On Joint Distributions for Arbitrary Variables

Richard G. Baraniuk and Leon Cohen

Abstract—**There has been considerable interest in the problem of joint representations for variables other than time and frequency. In this letter, we compare the methods of Cohen and of Baraniuk and Jones and show their equivalence for variables that have the same commutator as time and frequency. In addition, we report the following very general result: All pairs of variables connected by a unitary transformation have joint distributions that are functionally equivalent.**

I. Introduction and Conclusion

MARGENAU and Hill [1] were perhaps the first to attempt to extend the idea of the Wigner distribution to physical variables other than time and frequency (position and momentum in quantum mechanics), and since that time, a number of special cases have been considered. Recently Cohen [2]–[4] and Baraniuk and Jones [5], [6] have presented methods for obtaining joint representations for arbitrary physical quantities; both are based on associating variables with operators. In this note, we compare these two approaches and demonstrate their equivalence when the two new variables correspond to operators that are related to the time and frequency operators by a unitary transformation. In addition, we generalize this result by showing that all joint distributions whose variables are related by a unitary transformation have the same functional form. The importance of this result is that it shows how pairs of variables can be grouped into classes whose joint distributions are functionally equivalent and, therefore, share equivalent properties. We first briefly describe the two approaches.

Cohen Method: The procedure of Cohen [2]–[4], which is based on the methods of Cohen [7] and Scully and Cohen [8], generalizes the standard relation between a characteristic function and its corresponding density. Suppose we have two variables a and b associated with the Hermitian operators $\mathcal{A}$ and $\mathcal{B}$, respectively. We seek a joint distribution $P(a,b)$ that indicates the energy content of signals in terms of both a and b simultaneously.

The characteristic function for a and b is given by

$$M(\alpha,\beta) = \langle e^{j\alpha a+j\beta b}\rangle = \iint e^{j\alpha a+j\beta b} P(a,b)\, da\, db. \quad (1)$$

Hence, the distribution is obtained from $M(\alpha,\beta)$ by Fourier inversion

$$P(a,b) = \frac{1}{4\pi^2}\iint M(\alpha,\beta), e^{-j\alpha a-j\beta b}\, d\alpha\, d\beta. \quad (2)$$

The characteristic function is an average (the average of $e^{j\alpha a+j\beta b}$) and can be computed directly from the signal by averaging the characteristic function operator $\mathcal{M}(\alpha,\beta)$ by way of

$$M(\alpha,\beta) = \langle\mathcal{M}(\alpha,\beta)\rangle = \int s^*(t)\mathcal{M}(\alpha,\beta)s(t)\, dt. \quad (3)$$

These equations can be combined, as in the time frequency case [7], to yield the triple integral form

$$P(a,b) = \frac{1}{4\pi^2}\iiint s^*(t)\mathcal{M}(\alpha,\beta)s(t)e^{-j\alpha a-j\beta b}\, dt\, d\alpha\, d\beta. \quad (4)$$

The unitary characteristic function operator $\mathcal{M}(\alpha,\beta)$ is formed by combining exponentiated versions of the Hermitian operators $\mathcal{A}$ and $\mathcal{B}$ associated with the variables a and b. As in the time-frequency case, there are many possibilities for orderings of $e^{j\alpha\mathcal{A}}$ and $e^{j\beta\mathcal{B}}$ in $\mathcal{M}(\alpha,\beta)$. The way to handle the possible orderings is to choose one particular ordering and introduce a kernel function $\phi(\alpha,\beta)$ to generate the remaining ones [2]–[4], [7]. This leads to a general class of distributions for the variables a and b. Here, we choose the symmetrical ordering

$$\mathcal{M}(\alpha,\beta) = \phi(\alpha,\beta)e^{j\alpha\mathcal{A}+j\beta\mathcal{B}} \quad (5)$$

giving

$$M(\alpha,\beta) = \phi(\alpha,\beta)\langle e^{j\alpha\mathcal{A}+j\beta\mathcal{B}}\rangle \quad (6)$$

for the generalized characteristic function. The general class of distributions for the variables a and b is thus given by

$$P(a,b) = \frac{1}{4\pi^2}\iint \phi(\alpha,\beta)\langle e^{j\alpha\mathcal{A}+j\beta\mathcal{B}}\rangle e^{-j\alpha a-j\beta b}\, d\alpha\, d\beta \quad (7)$$

$$= \frac{1}{4\pi^2}\iiint \phi(\alpha,\beta)s^*(t)\, e^{j\alpha\mathcal{A}+j\beta\mathcal{B}} \times s(t)e^{-j\alpha a-j\beta b}\, dt\, d\alpha\, d\beta. \quad (8)$$

The marginal distributions

$$\int P(a,b)db = |F(a)|^2, \quad \int P(a,b)\, da = |F(b)|^2 \quad (9)$$

correspond to the true densities of a and b, provided the kernel satisfies $\phi(0,\beta) = \phi(\alpha,0) = 1$. These densities are obtained

Manuscript received June 13, 1994; approved September 27, 1994. This work was supported by the Texas Advanced Technology Program, the NSA HBCU/MI program, and the PSC-CUNY Research Award Program. The associate editor coordinating the review of this letter and approving it for publication was Prof. G. F. Boudreaux-Bartels.

R. Baraniuk is with the Department of Electrical and Computer Engineering, Rice University, Houston, TX 77005 USA.

L. Cohen is with Hunter College and Graduate Center of CUNY, New York, NY 10021 USA.

IEEE Log Number 9408346.

1070–9908/95$04.00

by projecting the signal onto the eigenfunctions $u(a,t)$ and $v(b,t)$ of the Hermitian operators $\mathcal{A}$ and $\mathcal{B}$

$$F(a) = \int s(t)u^*(a,t)\,dt, \quad F(b) = \int s(t)v^*(b,t)\,dt. \quad (10)$$

We point out that densities can be computed directly from the signal, thereby circumventing the eigenfunctions, using

$$|F(a)|^2 = \frac{1}{2\pi}\int \mathcal{M}_a(\theta), e^{-j\theta a}d\theta,$$
$$\mathcal{M}_a(\theta) = \int s^*(t)e^{j\theta\mathcal{A}}s(t)\,dt \quad (11)$$

for a and a similar result for b. The equivalence of the two methods for obtaining densities was shown by Margenau and Cohen in [9]. The structural similarity between (11), which apply to one variable, and (2) and (3), which apply to two variables, should be noted, because these results generalize to an arbitrary number of variables [2]–[4].

Baraniuk and Jones Method: As derived in [5] and [6], this approach begins with the generalized time-frequency representation [7]

$$C_s(t,\omega) = \frac{1}{4\pi^2}\iiint s^*(u)e^{j\alpha\mathcal{T}+j\beta\mathcal{W}} \times s(u)\phi(\alpha,\beta)e^{-j\alpha t-j\beta\omega}du\,d\alpha\,d\beta \quad (12)$$

which is obtained from Cohen's general method in (8) by employing the Hermitian time and frequency operators $\mathcal{T}$ and $\mathcal{W}$. The actions of $\mathcal{T}$ and $\mathcal{W}$ in the time domain are $\mathcal{T}s(t) = ts(t)$ and $\mathcal{W}s(t) = -j\frac{d}{dt}s(t)$. The time and frequency marginals

$$\int C_s(t,\omega)d\omega = |s(t)|^2, \quad \int C_s(t,\omega)\,dt = |S(\omega)|^2 \quad (13)$$

where $S(\omega)$ denotes the Fourier transform of $s(t)$, result from projecting the signal onto the eigenfunctions of $\mathcal{T}$ and $\mathcal{W}$, respectively.

Substituting a unitarily transformed signal $\mathcal{U}s(t)$ for $s(t)$ in (12) and replacing t, ω by a', b' yield a new distribution

$$C_{\mathcal{U}s}(a',b') = \frac{1}{4\pi^2}\iiint \mathcal{U}s^*(u)e^{j\alpha\mathcal{T}+j\beta\mathcal{W}} \times \mathcal{U}s(u)\phi(\alpha,\beta)e^{-j\alpha a'-j\beta b'}du\,d\alpha\,d\beta \quad (14)$$

having

$$\int C_{\mathcal{U}s}(a',b')db' = |\mathcal{U}s(a')|^2,$$
$$\int C_{\mathcal{U}s}(a',b')da' = |\mathcal{F}\mathcal{U}s(b')|^2 \quad (15)$$

for marginals, where $\mathcal{F}$ is teh Fouirier transfom operator, and $\mathcal{F}\mathcal{U}s(b')$ is the Fourier transform of the transformed signal. Note that in this procedure, we do not have complete control over both variables a' and b' because the unitary operator $\mathcal{U}$ provides only a single degree of freedom.

II. Comparison

We now show that the Baraniuk and Jones procedure can be used to obtain the same joint distributions as the Cohen procedure when the desired variables a and b are related to the time and frequency operators $\mathcal{T}$ and $\mathcal{W}$ by a unitary transformation $\mathcal{U}$; that is, when

$$\mathcal{A} = \mathcal{U}^\dagger\mathcal{T}\mathcal{U}, \quad \mathcal{B} = \mathcal{U}^\dagger\mathcal{W}\mathcal{U}, \quad (16)$$

where † denotes the adjoint. For a unitary transformation, recall that $\mathcal{U}^\dagger = \mathcal{U}^{-1}$. Before we start, it is important to observe that the commutator of $\mathcal{A}$ and $\mathcal{B}$ must equal j, which is the commutator of $\mathcal{T}$ and $\mathcal{W}$. This follows from

$$[\mathcal{A},\mathcal{B}] = \mathcal{U}^{-1}[\mathcal{T},\mathcal{W}]\mathcal{U} = \mathcal{U}^{-1}j\mathcal{U} = j. \quad (17)$$

Hence, since the commutator $[\mathcal{A},\mathcal{B}]$ commutes with both $\mathcal{A}$ and $\mathcal{B}$, we have that [2], [4]

$$e^{j\alpha\mathcal{A}+j\beta\mathcal{B}} = e^{j\alpha\beta/2}e^{j\beta\mathcal{A}}e^{j\beta\mathcal{B}}. \quad (18)$$

In addition, note that, in general, we have

$$e^{j\alpha\mathcal{A}} = \mathcal{U}^{-1}e^{j\alpha\mathcal{T}}\mathcal{U}, \quad e^{j\beta\mathcal{B}} = \mathcal{U}^{-1}e^{j\beta\mathcal{W}}\mathcal{U}. \quad (19)$$

We now start with the Cohen procedure and calculate the characteristic function for a and b from (6)

$$M(\alpha,\beta) = \phi(\alpha,\beta)\int s^*(t)e^{j\alpha\mathcal{A}+j\beta\mathcal{B}}s(t)dt \quad (20)$$
$$= \phi(\alpha,\beta)\int s^*(t)e^{j\alpha\beta/2}e^{j\alpha\mathcal{A}}e^{j\beta\mathcal{B}}s(t)dt \quad (21)$$
$$= \phi(\alpha,\beta)\int s^*(t)e^{j\alpha\beta/2}(\mathcal{U}^{-1}e^{j\alpha\mathcal{T}}\mathcal{U}) \times (\mathcal{U}^{-1}e^{j\beta\mathcal{W}}\mathcal{U})s(t)dt \quad (22)$$
$$= \phi(\alpha,\beta)\int \mathcal{U}s^*(t)e^{j\alpha\mathcal{T}+j\beta\mathcal{W}}\mathcal{U}s(t)dt. \quad (23)$$

The final step follows from (18) and the definition of the adjoint. Substituting this result into (2) to obtain the distribution, we find

$$P(a,b) = \frac{1}{4\pi^2}\iiint \mathcal{U}s^*(t)e^{j\alpha\mathcal{T}+j\beta\mathcal{W}}\mathcal{U}s(t) \times \phi(\alpha,\beta)e^{-j\alpha a-j\beta b}d\alpha\,d\beta\,dt \quad (24)$$

which is identical to (14) except for the notation of the variables of integration. This demonstrates the equivalence of the two procedures when (16) holds.

III. Example

As an example, consider the two operators $\mathcal{A} = \log t$ and $\mathcal{B} = \frac{1}{2}(\mathcal{T}\mathcal{W} + \mathcal{W}\mathcal{T})$ representing log time and scale, respectively. While joint distributions of $\mathcal{A}$ and $\mathcal{B}$ could be derived using the general Cohen approach, these operators satisfy $[\mathcal{A},\mathcal{B}] = j$ [2]–[4] and are therefore a valid operator pair for the simpler Baraniuk and Jones procedure.[1] In particular,

[1] One of us (L.C.) would like to express his appreciation to Drs. A. Papandreou and G. F. Boudreaux-Bartels for an insightful conversation regarding this example and for pointing out that the joint distributions for these variables may be obtained by the Cohen procedure. Such distributions have been considered in [2]–[6], [10]–[14]. We also point out that log time and log frequency operators have been studied by a number of authors [2]–[6], [10].

utilizing the unitary transformation

$$\mathcal{U}s(t) = e^{t/2}s(e^t), \quad \mathcal{U}^{-1}s(t) = \frac{s(\log t)}{\sqrt{t}}, \quad t > 0 \tag{25}$$

in (16) yields the operators $\mathcal{A} = \log t$ and $\mathcal{B} = \frac{1}{2}(\mathcal{TW} + \mathcal{WT})$, respectively. The marginals of the distribution $C_{\mathcal{U}s}(a,b)$ from (14), computed by employing this $\mathcal{U}$ in (15), are given by

$$\int P(a,b)db = |\mathcal{U}s(a)|^2 = e^a|s(e^a)|^2 \tag{26}$$

and

$$\int P(a,b)da = \left|\int \mathcal{U}s(t)e^{-jbt}dt\right|^2 = \left|\int s(u)e^{-jb\log u}\frac{du}{\sqrt{u}}\right|^2. \tag{27}$$

Equation (27) corresponds to the marginal for scale given in [2]–[4] and [11]–[14].

IV. Generalization

Suppose now that we have two totally arbitrary quantities a and b represented by two operators $\mathcal{A}$ and $\mathcal{B}$ and that we have obtained their joint distribution $P_{ab}(a,b)$ by way of the Cohen procedure. Suppose, also, that we have two new quantities a' and b' represented by the operators $\mathcal{A}'$ and $\mathcal{B}'$ and wish to obtain their joint distribution $P_{a'b'}(a',b')$. This can be simply done if the two new operators are related to the old operators via a unitary transformation of the form

$$\mathcal{A}' = \mathcal{U}^{\dagger}\mathcal{A}\mathcal{U}, \quad \mathcal{B}' = \mathcal{U}^{\dagger}\mathcal{B}\mathcal{U}. \tag{28}$$

We note that in such a case, the commutator relations are left invariant; that is, if $[\mathcal{A},\mathcal{B}] = \mathcal{C}$, then $[\mathcal{A}',\mathcal{B}'] = \mathcal{C}'$ with $\mathcal{C}' = \mathcal{U}^{\dagger}\mathcal{C}\mathcal{U}$.

To obtain the distribution $P_{a'b'}(a',b')$, we need merely transform the signal by the unitary transformation $\mathcal{U}$ before computing $P_{ab}(a,b)$. To see that this is the case, first recall that the distribution $P_{a'b'}(a',b')$ is given by

$$P_{a'b'}(a',b') = \frac{1}{4\pi^2}\iiint s^*(t)\mathcal{M}_{a'b'}(\alpha,\beta) \times s(t)e^{-j\alpha a' - j\beta b'}dt\,d\alpha\,d\beta. \tag{29}$$

However, the characteristic function operators $\mathcal{M}_{a'b'}$ and $\mathcal{M}_{ab}$ are related by

$$\mathcal{M}_{a'b'}(\alpha,\beta) = \mathcal{U}^{-1}\mathcal{M}_{ab}(\alpha,\beta)\,\mathcal{U}. \tag{30}$$

Substituting (30) into (29), we obtain

$$P_{a'b'}(a',b') = \frac{1}{4\pi^2}\iiint s^*(t)(\mathcal{U}^{-1}\mathcal{M}_{ab}(\alpha,\beta)\,\mathcal{U}) \times s(t)e^{-j\alpha a' - j\beta b'}dt\,d\alpha\,d\beta \tag{31}$$

$$= \frac{1}{4\pi^2}\iiint \mathcal{U}s^*(t)\mathcal{M}_{ab}(\alpha,\beta)\mathcal{U} \times s(t)e^{-j\alpha a' - j\beta b'}\,dt\,d\alpha\,d\beta \tag{32}$$

which is precisely the distribution $P_{ab}(a,b)$ computed using the preprocessed signal $\mathcal{U}s(t)$.

References

[1] H. Margenau and R. N. Hill, "Correlation between measurements in quantum theory," *Prog. Theoret. Phys.*, vol. 26, pp. 722–738, 1961.

[2] L. Cohen, "A general approach for obtaining joint representations in signal analysis and an application to scale," *Proc. SPIE Advanced Signal-Processing Algorithms, Architectures Implementation II* (F. T. Luk, Ed.). vol. 1566, pp. 109–133, 1991.

[3] ——, "Scale and inverse frequency representations," in *The Role of Wavelets in Signal Processing* (B. W. Sutter, M. E. Oxley and G. T. Warhola, Eds.). AFIT, Wright-Patterson Air Force Base, OH, 1992, pp. 118–128.

[4] ——, "A general approach for obtaining joint representations in signal analysis," to be published in *Proc. IEEE.*

[5] 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*, pp. 320–323, vol. III.

[6] ——, "Unitary equivalence: A new twist on signal processing," to be published in *IEEE Trans. Signal Processing*, 1993.

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

[8] M. Scully and L. Cohen, "Quasi-probability distributions for arbitrary operators," in *The Physics of Phase Space* (Y. S. Kim and W. W. Zachary, Eds.). New York: Springer Verlag, 1987.

[9] H. Margenau and L. Cohen, "Probabilities in quantum mechanics," in *Quantum Theory and Reality* (M Bunge, Ed.). New York: Springer-Verlag, 1967; L. Cohen, "Joint probability distributions in quantum mechanics," in *Contemporary Research in the Foundations of Quantum Theory* (C. A. Hooker and D. Reidel, Eds.), 1973.

[10] A. Papandreaou, F. Hlawatsch and G. F. Boudreaux-Bartels, "The hyperbolic class of quadratic time-frequency distributions," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 41, pp. 3425–3444, 1993.

[11] G. Eichmann and N. M. Marinovic, "Scale-invariant Wigner distribution and ambiguity functions," *Proc. Int. Soc. Opt. Eng. SPIE* vol. 519, pp. 18–24, 1984.

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

[13] 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.

[14] J. Bertrand and P. Bertrand, "Affine time-frequency distributions," in: *Time-Frequency Signal Analysis—Methods and Applications* (B. Boashash, Ed.). New York: Longman-Cheshire, 1992.