $\|g\|^2 \equiv \langle g, g \rangle$. For convenience, we will also employ nonsquare-integrable functions such as the Dirac delta $\delta(x)$ and the complex sinusoid $e^{j2\pi fx}$. Both unitary and Hermitian operator³ representations prove useful in our development. In keeping with the standard set by [1], we will obey the following notation convention: Lowercase letters will denote functions and physical quantities, script capitals will denote Hermitian operators, and boldface capitals will denote all other operators, including unitary transformations.⁴ As our only exception, we will use the special symbol \mathcal{F} to represent the Fourier transform.

Operator Representations of Physical Quantities: We follow the approach of Gabor [32], Ville [33], and Cohen [1], [14] by associating physical quantities such as time, frequency, and scale with operators. To each abstract physical quantity “ a ,” this formalism links three equivalent representations: the variable a (a real number), a unitary operator \mathbf{A}_a parameterized by the value a , and a Hermitian operator \mathcal{A} . In the time domain, the unitary representations of time, frequency, and scale correspond to the time-shift, frequency-shift, and dilation operators, defined as

$$\begin{aligned} \text{Time shift:} \quad & (\mathbf{T}_t g)(x) \equiv g(x - t) \\ \text{Frequency shift:} \quad & (\mathbf{F}_f g)(x) \equiv e^{j2\pi fx} g(x) \\ \text{Dilation:} \quad & (\mathbf{D}_d g)(x) \equiv e^{-d/2} g(e^{-d}x). \end{aligned} \quad (3)$$

We will drop the subscripts when our intent is clear from context. The Hermitian representations of time, frequency, and a quantity we will term logarithmic modulation (for reasons to be explained later) are defined in the time domain as [1], [10]

$$\begin{aligned} \text{Time:} \quad & (\mathcal{T}g)(x) \equiv xg(x) \\ \text{Frequency:} \quad & (\mathcal{F}g)(x) \equiv \frac{1}{j2\pi} \dot{g}(x) \\ \text{Log modulation:} \quad & (\mathcal{H}g)(x) \equiv \left(\frac{\mathcal{T}\mathcal{F} + \mathcal{F}\mathcal{T}}{2} g \right) (x) \end{aligned} \quad (4)$$

Operator representations in domains different from the time domain are easily obtained using a similarity transformation (as we will do in Section II-B).

While the unitary and Hermitian representations of time, frequency, and scale appear very different, they are, in fact, equivalent. In particular, Stone’s theorem [34, p. 614] states that for each parameterized unitary operator \mathbf{A}_a such that $\mathbf{A}_{a_1}\mathbf{A}_{a_2} = \mathbf{A}_{a_1+a_2}$, there exists a unique Hermitian operator \mathcal{B} such that

$$\mathbf{A}_a = e^{j2\pi a\mathcal{B}} = \sum_{n=0}^{\infty} \frac{(j2\pi a\mathcal{B})^n}{n!}$$

³A unitary operator \mathbf{U} is a linear transformation from one Hilbert space onto another that preserves energy, that is, $\|\mathbf{U}g\|^2 = \|g\|^2$. Unitary operators also preserve inner products (isometry), that is, $\langle \mathbf{U}g, \mathbf{U}h \rangle = \langle g, h \rangle$. An operator \mathcal{U} is Hermitian (self-adjoint) if $\langle \mathcal{U}g, h \rangle = \langle g, \mathcal{U}h \rangle$.

⁴Beware: The current notation does *not* match that of [24], [28]–[31].

and

$$\mathcal{B} = \frac{1}{j2\pi} \lim_{a \rightarrow 0} \frac{\mathbf{A}_a - \mathbf{I}}{a}.$$

(Here, \mathbf{I} denotes the identity operator.) For time, frequency, and scale, we have

$$\mathbf{T}_{-t} = e^{j2\pi t\mathcal{F}}, \quad \mathbf{F}_f = e^{j2\pi f\mathcal{T}}, \quad \mathbf{D}_d = e^{j2\pi d\mathcal{H}}. \quad (5)$$

Generalized Fourier Transforms: The second important class of unitary transformations we will utilize consists of the generalized Fourier transforms, which can be interpreted as densities measuring the content of a certain physical quantity in a time signal. Given a unitary operator \mathbf{A} , we define the \mathbf{A} -Fourier transform $\mathcal{F}_{\mathbf{A}}$ as the expansion onto the eigenfunctions $\mathbf{u}_{\alpha}^{\mathbf{A}}(x)$ of \mathbf{A} . The forward transform \hat{s} of a signal $s \in L^2(\mathbb{R})$ is given by⁵

$$\begin{aligned} \hat{s}(\alpha) &= (\mathcal{F}_{\mathbf{A}}s)(\alpha) \equiv \langle s(x), \mathbf{u}_{\alpha}^{\mathbf{A}}(x) \rangle \\ &= \int s(x)\mathbf{u}_{\alpha}^{\mathbf{A}*}(x) dx \end{aligned}$$

while the inverse transform is given by

$$\begin{aligned} s(x) &= (\mathcal{F}_{\mathbf{A}}^{-1}\hat{s})(x) \equiv \langle \hat{s}(\alpha), \mathbf{u}_{\alpha}^{\mathbf{A}*}(x) \rangle \\ &= \int \hat{s}(\alpha)\mathbf{u}_{\alpha}^{\mathbf{A}}(x) d\alpha. \end{aligned}$$

Both of these transforms are unitary. The \mathcal{A} -Fourier transform generated by the Hermitian operator \mathcal{A} is defined similarly in terms of the eigenfunctions of \mathcal{A} .

Since the eigenfunctions $\mathbf{u}_{\alpha}^{\mathbf{A}}$ and $\mathbf{u}_{\alpha}^{\mathcal{A}}$ do not coincide, the unitary and Hermitian operators \mathbf{A} and \mathcal{A} generate two different signal transforms. Both are fundamental for studying the physical quantity a . Invariance of the \mathbf{A} -Fourier transform to \mathbf{A} (up to a phase factor)

$$|(\mathcal{F}_{\mathbf{A}}\mathbf{A}_a s)(\alpha)| = |(\mathcal{F}_{\mathbf{A}}s)(\alpha)|$$

proves useful when the effect of \mathbf{A} is to be ignored, while covariance of the \mathcal{A} -Fourier transform to \mathbf{A} [35]

$$(\mathcal{F}_{\mathcal{A}}\mathbf{A}_a s)(\alpha) = (\mathcal{F}_{\mathcal{A}}s)(\alpha - a) \quad (6)$$

suggests that $\mathcal{F}_{\mathcal{A}}$ measures the “ a content” of the signal s [1], [24], [28], [29].⁶

Table I summarizes the eigenfunctions of the operators we have considered thus far and prompts the following identifications:

$\mathcal{F}_{\mathcal{F}} = \mathcal{F}_{\mathbf{T}}$ corresponds to the usual Fourier transform \mathcal{F} (invariant to time shifts, covariant to frequency shifts, measures frequency content).

$\mathcal{F}_{\mathcal{T}} = \mathcal{F}_{\mathbf{F}}$ corresponds to the identity transform $\mathbf{I}s = s$ (invariant to frequency shifts, covariant to time shifts, measures time content).

⁵In an abuse of notation, we retain the indices of the functions inside the inner product symbol to indicate the variable of integration.

⁶Covariance by translation in (6) corresponds to only one of a range of possible covariances for an arbitrary operator pair \mathbf{A}, \mathcal{A} . However, for other covariances ($\alpha \mapsto \alpha/a$, for example), only the details change and not the interpretation of $\mathcal{F}_{\mathcal{A}}$ as the \mathcal{A} content measuring transform. Since with our definitions of time, frequency, and scale all covariances turn out to be additive, we will emphasize only this case in this paper. For more information on the general case, see [29].

TABLE I
EIGENFUNCTIONS OF THE OPERATORS DEFINED IN (3) AND (4)

Operators	Eigenfunctions
\mathbf{T}, \mathcal{F}	$\mathbf{e}_k^{\mathbf{T}}(x) = \mathbf{e}_k^{\mathcal{F}}(x) = e^{j2\pi kx}$
\mathbf{F}, \mathcal{T}	$\mathbf{e}_k^{\mathbf{F}}(x) = \mathbf{e}_k^{\mathcal{T}}(x) = \delta(x - k)$
\mathbf{D}, \mathcal{H}	$\mathbf{e}_k^{\mathbf{D}}(x) = \mathbf{e}_k^{\mathcal{H}}(x) = \frac{1}{\sqrt{x}} e^{j2\pi k \log x}, x > 0$

$\mathbf{F}_{\mathcal{H}} = \mathbf{F}_{\mathbf{D}}$ corresponds to the Mellin transform (invariant to scale changes, measures “logarithmic modulation” content) [1], [4], [5], [14].

Coordinate Transformations: Nonparameterized unitary coordinate transformations comprise the third and final class of transformations we will consider. A general formula for representing a coordinate transformation \mathbf{U} on $L^2(\mathbb{R})$ employs the linear superposition

$$(\mathbf{U}s)(x) = \int K_{\mathbf{U}}(x, v)s(v) dv$$

with the requirement that the “rows” and “columns” of the integration kernel $K_{\mathbf{U}}(x, v)$ must both form complete orthonormal sets for $L^2(\mathbb{R})$. Axis warpings of the form

$$(\mathbf{U}s)(x) = |\dot{w}(x)|^{1/2} s[w(x)] \quad (7)$$

with w a smooth, one-to-one function, comprise a large subclass of unitary transformations [15], [16], [24], [36]–[39]. The functions $w(x) = |x|^k \text{sgn}(x)$, $k \neq 0$, and $w(x) = e^x$ provide examples of simple yet useful warpings.

The inner product of two unitary operators can be defined as the trace of their composition or, equivalently, as the inner product between their respective kernel functions

$$\langle \mathbf{A}, \mathbf{B} \rangle \equiv \text{tr } \mathbf{A}\mathbf{B} = \iint K_{\mathbf{A}}(x, v)K_{\mathbf{B}}^*(x, v) dx dv.$$

This definition prompts a notion of the “angle” between two unitary operators and their associated physical concepts. For example, it is easily verified that time and frequency are orthogonal⁷

$$\langle \mathbf{T}_t, \mathbf{F}_f \rangle = 0 \quad \forall t, f \neq 0. \quad (8)$$

However, scale is orthogonal to neither time nor frequency

$$\langle \mathbf{T}_t, \mathbf{D}_d \rangle = \langle \mathbf{F}_f, \mathbf{D}_d \rangle = \frac{e^{d/2}}{1 - e^d} \quad \forall t, f, d \neq 0. \quad (9)$$

B. Unitary Equivalence as a Coordinate Transformation

To change the underlying basis of a signal s , we execute the unitary transformation $s \mapsto \mathbf{U}s$. To change the underlying basis of an operator \mathbf{A} , we execute the similarity transformation $\mathbf{A} \mapsto \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$. The concept of operators equivalent modulo such a change of basis motivates the following definition.

⁷The angular Fourier transform [40] provides an equivalent interpretation.

TABLE II
EFFECT OF A UNITARY PREPROCESSING TRANSFORMATION \mathbf{U} ON THE OPERATORS REPRESENTING PHYSICAL QUANTITIES IN A SYSTEM \mathbf{P} . THE FIRST THREE ENTRIES APPLY TO BOTH UNITARY AND HERMITIAN OPERATORS. THE LAST ENTRY APPLIES TO ALL PAIRS OF UNITARY OPERATORS

System \mathbf{P}	System $\mathbf{P}\mathbf{U}$
\mathbf{A}	$\tilde{\mathbf{A}} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$
$\mathbf{e}_\alpha^{\mathbf{A}}$	$\mathbf{e}_\alpha^{\tilde{\mathbf{A}}} = \mathbf{U}^{-1}\mathbf{e}_\alpha^{\mathbf{A}}$
$\mathbb{F}_{\mathbf{A}}$	$\mathbb{F}_{\tilde{\mathbf{A}}} = \mathbb{F}_{\mathbf{A}}\mathbf{U}$
$\langle \mathbf{A}_a, \mathbf{B}_b \rangle$	$\langle \tilde{\mathbf{A}}_a, \tilde{\mathbf{B}}_b \rangle = \langle \mathbf{A}_a, \mathbf{B}_b \rangle$

Definition: Two operators $\tilde{\mathbf{A}}$ and \mathbf{A} are *unitarily equivalent* if $\tilde{\mathbf{A}} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$, with \mathbf{U} a unitary transformation.

While the operator \mathbf{A} represents a single physical quantity, the unitarily equivalent operator $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$ can represent an infinite spectrum of *different* physical quantities, each corresponding to a particular choice of transformation \mathbf{U} .⁸

The recurring theme of this paper is the application of unitary preprocessing transformations to conventional signal analysis and processing systems \mathbf{P} to create new systems $\mathbf{P}\mathbf{U}$ based on alternative physical quantities and coordinate systems. As we will see, the unitary transform \mathbf{U} maps the physical quantities in \mathbf{P} represented by the unitary operators $\mathbf{A}, \mathbf{B}, \dots$ to new quantities in $\mathbf{P}\mathbf{U}$ represented by the unitarily equivalent operators

$$\tilde{\mathbf{A}} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}, \quad \tilde{\mathbf{B}} = \mathbf{U}^{-1}\mathbf{B}\mathbf{U}, \quad \dots \quad (10)$$

The corresponding Hermitian representations undergo identical transformations, from $\mathcal{A}, \mathcal{B}, \dots$ to $\tilde{\mathcal{A}}, \tilde{\mathcal{B}}, \dots$. Table II summarizes the primary effects on these operator representations. Note that while all eigenfunctions and content measuring transforms change with \mathbf{U} , the relative angles between operators and concepts remain unchanged. Roughly speaking, \mathbf{U} “rotates” all physical quantities through the same angle.

To illustrate, consider time and frequency. These quantities are unitarily equivalent, with the Fourier transform operator \mathbb{F} as link

$$\mathbf{F}_k = \mathbb{F}^{-1} \mathbf{T}_k \mathbb{F}. \quad (11)$$

Application of a similarity transform of the form (10) to \mathbf{T} and \mathbf{F} results in two new operators

$$\tilde{\mathbf{T}}_k = \mathbf{U}^{-1}\mathbf{T}_k\mathbf{U}, \quad \tilde{\mathbf{F}}_k = \mathbf{U}^{-1}\mathbf{F}_k\mathbf{U}. \quad (12)$$

The relationships between these four operators can be summarized in the following diagram:

$$\begin{array}{ccc} \mathbf{T}_k & \xrightarrow{\mathbb{F}} & \mathbf{F}_k \\ \mathbf{U} \downarrow & & \mathbf{U} \downarrow \\ \tilde{\mathbf{T}}_k & \xrightarrow{\mathbf{U}^{-1}\mathbb{F}\mathbf{U}} & \tilde{\mathbf{F}}_k \end{array} \quad (13)$$

where the quantity \mathbf{B} at the head of an arrow labeled \mathbf{Z} is obtained from the quantity \mathbf{A} at the tail as $\mathbf{B} = \mathbf{Z}^{-1}\mathbf{A}\mathbf{Z}$. Results identical to (11)–(13) hold for the Hermitian time and

⁸A group theoretic interpretation of unitary equivalence in the context of signal processing is given in [29].

frequency operators: $\mathcal{T} \mapsto \tilde{\mathcal{T}} = \mathbf{U}^{-1}\mathcal{T}\mathbf{U}$ and $\mathcal{F} \mapsto \tilde{\mathcal{F}} = \mathbf{U}^{-1}\mathcal{F}\mathbf{U}$ such that $\tilde{\mathbf{T}}_{-k} = e^{j2\pi k\tilde{\mathcal{F}}}$ and $\tilde{\mathbf{F}}_k = e^{j2\pi k\tilde{\mathcal{T}}}$.

The generalized Fourier transforms for the transformed time and frequency operators follow directly from (12), Table II, and the corresponding results for time and frequency from the previous section. The $\tilde{\mathbf{T}}$ -covariant, $\tilde{\mathbf{F}}$ -invariant transform

$$\mathbf{F}_{\tilde{\mathcal{T}}} = \mathbf{F}_{\tilde{\mathcal{F}}} = \mathbf{F}_{\mathcal{T}}\mathbf{U} = \mathbf{U} \quad (14)$$

measures $\tilde{\mathbf{T}}$ (transformed time) content in signals, while the $\tilde{\mathbf{F}}$ -covariant, $\tilde{\mathbf{T}}$ -invariant transform

$$\mathbf{F}_{\tilde{\mathcal{F}}} = \mathbf{F}_{\tilde{\mathcal{T}}} = \mathbf{F}_{\mathcal{T}}\mathbf{U} = \mathbf{F}\mathbf{U} \quad (15)$$

measures $\tilde{\mathbf{F}}$ (transformed frequency) content in signals. We will return to these results often in the sequel.

The balance of this paper studies the equivalence classes of unitarily equivalent signal processing tools spawned by various choices for the system \mathbf{P} and the unitary preprocessing transformation \mathbf{U} . We will find the unitary equivalence concept very useful for generalizing the concepts of time, frequency, and scale, mapping them to new concepts that can better match certain classes of signals to be analyzed or processed.

III. UNITARY EQUIVALENCE AND LINEAR TIME-INVARIANT SYSTEMS

Linear time-invariant⁹ (LTI) systems provide a simple setting in which to illustrate the primary effects of unitary coordinate transformations. An LTI system can be interpreted as computing the inner product of the input signal s with a reversed and time-shifted version of the impulse response function g

$$(\mathbf{P}s)(t) = \langle s, \mathbf{T}_t g_r \rangle = \int s(x)g^*(t-x)dx \quad (16)$$

with $g_r(x) = g(-x)$. Time and frequency are the fundamental quantities for LTI systems. Covariance of \mathbf{P} to time shifts (the defining property of an LTI system) follows from the isometry of \mathbf{T} since

$$\begin{aligned} (\mathbf{P}\mathbf{T}_k s)(t) &= \langle \mathbf{T}_k s, \mathbf{T}_t g_r \rangle = \langle s, \mathbf{T}_{-k}\mathbf{T}_t g_r \rangle \\ &= \langle s, \mathbf{T}_{t-k} g_r \rangle = (\mathbf{T}_k \mathbf{P}s)(t) \\ &= (\mathbf{P}s)(t-k). \end{aligned} \quad (17)$$

The expansion $\mathbf{F}_{\mathcal{T}} = \mathbf{F}_{\mathcal{P}} = \mathbf{F}$ onto the eigenfunctions of \mathbf{T} and \mathbf{P} (since \mathbf{T} and \mathbf{P} commute, they share a common set of eigenfunctions) measures frequency content.

As anticipated in (12), the application of a unitary preprocessing transformation \mathbf{U} to the input of an LTI system \mathbf{P} maps the concepts of time and frequency to the concepts associated with the new operators $\tilde{\mathbf{T}} = \mathbf{U}^{-1}\mathbf{T}\mathbf{U}$ and $\tilde{\mathbf{F}} = \mathbf{U}^{-1}\mathbf{F}\mathbf{U}$. The fundamental coordinate system of $\mathbf{P}\mathbf{U}$ changes to \tilde{t} (remapped time) since from (16), we have

$$\begin{aligned} (\mathbf{P}\mathbf{U}s)(\tilde{t}) &= \langle \mathbf{U}s, \mathbf{T}_{\tilde{t}} g_r \rangle = \langle s, \mathbf{U}^{-1}\mathbf{T}_{\tilde{t}} g_r \rangle \\ &= \langle s, (\mathbf{U}^{-1}\mathbf{T}_{\tilde{t}}\mathbf{U})(\mathbf{U}^{-1}g_r) \rangle. \end{aligned}$$

⁹Time-covariant is a more accurate term.

Moreover, a simple calculation similar to (17) demonstrates that the system $\mathbf{P}\mathbf{U}$ is covariant by translation not to \mathbf{T} but to $\tilde{\mathbf{T}}$

$$(\mathbf{P}\mathbf{U}\tilde{\mathbf{T}}_k s)(\tilde{t}) = (\mathbf{P}\mathbf{U}s)(\tilde{t}-k).$$

Thus, unitary preprocessing creates a $\tilde{\mathbf{T}}$ -“invariant” system from a \mathbf{T} -“invariant” one. The fundamental transform for the system $\mathbf{P}\mathbf{U}$ is the $\tilde{\mathbf{T}}$ -Fourier transform $\mathbf{F}_{\tilde{\mathcal{T}}} = \mathbf{F}\mathbf{U}$. This transform measures in time signals the physical quantity associated with the new operator $\tilde{\mathbf{F}}$. Using group theoretic arguments such as in [29], it is simple to show that unitary equivalence can construct LTI-based, \mathbf{A} -invariant systems for all physical quantities a taking on values isomorphic to the real number system \mathbb{R} .

Example—Logarithmic Time-Axis Warping: Consider the unitary warping operator¹⁰ [6], [15]

$$(\mathbf{U}_{\log} s)(x) \equiv e^{x/2} s(e^x) \quad (18)$$

that takes functions in $L^2(\mathbb{R}_+)$ and stretches them into functions in $L^2(\mathbb{R})$. This transformation maps the unitary time operator to the scale operator

$$\mathbf{U}_{\log}^{-1}\mathbf{T}_k\mathbf{U}_{\log} = \mathbf{D}_k$$

on $L^2(\mathbb{R}_+)$ and therefore converts linear time-invariant systems \mathbf{P} into linear *scale*-invariant systems $\mathbf{P}\mathbf{U}_{\log}$ [1], [14], [24], [41]. Scale changes (dilations or compressions) at the inputs to such systems manifest themselves as translations at the outputs. Scale-invariant systems find application wherever scale changes are of fundamental importance; wideband Doppler processing and the detection and estimation of signals of unknown size furnish two examples.

The remaining transformed time and frequency operators are given by

$$\begin{aligned} \mathbf{U}_{\log}^{-1}\mathcal{F}\mathbf{U}_{\log} &= \mathcal{H}, & \mathbf{U}_{\log}^{-1}\mathbf{F}_f\mathbf{U}_{\log} &= \mathbf{H}_f \\ \mathbf{U}_{\log}^{-1}\mathcal{T}\mathbf{U}_{\log} &= \mathcal{D}. \end{aligned}$$

The operators \mathcal{H} and \mathbf{H} correspond to logarithmic modulation; \mathcal{H} was defined in (4), whereas \mathbf{H} is given by

$$(\mathbf{H}_h g)(x) \equiv e^{j2\pi h \log x} g(x), \quad x > 0.$$

The operator \mathcal{D} , the Hermitian scale operator, corresponds to the “logarithmic time” operator [42]

$$(\mathcal{D}g)(x) \equiv (\log x)g(x), \quad x > 0. \quad (19)$$

Note that $\mathbf{H}_h = e^{j2\pi h\mathcal{D}}$. The scale-invariant transform $\mathbf{F}_{\mathcal{D}} = \mathbf{F}\mathbf{U}_{\log}$ coincides with the Mellin transform on $L^2(\mathbb{R}_+)$, as discussed in Section II-A. This transform was named the “scale transform” by Cohen in [1] and [14], but it must be emphasized that its invariance to scale changes prohibits it from measuring scale content in signals. The scale-covariant signal transform,

¹⁰Strictly speaking, (18) is inaccurate in terms of dimensional analysis: The index x in the exponentials of (18) must be dimensionless, yet the function s expects an index with units of seconds. Correct notation for such an expression would involve factors of the form e^{x/x_0} , where x_0 is some arbitrary reference time. We will adopt a much cleaner (although cavalier) notation in our development by suppressing these normalization factors.

which does measure scale content, is $F_{\mathcal{D}} = \mathbf{U}_{\log}$ [28], [29], [35].¹¹

An example of warping an LTI comb filter to match and denoise a warped version $h[m(x)]$ of a harmonic signal $h(x)$ was sketched in the Introduction (see (1), (2), and Fig. 2).

IV. UNITARY EQUIVALENCE AND BASIS AND FRAME DECOMPOSITIONS

In Sections II and III, we found that a unitary similarity transformation takes the Fourier basis $\{\mathbf{u}_f^{\mathcal{F}}\}$, which is natural for LTI systems, to the new basis $\{\mathbf{U}^{-1}\mathbf{u}_f^{\mathcal{F}}\}$, which is natural for signals and systems related to the operators $\tilde{\mathbf{T}} = \mathbf{U}^{-1}\mathbf{T}\mathbf{U}$ and $\tilde{\mathbf{F}} = \mathbf{U}^{-1}\mathbf{F}\mathbf{U}$. This coordinate changing procedure is not limited solely to the Fourier basis, however; it applies equally well to an arbitrary basis or frame. In this section, we will investigate several unitarily equivalent bases for $L^2(\mathbb{R})$. We choose the wavelet, Gabor, and Wilson orthonormal bases to illustrate the procedure; analogous constructions hold for biorthogonal bases and nonorthogonal frames [2]. Since time, frequency, and scale form the foundation for these bases, their unitarily equivalent counterparts utilize transformed time, frequency, and scale.

Given a doubly indexed orthonormal basis $\{\mathbf{b}_{m,n}\}$ for $L^2(\mathbb{R})$, we can decompose any signal $s \in L^2(\mathbb{R})$ as

$$s = \sum_{m,n} \langle s, \mathbf{b}_{m,n} \rangle \mathbf{b}_{m,n}. \quad (20)$$

The elements of a wavelet basis are obtained by translating and scaling a nonarbitrary but fixed wavelet function g_{wavelet} [2]

$$\begin{aligned} \mathbf{b}_{m,n}^{\text{wavelet}}(x) &\equiv (\mathbf{D}'_{d_0^m} \mathbf{T}_{nt_0} g_{\text{wavelet}})(x) \\ &= d_0^{-m/2} g_{\text{wavelet}}(d_0^{-m} x - nt_0) \end{aligned}$$

with $t_0 d_0 = 2$, $m, n \in \mathbb{Z}$, and $\mathbf{D}'_d = \mathbf{D}_{\log d}$. These basis elements "tile" the time-frequency plane in a proportional-bandwidth (constant-Q) fashion; Fig. 3(a) depicts the tiling for an idealized wavelet basis. The elements of a Gabor basis are obtained by translating and modulating a nonarbitrary but fixed window function g_{Gabor} [32]

$$\begin{aligned} \mathbf{b}_{m,n}^{\text{Gabor}}(\tau) &\equiv (\mathbf{F}_{mf_0} \mathbf{T}_{nt_0} g_{\text{Gabor}})(x) \\ &= g_{\text{Gabor}}(x - nt_0) e^{j2\pi m f_0 x} \end{aligned} \quad (21)$$

with $t_0 f_0 = 1$ and $m, n \in \mathbb{Z}$. These basis elements tile the time-frequency plane in a constant-bandwidth fashion; Fig. 3(b) depicts the tiling for an idealized Gabor basis. Recently, constant-bandwidth bases of the Wilson type have been proposed as a well-localized alternative to the Gabor bases [3]. Their construction is similar to (21) but employs sine and cosine modulations.

The wavelet, Gabor, and Wilson bases provide two disparate tilings of the time-frequency plane well suited to

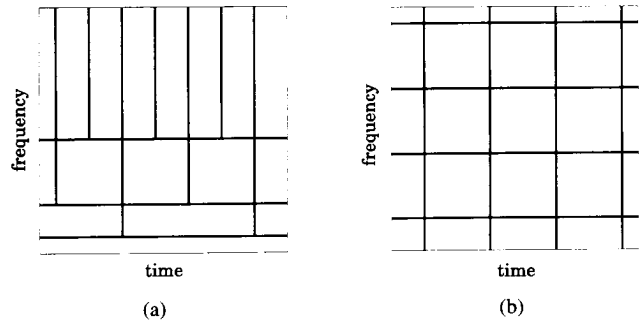


Fig. 3. Idealized tiling of the time-frequency plane by (a) the elements $\{\mathbf{b}_{m,n}^{\text{wavelet}}\}$ of a wavelet orthonormal basis, and (b) the elements $\{\mathbf{b}_{m,n}^{\text{Gabor}}\}$ of a Gabor orthonormal basis.

representing certain classes of signals. However, not all signals are well matched by proportional-bandwidth or constant-bandwidth analyses. For example, the energy of a frequency modulated signal will be spread over many basis coefficients in these types of expansions since such a signal traces a path in the time-frequency plane that matches neither of the basis tilings of Fig. 3.

While the wavelet, Gabor, and Wilson bases may not be the most appropriate for representing certain classes of signals, there may exist unitarily equivalent bases that are appropriate. Consider the effect of operating on these bases with a unitary transformation \mathbf{U}^{-1} . Since \mathbf{U}^{-1} is unitary, the sets $\{\mathbf{U}^{-1}\mathbf{b}_{m,n}^{\text{wavelet}}\}$, $\{\mathbf{U}^{-1}\mathbf{b}_{m,n}^{\text{Gabor}}\}$, and $\{\mathbf{U}^{-1}\mathbf{b}_{m,n}^{\text{Wilson}}\}$ are also valid orthonormal bases for $L^2(\mathbb{R})$. However, while the index parameters n and m represent time and scale in the wavelet bases and time and frequency in the Gabor and Wilson bases, they change to new concepts in these unitarily equivalent bases.

To see this, we can write for the wavelet bases

$$\begin{aligned} \mathbf{U}^{-1}\mathbf{b}_{m,n}^{\text{wavelet}} &= \mathbf{U}^{-1}(\mathbf{D}'_{d_0^m} \mathbf{T}_{nt_0} g_{\text{wavelet}}) \\ &= (\mathbf{U}^{-1}\mathbf{D}'_{d_0^m} \mathbf{U})(\mathbf{U}^{-1}\mathbf{T}_{nt_0} \mathbf{U})(\mathbf{U}^{-1}g_{\text{wavelet}}) \end{aligned} \quad (22)$$

and for the Gabor bases

$$\begin{aligned} \mathbf{U}^{-1}\mathbf{b}_{m,n}^{\text{Gabor}} &= \mathbf{U}^{-1}(\mathbf{F}_{mf_0} \mathbf{T}_{nt_0} g_{\text{Gabor}}) \\ &= (\mathbf{U}^{-1}\mathbf{F}_{mf_0} \mathbf{U})(\mathbf{U}^{-1}\mathbf{T}_{nt_0} \mathbf{U})(\mathbf{U}^{-1}g_{\text{Gabor}}). \end{aligned} \quad (23)$$

(The elements of a unitarily equivalent Wilson basis resemble (23).) Thus, \mathbf{U}^{-1} transforms the original time, frequency, and scale operators to three new unitarily equivalent operators $\tilde{\mathbf{T}}$, $\tilde{\mathbf{F}}$, and $\tilde{\mathbf{D}}$ and maps the original wavelet g_{wavelet} and window g_{Gabor} to two new functions. As with the linear systems of Section III, the key to this procedure is clearly to choose \mathbf{U} so that the unitarily equivalent basis better matches the characteristics of the signals at hand.

Example—Fan Bases: Unitarily equivalent bases were first introduced in [22], [24], [36], and [37], where special polynomial warping operators manufactured bases employing chirp functions. For example, the frequency domain warping transformation $\hat{\mathbf{U}}_c \equiv \mathbf{F}^{-1}\mathbf{U}_c\mathbf{F}$, with

$$(\mathbf{U}_c s)(x) \equiv (|c|^{-1/2} |x|^{(1-c)/2c}) s[|x|^{1/c} \text{sgn}(x)] \quad (24)$$

¹¹ Cohen's definition of $F_{\mathcal{D}}$ as the scale transform follows from his identification of \mathcal{H} in (4) as the Hermitian scale representation [1], [14]. However, closer scrutiny of the relationship between the operator pairs \mathbf{T} , $\tilde{\mathbf{T}}$ and \mathbf{D} , \mathcal{D} reveals that \mathcal{D} defined in (19), rather than \mathcal{H} , is the correct Hermitian scale representation.

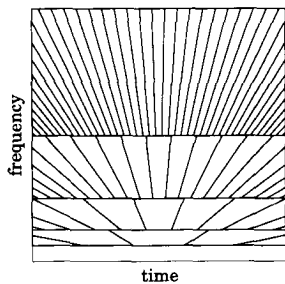


Fig. 4. Idealized tiling of the time-frequency plane by the elements $\{\mathbf{b}_{m,n}^{\text{fan}}\}$ of a fan orthonormal basis for the case $c = 2$ (linear chirp) (see (27)). The tiling is generated from a single, fixed function via scaling and convolution with chirp functions.

for $c \neq 0$, warps the wavelet bases to the equivalence class of *fan bases*. This transformation maps the time operator to a chirp convolution operator of order c and chirp rate t

$$\begin{aligned} (\mathbf{Y}_t^c s)(x) &\equiv (\hat{\mathbf{U}}_c^{-1} \mathbf{T}_t \hat{\mathbf{U}}_c s)(x) \\ &= \mathbb{F}_{\omega \rightarrow x}^{-1} \left[(\mathbb{F} s)(\omega) e^{-j2\pi t |\omega|^c \text{sgn}(\omega)} \right] (x) \end{aligned} \quad (25)$$

($\mathbb{F}_{\omega \rightarrow x}^{-1}$ denotes the inverse Fourier transform from variable ω to x) and, interestingly, essentially commutes with the scale operator

$$\hat{\mathbf{U}}_c^{-1} \mathbf{D}'_d \hat{\mathbf{U}}_c = \mathbf{D}'_{d^{1/c}}. \quad (26)$$

Thus, the elements of a fan basis employ scaling and convolution with chirp functions

$$\mathbf{b}_{m,n}^{\text{fan}} \equiv \hat{\mathbf{U}}_c^{-1} \mathbf{b}_{m,n}^{\text{wavelet}} = \mathbf{D}'_{d_0^{m/c}} \mathbf{Y}_{nt_0}^c g_{\text{fan}}. \quad (27)$$

Here $g_{\text{fan}} \equiv \hat{\mathbf{U}}_c^{-1} g_{\text{wavelet}}$. The chirp convolution causes the basis elements to *shear* in the time direction in the time-frequency plane. The idealized plane tiling for the fan basis, shown in Fig. 4 for the case $c = 2$ (linear chirp), differs greatly from the wavelet and Gabor tilings of Fig. 3 and could be useful for representing linear chirp functions or signals dispersed along linear group delays. Other unitarily equivalent bases resulting from warping transformations have been proposed, including the FAM bases of Laine [38] and the warped polynomial bases of Philips (see [43] and the references therein).

Other choices of the unitary warping operator \mathbf{U} yield bases employing different building blocks and, thus, radically different time-frequency plane tilings. However, note from (22) and (23) that since \mathbf{U} introduces only a single degree of freedom to a unitarily equivalent basis (it works on both the time and scale or time and frequency operators simultaneously), there exists an important tradeoff in its choice: Any changes that we make to one operator are also imparted in the other. Thus, special choices of unitary transformations can prove useful, such as the warp $\hat{\mathbf{U}}_c$, which maps time to chirp convolution, yet essentially commutes with scale.

The coefficients of unitarily equivalent basis expansions may seem cumbersome to compute. However, the isometry of \mathbf{U}^{-1} allows us to calculate the unitarily equivalent basis coefficients by first preprocessing the signal and then computing the usual basis coefficients. That is, the coefficients of a

unitarily equivalent wavelet basis expansion can be obtained as

$$\langle s, \mathbf{b}_{m,n}^{\text{ue}} \rangle = \langle s, \mathbf{U}^{-1} \mathbf{b}_{m,n}^{\text{wavelet}} \rangle = \langle \mathbf{U} s, \mathbf{b}_{m,n}^{\text{wavelet}} \rangle$$

with a similar calculation for the Gabor and Wilson bases. While indicating that unitarily equivalent basis expansions can be implemented just as efficiently (modulo the unitary preprocessing) as the original basis expansions, this computation also emphasizes the equivalence of changing bases to match signals and changing signals to match bases.

It has been recently demonstrated that *regularity* (roughly, smoothness) plays an important role in basis expansions since the degree of regularity controls the extent to which errors in the basis expansion coefficients propagate into the resulting signal expansion (20) [44]. While results for arbitrary unitary transformations have thus far eluded us, we demonstrate in Appendix A that for a large class of frequency axis warping operators, the degree of regularity of a warped basis matches or exceeds that of the original basis.

V. UNITARY EQUIVALENCE AND TIME-FREQUENCY ANALYSIS

The utility of unitary equivalence for joint time-frequency analysis does not stop with orthonormal basis and frame expansions. We now turn our attention to unitary transformations of bilinear time-frequency distributions (TFD's). TFD's map 1-D signals to a 2-D time-frequency plane that indicates the joint time versus frequency content of signals; hence, unitarily equivalent TFD's will indicate the joint $\tilde{\mathbf{T}}$ versus $\tilde{\mathbf{F}}$ energy content of signals. Unitarily equivalent TFD's have numerous potential applications, including matching distributions to special signals and systems and constructing TFD's with nonuniform time and frequency resolutions that match families of group delay and instantaneous frequency characteristics.

A. U-Cohen's Classes

Most TFD's of current interest belong to Cohen's bilinear class [1]. Examples include the Wigner distribution

$$\begin{aligned} (\mathbf{W} s)(t, f) &\equiv 2 \langle \mathbf{F}_{-f} \mathbf{T}_{-t} s, \mathbf{F}_f \mathbf{T}_t s \rangle \\ &= \int s \left(t + \frac{v}{2} \right) s^* \left(t - \frac{v}{2} \right) e^{-j2\pi f v} dv \end{aligned}$$

where $s_r(u) = s(-u)$ and (t, f) represent the time-frequency plane coordinates, as well as the spectrogram, the Choi-Williams distribution, and the cone-kernel representation [1]. Cohen's class can be defined as essentially the set of all bilinear functionals $\mathbf{C}: L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}^2)$ covariant by translation to both time and frequency shifts

$$(\mathbf{C} \mathbf{F}_\nu \mathbf{T}_\tau s)(t, f) = (\mathbf{C} s)(t - \tau, f - \nu).$$

We now investigate the effect of a unitary preprocessing transformation \mathbf{U} on Cohen's class and demonstrate that the operators $\tilde{\mathbf{T}} = \mathbf{U}^{-1} \mathbf{T} \mathbf{U}$ and $\tilde{\mathbf{F}} = \mathbf{U}^{-1} \mathbf{F} \mathbf{U}$ characterize the resulting distributions. We will call the unitary equivalence class $\mathbf{C} \mathbf{U}$ of bilinear " $\tilde{\mathbf{T}}\text{-}\tilde{\mathbf{F}}$ " distributions the *U-Cohen's class*. Our approach is inspired by and generalizes the work of Altes [6] and Papandreou *et al.* [15], who obtained special cases of unitarily equivalent TFD's using warping arguments.

Unitary preprocessing transforms every property of Cohen's class into a dual property for each U-Cohen's class (we need simply make the transformation $s \mapsto \mathbf{U}s$ in each case). We now summarize some of the more important attributes of a U-Cohen's class; note that setting $\mathbf{U} = \mathbf{I}$ yields the properties of the original Cohen's class, which are detailed in [1]:

- 1) U-Cohen's class distributions are covariant by translation to the operators $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{F}}$:

$$(\mathbf{CU}\tilde{\mathbf{F}}_\nu\tilde{\mathbf{T}}_\tau s)(a, b) = (\mathbf{CU}s)(a - \tau, b - \nu). \quad (28)$$

Therefore, U-Cohen's class distributions measure not joint time \mathbf{T} and frequency \mathbf{F} content but joint $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{F}}$ content. Note that like \mathbf{T} and \mathbf{F} , $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{F}}$ remain orthogonal concepts (see (8) and Table II). We will use the coordinates (a, b) to represent the $\tilde{\mathbf{T}}\text{-}\tilde{\mathbf{F}}$ plane.

- 2) The natural functions for U-Cohen's class distributions are the eigenfunctions of $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{F}}$ (equivalently $\tilde{\mathcal{F}}$ and $\tilde{\mathcal{T}}$) because these distributions are *fixed* for only these functions. That is, for $\tilde{\mathbf{T}}$, the unique solution to the equation $(\mathbf{CU}\tilde{\mathbf{T}}_\tau s_1)(a, b) = (\mathbf{CU}s_1)(a, b) \forall \tau$ is the signal $s_1 = \mathbf{u}_\nu^{\tilde{\mathbf{T}}} = \mathbf{U}^{-1}\mathbf{u}_\nu^{\mathbf{T}}$, whereas for $\tilde{\mathbf{F}}$, the unique solution to the equation $(\mathbf{CU}\tilde{\mathbf{F}}_\nu s_2)(a, b) = (\mathbf{CU}s_2)(a, b) \forall \nu$ is the signal $s_2 = \mathbf{u}_\tau^{\tilde{\mathbf{F}}} = \mathbf{U}^{-1}\mathbf{u}_\tau^{\mathbf{F}}$. The U-Wigner distribution

$$(\mathbf{WU}s)(a, b) = 2\langle \tilde{\mathbf{F}}_{-b}\tilde{\mathbf{T}}_{-a}s, \tilde{\mathbf{F}}_b\tilde{\mathbf{T}}_a s \rangle$$

maps both of these eigenfunctions to impulse ridges in the $\tilde{\mathbf{T}}\text{-}\tilde{\mathbf{F}}$ plane since $(\mathbf{WU}\mathbf{U}^{-1}\mathbf{u}_\nu^{\mathbf{T}})(a, b) = \delta(b - \nu)$ and $(\mathbf{WU}\mathbf{U}^{-1}\mathbf{u}_\tau^{\mathbf{F}})(a, b) = \delta(a - \tau)$. Therefore, we can interpret the functions $\mathbf{u}_0^{\tilde{\mathbf{T}}}$ and $\mathbf{u}_0^{\tilde{\mathbf{F}}}$ as lying along the axes of the $\tilde{\mathbf{T}}\text{-}\tilde{\mathbf{F}}$ plane.

- 3) Each U-Cohen's class distribution \mathbf{CU} can be obtained by convolving the U-Wigner distribution with a kernel function $\phi_{\mathbf{C}}$

$$\begin{aligned} (\mathbf{CU}s)(a, b) &= \iint (\mathbf{WU}s)(x, y)\phi_{\mathbf{C}}(a - x, b - y) dx dy. \end{aligned} \quad (29)$$

- 4) The U-ambiguity function (U-AF) is based on $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{F}}$, with

$$\begin{aligned} (\mathbf{AU}s)(a, \beta) &= \langle \mathbf{F}_{\alpha/2}\mathbf{T}_{\beta/2}\mathbf{U}s, \mathbf{F}_{-\alpha/2}\mathbf{T}_{-\beta/2}\mathbf{U}s \rangle \\ &= \langle \tilde{\mathbf{F}}_{\alpha/2}\tilde{\mathbf{T}}_{\beta/2}s, \tilde{\mathbf{F}}_{-\alpha/2}\tilde{\mathbf{T}}_{-\beta/2}s \rangle \end{aligned} \quad (30)$$

where $\mathbf{A}s$ denotes the narrowband AF [1]. All U-Cohen's class distributions can be computed via the 2-D Fourier transform of a weighted U-AF

$$\begin{aligned} (\mathbf{CU}s)(a, b) &= \mathbf{F}_{\alpha \rightarrow a}\mathbf{F}_{\beta \rightarrow b}[(\mathbf{AU}s)(\alpha, \beta)\Phi_{\mathbf{C}}(\alpha, \beta)](a, b). \end{aligned}$$

Here, the kernel $\Phi_{\mathbf{C}}$ corresponds to the 2-D inverse Fourier transform of the function $\phi_{\mathbf{C}}$ from (29). Note that the Cohen's class TFD \mathbf{C} and the U-Cohen's class distribution \mathbf{CU} share exactly the same kernel function $\Phi_{\mathbf{C}}$. The U-AF is the inverse Fourier transform of the U-Wigner distribution, which implies that

$\Phi_{\mathbf{W}}(\alpha, \beta) = 1$. The U-AF maps the eigenfunctions of $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{F}}$ to impulse ridges in the warped ambiguity plane since $(\mathbf{AU}\mathbf{U}^{-1}\mathbf{u}_\nu^{\mathbf{T}})(\alpha, \beta) = e^{-j2\pi\nu\beta}\delta(\alpha)$ and $(\mathbf{AU}\mathbf{U}^{-1}\mathbf{u}_\tau^{\mathbf{F}})(\alpha, \beta) = e^{j2\pi\tau\alpha}\delta(\beta)$. Gray and Helmick derive an alternative warped AF in [45].

- 5) U-distributions \mathbf{CU} whose kernels satisfy the constraint $\Phi_{\mathbf{C}}(\alpha, 0) = \Phi_{\mathbf{C}}(0, \beta) = 1 \forall \alpha, \beta$ possess marginal distributions that measure the $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{F}}$ content of the time signal s (see (14) and (15))

$$\int (\mathbf{CU}s)(a, b) db = |\mathbf{F}_{\tilde{\mathcal{T}}}s(a)|^2 = |(Us)(a)|^2 \quad (31)$$

$$\int (\mathbf{CU}s)(a, b) da = |\mathbf{F}_{\tilde{\mathcal{F}}}s(b)|^2 = |(Us)(b)|^2. \quad (32)$$

- 6) A U-Cohen's class distribution \mathbf{CU} is unitary—and therefore satisfies Moyal's formula $\langle \mathbf{CU}s_1, \mathbf{CU}s_2 \rangle = \langle s_1, s_2 \rangle^2$ —if its kernel function is allpass: $|\Phi_{\mathbf{C}}(\alpha, \beta)| = 1 \forall \alpha, \beta$ [1], [46].
- 7) The interference structure of a U-Cohen's class is based on a 2-D arithmetic mean. Considering the U-Wigner distribution, the level of the interference term at the point (a, b) in the $\tilde{\mathbf{T}}\text{-}\tilde{\mathbf{F}}$ plane is determined by the contributions from all other pairs of points in the plane that have arithmetic mean (a, b) [47], [48]

$$\begin{aligned} |(\mathbf{WU}s)(a, b)|^2 &= \iint (\mathbf{WU}s)\left(a + \frac{x}{2}, b + \frac{y}{2}\right) \\ &\quad \cdot (\mathbf{WU}s)\left(a - \frac{x}{2}, b - \frac{y}{2}\right) dx dy. \end{aligned}$$

In Cohen's class, the choice of TFD—and thus the kernel function $\Phi_{\mathbf{C}}$ —is in many cases critical for accurately representing the time-varying frequency content of a signal. Therefore, optimal-kernel TFD's that automatically select the "best" kernel function for a given signal have been developed [49], [50]. Since the Cohen's class TFD \mathbf{C} and the U-Cohen's class distribution \mathbf{CU} share exactly the same kernel function $\Phi_{\mathbf{C}}$, kernel design, and optimal kernel design in particular, proceeds in exactly the same fashion in a U-Cohen's class as in Cohen's class. The sole difference is that kernels favoring signals resembling the time and frequency eigenfunctions in Cohen's class will now favor signals resembling the transformed eigenfunctions. Kernel constraints yielding desirable U-Cohen's class properties can be determined simply by warping the properties back to Cohen's class with the operator \mathbf{U}^{-1} .

Each choice of unitary transformation \mathbf{U} results in a U-Cohen's class with radically different properties.

Example—Scale and Logarithmic Modulation: As we saw in Section III, the time-domain logarithmic warping operator (18) maps time to scale ($\mathbf{T}_t \mapsto \mathbf{D}_t$) and frequency to logarithmic modulation ($\mathbf{F}_f \mapsto \mathbf{H}_f$), both defined on $L^2(\mathbb{R}_+)$. The \mathbf{U}_{\log} -Wigner distribution corresponds to the time-domain Q distribution [6] derived by Altes using an identical warping argument. The covariance properties and marginals of \mathbf{U}_{\log} -Cohen's class distributions (covariance to \mathbf{D} and \mathbf{H} , scale and Mellin transforms as marginals when satisfied) make them ideally suited to studying signals resembling the single-sided

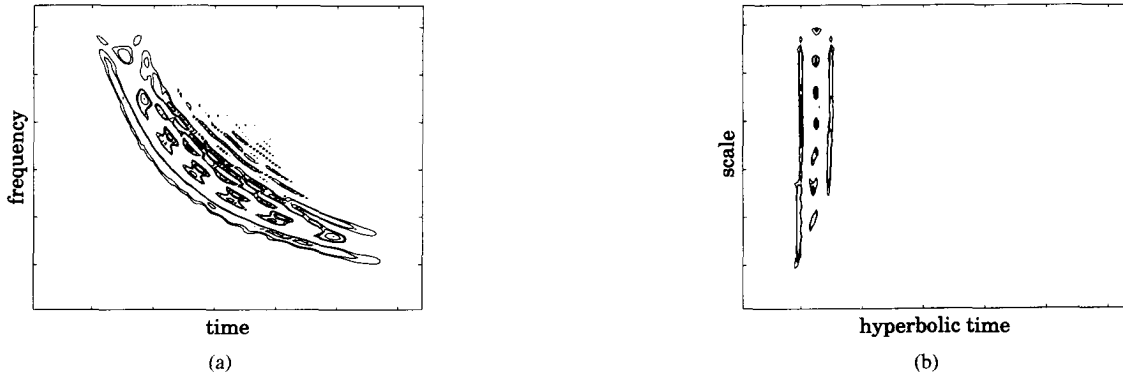


Fig. 5. Joint distributions of a windowed sum of two hyperbolic chirp functions $\mathbf{u}_{h_1}^D + \mathbf{u}_{h_2}^D$ from (33) [51]: (a) Wigner distribution, with axes of time and frequency; (b) $\mathbb{F}^{-1}\mathbf{U}_{\log}\mathbb{F}$ -Wigner distribution (Altes–Marinovich distribution [6], [7]), with axes of hyperbolic time and scale.

eigenfunctions of \mathbf{D} and \mathbf{H}

$$\mathbf{u}_h^D(x) = \frac{1}{\sqrt{x}} e^{j2\pi h \log x}, \quad \mathbf{u}_d^H(x) = e^{d/2} \delta(x - e^d), \quad x > 0.$$

Example—Hyperbolic Time and Scale: The frequency-domain logarithmic warping operator $\mathbb{F}^{-1}\mathbf{U}_{\log}\mathbb{F}$ maps frequency to scale ($\mathbf{F}_f \mapsto \mathbf{D}_f$) and time to “hyperbolic time” ($\mathbf{T}_t \mapsto \mathbf{L}_t$), where the hyperbolic time-shift operator $\mathbf{L}_t \equiv \mathbb{F}^{-1}\mathbf{H}_t\mathbb{F}$ [13], [15]. Both \mathbf{D} and \mathbf{L} operate on the Hardy space of analytic signals, with \mathbf{L} implementing a dispersive time shift that translates high-frequency signal components less than low-frequency signal components. We will call the $\mathbb{F}^{-1}\mathbf{U}_{\log}\mathbb{F}$ -Cohen’s class the *prehyperbolic class* because it foreshadows the hyperbolic class we will see in Section V-B. The prehyperbolic class contains the Altes–Marinovich distribution [6], [7] as the $\mathbb{F}^{-1}\mathbf{U}_{\log}\mathbb{F}$ -Wigner distribution.

The covariance and marginal properties of $\mathbb{F}^{-1}\mathbf{U}_{\log}\mathbb{F}$ -Cohen’s class distributions (covariance to \mathbf{D} and \mathbf{L} , exponentiated Fourier and Mellin transforms as marginals when satisfied) are matched to signals resembling the analytic eigenfunctions [9]

$$\mathbf{u}_h^D(x) = \mathbb{F}_{\omega \mapsto x}^{-1} \frac{1}{\sqrt{\omega}} e^{j2\pi h \log \omega}, \quad \mathbf{u}_d^L(x) = e^{d/2} e^{j2\pi e^d x}. \quad (33)$$

The terms hyperbolic time shift and prehyperbolic class stem from the concentration of the \mathbf{u}_h^D along hyperbolas in the time–frequency plane. The Doppler tolerance of these so-called hyperbolic chirp functions is of fundamental importance in applications such as wideband sonar and dispersive signal processing; in fact, several species of echolocating bats employ similar waveforms [4]. Fig. 5(a) and (b) illustrate the Wigner distribution and Altes–Marinovich distribution of the sum of two hyperbolic chirps [51]. The dramatic increase in clarity from (a) to (b) is indicative of the improvements obtainable by matching a joint distribution coordinate system to the data being analyzed. Ristic and Boashash perform a similar analysis in [52].

Example—Power Time Shift and Chirp Modulation: The time-domain power warping operator \mathbf{U}_c from (24) in Section IV maps the time operator to a power time shift [24]

$$(\mathbf{M}_t s)(x) \equiv s\{m_c[m_{1/c}(x) - t]\} |x|^{(1-c)/2c} |m_{1/c}(x) - t|^{(t-1)/2}$$

where $m_c(u) \equiv |u|^c \operatorname{sgn}(u)$ and maps the frequency operator to a chirp modulation

$$(\mathbf{N}_f s)(x) \equiv e^{j2\pi f |x|^c \operatorname{sgn}(x)} s(x).$$

Marginals for \mathbf{U}_c -Cohen’s class distributions thus correspond to the power time transform $|\mathbf{U}_c s|^2$ and the chirp transform $|\mathbb{F}\mathbf{U}_c s|^2$ obtained by projecting the signal onto chirp functions of order c . All \mathbf{U}_c -Cohen’s class distributions enjoy covariance to \mathbf{M} and \mathbf{N} , making them ideal for studying (power) chirping signals. The frequency-domain power warping operator $\mathbb{F}^{-1}\mathbf{U}_c\mathbb{F}$ works similarly and constructs distributions of chirp time shift $\mathbb{F}^{-1}\mathbf{N}_t\mathbb{F}$ and power frequency shift $\mathbb{F}^{-1}\mathbf{M}_f\mathbb{F}$.

B. VU-Cohen’s Classes

Up to this point, we have concentrated exclusively on preprocessing transformations; however, postprocessing transformations also merit consideration. We define the *VU-Cohen’s class* as the set of bilinear functionals \mathbf{VCU} , where $\mathbf{C}s$ is a Cohen’s class TFD, \mathbf{U} is a unitary preprocessing signal transformation, and $\mathbf{V}: L^2(\mathbb{R}^2) \rightarrow L^2(\mathbb{R}^2)$ is a postprocessing transformation defined on the (a, b) plane.

The extra degree of freedom gained through postprocessing adds considerable richness to the already powerful \mathbf{U} -Cohen’s classes. \mathbf{VU} -Cohen’s classes can have covariance and marginals much more general than (28) and (31), (32). For example, setting \mathbf{U} to the identity operator and choosing \mathbf{V} such that $(\mathbf{VC}s)(t, r) = r^{-2}(\mathbf{C}s)(t, r^{-1})$ yields one of the time and inverse frequency classes of Cohen [1, p. 238], which contains some of the important distributions from the affine class [11]–[13]. This new class and its relationship with the affine class is studied in more detail in [31].

While \mathbf{U} - and \mathbf{VU} -Cohen’s classes show considerable promise for matching signals different from simple pulses, sinusoids, and linear chirps, certain applications demand representations with explicit time and frequency axes. Therefore, rather than exploring postprocessing in complete generality (see [30] for more), we will focus now on using postprocessing to invert the action of \mathbf{U}^{-1} on the $\tilde{\mathbf{T}}\text{--}\tilde{\mathbf{F}}$ plane, warping the (a, b) axes of \mathbf{U} -Cohen’s class distributions to new axes (t, f) providing correct time–frequency localization. While these \mathbf{VU} -Cohen’s class distributions measure time–frequency

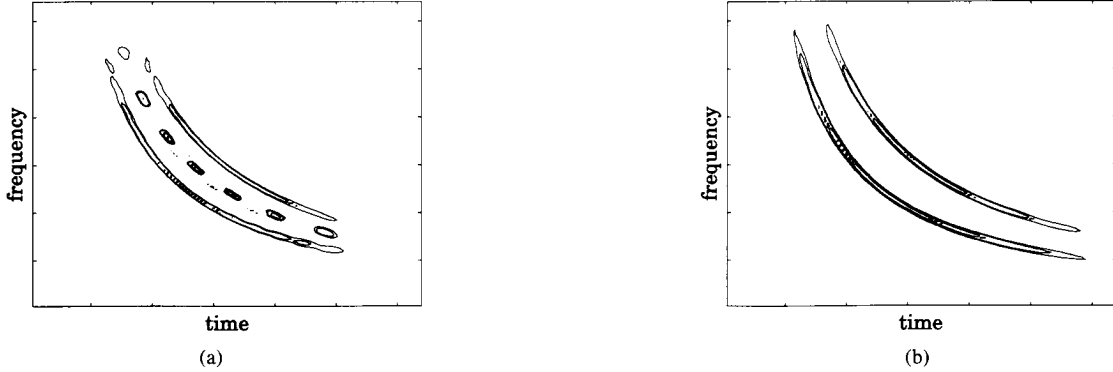


Fig. 6. (a) $\mathbf{V}\mathbf{U}$ -Wigner distribution (Altes Q TFD [6] from the hyperbolic class [15]) obtained by applying the axis warping transformation \mathbf{V} from (36) and (42) to the \mathbf{U} -Wigner distribution of Fig. 5(b). The axes of this distribution correspond to (nonuniform resolution) time and frequency. Comparison with the Wigner distribution of Fig. 5(a) reveals the hyperbolic geometry underlying the localization of this $\mathbf{V}\mathbf{U}$ -Cohen's class; (b) optimal-kernel hyperbolic TFD [51].

content in signals, their underlying fundamental operators remain $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{F}}$.

Our procedure for inverting the effect of \mathbf{U} is best interpreted graphically on the warped time-frequency plane. A \mathbf{U} -Cohen's class distribution $\mathbf{C}\mathbf{U}s$ maps the time eigenfunction $\mathbf{u}_t^T(x) = \delta(x - t)$ localized at time t and the frequency eigenfunction $\mathbf{u}_f^F(x) = e^{j2\pi fx}$ localized at frequency f to distributions localized on curves in the (a, b) plane. The equations of these curves can be obtained from the \mathbf{U} -Cohen's class analogues of the time-frequency concepts of group delay and instantaneous frequency. Modifying the usual definitions of these quantities [1] to account for the preprocessing by \mathbf{U} , we obtain a curve for the localization of \mathbf{u}_t^T (its " \mathbf{U} -group delay")

$$\rho(b, t) \equiv -\frac{1}{2\pi} \frac{\partial}{\partial b} \arg(\mathbf{F}\mathbf{U}\mathbf{u}_t^T)(b) \quad (34)$$

and a curve for the localization of \mathbf{u}_f^F (its " \mathbf{U} -instantaneous frequency")

$$\theta(a, f) \equiv \frac{1}{2\pi} \frac{\partial}{\partial a} \arg(\mathbf{U}\mathbf{u}_f^F)(a). \quad (35)$$

If \mathbf{V} reparameterizes the axes of the (a, b) plane in terms of these curves, then \mathbf{u}_t^T and \mathbf{u}_f^F will be localized along straight lines at time t and frequency f , and correct time-frequency localization will result, albeit with nonuniform resolution. The requisite reparameterization is given by

$$(\mathbf{V}\mathbf{C}\mathbf{U}s)(t, f) = (\mathbf{C}\mathbf{U}s)[A(t, f), B(t, f)] \quad (36)$$

where the functions $A(t, f)$ and $B(t, f)$ denote the solution (when one exists) to the system of equations

$$a = \rho(b, t), \quad b = \theta(a, f) \quad (37)$$

for a and b in terms of t and f . The reverse reparameterization is also useful and is given by

$$(\mathbf{V}\mathbf{C}\mathbf{U}s)[T(a, b), F(a, b)] = (\mathbf{C}\mathbf{U}s)(a, b) \quad (38)$$

where the functions $T(a, b)$ and $F(a, b)$ denote the solution to the system (37) for t and f in terms of a and b . Since these functions are nonlinear in general, they have the effect of warping the geometry of the $\tilde{\mathbf{T}}\text{-}\tilde{\mathbf{F}}$ plane so that impulses

and sinusoids lie concentrated along lines parallel to the (t, f) axes.

For preprocessing transformations \mathbf{U} of the warp type (7), analytic solution of (37) yields

$$A(t, f) = w^{-1}(t), \quad B(t, f) = f\dot{w}[w^{-1}(t)] \quad (39)$$

$$T(a, b) = w(a), \quad F(a, b) = \frac{b}{\dot{w}(a)} \quad (40)$$

while for frequency axis warping transformations of the form $\mathbf{F}^{-1}\mathbf{U}\mathbf{F}$, with \mathbf{U} of the warp type, we have

$$A(t, f) = t\dot{w}[w^{-1}(f)], \quad B(t, f) = w^{-1}(f) \\ T(a, b) = \frac{a}{\dot{w}(b)}, \quad F(a, b) = w(b).$$

The properties of a $\mathbf{V}\mathbf{U}$ -Cohen's class are summarized in Appendix B.

Example—Hyperbolic Time and Scale: Reconsidering the $\mathbf{F}^{-1}\mathbf{U}_{\log}\mathbf{F}$ -Cohen's class example from Section V-A, the formulas (39) and (40) give

$$A(t, f) = tf, \quad B(t, f) = \log f \quad (41)$$

$$T(a, b) = ae^{-b}, \quad F(a, b) = e^b \quad (42)$$

as the required axis reparameterization. Substitution of these quantities in (36) yields the *hyperbolic class* of TFD's [15], which includes the frequency-domain Q TFD of Altes [6]. The hyperbolic geometry of the warping \mathbf{V} defined by (42) is evident in the Q TFD of the sum of two hyperbolic chirp functions shown in Fig. 6(a). Fig. 6(b) illustrates an optimal kernel hyperbolic TFD [51] computed by applying the Cohen's class optimal-kernel design procedure of [50] to the $\mathbf{F}^{-1}\mathbf{U}_{\log}\mathbf{F}$ -AF of the signal.

Comparison of Figs. 5(a) and 6(a) reveals the salient feature distinguishing Cohen's class from a $\mathbf{V}\mathbf{U}$ -Cohen's class, both of which measure the time-frequency content of signals: The geometry of Cohen's class localizes perfectly the impulses \mathbf{u}_t^T and sinusoids \mathbf{u}_f^F , whereas the geometry of a $\mathbf{V}\mathbf{U}$ -Cohen's class perfectly localizes the functions $\mathbf{U}^{-1}\mathbf{u}_t^T$ and $\mathbf{U}^{-1}\mathbf{u}_f^F$. Thus, for example, the hyperbolic class contains the TFD's of choice for studying in time-frequency signals resembling the scale and hyperbolic time eigenfunctions from (33).

C. Warp Synthesis

The property of the hyperbolic class of localizing signals on hyperbolic paths in the time-frequency plane can be extended to arbitrary paths, as long as they are smooth and monotonic. We now consider the problem of *warp synthesis*, which is stated as follows: Find the unitary operator \mathbf{U} generating a \mathbf{VU} -Cohen's class whose (t, f) plane perfectly localizes two types of signals—those whose group delay lies along the curve $\Gamma_{gd}(f)$ and those whose instantaneous frequency lies along the curve $\Gamma_{if}(t)$.

When a solution exists to the warp synthesis problem, it can be found by reversing the procedure of Section V-B. The two paths Γ_{if} and Γ_{gd} specify $\rho(b, t)$ and $\theta(a, f)$; integration yields functions from which the form of \mathbf{U} can be inferred. Clearly, a key consideration is the invertibility of the system of (34) and (35). The solution is straightforward when \mathbf{U} is constrained to be a time or frequency axis warping, however. In the time domain case, with \mathbf{U} of the form (7), only the instantaneous frequency localization can be adjusted, the group delay localization being fixed along lines parallel to the frequency axis. Given a desired, monotonic instantaneous frequency function $\Gamma_{if}(t)$, the corresponding warping function w for \mathbf{U} is given by the inverse of the indefinite integral of Γ_{if}

$$w = z^{-1}, \quad z(u) = \int \Gamma_{if}(u) du.$$

For frequency warping operators of the form $\mathbf{F}^{-1}\mathbf{U}\mathbf{F}$, an identical calculation yields TFD's that localize along group delays of the form $\Gamma_{gd}(f)$ and along instantaneous frequencies parallel to the time axis.

Example—Hyperbolic Instantaneous Frequency: A desire to localize signals having hyperbolic instantaneous frequencies would prompt the choice $\Gamma_{if}(t) = C/t$. Integration and inversion of Γ_{if} yield the warping function $w(x) = e^{x/C}$ for use in the preprocessing operator \mathbf{U} . The resulting \mathbf{U}_{\log} -Cohen's class contains the scale versus logarithmic modulation distributions from Section V-A.

Example—Hyperbolic Group Delay: Alternatively, a desire to localize signals having hyperbolic group delay would prompt the selection of $\Gamma_{gd}(f) = C/f$. Integration and inversion of Γ_{gd} yield the warping function $w(x) = e^{x/C}$ for use in the preprocessing operator $\mathbf{F}^{-1}\mathbf{U}\mathbf{F}$. The prehyperbolic class of hyperbolic time and scale distributions results from this preprocessing.

D. U-Affine and VU-Affine Classes

Unitary equivalence remains equally valid for other joint signal representations, including the *affine class* of TFD's \mathbf{Q} covariant to time shifts and scale changes [11]–[13]

$$(\mathbf{Q}\mathbf{T}_\tau\mathbf{D}_d s)(t, f) = (\mathbf{Q}s)[e^{-d}(t - \tau), e^d f]. \quad (43)$$

Members of the affine class include the Wigner, Choi–Williams, and Bertrand distributions, as well as the scalogram, the squared magnitude of the continuous wavelet transform. Clearly, as the operators \mathbf{T} and \mathbf{D} form the basis for the affine class, the unitarily equivalent operators $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{D}}$ will form the basis for the preprocessed *U-affine*

class of distributions \mathbf{QU} . The properties of a \mathbf{U} -affine class parallel those of the affine class and can be obtained as readily as we obtained the \mathbf{U} -Cohen's class properties in Section V-A. Furthermore, a warping procedure identical to that of Section V-B will restore correct time-frequency localization to a \mathbf{U} -affine class.

Each example unitary operator introduced above also spawns a \mathbf{U} -affine class. Of particular recent interest is the frequency domain power warping function $\tilde{\mathbf{U}}_c = \mathbf{F}^{-1}\mathbf{U}_c\mathbf{F}$ (see (24)–(26)). Since $\tilde{\mathbf{U}}_c$ essentially commutes with the scale operator, the resulting $\tilde{\mathbf{U}}_c$ -affine class contains distributions measuring joint scale and “chirp time” content [24]. This class has been studied independently in [16] and contains, for the case $c = 2$, the scale-shear transform proposed in [21] and [22].

The intersection of Cohen's class and the affine class contains distributions covariant not only to time and frequency shifts but also to scale changes [11], [53]. Unitary preprocessing preserves this structure: Distributions lying in the intersection of the \mathbf{U} -Cohen's class and \mathbf{U} -affine class are covariant to the transformed operators $\tilde{\mathbf{T}}$, $\tilde{\mathbf{F}}$, and $\tilde{\mathbf{D}}$. Similar results hold for the affine class-hyperbolic class intersection studied in [54].

Finally, the concept of the angle between two linear operators introduced in Section II reveals a fundamental difference between Cohen's class and the affine class: Cohen's class is based on orthogonal concepts (see (8)), whereas the affine class is based on nonorthogonal concepts (see (9)). The nonorthogonality of the affine class manifests itself, for example, in the “coupling” of the scale change d into both arguments in the covariance formula (43). Identical comments apply to \mathbf{U} -Cohen's and \mathbf{U} -affine classes.

VI. CONCLUSIONS

Unitary equivalence provides a simple means of developing an infinite number of new signal analysis and processing tools tailored to different classes of signals or systems. The benefits of this general approach are twofold. First, it provides a general theoretical framework for deriving new tools with desired properties, instead of the piecemeal approach adopted in the past. Second, because the new tools can all be implemented by applying standard algorithms to a preprocessed signal, efficient, robust, and well-understood implementations are immediately available. Perhaps surprisingly, signal processing tools with characteristics very different from known methods often result from very simple transformations.

We started this paper with the observation that only signal analysis and processing tools that are matched to the signal can provide maximum performance. Our development of the unitary equivalence principle elicits a two-step approach to matching systems to signals: First, generalize current tools by introducing extra degrees of freedom; then, within each class of generalized tools, select the one tool best suited to the given data. We have emphasized only the first step of this procedure in this paper; determining the unitary transform that achieves the desired goal remains the most challenging part of the problem. Techniques for automatically optimizing

the preprocessing are currently under investigation and could lead to substantial benefits in many applications. Nevertheless, the unitary equivalence principle has already proven its worth as a generalizing and unifying tool, with its application to the fan and chevron bases [24], [36], [37], the aforementioned hyperbolic and power classes [6], [15], [16], new signal transforms [28], [35], [38], and the relationships between joint distributions of arbitrary variables [29], [42].

The primary limitation of unitary equivalence is that the effect of a unitary transformation is distributed evenly on all variables in a system (since it introduces only a single degree of freedom). Thus, while a unitary transformation may result in desirable properties in a new system (scale invariance, for example), it may also result in a loss of some of the desirable properties of the original system (time-shift invariance, for example). We have also dealt exclusively with linear operators in this paper; connections with nonlinear transformation techniques such as homomorphic signal processing [55] could yield new insight into unitary equivalence.

APPENDIX A

REGULARITY OF UNITARILY EQUIVALENT BASES AND FRAMES

The regularity of a unitarily equivalent basis depends on the regularity of the generalized wavelet/window function $U^{-1}g$ from (22) or (23). Complete results have not yet been determined, since in general, it is necessary to know both the operator U and the function g before a calculation can be made. However, for a special class of frequency axis warping operators, results follow easily. The following theorem has important implications, for it shows that warped bases can be more regular than the wavelet, Gabor, or Wilson bases from which they are derived. One example is the fan basis set (27) for $c > 1$.

Theorem: Let g have α continuous derivatives, and let U^{-1} be a unitary transformation of the warping type (7), whose warping function w is continuously differentiable and grows as $|w(v)| > C|v|^k$, $k > 0$, for large $|v|$. Then, the function $F^{-1}U^{-1}Fg$ has at least $k(\alpha + \frac{1}{2}) - \frac{1}{2}$ continuous derivatives.

Proof: The assertion that the function g has α continuous derivatives is equivalent to stating that its Sobolev regularity is α or that the function $u^{\alpha+(1/2)}\hat{g}(u) \in L^2(\mathbb{R})$, where \hat{g} denotes the Fourier transform of g [44]. Therefore, we need only show that the function $u^{k[\alpha+(1/2)]}(U^{-1}\hat{g})(u) \in L^2(\mathbb{R})$. Let $\beta = \alpha + \frac{1}{2}$. Using the change of variable $v = w(u)$ and the bound on w from the statement, we have

$$\begin{aligned} \int |u^{k\beta}(U\hat{g})(u)|^2 du &= \int |u|^{2k\beta} |\hat{g}[w(u)]| |\dot{w}(u)|^{1/2} |^2 du \\ &= \int |w^{-1}(v)|^{2k\beta} |\hat{g}(v)|^2 dv \\ &< \frac{1}{C} \int |v^\beta \hat{g}(v)|^2 dv. \end{aligned}$$

Since \hat{g} has regularity order $\alpha = \beta - \frac{1}{2}$, this last term is finite, and the result follows. \square

A simple explanation of the theorem is that for $k > 1$ the frequency axis warping compresses the function \hat{g} in the

frequency domain, reducing the high frequency content of the resulting time domain function g , and thus making it smoother.

APPENDIX B

VU-COHEN'S CLASS PROPERTIES

While leaving many of the desirable properties of a U-Cohen's class intact, the postprocessing transformation V from Section V-B has the effect of changing the geometry of the \tilde{T} - \tilde{F} analysis plane. The salient properties of a VU-Cohen's class can be summarized as follows:¹²

- 1) VU-Cohen's class distributions remain covariant to \tilde{T} and \tilde{F} . However, the covariance relationship is no longer a simple 2-D translation; instead, we have

$$\begin{aligned} (VCU_{\tilde{T}, \tilde{F}})(t, f) &= \\ (VCU_s)\{T[A(t, f) - \tau, B(t, f) - \nu], \\ F[A(t, f) - \tau, B(t, f) - \nu]\}. \end{aligned}$$

- 2) VU-Cohen's class TFD's are fixed for the eigenfunctions $U^{-1}u_\nu^T$ and $U^{-1}u_\tau^F$ of \tilde{T} and \tilde{F} . Thus, these functions remain the natural functions for a VU-Cohen's class. Moreover, the VU-Wigner distribution VWU maps both of these functions to delta ridges lying along the curves A and B in the (t, f) plane since $(VWUU^{-1}u_\nu^T)(t, f) = \delta[B(t, f) - \nu]$ and $(VWUU^{-1}u_\tau^F)(t, f) = \delta[A(t, f) - \tau]$.
- 3) Each VU-Cohen's class TFD can be written in terms of the VU-Wigner distribution as

$$\begin{aligned} (VCU_s)(t, f) &= \iint (VWU_s)(x, y) \\ &\cdot \phi_C[A(t, f) - A(x, y), B(t, f) - B(x, y)] dx dy. \end{aligned}$$

Note that the relationship is no longer a simple 2-D convolution.

- 4) All VU-Cohen's class TFD's can be computed by warping the 2-D Fourier transform of a weighted U-AF (30) of the signal

$$\begin{aligned} (VCU_s)(t, f) &= F_{\alpha \rightarrow x} F_{\beta \rightarrow y} [(AU_s)(\alpha, \beta) \Phi_C(\alpha, \beta)](x, y), \\ x &= A(t, f), y = B(t, f). \end{aligned}$$

This suggests a simple four-step procedure for computing any VU-Cohen's class TFD:

- 1) preprocess the signal by U
- 2) compute the usual AF of the processed signal
- 3) weight the AF by the kernel function Φ_C and compute the Fourier transform of the product
- 4) warp the axes using V .

¹²Note that while we write all formulas below in complete generality, considerable simplifications occur for specific unitary prewarping operators U , especially those from the axis warping subclass.

Note that the Cohen's class, U-Cohen's class, and VU-Cohen's class distributions C , CU , and VCU all share the same kernel function Φ_C .

- 5) VU-Cohen's class distributions whose kernels satisfy the constraint $\Phi_C(\alpha, 0) = \Phi_C(0, \beta) = 1 \forall \alpha, \beta$ possess the marginal properties (see (31), (32), and (38))

$$\int (\mathbf{VCU}_s)[T(a, b), F(a, b)] db = |(\mathbf{U}_s)(a)|^2$$

$$\int (\mathbf{VCU}_s)[T(a, b), F(a, b)] da = |(\mathbf{FU}_s)(b)|^2.$$

These path integrals implement in the (t, f) plane the tomographic equivalent of projecting the signal onto the eigenfunctions \mathbf{u}_a^T and \mathbf{u}_b^F , respectively.

- 6) A VU-Cohen's class distribution is unitary if its kernel function is allpass.
- 7) The interference structure of the VU-Wigner distribution is based on a generalized mean:

$$|(\mathbf{VWU}_s)(t, f)|^2 =$$

$$\iint (\mathbf{VWU}_s) \left\{ T \left[A(t, f) - \frac{x}{2}, B(t, f) - \frac{y}{2} \right], \right.$$

$$F \left[A(t, f) - \frac{x}{2}, B(t, f) - \frac{y}{2} \right] \left. \right\}$$

$$\cdot (\mathbf{VWU}_s) \left\{ T \left[A(t, f) + \frac{x}{2}, B(t, f) + \frac{y}{2} \right], \right.$$

$$F \left[A(t, f) + \frac{x}{2}, B(t, f) + \frac{y}{2} \right] \left. \right\} dx dy$$

and, therefore, interference terms will form along the curves A and B in the (t, f) plane. (See (42) and Fig. 6(a) for an example with the hyperbolic class.) Gonçalves and Flandrin have studied similar generalized means in the context of the affine class [56].

Because \mathbf{V} warps the geometry of the (a, b) plane, there exist important differences between a U-Cohen's class and a VU-Cohen's class. First, in general, the distribution \mathbf{VCU}_s does not correspond to a Cohen's class TFD of some pre-processed signal. Second, although VU-Cohen's class TFD's provide correct time-frequency localization of signals, they cannot in general be simultaneously covariant to both time and frequency shifts. Note, however, that there do exist VU-Cohen's class TFD's covariant to one or the other of these shifts. For example, the unitary Bertrand distribution [8]–[10] from the affine and hyperbolic classes is covariant to not only hyperbolic-time shifts and scale changes but also to regular time shifts. In general, VU-TFD's covariant to \mathbf{T} or \mathbf{F} are characterized by kernels corresponding to Cohen's class TFD's covariant to the inversely transformed operators $\mathbf{U}\mathbf{T}\mathbf{U}^{-1}$ or $\mathbf{U}\mathbf{F}\mathbf{U}^{-1}$ (note the reversed order of \mathbf{U} and \mathbf{U}^{-1}).

ACKNOWLEDGMENT

The authors wish to thank G. F. Boudreaux-Bartels, P. Flandrin, P. Gonçalves, F. Hlawatsch, and M. Stewart for stimulating discussions, the reviewers for inspiring several changes to the manuscript, and K. Canfield for providing Figs. 5 and 6.

REFERENCES

- [1] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [2] C. Heil and D. Walnut, "Continuous and discrete wavelet transforms," *SIAM Rev.*, vol. 31, pp. 626–666, Dec. 1989.
- [3] I. Daubechies, S. Jaffard, and J. L. Journé, "A simple Wilson basis with exponential decay," *SIAM J. Math. Anal.*, vol. 22, pp. 554–572, Mar. 1991.
- [4] R. A. Altes, "The Fourier–Mellin transform and mammalian hearing," *J. Acoust. Soc. Amer.*, vol. 63, pp. 174–183, 1978.
- [5] J. P. Ovarlez, "La Transformation de Mellin: Un outil pour l'analyse des signaux à large bande," Ph.D. dissertation, Université de Paris VI, 1992.
- [6] R. A. Altes, "Wideband, proportional-bandwidth Wigner–Ville analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1005–1012, June 1990.
- [7] G. Eichmann and N. M. Marinovich, "Scale-invariant Wigner distribution," in *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 519, 1984, pp. 18–25.
- [8] J. Bertrand and P. Bertrand, "Représentation temps-fréquence des signaux," *C. R. Acad. Sci. I*, vol. 299, pp. 635–638, 1984.
- [9] ———, "A class of affine Wigner functions with extended covariance properties," *J. Math. Phys.*, vol. 33, pp. 2515–2527, July 1992.
- [10] R. G. Shenoy and T. W. Parks, "Wide-band ambiguity functions and affine Wigner distributions," *Signal Processing*, vol. 41, no. 3, pp. 339–363, 1995.
- [11] O. Rioul and P. Flandrin, "Time-scale energy distributions: A general class extending wavelet transforms," *IEEE Trans. Signal Processing*, vol. 40, pp. 1746–1757, July 1992.
- [12] J. Bertrand and P. Bertrand, "Time-frequency representations of wideband signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing—ICASSP'88*, 1988, pp. 2196–2199.
- [13] ———, "Affine time-frequency distributions," in *Time-Frequency Signal Analysis-Methods and Applications*, B. Boashash, Ed. Melbourne: Longman Cheshire, 1992, pp. 118–140.
- [14] L. Cohen, "The scale representation," *IEEE Trans. Signal Processing*, vol. 41, pp. 3275–3292, Dec. 1993.
- [15] A. Papandreou, F. Hlawatsch, and G. F. Boudreaux-Bartels, "The hyperbolic class of quadratic time-frequency representations. Part I: Constant-Q warping, the hyperbolic paradigm, properties, and members," *IEEE Trans. Signal Processing*, vol. 41, pp. 3425–3444, Dec. 1993.
- [16] F. Hlawatsch, A. Papandreou, and G. F. Boudreaux-Bartels, "The power classes of quadratic time-frequency representations: A generalization of the affine and hyperbolic classes," in *Proc. 27th Asilomar Conf.*, Pacific Grove, CA, 1993, pp. 1265–1270.
- [17] A. Berthon, "Operator groups and ambiguity functions in signal processing," in *Wavelets: Time-Frequency Methods and Phase Space*, J. M. Combes, A. Grossman, and P. Tchamitchian, Eds. New York: Springer-Verlag, 1989, pp. 172–180.
- [18] A. Grossman and T. Paul, "Wave functions on subgroups of the group of affine canonical transformations," in *Lecture Notes in Physics no. 211: Resonances-Models and Phenomena*, L. Strein, Ed. Springer-Verlag, 1984, pp. 128–138.
- [19] S. Mann and S. Haykin, "The chirplet transform—A generalization of Gabor's logon transform," in *Vision Interface '91*, Calgary, Canada, June 1991.
- [20] ———, "Time-frequency perspectives: The chirplet transform," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing—ICASSP'92*, 1992, vol. 3, pp. 417–420.
- [21] R. G. Baraniuk and D. L. Jones, "New dimensions in wavelet analysis," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing—ICASSP '92*, vol. V, 1992, pp. 137–140.
- [22] R. G. Baraniuk, "Shear madness: Signal-dependent and metaplectic time-frequency representations," Coord. Sci. Lab., Univ. Illinois, Urbana-Champaign, Tech. Rep. no. UILU-ENG-92-2226, 1992.
- [23] B. Boashash and P. O'Shea, "Polynomial Wigner–Ville distributions and their relationship to time-varying higher order spectra," *IEEE Trans. Signal Processing*, vol. 42, pp. 216–222, Jan. 1994.
- [24] R. G. Baraniuk and D. L. Jones, "Warped wavelet bases: Unitary equivalence and signal processing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing—ICASSP'93*, vol. III, 1993, pp. 320–323.
- [25] G. F. Boudreaux-Bartels and T. W. Parks, "Time-varying filtering and signal estimation using Wigner distribution synthesis techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 442–451, June 1986.
- [26] W. Krattenthaler and F. Hlawatsch, "Time-frequency design and processing of signals via smoothed Wigner distributions," *IEEE Trans. Signal Processing*, vol. 41, pp. 278–287, Jan. 1993.

- [27] M. J. Arnold, M. Roessgen, and B. Boashash, "Filtering real signals through frequency modulation and peak detection in the time-frequency plane," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing-ICASSP'94*, vol. III, 1994, pp. 345-348.
- [28] R. G. Baraniuk, "A signal transform covariant to scale changes," *Electron. Lett.*, vol. 29, pp. 1675-1676, Sept. 17, 1993.
- [29] ———, "Beyond time-frequency analysis: Energy densities in one and many dimensions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing-ICASSP'94*, vol. III, 1994, pp. 357-360.
- [30] ———, "Warped perspectives in time-frequency analysis," in *IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, Oct. 1994, pp. 528-531.
- [31] ———, "Marginals versus covariance in joint distribution theory," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing-ICASSP'95*, vol. 2, 1995, pp. 1021-1024.
- [32] D. Gabor, "Theory of communication," *J. IEE*, vol. 93, pp. 429-457, 1946.
- [33] J. Ville, "Théorie et applications de la notion de signal analytique," *Cables et Transmission*, vol. 2A, pp. 61-74, 1948.
- [34] N. Dunford and J. T. Schwartz, *Linear Operators, Part I: General Theory*. New York: Interscience, 1958.
- [35] A. M. Sayeed and D. L. Jones, "Signal transforms covariant and invariant to operators," *IEEE Trans. Signal Processing*, 1994, preprint.
- [36] R. G. Baraniuk and D. L. Jones, "New signal-space orthonormal bases via the metaplectic transform," in *IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, Victoria, Canada, Oct. 1992, pp. 339-342.
- [37] ———, "Shear madness: New orthonormal bases and frames using chirp functions," *IEEE Trans. Signal Processing*, vol. 41, pp. 3543-3548, Dec. 1993.
- [38] U. Laine, "Famlet, to be or not to be a wavelet?" in *IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, Victoria, Canada, Oct. 1992, pp. 335-338.
- [39] A. Berthon, "Représentations et changements d'horloge," in *Proc. INSA TOM Workshop: Temps-Fréquence, Ondelettes et Multiresolution*, Lyon, France, Mar. 1994, pp. 19.1-19.4.
- [40] L. B. Almeida, "The fractional Fourier transform and time-frequency representations," *IEEE Trans. Signal Processing*, vol. 42, pp. 3084-3091, Nov. 1994.
- [41] C. Braccini and G. Gambardella, "Form-invariant linear filtering: Theory and applications," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 1612-1628, Dec. 1986.
- [42] R. G. Baraniuk and L. Cohen, "On joint distributions of arbitrary variables," *IEEE Signal Processing Letters*, vol. 2, pp. 10-12, Jan. 1995.
- [43] W. Philips, "Adaptively subsampled image coding with warped polynomials," in *Proc. IEEE Int. Conf. Image Processing-ICIP'94*, vol. II, Nov. 1994, pp. 366-369.
- [44] O. Rioul, "Simple regularity criteria for subdivision schemes," *SIAM J. Math. Anal.*, vol. 23, pp. 1544-1576, Nov. 1992.
- [45] J. E. Gray and R. E. Helmick, "The ambiguity function for broadband signals with application to objects undergoing nonuniform motion," in *IEEE Int. Symp. Time-Frequency Time-Scale Anal.*, Oct. 1994, pp. 493-496.
- [46] F. Hlawatsch, "Regularity and unitarity of bilinear time-frequency signal representations," *IEEE Trans. Inform. Theory*, vol. IT-38, pp. 82-94, Jan. 1992.
- [47] A. J. E. M. Janssen, "On the locus and spread of pseudo-density functions in the time-frequency plane," *Philips J. Res.*, vol. 37, no. 3, pp. 79-110, 1982.
- [48] F. Hlawatsch and P. Flandrin, "The interference structure of the Wigner distribution and related time-frequency signal representations," in *The Wigner Distribution-Theory and Applications in Signal Processing*, W. Mecklenbräuker, Ed. New York: Elsevier, 1995, to appear.
- [49] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: Optimal kernel design," *IEEE Trans. Signal Processing*, vol. 41, pp. 1589-1602, Apr. 1993.
- [50] ———, "A radially Gaussian, signal-dependent time-frequency representation," *Signal Processing*, vol. 32, pp. 263-284, June 1993.
- [51] K. G. Canfield and D. L. Jones, "Implementing time-frequency representations for non-Cohen classes," in *Proc. 27th Asilomar Conf.*, Pacific Grove, CA, 1993, pp. 1464-1468.
- [52] B. Ristic and B. Boashash, "Scale domain analysis of a bat sonar signal," in *IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, Oct. 1994, pp. 373-376.
- [53] F. Hlawatsch and R. Urbanke, "Bilinear time-frequency representations of signals: The shift-scale invariant class," *IEEE Trans. Signal Processing*, vol. 42, pp. 357-366, Feb. 1994.
- [54] A. Papandreou, F. Hlawatsch, and G. F. Boudreaux-Bartels, "Quadratic time-frequency distributions: The new hyperbolic class and its intersection with the affine class," in *Proc. Sixth SP Workshop Statistical Signal Array Processing*, Victoria, Canada, Oct. 1992, pp. 26-29.
- [55] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.
- [56] P. Flandrin and P. Gonçalves, "Geometry of affine time-frequency distributions," *J. Appl. Comp. Harmonic Anal.*, Oct. 1995.

Richard G. Baraniuk (M'93), for photograph and biography, please see p. 2371 of this issue of this TRANSACTIONS.

Douglas L. Jones (M'87), for photograph and biography, please see p. 2371 of this issue of this TRANSACTIONS.