BEYOND TIME-FREQUENCY ANALYSIS: ENERGY DENSITIES IN ONE AND MANY DIMENSIONS

*Richard G. Baraniuk**
Member, IEEE

Department of Electrical and Computer Engineering
Rice University
6100 Main Street, Houston, TX 77005–1892, USA
E-mail: richb@rice.edu, Fax: (713) 524–5237
Web: http://www-dsp.rice.edu

Submitted to *IEEE Transactions on Signal Processing*, November 1996
EDICS Number SP–2.3.1 Time-Frequency Signal Analysis

Abstract— **Given a unitary operator A representing a physical quantity of interest, we employ concepts from group representation theory to define two natural signal energy densities for A. The first is invariant to A and proves useful when the effect of A is to be ignored; the second is covariant to A and measures the "A" content of signals. We also consider joint densities for multiple operators. In the process, we provide an alternative interpretation of Cohen's general construction for joint distributions of arbitrary variables.**

*This work was supported by the National Science Foundation, grant no. MIP–9457438, and the Office of Naval Research, grant no. N00014–95–1–0849.

I Introduction

Time-frequency distributions, which indicate the energy content of signals simultaneously in both time and frequency, have proven indispensable for the study of the nonstationary signals appearing in many applications, including speech, radar, geophysical, biological, and transient signal analysis and processing. To date, most time-frequency distributions (the spectrogram and the Wigner distribution are two examples) have been matched to the time and frequency shift operators and, consequently, they perform well in applications where time and frequency shifts are fundamental [1–3]. In other problems, such as sonar and image processing, the concept of scale (compression/dilation) is more relevant than frequency shift; hence joint time-scale distributions have been developed [2, 4–6], the most popular of which is certainly the wavelet transform.

While time-frequency and time-scale distributions are the natural analysis and processing tools for large classes of signals, they are not perfectly matched to all signals, just as time shifts, frequency shifts, and scale changes are not the fundamental transformations appearing in all applications. For these different classes of signals, joint distributions based on concepts other than time, frequency, and scale have been developed. Among these new classes, we find the generalized Wigner distributions for the extended affine group [5] and arbitrary Lie groups [6], the hyperbolic class [7, 8], the power classes [8], the exponential class [9], the covariant classes [10–15], and the unitarily equivalent classes [16–18].

Operator methods have played a central role in the development of these new classes of joint signal distributions. Using ideas from quantum mechanics, these methods associate to a physical quantity a an operator $\mathbf{A}$ that may be unitary or Hermitian (symmetric). Operator representations can be manipulated, and typically, we seek transformations that measure the a or joint a vs. b energy content in a signal through the corresponding operator representations $\mathbf{A}$ and $\mathbf{B}$.

The endless variety of physical quantities necessitates general methods for generating and characterizing joint signal representations. In [1,19–22] Cohen pioneered such a method. At once simple and powerful, it is based on Hermitian operators, averages, and characteristic functions. In this paper, we take an alternate, yet equivalent, path to signal representations in arbitrary variables using unitary operators and the theory of commutative groups. Our objectives are to:

1. demonstrate that signal representations can be derived from unitary, as well as Hermitian operators. Unitary operators may be more natural in certain applications.
2. clarify the key differences between *covariant* and *invariant* signal representations, in particular for the case of "scale."
3. emphasize the rôle of the group Fourier transform in deriving and manipulating covariant and invariant signal representations (in the one-dimensional, commutative case).
4. extend Cohen's method to allow choice of either covariant or invariant marginals.
5. briefly indicate how Cohen's results generalize to discrete quantities typical in signal processing applications.

The ideas developed in this paper were announced in [23] and have already been been put to use and expanded upon in several papers, including [10–13,24–26]. In particular, Sayeed and Jones also studied the relationships between Hermitian and unitary operator representations and covariant and invariant marginals in [24–26].

The close relationship between unitary and Hermitian operator representations (by Stone's celebrated theorem [27, p. 614]) enforces an equivalence between our approach and that of Cohen. (This was proved by Sayeed and Jones in [24,25] and is sketched in Section IV-E below). Nevertheless, we feel that the insight into operator representations gained by looking at the "flip side" of the Hermitian operator method makes the development worth the trip.

Following a background section that reviews some basic results from functional analysis and group theory, we will proceed in steps of increasing complexity, beginning with one-dimensional distributions that measure the **A** content of signals (in Section III) and culminating in the construction of joint **A** vs. **B** distributions (in Section IV). While our approach to this problem is quite general, it is also very simple, owing to the group theoretic arguments employed.

II Unitary Operator Methods

In this paper, we set into correspondence physical quantities with parameterized unitary operators on a Hilbert function space H. In signal processing applications, H is typically one of the $L^2(G, d\mu_G)$ spaces, with G the set of function indices and $d\mu_G$ a measure on that set. These spaces have inner product $\langle g, h\rangle = \int_G g(x)\,\overline{h(x)}\,d\mu_G(x)$ for $g, h \in L^2(G, d\mu_G)$ and norm $\|h\|^2 = \langle h, h\rangle$.

We will adopt the relaxed mathematical tone that has become standard in the literature on joint distributions. For example, we will have occasion to employ generalized functions such as the Dirac delta $\delta(x)$ and sinusoid $e^{j2\pi fx}$ that exist only as the limit of functions in $L^2(G, d\mu_G)$. All arguments can be rigorized by working in a suitable Schwartz space or by using projection-valued measures for eigenfunctions [28, 29].

II-A Unitary representations of physical quantities

On $L^2(\mathbb{R}, dx)$ (continuous-time signals), we define the time and frequency operators as the unitary[1] time shift

$$(\mathbf{T}_t\, g)(x) \;\equiv\; g(x-t)$$

and unitary frequency shift

$$(\mathbf{F}_f\, g)(x) \;\equiv\; e^{j2\pi fx} g(x),$$

[1]A *unitary* operator $\mathbf{U}$ is a linear transformation from one Hilbert space H onto another that preserves energy; that is, $\|\mathbf{U}g\|^2 = \|g\|^2$ for all $g \in H$. Unitary operators also preserve inner products (isometry); that is, $\langle \mathbf{U}g, \mathbf{U}h\rangle = \langle g, h\rangle$ for any unitary $\mathbf{U}$ and for all $g, h \in H$.

with $x, t, f \in \mathbb{R}$. On $L^2(\mathbb{R}_+, dx)$ (one-sided continuous-time signals), we define the scale operator as the unitary dilation

$$(\mathbf{D}'_d\, g)(x) \;\equiv\; d^{-1/2}\, g(x/d),$$

with $x, d \in \mathbb{R}_+$. On $L^2(\mathbb{R}_+, dx/x)$ unitary dilation becomes

$$(\mathbf{D}_d\, g)(x) \equiv g(x/d).$$

On both $L^2(\mathbb{R}, dx)$ and $L^2(\mathbb{R}_+, dx)$, we define the "Mellin" operator as

$$(\mathbf{H}_h s)(x) \;\equiv\; e^{j2\pi h \log x}\, s(x) \;=\; x^{j2\pi h}\, s(x),$$

with $h \in \mathbb{R}$. Scaling can also be defined in terms of the (one-sided) Fourier transform of the signal, in which case the signal time domain becomes that of analytic signals.

Throughout this paper, we will assume that the function index $x \in G$ represents a time coordinate. Otherwise, the interpretations of $\mathbf{T}$, $\mathbf{F}$, $\mathbf{D}$, and $\mathbf{H}$ and other operators must be adjusted via a similarity transform. For example, if x represents frequency, then $\mathbf{T}$ corresponds to a frequency shift and $\mathbf{F}$ corresponds to a time shift. (More on this in Section II-C below.)

II-B Eigenanalysis and generalized Fourier transforms

Given a parameterized linear operator $\mathbf{A}_a$ on H, solution of the formal eigenequation

$$\left(\mathbf{A}_a \mathbf{u}^{\mathbf{A}}_\alpha\right)(x) \;=\; \lambda^{\mathbf{A}}_{a,\alpha}\; \mathbf{u}^{\mathbf{A}}_\alpha(x)$$

yields the formal eigenfunctions $\left\{\mathbf{u}^{\mathbf{A}}_\alpha(x)\right\}$ and the eigenvalues $\left\{\lambda^{\mathbf{A}}_{a,\alpha}\right\}$ of $\mathbf{A}_a$, both of which are indexed by the parameter α.[2] The eigenvalues and eigenfunctions of the time and frequency operators on $L^2(\mathbb{R}, dx)$ are easily shown to be

$$\begin{aligned} \lambda^{\mathbf{T}}_{t,\nu} &= e^{-j2\pi\nu t}, & \mathbf{u}^{\mathbf{T}}_\nu(x) &= e^{j2\pi\nu x} \\ \lambda^{\mathbf{F}}_{f,\tau} &= e^{-j2\pi f\tau}, & \mathbf{u}^{\mathbf{F}}_\tau(x) &= \delta(x+\tau) \end{aligned} \qquad (1)$$

[2]When the operator $\mathbf{A}_a$ is unitary, the eigenequation is merely algebraic, and the eigenfunctions are actually tempered distributions. More rigorously, we could employ projection-valued measures for the eigenfunctions [28, 29].

with $x, t, f, \nu, \tau \in \mathbb{R}$. The eigenvalues and eigenfunctions of the scale operators are given by

$$\lambda^{\mathbf{D}}_{d,\eta} = e^{-j2\pi\eta \log d}, \qquad \mathbf{u}^{\mathbf{D}}_{\eta}(x) = e^{j2\pi\eta \log x}$$

on $L^2(\mathbb{R}_+, dx/x)$ and by $\lambda^{\mathbf{D}'}_{d,\eta} = \lambda^{\mathbf{D}}_{d,\eta}$, $\mathbf{u}^{\mathbf{D}'}_{\eta}(x) = x^{-1/2}\,\mathbf{u}^{\mathbf{D}}_{\eta}(x)$ on $L^2(\mathbb{R}, dx)$, with $d, x \in \mathbb{R}_+, \eta \in \mathbb{R}$.

If $\mathbf{A}_a$ is unitary, then $\left|\lambda^{\mathbf{A}}_{a,\alpha}\right| = 1$ and the formal eigenfunctions form a complete orthonormal set in H [30]. The expansion onto these eigenfunctions then yields another operator that we will refer to as the **A**–*Fourier transform* $\mathbb{F}_{\mathbf{A}}$ [16]

$$(\mathbb{F}_{\mathbf{A}} s)(\alpha) \;\equiv\; \left\langle s,\, \mathbf{u}^{\mathbf{A}}_{\alpha} \right\rangle \;=\; \int_G s(x)\, \overline{\mathbf{u}^{\mathbf{A}}_{\alpha}(x)}\, d\mu_G(x).$$

Since $\mathbb{F}_{\mathbf{A}}$ is unitary, it is also invertible.

To continue the examples from above, $\mathbb{F}_{\mathbf{T}}$ is the usual Fourier transform (unitary on $L^2(\mathbb{R})$), $(\mathbb{F}_{\mathbf{F}} s)(t) = s(-t)$, and $\mathbb{F}_{\mathbf{D}}$ is a Mellin transform [1,5,6,31]

$$(\mathbb{F}_{\mathbf{D}} s)(\eta) \;=\; \left\langle s,\, \mathbf{u}^{\mathbf{D}}_{\eta} \right\rangle \;=\; \int_0^{\infty} s(x)\, e^{-j2\pi\eta \log x}\, \frac{dx}{x}, \qquad (2)$$

shown here for $s \in L^2(\mathbb{R}_+, dx/x)$. We will use simply $\mathbb{F}$ to denote the usual continuous-time Fourier transform $\mathbb{F}_{\mathbf{T}}$.

II-C Unitary equivalence

Because a unitary operator maps one Hilbert space onto another in a manner that exactly preserves its structure — it does not change the distances or angles between vectors — unitary operators can be interpreted as basis transformations, prompting the following definition of operators that are equivalent modulo a change of basis [16]:

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

The time and frequency operators are unitarily equivalent, since $\mathbf{F}_k = \mathbb{F}^{-1}\mathbf{T}_k\mathbb{F}$. The time operator on $L^2(\mathbb{R}, dx)$ is also equivalent to the exponentiated scale operator $\mathbf{D}'_{e^k}$ on $L^2(\mathbb{R}_+, dx)$, since $\mathbf{D}'_{e^k} = \mathbf{U}_{\text{hyp}}^{-1}\mathbf{T}_k\mathbf{U}_{\text{hyp}}$ with

$$(\mathbf{U}_{\text{hyp}}\, g)(x) \;=\; \sqrt{e^x x_0}\;\, g(e^x x_0), \qquad x_0 > 0. \tag{3}$$

This operator takes functions in $L^2(\mathbb{R}_+, dx)$ and stretches them into functions in $L^2(\mathbb{R})$. The time operator is *not* equivalent to either of the scale operators $\mathbf{D}_d$ or $\mathbf{D}'_d$, however.

The underlying structure of an operator is unchanged by a unitary equivalence transformation [16]. Given operators $\mathbf{A}_a$ and $\widetilde{\mathbf{A}}_a = \mathbf{U}^{-1}\mathbf{A}_a\mathbf{U}$, it is straightforward to show that

$$\lambda_{a,\alpha}^{\widetilde{\mathbf{A}}} = \lambda_{a,\alpha}^{\mathbf{A}}, \qquad \mathbf{u}_\alpha^{\widetilde{\mathbf{A}}} = \mathbf{U}^{-1}\mathbf{u}_\alpha^{\mathbf{A}}, \qquad \mathbb{F}_{\widetilde{\mathbf{A}}} = \mathbb{F}_{\mathbf{A}}\mathbf{U}. \tag{4}$$

II-D Unitary operators as group representations

A powerful tool for studying invariants of any kind is group theory [32,33].[3] Each parameterized operator that we use to represent a physical quantity can be interpreted as a unitary *representation* of some group. The particular group and group operation corresponding to a representation $\mathbf{A}_a$ are given by the domain of the parameter a and its behavior when $\mathbf{A}_a$ is composed with itself:

$$\mathbf{A}_x\mathbf{A}_y \;=\; \mathbf{A}_{x\bullet y}\,.$$

It follows immediately that all unitarily equivalent operators are representations of the same group, since if $\widetilde{\mathbf{A}} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$, then

$$\widetilde{\mathbf{A}}_{x\bullet y} \;=\; \widetilde{\mathbf{A}}_x\widetilde{\mathbf{A}}_y \;=\; (\mathbf{U}^{-1}\mathbf{A}_x\mathbf{U})(\mathbf{U}^{-1}\mathbf{A}_y\mathbf{U}) \;=\; \mathbf{U}^{-1}\mathbf{A}_{x\bullet y}\mathbf{U}.$$

Thus, the group concept partitions parameterized unitary operators into mutually exclusive equivalence classes.

[3] A set G with binary operation $\bullet$ is called a *group* if: 1) G is closed under $\bullet$; that is, $x \bullet y \in G \;\; \forall\, x, y \in G$; 2) the operation $\bullet$ is associative; that is $x \bullet (y \bullet z) = (x \bullet y) \bullet z \;\; \forall\, x, y, z \in G$; 3) there exists an identity element $\theta \in G$ such that $x \bullet \theta = \theta \bullet x = x \;\; \forall\, x \in G$; 4) for each $x \in G$ there exists an inverse element x^{-1} such that $x^{-1} \bullet x = x \bullet x^{-1} = \theta$. A group G is *abelian* (commutative) if $x \bullet y = y \bullet x \;\; \forall\, x, y \in G$ [32].

While we started out this section referring to the time, frequency, and scale operators as "parameterized unitary operators," we now see them for that they really are: The time and frequency operators are both representations of the Lie group $(\mathbb{R}, +)$ of real numbers with addition as the group operation, whereas the scale operators are representations of the Lie group $(\mathbb{R}_+, *)$ of positive real numbers with multiplication as the group operation. Although, for the sake of brevity, we will concentrate mainly these two groups for examples in this paper, note that all results are valid for arbitrary locally compact abelian (LCA) groups [32, 33]. In Section II-G, we will consider the integer and circle groups relevant to discrete-time signal analysis.

Given an LCA group G with group operation $\bullet$, the simplest possible representation of G on $L^2(G, d\mu_G)$ is the *group translation operator* $\mathbb{T}^G$, defined as

$$\left(\mathbb{T}^G_r s\right)(x) \;=\; s\left(x \bullet r^{-1}\right), \qquad s \in L^2(G, d\mu_G), \quad x, r \in G.$$

Translation is unitary if the measure $d\mu_G$ is chosen to be invariant to $\mathbb{T}^G$; that is, if

$$\int_G \left(\mathbb{T}^G_r s\right)(x)\, d\mu_G(x) \;=\; \int_G s(x)\, d\mu_G(x) \qquad \forall\, r \in G.$$

This invariant measure, usually referred to as the *Haar measure* [32, 33], can be regarded as the natural measure for the group G, singling out the space $L^2(G, d\mu_G)$ as the natural space on which the operators unitarily equivalent to $\mathbb{T}^G$ act. From this point onwards, we will employ only the Haar measure and write $L^2(G)$ for $L^2(G, d\mu_G)$.

Continuing the examples from above, for the group $(\mathbb{R}, +)$, the invariant measure is $d\mu_G(x) = dx$ and the group translation operator is $\mathbf{T}_r$, while for the group $(\mathbb{R}_+, *)$, the invariant measure is $d\mu_G(x) = dx/x$ and the group translation operator is $\mathbf{D}_r$.

The unitary equivalence transformation $\Theta_{\mathbf{A}}$ linking the representation $\mathbf{A}_a$ of the group G with its respective translation $\mathbb{T}^G$, through

$$\mathbf{A}_a \;=\; \Theta_{\mathbf{A}}^{-1}\, \mathbb{T}^G_a\, \Theta_{\mathbf{A}}, \tag{5}$$

plays a special role in the sequel. An illuminating interpretation of $\Theta_{\mathbf{A}}$ is as a coordinate transformation from "time" coordinates to "a" coordinates, where the action of $\mathbf{A}$ is simpler than any other unitary representation of the group G.

II-E The group Fourier transform for LCA groups

The fact that all unitary representations of a group G share common eigenvalues (see (4)) inspires the *group Fourier transform* $\mathbb{F}_G$ [32,33]

$$(\mathbb{F}_G s)(\chi) \;=\; \int_G s(x)\,\overline{\lambda^G_{x,\chi}}\,d\mu_G(x).$$

Here, λ^G denotes the common eigenvalues, which are called the *characters* of the group. The transformation $\mathbb{F}_G$ is a unitary map from the space $L^2(G)$ onto the space $L^2(G^*, d\mu_{G^*})$ based on the *dual group* G^* of G [30].

The characters form one-dimensional representations of both groups G and G^*. For G, we have

$$\lambda^G_{x_1,\chi}\lambda^G_{x_2,\chi} \;=\; \lambda^G_{x_1\bullet x_2,\chi}\,, \qquad x\in G,\;\chi\in G^*.$$

For the dual group G^*, with group operation $\circ$, we have

$$\lambda^G_{x,\chi_1}\lambda^G_{x,\chi_2} \;=\; \lambda^G_{x,\chi_1\circ\chi_2}\,, \qquad x\in G,\;\chi\in G^*.$$

The dual group is also LCA, but in general $\circ$ differs from $\bullet$. The string of identities

$$\lambda^G_{x^{-1},\chi} \;=\; \lambda^G_{x,\chi^{-1}} \;=\; \left(\lambda^G_{x,\chi}\right)^{-1} \;=\; \overline{\lambda^G_{x,\chi}}$$

echos the properties of the complex sinusoids employed in the traditional Fourier transform. Throughout this paper, we will use roman letters to denote elements of G and greek letters to denote elements of G^*.

Since $\mathbb{F}_G$ is unitary, the inverse group transform of a function $S(\chi)\in L^2(G^*)$ is defined similarly to the above

$$\left(\mathbb{F}_G^{-1}S\right)(x) \;=\; \int_{G^*} S(\chi)\,\lambda^G_{x,\chi}\,d\mu_{G^*}(\chi), \qquad x\in G.$$

Since $\lambda^G_{x,\chi} = \overline{\lambda^{G*}_{\chi,x}}$, this inverse transform can be interpreted as a group Fourier transform in its own right.

On $L^2(G)$ (using the invariant measure), the group Fourier transform coincides with the $\mathbb{T}^G$–Fourier transform; therefore, the $\mathbf{A}$–Fourier transforms of all operators unitarily equivalent to $\mathbb{T}^G$ can be obtained (up to reversal) as

$$\mathbb{F}_{\mathbf{A}} \;=\; \mathbb{F}_G\,\Theta_{\mathbf{A}},$$

with $\Theta_{\mathbf{A}}$ the transformation to "a" coordinates defined by (5).

For the two groups we have focused on thus far, we have $(\mathbb{R},+)^* = (\mathbb{R},+)$ and $(\mathbb{R}_+,*)^* = (\mathbb{R},+)$. The $(\mathbb{R},+)$ group Fourier transform is the usual Fourier transform mapping $L^2(\mathbb{R},+)$ onto $L^2(\mathbb{R},+)$. The $(\mathbb{R}_+,*)$ group Fourier transform is the Mellin transform mapping $L^2(\mathbb{R}_+, dx/x)$ onto $L^2(\mathbb{R},+)$.

II-F Dual Operators

After group translation $\mathbb{T}^G$, the simplest nontrivial unitary transformation on $L^2(G)$ is the phase shift

$$(\Lambda^{G^*}_\alpha s)(x) \;\equiv\; \overline{\lambda^{G*}_{\alpha,x}}\, s(x) \;=\; \lambda^G_{x,\alpha}\, s(x), \qquad x \in G,\ \alpha \in G^*.$$

A representation of the dual group G^*, this transformation induces a translation in the group Fourier transform:

$$\mathbb{F}_G\,\Lambda^{G^*}_\alpha \;=\; \mathbb{T}^{G^*}_\alpha\,\mathbb{F}_G.$$

Because of the close connection between translation and phase shift, we refer to $\mathbb{T}^G$ and Λ^{G^*} as *dual operators* on $L^2(G)$ and write $\Lambda^{G^*}_\alpha = \mathrm{d}(\mathbb{T}^G)_\alpha$. Dual operators were introduced in this context in [34] and studied in more detail in [26, 35]. Hlawatsch et al refer to dual operators as "conjugate operators" [10–13].

If the operator $\mathbf{A}_a$ translates in a coordinates, then its dual operator will phase shift in these same coordinates. To find the operator dual to an arbitrary unitary $\mathbf{A}_a$, we can employ the unitary equivalence principle. Given that $\mathbf{A}_a$ represents the group G, we have the following diagram:

$$\begin{array}{ccc} \mathbb{T}^G_a & \stackrel{\text{dual}}{\longleftrightarrow} & \Lambda^{G^*}_\alpha \\ \Theta_{\mathbf{A}}\downarrow & & \Theta_{\mathbf{A}}\downarrow \\ \mathbf{A}_a & \stackrel{\text{dual}}{\longleftrightarrow} & \mathrm{d}(\mathbf{A})_\alpha \end{array} \tag{6}$$

and the conclusion that

$$\begin{aligned} \mathbf{A}_a &= \Theta_{\mathbf{A}}^{-1}\, \mathbb{T}^G_a\, \Theta_{\mathbf{A}}, \qquad && a \in G \\ \Rightarrow \quad \mathrm{d}(\mathbf{A})_\alpha &= \Theta_{\mathbf{A}}^{-1}\, \Lambda^{G^*}_\alpha\, \Theta_{\mathbf{A}}, \qquad && \alpha \in G^*. \end{aligned} \tag{7}$$

Both $\mathbf{A}_a$ and $\mathrm{d}(\mathbf{A})_\alpha$ operate on $L^2(G)$; however, unless $G = G^*$, $\mathbf{A}$ and $\mathrm{d}(\mathbf{A})$ are *not* unitarily equivalent.

Time $\mathbf{T}$ and frequency $\mathbf{F}$ provide the classical example of dual operators, with $\mathbf{T} = \mathbb{T}^G$ and $\mathbf{F} = \Lambda^{G^*}$ for $G = (\mathbb{R}, +)$. Since $G^* = G$ in this case, these operators are also unitarily equivalent, with the usual Fourier transform $\mathbb{F}$ providing the link.

Scale $\mathbf{D}$ and Mellin $\mathbf{H}$ are also dual operators, with $\mathbf{D} = \mathbb{T}^G$ and $\mathbf{H} = \Lambda^{G^*}$ for $G = (\mathbb{R}_+, *)$. However, since $G^* = (\mathbb{R}, +) \neq G$, $\mathbf{H}$ is not unitarily equivalent to $\mathbf{D}$, but rather to $\mathbb{T}^{G^*} = \mathbf{T}$, with the Mellin transform as link. The interpretation of dual operators is complicated by the fact that Λ^{G^*} ($\mathbf{H}$ here) can operate on both $L^2(G)$ and $L^2(G^*)$.

II-G Further examples

Group $G = (\mathbb{Z}, +)$: Discrete-time signal analysis involves a set of groups different from $(\mathbb{R}, +)$ and $(\mathbb{R}_+, *)$. On $L^2(\mathbb{Z})$ (discrete-time signals with the counting measure), we define the time shift operator as $(\mathbf{T}_n\, g)(x) = g(x - n)$, with $x, n \in \mathbb{Z}$. This operator represents the group $(\mathbb{Z}, +)$ of integers with addition as the group operation; in fact, it is the group translation operator. Its eigenvalues and eigenfunctions are given by

$$\lambda^{\mathbf{T}}_{t,\xi} = e^{-j2\pi\xi t}, \qquad \mathbf{u}^{\mathbf{T}}_\xi(x) = e^{j2\pi\xi x}, \qquad x, n \in \mathbb{Z},\ \ \xi \in [0, 1). \tag{8}$$

The dual group of $(\mathbb{Z}, +)$ is the circle group $([0, 1), +_1)$ having addition mod 1 as group operation. The group Fourier transform mapping $L^2(\mathbb{Z})$ onto $L^2[0, 1)$ is the classical discrete-time Fourier transform. The frequency shift operator $(\mathbf{F}_m\, g)(x) = e^{j2\pi m x} g(x)$, with $m \in [0,1)$, operates on $L^2(\mathbb{Z})$ as a representation of the circle group. In addition to being unitarily equivalent, $\mathbf{T}_m$ and $\mathbf{F}_m$ are dual operators.

Group $G = (\mathbb{Z}_N, +_N)$: Finite data sets dictate the group of integers modulo N, $(\mathbb{Z}_N, +_N)$, having addition modulo N as group operation. With this group, translation takes the form of a cyclic shift, while the group Fourier transform corresponds to the discrete Fourier transform mapping $(\mathbb{Z}_N, +_N)$ onto itself.

The results of this paper spring immediately from the characterization of physical quantities (time, frequency, scale, and so on) as group objects, specifically, as unitary representations of groups on certain Hilbert signal spaces. We will now apply this powerful mathematical machinery to the problem at hand: energy densities in one and many dimensions.

III Energy Densities

In this section, we define two natural transforms for a unitary representation $\mathbf{A}$ of an LCA group [23]. The first is invariant to $\mathbf{A}$, while the second measures the "a" content of signals. It is useful to keep in mind as models the operators $\mathbf{F}$ and $\mathbf{T}$ and the transformations $|s(t)|^2$ and $|\mathbb{F}s(f)|^2$, which indicate the time and frequency energy content of a time signal s. While the results of this section are general, when interpreting them we will assume that the signal $s(x)$ has been expressed in the "time domain."

III-A Invariant energy densities

Squaring the $\mathbf{A}$–Fourier transform yields an energy density that is invariant to the operator $\mathbf{A}$

$$|(\mathbb{F}_{\mathbf{A}} \mathbf{A}_a s)(\alpha)|^2 \;=\; |(\mathbb{F}_{\mathbf{A}} s)(\alpha)|^2 \,.$$

Thus, we will refer to $|\mathbb{F}_{\mathbf{A}} s|^2$ as the $\mathbf{A}$–*invariant energy density* ($\mathbf{A}$–IED). It is extremely important to note that the $\mathbf{A}$–IED does *not* indicate the $\mathbf{A}$ content of the signal s, precisely because of this invariance. In fact, the **A**–IED is the transform of choice when the action of **A** is to be ignored!

As examples, recall that the **T**–Fourier transform reduces to the usual Fourier transform, which indicates not time but frequency content. Similarly, the **D**–Fourier (Mellin) transform of (2) indicates a sort of "logarithmic chirp" content. (Cohen refers to this quantity as "scale" [1, 21, 22, 36].)

III-B Covariant energy densities

Although it is clear that the energy density indicating the **A** content of a signal must change when **A** is applied to the signal, it is unclear what type of change should occur.

> **Definition:** A transformation Θ is **Z–covariant** to an operator **A** if $\Theta \mathbf{A}_a = \mathbf{Z}_a \Theta$.

The essence of **Z**–covariance is that the operator **Z** describes the effect of "pulling" **A** out through the transformation Θ. It follows directly from the definition that if we desire **Z**–covariance in an **A**–content energy density, then the density of choice is $|\Theta s|^2$. Note that **Z** cannot be arbitrary, since **A** and **Z** must be unitarily equivalent, with $\mathbf{A}_a = \Theta^{-1}\mathbf{Z}_a\Theta$, and, hence, must be representations of the same group G.

Since the natural choice for Θ should lead to the simplest possible covariance, a group-theoretic argument suggests we choose $\mathbf{Z} = \mathbb{T}^G$, the translation operator of the group G. With this choice, the **A**–*covariant energy density* (**A**–CED) becomes the square $|(\Theta_{\mathbf{A}} s)(a)|^2$ of the transformation (5) that maps signals from time to a coordinates.

It is straightforward to demonstrate that the d(**A**)–IED, $\mathsf{IF}_{\mathrm{d}(\mathbf{A})}$, coincides precisely with the **A**–CED and vice versa.

III-C Examples

While the **A**–IED does not indicate the **A** content of signals, the **A**–CED does.

Group $G = (\mathbb{R}, +)$: Translation corresponds to $\mathbb{T}_k^G = \mathbf{T}_k$. On $L^2(\mathbb{R})$, the time and frequency operators produce the expected results: The **T**–IED is the square of the usual Fourier transform $|(\mathbb{F}s)(f)|^2$ and the **T**–CED is $|s(t)|^2$, whereas for **F** these densities are reversed.

The chirp modulation operator $(\mathbf{C}_c s)(x) = e^{j2\pi c|x|^p \mathrm{sgn}(x)} s(x)$, $p \neq 0$ on $L^2(\mathbb{R})$ is related to $\mathbb{T}^G$ by $\mathbf{C}_c = \Theta_{\mathbf{C}}^{-1} \mathbb{T}_c^G \Theta_{\mathbf{C}}$ with

$$(\Theta_{\mathbf{C}} s)(c) \;=\; \int s(x)\, e^{-j2\pi c|x|^p \mathrm{sgn}(x)}\, |x|^{(p-1)/2}\, |p|^{1/2}\, dx. \tag{9}$$

Since the **C**–CED $|(\Theta_{\mathbf{C}} s)(c)|^2$ indicates the "chirp" content of signals, it has been called the *chirp transform* [8,16]. The **C**–IED is given by $|(\mathbb{F}_{\mathbf{C}} s)(\gamma)|^2 = |p|^{-1}\, |\gamma|^{(1-p)/p}\, |s(|\gamma|^{1/p} \mathrm{sgn}(\gamma))|^2$. This transform demonstrates the nonuniqueness of IEDs, since the density $|s(\gamma)|^2$ is also invariant under **C**.

The Mellin operator $\mathbf{H}_h$ represents G on $L^2(\mathbb{R}_+)$ and $L^2(\mathbb{R}_+, dx/x)$. The **H**–CED is the Mellin transform $\Theta_{\mathbf{H}} = \mathbb{F}_{\mathbf{D}}$; an **H**–IED is $|s(x)|^2$.

Group $G = (\mathbb{R}_+, *)$: Translation corresponds to $\mathbb{T}_k^G = \mathbf{D}_k$. The **D**–IED is the Mellin transform (Cohen's "scale transform" [1,36]).

The **D**–CED is $|s(d)|^2$, $d > 0$ This result may be surprising (and perhaps a little disappointing), but it is also correct, consistent, and reasonable [34]. First, this density has the correct covariance property: $|(\mathbf{D}_k s)(d)|^2 = |s(d/k)|^2$. Second, this density indicates the scale content of the signal s: just as the **T**-CED $|s(t)|^2$ indicates the amount of group translation ($\mathbb{T}^G = \mathbf{T}$ in the additive group) required to bring the signal energy at the point t to the identity element $\theta = 0$ of that group, the **D**–CED $|s(d)|^2$ indicates the amount of group translation ($\mathbb{T}^G = \mathbf{D}$ in the multiplicative group) required to bring the signal energy at the point d to the identity element $\theta = 1$ of that group. Third, since scale is a ratio — "this thing is twice as big as that thing" — the features on which the scale content of a signal can be judged exist in the time domain, and so no elaborate coordinate transformations are necessary. An enlightening interpretation of the **D**–CED is provided by considering the index d of $s(d)$ as a spatial variable in an imaging system, with $|s(d)|^2$ the distribution of an object along the image axis. In this case, d represents a "zoom" parameter, and the spread of the **D**–CED indicates the amount of focus change required to successively bring all of the object into focus at an image plane at $d = 1$.

Discrete groups: In discrete-time, over the groups $(\mathbb{Z}, +)$ and $(\mathbb{Z}_N, +_N)$, the time signal $|s[n]|^2$ and the square of the (discrete) Fourier transform are the time and frequency IEDs and CEDs.

III-D From ambiguity functions to energy densities

The correlation, characteristic, or *ambiguity function*

$$(\mathbf{Q}s)(a) \;\equiv\; \langle s, \mathbf{A}_a s\rangle \tag{10}$$

measures the similarity between s and $\mathbf{A}_a s$ as a function of a. When $\mathbf{A} = \mathbb{T}^G$, (10) corresponds to a group convolution — the usual convolution on $(\mathbb{R}, +)$ and multiplicative convolution on $(\mathbb{R}_+, *)$ [5, 6, 31]. The Fourier transform takes an ambiguity function from the correlative domain to the energetic domain [3]; two cases will be of interest in the sequel.

The group Fourier transform of (10) yields the **A**–IED $|(\mathbb{F}_{\mathbf{A}} s)(\alpha)|^2$. To see this, we recall the identity $\mathbf{A}_a = \Theta_{\mathbf{A}}^{-1} \mathbb{T}_a^G \Theta_{\mathbf{A}}$ and write

$$\begin{aligned}
\mathbb{F}_{G,\, a\mapsto\alpha} \langle s,\ \mathbf{A}_a s\rangle &= \mathbb{F}_{G,\, a\mapsto\alpha} \left\langle \Theta_{\mathbf{A}} s,\ \mathbb{T}_a^G \Theta_{\mathbf{A}} s\right\rangle \\
&= \int_G (\Theta_{\mathbf{A}} s)(u) \left[\int_G \overline{(\Theta_{\mathbf{A}} s)(u \bullet a^{-1})}\ \overline{\lambda_{a,\alpha}^G}\, d\mu_G(a)\right] d\mu_G(u) \\
&= \int_G (\Theta_{\mathbf{A}} s)(u) \left[\int_G \overline{(\Theta_{\mathbf{A}} s)(v)}\ \overline{\lambda_{u\circ v^{-1},\alpha}^G}\, d\mu_G(v)\right] d\mu_G(u) \\
&= \left[\int_G (\Theta_{\mathbf{A}} s)(u)\ \overline{\lambda_{u,\alpha}^G}\, d\mu_G(u)\right] \left[\int_G \overline{(\Theta_{\mathbf{A}} s)(v)}\ \lambda_{v,\alpha}^G\, d\mu_G(v)\right]
\end{aligned}$$

$$= |(\mathbb{F}_G \Theta_{\mathbf{A}} s)(\alpha)|^2 = |(\mathbb{F}_{\mathbf{A}} s)(\alpha)|^2. \tag{11}$$

The inverse group Fourier transform of (10) with the dual operator $\mathrm{d}(\mathbf{A})_\alpha$ in place of $\mathbf{A}_a$ yields the $\mathbf{A}$–CED $|(\Theta_{\mathbf{A}} s)(a)|^2$, as we see from

$$\begin{aligned}
\mathbb{F}^{-1}_{G,\,\alpha\mapsto a} \langle s,\, \mathrm{d}(\mathbf{A})_\alpha s\rangle &= \mathbb{F}^{-1}_{G,\,\alpha\mapsto a} \left\langle \Theta_{\mathbf{A}} s,\, \Lambda^{G^*}_\alpha \Theta_{\mathbf{A}} s\right\rangle \\
&= \int_G |(\Theta_{\mathbf{A}} s)(u)|^2 \left[\int_{G^*} \overline{\lambda^G_{u,\alpha}}\, \lambda^G_{a,\alpha}\, d\mu_{G^*}(\alpha)\right] d\mu_G(u) \\
&= \int_G |(\Theta_{\mathbf{A}} s)(u)|^2 \left[\int_{G^*} \lambda^G_{a\circ u^{-1},\alpha}\, d\mu_{G^*}(\alpha)\right] d\mu_G(u) \\
&= |(\Theta_{\mathbf{A}} s)(a)|^2.
\end{aligned} \tag{12}$$

While merely an intermediate step in the single operator case, ambiguity functions form the foundation for joint energy densities for multiple operators, to which we now turn.

IV Joint Energy Densities

IV-A Motivation

In the previous section, we defined two energy densities that are natural for any unitary representation $\mathbf{A}$ of an LCA group. The $\mathbf{A}$–IED is invariant to $\mathbf{A}$, while the $\mathbf{A}$–CED is covariant to $\mathbf{A}$ (and thus measures the $\mathbf{A}$ content in signals). In many applications, these densities and their linear equivalents will be more than adequate for characterizing, analyzing, and processing signals. One obvious example of an IED/CED is the square of the Fourier transform, which has found a multitude of applications.

For more complicated signals and systems, several physical quantities may be meaningful simultaneously, making joint densities based on several operators a necessity. Time-frequency distributions, for example, were developed for problems where both the time shift $\mathbf{T}$ and frequency shift $\mathbf{F}$ are important. In this section, we build on the above results to construct joint energy densities for combinations of physical quantities [23].

The basis for our method lies in Cohen's pioneering construction for joint densities of arbitrary variables [1, 19]. Like Cohen, we will employ ambiguity (characteristic) functions to extend the one-dimensional construction of Section III-D to two and higher dimensions. Unlike Cohen, we will begin with unitary operator representations and finish with an extended construction that supports groups beyond $(\mathbb{R}, +)$ and offers a choice of either IED or CED marginals.

A relaxable marginalization constraint will ensure that the axes of our joint distributions lie oriented along the right quantities. We will say that a two-dimensional mapping $(\mathbf{P}s)(u, v)$ is a joint density for $\mathbf{A}$ and $\mathbf{B}$ if its marginal distributions $\mathbf{M_A}$ and $\mathbf{M_B}$, given by

$$\begin{aligned} \int (\mathbf{P}s)(u,v)\, d\mu_{\mathbf{B}}(v) &= |(\mathbf{M_A}s)(u)|^2 \\ \int (\mathbf{P}s)(u,v)\, d\mu_{\mathbf{A}}(u) &= |(\mathbf{M_B}s)(v)|^2, \end{aligned}$$

are energy densities for $\mathbf{A}$ and $\mathbf{B}$. Natural possibilities for $\mathbf{M_A}$ and $\mathbf{M_B}$ are the IEDs and CEDs for $\mathbf{A}$ and $\mathbf{B}$. If CEDs are chosen for both $\mathbf{A}$ and $\mathbf{B}$ (one of the four options), then $\mathbf{P}s$ can be interpreted as indicating the joint $\mathbf{A}$–$\mathbf{B}$ content of the signal s.

IV-B Construction

We will explicitly formulate joint energy densities only for pairs of unitary operators; the jump from two to higher dimensions is straightforward. For the moment, we will assume that both operators represent the same LCA group G. This assumption will be relaxed in Section IV-E. The construction of joint densities for two operators $\mathbf{A}$ and $\mathbf{B}$ proceeds as follows:

Step 1: Choose either IED or CED marginals for $\mathbf{A}$ and $\mathbf{B}$. For an $\mathbf{A}$–IED marginal, set $\mathbf{A}^{\star}{}_a \equiv \mathbf{A}_a$, while for an $\mathbf{A}$–CED marginal, set $\mathbf{A}^{\star}{}_\alpha \equiv \mathrm{d}(\mathbf{A})_\alpha$, the dual operator. Form $\mathbf{B}^{\star}$ in the same manner.

To illustrate, we will work towards **A**-CED and **B**-IED marginals in the following. Hence, we have $\mathbf{A}^\star = \mathrm{d}(\mathbf{A})$ and $\mathbf{B}^\star = \mathbf{B}$.

Step 2: Form the ambiguity function

$$(\mathbf{Q}s)(\alpha, b) \;=\; \langle s,\; \mathbf{A}^\star{}_\alpha \mathbf{B}^\star{}_b\, s\rangle\,, \tag{13}$$

with the inner product taken in $L^2(G)$.

Different orderings of $\mathbf{A}^\star$ and $\mathbf{B}^\star$ yield different joint distributions [1,20]. In general, any function $\mathrm{ord}(\mathbf{A}^\star, \mathbf{B}^\star; \alpha, b)$ of $\mathbf{A}^\star$ and $\mathbf{B}^\star$ is permissible as long as[4]

$$\mathrm{ord}(\mathbf{A}^\star, \mathbf{B}^\star;\, \alpha, \theta) \;=\; \mathbf{A}^\star{}_\alpha \tag{14}$$

$$\mathrm{ord}(\mathbf{A}^\star, \mathbf{B}^\star;\, \theta, b) \;=\; \mathbf{B}^\star{}_b \tag{15}$$

with θ the identity element appropriate for the group. (With $\mathbf{A}^\star = \mathrm{d}(\mathbf{A})$ and $\mathbf{B}^\star = \mathbf{B}$, $\theta \in G$ in (14) and $\theta \in G^*$ in (15).)

Rather than enumerating all orderings, we can fix a single valid ordering in (13) and obtain a large class of others using a kernel function $\phi(\alpha, b)$

$$(\mathbf{Q}_\phi s)(\alpha, b) \;=\; \phi(\alpha, b)\, \langle s,\; \mathbf{A}^\star{}_\alpha \mathbf{B}^\star{}_b\, s\rangle\,, \tag{16}$$

Kernels satisfying the constraint $\phi(\alpha, \theta) = \phi(\theta, b) = 1\ \forall\ \alpha, b$ will generate distributions satisfying the correct marginals. Distributions generated by kernels that violate this constraint will not marginalize correctly, but are nonetheless useful. Note that the kernel method is not foolproof, however; see [35, 37, 38].

Step 3: Compute the **A**–**B** distribution as the double group Fourier transform of $\mathbf{Q}s$. For IED marginals, use the forward transform; for CED marginals, use the inverse transform. Given our choice of **A**-CED and **B**-IED marginals, we have

$$(\mathbf{P}_\phi s)(a, \beta) \;=\; \mathbb{F}^{-1}_{G,\, \alpha \to a}\, \mathbb{F}_{G,\, b \to \beta}\ \phi(\alpha, b)\ \langle s,\; \mathrm{d}(\mathbf{A})_\alpha \mathbf{B}_b\, s\rangle\,. \tag{17}$$

[4]Using unitary equivalence, we can expand the class of operators that when inserted into (13) yield distributions that marginalize appropriately. See [25] for more details.

To obtain distributions of more than two variables, simply place all of the relevant operators into the ambiguity function (16) and Fourier transform accordingly.

IV-C Marginal and covariance properties

The marginal properties of $\mathbf{P}s$,

$$\int (\mathbf{P}s)(a,\beta)\, d\mu_{G^*}(\beta) = |(\Theta_{\mathbf{A}} s)(a)|^2 \tag{18}$$

$$\int (\mathbf{P}s)(a,\beta)\, d\mu_{G}(a) = |(\mathbb{F}_{\mathbf{B}} s)(\beta)|^2, \tag{19}$$

follow directly from the properties of the group Fourier transform. We begin with the **A**–CED marginal (18). Integrating $\mathbf{P}s$ with respect to $d\mu_{G^*}(\beta)$ yields

$$\int_{G^*} (\mathbf{P}s)(a,\beta)\, d\mu_{G^*}(\beta) = \mathbb{F}^{-1}_{G,\,\alpha\mapsto a} \int_G \phi(\alpha,b)\, \langle s,\, \mathrm{d}(\mathbf{A})_\alpha\, \mathbf{B}_b s\rangle \left[\int_{G^*} \overline{\lambda^G_{b,\beta}}\, d\mu_{G^*}(\beta)\right] d\mu_G(b).$$

Formally, the term in brackets equals the Dirac impulse $\delta(\beta)$; hence, the integration over b sifts the value $b=\theta$ into $\mathbf{B}$ and ϕ. Since $\mathbf{B}_\theta = \mathbf{I}$, the identity operator, the analysis of (12) confirms that the $\mathbf{A}$ marginal of $\mathbf{P}s$ equals $|(\Theta_{\mathbf{A}} s)(a)|^2$ provided $\phi(\alpha,\theta) = 1\ \forall\, \alpha$.

The **B**–IED marginal (19) follows similarly. Integrating $\mathbf{P}s$ with respect to $d\mu_G(a)$ (and skipping the disappearance of $\mathrm{d}(\mathbf{A})$, which occurs as above for $\mathbf{B}$), an analysis similar to (11) confirms that the $\mathbf{B}$ marginal of $\mathbf{P}s$ equals $|(\mathbb{F}_{\mathbf{B}} s)(\beta)|^2$ provided $\phi(\theta,b) = 1\ \forall\, b$.

The covariance properties of $\mathbf{P}s$ are completely determined by the operator $\mathbf{A}^\star\,\mathbf{B}^\star$ that rules over the ambiguity function (16). Here the theory departs from that of Section III, because in general the composition $\mathbf{A}^\star\,\mathbf{B}^\star$ is neither commutative nor even a group representation. Thus, even if we select **A**-CED and **B**-CED marginal distributions in $\mathbf{P}s$, it might not exhibit any covariance aside from in those marginals. A complete discussion of covariance lies beyond the scope of this paper; we refer the reader to [5, 6, 14, 15, 18, 24, 25, 37, 38] for results applicable to continuous groups.

IV-D Examples

Time-frequency: By setting $\mathbf{A}^\star = \mathbf{F}_\nu$ and $\mathbf{B}^\star = \mathbf{T}_\tau$ in (17), we obtain the classical Cohen's class of time-frequency distributions. The resulting distributions possess a special symmetry: each marginal plays a dual role as IED for one variable and CED for the other. Distributions generated by fixed kernels sport covariance to both time and frequency shifts. We assume $\phi \equiv 1$ throughout this section.

To analyze continuous-time signals on $L^2(\mathbb{R})$, we employ $\nu, \tau \in \mathbb{R}$ and construct distributions on $(\mathbb{R}, +) \times (\mathbb{R}, +)$. The central Wigner distribution [1]

$$(\mathbf{W}s)(t,f) \;=\; \int s\left(t+\frac{\tau}{2}\right) s^*\left(t-\frac{\tau}{2}\right) e^{-j2\pi f\tau}\, d\tau$$

results from plugging $\mathbf{T}_{\tau/2}\mathbf{F}_\nu\mathbf{T}_{\tau/2} = e^{-j\pi\nu\tau}\,\mathbf{F}_\nu\mathbf{T}_\tau$ into (16).

To analyze discrete-time signals on $L^2(\mathbb{Z})$, we employ $\nu \in [0,1)$, $\tau \in \mathbb{Z}$ and construct distributions on $(\mathbb{Z},+) \times [0,1)$. Note that although $\mathbf{F}_\nu$ and $\mathbf{T}_\tau$ represent different groups in this case, they both act on the same signal space, $L^2(\mathbb{Z})$, and so the construction remains valid. Use of the ordering $e^{-j\pi\nu\tau}\,\mathbf{F}_\nu\mathbf{T}_\tau$ in (16) with double-oversampled signals results in the discrete-time Wigner distribution of Claasen and Mecklenbräuker [39]

$$(\mathbf{W}_{\mathrm{d}}s)(n,m) \;=\; \sum_{\tau=-\infty}^{\infty} s(n+\tau)\, s^*(n-\tau)\, e^{-j4\pi m\tau}.$$

Other orderings yield alternative definitions of the discrete-time Wigner distribution (see [40] and the references in [41]).

To analyze finite length discrete-time signals on $L^2(\mathbb{Z}_N)$, we employ $\nu, \tau \in \mathbb{Z}_N$ and construct distributions on $(\mathbb{Z}_N, +_N) \times (\mathbb{Z}_N, +_N)$. The ordering $e^{-j\pi\frac{(\nu\tau)_N}{N}}\,\mathbf{F}_{\nu/N}\mathbf{T}_\tau$ results in the discrete-time, discrete-frequency Wigner distribution of Richman, Parks, and Shenoy [41]

$$(\mathbf{W}_N s)(n,k) \;=\; \frac{1}{N}\sum_{\tau,\nu,l=0}^{N-1} s(l)\, s^*((l+\tau)_N)\, e^{\frac{-j\pi}{N}[2(n\nu+k\tau)-(\nu\tau)_N-2\nu l]}.$$

(Operations $(\cdot)_N$ are carried out modulo N.)

Time-scale: Since time $\mathbf{T}$ and scale $\mathbf{D}$ are not dual variables, there exist several species of time-scale distributions.

Time-scale CED distributions are of limited interest, as both the $\mathbf{T}$–CED and $\mathbf{D}$–CED marginal will be of the form $|s(t)|^2$.

The $\mathbf{T}$–IED vs. $\mathbf{D}$–IED distributions have Fourier and Mellin marginals. We will work with continuous-time analytic signals whose Fourier transforms live in $L^2(\mathbb{R}_+)$. The Altes Q distribution results from (16) with the ordering $\mathbf{D}'_{d/2}\mathbf{T}_t\mathbf{D}'_{d/2}$, with $t \in \mathbb{R}$, $d \in \mathbb{R}_+$ [42]. The ordering $\mathbf{T}_{\frac{d-1}{\log d}t}\mathbf{D}'_d$ [6] yields the Bertrand "tomographic" distribution [5], which has also been obtained by Shenoy and Parks [6] and by Cohen [1, p. 257]. The Q distribution is covariant to scale changes, but not to time shifts. The Bertrand distribution is covariant to both scale changes and time shifts.

The $\mathbf{T}$–CED vs. $\mathbf{D}$–IED distributions have time and Mellin marginals. The ordering $\mathbf{D}'_{d/2}\mathbf{F}_f\mathbf{D}'_{d/2}$ yields the time-scale distribution of Eichmann and Marinovich [43] and Altes [42] for single-sided signals in $L^2(\mathbb{R}_+)$.

Time-A: For analyzing time signals, joint distributions with a time marginal are fundamental. Joint distributions of $\mathbf{T}$ and $\mathbf{A}$ measure joint time and $\mathbf{A}$ content. For the distributed ordering $\mathbf{A}_{a^{-1}/2}\mathbf{F}_{-\nu}\mathbf{A}_{a^{-1}/2}$,[5] $\mathbf{T}$–CED vs. $\mathbf{A}$–IED distributions assume a symmetrical form

$$\begin{aligned}(\mathbf{P}s)(t,\alpha) &= \mathbb{F}_{G,a\mapsto\alpha}\left(\mathbf{A}_{a/2}s\right)(t)\,\overline{\left(\mathbf{A}_{a^{-1}/2}s\right)(t)}\\ &= \int_G \left(\mathbf{A}_{a/2}s\right)(t)\,\overline{\left(\mathbf{A}_{a^{-1}/2}s\right)(t)}\;\overline{\lambda^G_{a,\alpha}}\,d\mu_G(a)\end{aligned}$$

Distributions measuring the $\mathbf{A}$–CED use the dual operator and the inverse group transform. This prescription generalizes the time-scale distribution formulation of Eichmann and Marinovich [43].

For example, setting $\mathbf{A}$ to

$$[(\mathrm{d}(\mathbf{C})_\gamma s](x) \;=\; s\Big(m_p(m_{1/p}(x)-\gamma)\Big)\;|x|^{(1-p)/2p}\,\Big|m_{1/p}(x)-\gamma\Big|^{(p-1)/2},$$

[5]The $a^{-1}/2$ factor should be interpreted as the group element such that $(a^{-1}/2)\bullet(a^{-1}/2) = a^{-1}$.

with $m_p(u) = |u|^p \mathrm{sgn}(u)$, results in a Wigner-like time-chirp distribution having marginals of time $|s(t)|^2$ and chirp $|(\Theta_{\mathbf{C}} s)(c)|^2$ from (9).

Unitarily equivalent distributions: The powerful unitary equivalence concept applies also to joint distributions. Given a joint distribution $\mathbf{P}_{a,b}$ matched to two operators $\mathbf{A}$ and $\mathbf{B}$, we can easily obtain a distribution $\mathbf{P}_{\widetilde{a},\widetilde{b}}$ matched to the unitarily equivalent operators $\widetilde{\mathbf{A}} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U}$ and $\widetilde{\mathbf{B}} = \mathbf{U}^{-1}\mathbf{B}\mathbf{U}$ by simply preprocessing the signal by the unitary transformation $\mathbf{U}$ [16,17]

$$(\mathbf{P}_{\widetilde{a},\widetilde{b}}\, s)(\widetilde{a},\widetilde{b}) \;=\; (\mathbf{P}_{a,b}\mathbf{U}s)(\widetilde{a},\widetilde{b}).$$

The analysis of unitarily equivalent distributions is straightforward [16, 17]. Hlawatsch, Bölcskei, and Twaroch have also studied a closely related set of classes in great detail [10–13].

Distributions generated by operator pairs that are unitarily equivalent to the time and frequency operators share the total marginal symmetry of Cohen's class time-frequency distributions. The hyperbolic class of Papandreou, Hlawatsch, and Boudreaux-Bartels [7] provides a prime example, with $\mathbf{U} = \mathbf{U}_{\mathrm{hyp}}$ from (3).

IV-E Relationship to Cohen's Method

Cohen pioneered the construction of joint distributions of arbitrary variables [1,19,36,44]. In this section, we show that the steps of Section IV-B and Cohen's method yield equivalent distributions, assuming we are interested in CED marginals over the real numbers. (This fact was originally proved by Sayeed and Jones in [24,25], so we will merely sketch the result. See [1] for more information on Cohen's method.)

Cohen's method: A physical quantity a has a Hermitian (self-adjoint) representation $\mathcal{A}$ in addition to a unitary representation $\mathbf{A}$ [1]. Projection onto the eigenfunctions of $\mathcal{A}$ measures the a content in a signal. In fact, the square $|(\mathbb{F}_{\mathcal{A}} s)(a)|^2$ coincides with the $\mathbf{A}$–CED.

Exponentiating a Hermitian operator yields a unitary operator

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

sharing the same eigenfunctions. For example, the Hermitian representations for time, $(\mathcal{T}g)(x) \equiv x\,g(x)$, and frequency, $(\mathcal{F}g)(x) \equiv \frac{1}{j2\pi}\,\dot{g}(x)$, exponentiate to the unitary representations for frequency, $\mathbf{F}_f$ and time $\mathbf{T}_{-t}$, respectively. (Note carefully the reversed order of the unitary operators.)

Cohen constructs joint a–b distributions by plugging the corresponding Hermitian operator representations $\mathcal{A}$ and $\mathcal{B}$ into a characteristic function formula [1,19,36,44]

$$(\mathbf{P}^{\text{cohen}}s)(a,b) \;=\; \iint \left\langle e^{j2\pi(\alpha\mathcal{A}+\beta\mathcal{B})}s, s\right\rangle\, e^{-j2\pi\alpha a}\, e^{-j2\pi\beta b}\, d\alpha\, d\beta. \tag{20}$$

The exponentiation $e^{j2\pi(\alpha\mathcal{A}+\beta\mathcal{B})}$ can be performed many different ways — as long as $e^{j2\pi(\alpha\mathcal{A}+\beta\mathcal{B})}|_{\beta=0} = e^{j2\pi\alpha\mathcal{A}}$ and $e^{j2\pi(\alpha\mathcal{A}+\beta\mathcal{B})}|_{\alpha=0} = e^{j2\pi\beta\mathcal{B}}$, distributions of this form will correctly marginalize to the **A**–CED and **B**–CED. To generate a class of distributions corresponding to many possible exponentiations, Cohen fixes one and introduced a kernel function $\phi(\alpha,\beta)$ into (20) to take care of the others.

We now will sketch the equivalence between Cohen's formulation (20) and the formulation of (17) for the case of CED marginals.

Equivalence on $(\mathrm{I\!R},+)$: When the quantities of interest a and b live in the group $(\mathrm{I\!R},+)$, the equivalence between (17) and (20) is direct and immediate. The key in this case lies in the fact that exponentiating a Hermitian operator representation $\mathcal{A}$ yields the dual unitary operator $\mathrm{d}(\mathbf{A})$ (by Stone's theorem, see [27, p. 614] and [24,25]). Therefore, we have

$$\mathrm{d}(\mathbf{A})_\alpha = e^{j2\pi\alpha\mathcal{A}}, \qquad \lambda^{G^*}_{\alpha,a} = e^{j2\pi\alpha a}, \qquad d\mu_{G^*}(\alpha) = d\alpha,$$

and simple substitution equates (17) to (20).

Equivalence on more groups related to $(\mathrm{I\!R},+)$: Matters are hardly more complicated when the quantities of interest lie in arbitrary unbounded groups over the reals. Key in this case is that all groups of this kind are *equivalent*, meaning there exists a one-to-one warping function $w : G^* \to (\mathbb{R}, +)$ such that (we illustrate with the dual group)

$$\alpha_1 \circ \alpha_2 \;=\; w\Big[w^{-1}(\alpha_1') + w^{-1}(\alpha_2')\Big], \qquad \alpha_1, \alpha_2 \in G^*, \;\; \alpha_1', \alpha_2' \in (\mathbb{R}, +).$$

with $\circ$ the group operation on G^*. For example, $w(x) = \log x$ takes $(\mathbb{R}_+, *)$ to $(\mathbb{R}, +)$, while $w(\alpha) = 1/\alpha$ takes the group $(\mathbb{R}\backslash\{0\}, \diamond)$, with $\alpha_1 \diamond \alpha_2 = 1/(1/\alpha_1 + 1/\alpha_2)$, to $(\mathbb{R}, +)$.

The warping w remaps the operators, eigenfunctions, and measures in a predictable way. By Stone's theorem again and the equivalence of the groups, we have the Hermitian-unitary operator equivalence

$$\mathrm{d}(\mathbf{A})_\alpha \;=\; e^{j2\pi w(\alpha)\mathcal{A}}, \qquad \alpha \in G^*$$

or

$$\mathrm{d}(\mathbf{A})_{w^{-1}(\alpha')} \;=\; e^{j2\pi\alpha'\mathcal{A}}, \qquad \alpha' \in (\mathbb{R}, +).$$

These operators still share common eigenfunctions, but with a different indexing scheme, since

$$\mathrm{d}(\mathbf{A})_{w^{-1}(\alpha')}\, \mathbf{u}^{\mathrm{d}(\mathbf{A})}_{w^{-1}(\alpha')} \;=\; \lambda^{G^*}_{w^{-1}(\alpha'),a}\, \mathbf{u}^{\mathrm{d}(\mathbf{A})}_{w^{-1}(\alpha')}.$$

Under the warping, the Haar measure changes to

$$d\mu^{G^*}(w^{-1}(\alpha')) \;=\; d\mu^{(\mathbb{R},+)}(\alpha) \;=\; d\alpha.$$

With these three changes in place, the general formula (17) would yield $(\mathbf{P}^{\mathrm{cohen}}s)(a, b)$, but with a, b with not necessarily in $(\mathbb{R}, +)$. In order to assure this (and link our result to Cohen's), we use the warping transformation v that takes $a \in G$ to $a' \in (\mathbb{R}, +)$ (plus a similar transformation for b). Under this warping, the characters of G become additive and we have

$$\lambda^{G^*}_{w^{-1}(\alpha'),v^{-1}(a')} \;=\; e^{j2\pi\alpha' a'}, \qquad \alpha', a' \in (\mathbb{R}, +).$$

Thus, for real-valued variables, we can interpret (17) for CED marginals as equivalent to a warped version of Cohen's construction (20). (See [24, 25] for more details.)

The utility of warping extends beyond demonstrating this equivalence. Consider the case in Section IV-B where **A** and **B** represent different, but equivalent groups. By warping one of the variables, we can base (17) in a common group and signal space. A simple dewarping procedure on the resulting distribution can then restore the variable to its natural state.

Discretized versions of (20) can be implemented by an FFT-based algorithm. Discretization of (17) will lead to number theoretic transforms [45], which also have fast implementations.

V Conclusions

Viewing invariant and covariant signal energy densities from a group theory perspective has proven illuminating. These simple concepts are central for studying both single- and multi-operator energy distributions. Our approach to joint distributions — simply combining two one-dimensional energy densities to form one joint density — is simple, but effective. Somewhat surprisingly, this method is equivalent to Cohen's general prescription.

In constructing joint densities for multiple operators, we have ignored the important fact that in certain special cases, operator pairs can be representations of (noncommutative) higher-dimensional groups. For general constructions that utilize the resulting noncommutative group theory, see [5, 6, 41].

Acknowledgements

Thanks to Jacqueline and Pierre Bertrand, Faye Boudreaux-Bartels, Paulo Gonçalvés, Franz Hlawatsch, Akbar Sayeed, and Ram Shenoy for sharing their thoughts with me on joint distribution theory. Special thanks to Leon Cohen for many stimulating conversations on the subject of "scale."

References

[1] L. Cohen, *Time-Frequency Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[2] P. Flandrin, *Temps-Fréquence.* Paris: Hermes, 1993.

[3] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency representations," *IEEE Signal Processing Mag.*, vol. 9, pp. 21–67, Apr. 1992.

[4] 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.

[5] J. Bertrand and P. Bertrand, "A class of affine Wigner functions with extended covariance properties," *J. Math. Phys.*, vol. 33, pp. 2515–2527, July 1992.

[6] 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.

[7] 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.

[8] 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), pp. 1265–1270, 1993.

[9] A. Papandreou-Suppappola and G. F. Boudreaux-Bartels, "The exponential class and generalized time-shift covariant quadratic time-frequency representations," in *IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, (Paris, France), pp. 429–432, June 1996.

[10] F. Hlawatsch and H. Bölcskei, "Displacement-covariant time-frequency distributions based on conjugate operators," *IEEE Signal Processing Letters*, vol. 3, pp. 44–46, Feb. 1996.

[11] F. Hlawatsch, T. Twaroch, and H. Bölcskei, "Displacement-covariant time-frequency analysis," *IEEE Trans. Signal Processing*, 1996. Preprint.

[12] F. Hlawatsch, T. Twaroch, and H. Bölcskei, "Wigner-type a-b and time-frequency analysis based on conjugate operators," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '96*, vol. 3, pp. 1395–1398, 1996.

[13] F. Hlawatsch and T. Twaroch, "Covariant (α, β), time-frequency, and (a, b) representations," in *IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, (Paris, France), pp. 437–440, June 1996.

[14] A. M. Sayeed and D. L. Jones, "A canonical covariance-based method for generalized joint signal representations," *IEEE Signal Processing Letters*, vol. 3, pp. 121–123, April 1996.

[15] A. M. Sayeed and D. L. Jones, "A simple covariance-based characterization of joint signal representations of arbitrary variables," *IEEE Int. Symp. Time-Frequency and Time-Scale Analysis*, pp. 433–436, June 1996.

[16] R. G. Baraniuk and D. L. Jones, "Unitary equivalence: A new twist on signal processing," *IEEE Trans. Signal Processing*, vol. 43, pp. 2269–2282, Oct. 1995.

[17] R. G. Baraniuk and L. Cohen, "On joint distributions of arbitrary variables," *IEEE Signal Processing Letters*, vol. 2, pp. 10–12, Jan. 1995.

[18] R. G. Baraniuk, "Covariant time-frequency representations through unitary equivalence," *IEEE Signal Processing Letters*, vol. 3, pp. 79–81, Mar. 1996.

[19] M. O. Scully and L. Cohen, "Quasi-probability distributions for arbitrary operators," in *Proceedings of Nonlinear Dynamics and Chaos, Geometric Quantization, and Wigner Functions* (Y. S. Kim and W. W. Zachary, eds.), (Berlin), pp. 253–263, Springer-Verlag, 1987.

[20] L. Cohen, "Generalized phase space distribution functions," *J. Math. Phys.*, vol. 7, pp. 781–786, 1966.

[21] L. Cohen, "A general approach to obtaining joint representations in signal analysis — Part I: Characteristic function operator method," *IEEE Trans. Signal Processing*, vol. 44, pp. 1080–1090, May 1996.

[22] L. Cohen, "A general approach to obtaining joint representations in signal analysis — Part II: General class, mean and local values, and bandwidth," *IEEE Trans. Signal Processing*, vol. 44, pp. 1091–1099, May 1996.

[23] R. G. Baraniuk, "Beyond time-frequency analysis: Energy densities in one and many dimensions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '94*, vol. III, pp. 357–360, 1994.

[24] A. M. Sayeed and D. L. Jones, "On the equivalence of generalized joint signal representations," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '95*, vol. 3, pp. 1533–1536, 1995.

[25] A. M. Sayeed and D. L. Jones, "Equivalence of generalized joint signal representations of arbitrary variables," *IEEE Trans. Signal Processing*, vol. 44, Dec. 1996.

[26] A. M. Sayeed and D. L. Jones, "Integral transforms covariant to unitary operators and their implications for joint signal representations," *IEEE Trans. Signal Processing*, vol. 44, pp. 1365–1377, June 1996.

[27] N. Dunford and J. T. Schwartz, *Linear Operators, Part I: General Theory.* New York: Interscience, 1958.

[28] F. Riesz and B. Sg.-Nagy, *Functional Analysis.* New York: Frederick Ungar Co., 1955.

[29] G. B. Folland, *Harmonic Analysis in Phase Space.* Princeton, NJ: Princeton University Press, 1989.

[30] W. Rudin, *Functional Analysis.* New York: McGraw-Hill, second ed., 1991.

[31] J. Bertrand, P. Bertrand, and J. P. Ovarlez, "The Mellin transform," in *Electrical Engineering Handbook* (A. D. Poularikas, ed.), ch. 12, Boca Raton, FL: CRC Press, 1994.

[32] W. Rudin, *Fourier Analysis on Groups.* New York: McGraw-Hill, 1962.

[33] A. A. Kirillov, *Elements of the Theory of Representations.* New York: Springer-Verlag, 1976.

[34] R. G. Baraniuk, "A signal transform covariant to scale changes," *IEE Electronics Letters*, vol. 29, pp. 1675–1676, Sept. 17, 1993.

[35] A. M. Sayeed, "Equivalence of the operator and kernel methods for joint distributions of arbitrary variables," *IEEE Trans. Signal Processing*, vol. 44, Dec. 1996.

[36] L. Cohen, "The scale representation," *IEEE Trans. Signal Processing*, vol. 41, pp. 3275–3292, Dec. 1993.

[37] R. G. Baraniuk, "A limitation of the kernel method for joint distributions of arbitrary variables," *IEEE Signal Processing Letters*, vol. 3, pp. 51–53, Feb. 1996.

[38] R. G. Baraniuk, "Marginals vs. covariance in joint distribution theory," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '95*, vol. 2, pp. 1021–1024, 1995.

[39] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, "The Wigner distribution — A tool for time-frequency signal analysis — Part II: Discrete-time signals," *Philips J. Res.*, vol. 35, no. 4/5, pp. 277–300, 1980.

[40] J. C. O'Neill and W. J. Williams, "A new direction for discrete time-frequency distributions," *IEEE Trans. Signal Processing*, 1996. Submitted.

[41] M. S. Richman, T. W. Parks, and R. G. Shenoy, "Discrete-time, discrete-frequency time-frequency representations," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '95*, vol. 2, pp. 1029–1032, 1995.

[42] R. A. Altes, "Wideband, proportional-bandwidth Wigner-Ville analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1005–1012, June 1990.

[43] G. Eichmann and N. M. Marinovich, "Scale-invariant Wigner distribution," in *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 519, pp. 18–25, 1984.

[44] L. Cohen, "What is a multicomponent signal?," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing — ICASSP '92*, vol. V, pp. 113–116, 1992.

[45] R. Shenoy. Personal communication.