Geometry of Affine Time–Frequency Distributions

Patrick Flandrin and Paulo Gonçalvès[1]

Ecole Normale Supérieure de Lyon, Laboratoire de Physique, URA 1325 CNRS, 46 allée d'Italie 69364 Lyon Cedex 07, France

Communicated by Stephane Mallat

Bertrands' bilinear affine time–frequency distributions are considered from the point of view of their geometry in the time–frequency plane. General construction rules are established for interference terms, with further interpretations in terms of localization properties, generalized means and symmetries. In the case of frequency modulated signals, it is shown how the pointwise application of these rules can be refined by the study of a critical manifold and stationary phase-type approximations. Theoretical results are supported by both analytical and numerical examples. © 1996 Academic Press, Inc.

Contents.

1. INTRODUCTION

Recently, a number of new classes of joint energy distributions (affine class [6, 35], hyperbolic class [31], power classes [25], "warped" classes [3]) have emerged, which supplement the well-known Cohen's class [8, 9] and offer new time–frequency analysis tools [14, 23]. All of the distributions of these classes are *bilinear* in the signal and thus create "interference terms" which, depending on the situation, can be considered as troublesome or as informative. A better knowledge of the creation of those interference terms is therefore highly desirable, for at least three reasons:

1. (direct problem) given a known signal, it allows us to associate to it a well-defined time–frequency signature;

2. (inverse problem) given an unknown signal, it allows us to propose an interpretation for the structure of its components;

3. (improved readability) understanding the way interference terms are created by a distribution is the key for proposing accurate modifications aimed at their reduction.

Thorough studies have already been devoted to the geometry of the Wigner–Ville distribution [4, 13, 14, 16, 22, 24]. It is the purpose of this paper to provide results along the same lines, but in the case of Bertrands' affine distributions, for which very little is known [11, 15, 30]. Note that we will focus here on the class of affine distributions with extended covariance properties, as it is defined in [6], and that we will not consider other types of affine distributions such as the scalogram (squared modulus of the wavelet transform), affine smoothed Wigner–Ville distributions [35] or the Choj-Williams distribution [23]. (Discussions about the geometry of their interference terms can be found, e.g., in [26] or [28].)

The paper is organized as follows. In Section 2, we begin with the Wigner–Ville distribution and recall the main results that are known about its geometry, its localization properties on linear "chirps," and the construction rules of its interference terms. Section 3 is concerned with the general class of Bertrands' affine time–frequency distributions. Basics of affine distributions are recalled, with emphasis on their localization properties on specific nonlinear curves (namely, power-law group delays) in the time–frequency plane. This localization is given a geometric interpretation in terms of interference construction rules, which are expressed in closed form and whose main features are detailed. In particular, it is discussed how the creation of interference terms may obey a simple rule based on an idea of *generalized mean,* thus generalizing the classical "midpoint constructions" (arithmetic mean) encountered in the

[1] E-mail: flandrin@physique. ens-lyon.fr, gpaulo@physique.ens-lyon.fr.

case of the Wigner–Ville distribution and providing us with some insightful understanding of notable symmetries underlying these affine distributions. In the case of modulated signals, the fine structure of interference terms is investigated in further detail in Section 4. There, our approach is based on the study of a critical manifold, whose singularities and their projection onto the time–frequency plane determine the overall structure of the distributions. Using a terminology borrowed from catastrophe theory [33], this leads to a classification of typical behaviors which involves time–frequency regions, fold lines, and cusp points. Local approximations, by means of stationary phase-type methods, are also provided, and the theoretical predictions are compared to numerical experimentations.

2. THE WIGNER–VILLE DISTRIBUTION AS A STARTING POINT

2.1. Definition

Let us first recall some basic facts and results pertaining to the Wigner–Ville distribution and its geometry.[2] By definition, the Wigner–Ville distribution of a signal $x(t)$ is given by

$$W_x(t, f) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-i2\pi f\tau} d\tau \quad (1)$$

or equivalently, in the frequency domain, by

$$W_X(t, f) = \int_{-\infty}^{\infty} X\left(f + \frac{\nu}{2}\right) X^*\left(f - \frac{\nu}{2}\right) e^{i2\pi\nu t} d\nu, \quad (2)$$

where $X(f)$ stands for the Fourier transform of $x(t)$.

2.2. Interference Terms

Because it is bilinear (to be more precise, sesquilinear), the Wigner–Ville distribution of the sum of two signals does not reduce to the sum of the individual distributions. In fact, given two signals $x_1(t)$ and $x_2(t)$, we have

$$W_{x_1+x_2}(t, f) = W_{x_1}(t, f) + W_{x_2}(t, f) + I_{x_1, x_2}(t, f), \quad (3)$$

where $I_{x_1, x_2}(t, f)$ is an extra term, also referred to as a "cross-term" or "interference term," given by

$$I_{x_1, x_2}(t, f) = 2 \text{ Re } \left\{ \int_{-\infty}^{\infty} x_1\left(t + \frac{\tau}{2}\right) x_2^*\left(t - \frac{\tau}{2}\right) e^{-i2\pi f\tau} d\tau \right\}. \quad (4)$$

It is clear that the geometry of the Wigner–Ville distribution depends heavily on the structure of such interference terms. Previous detailed studies [16, 22, 24] have shown that Wigner–Ville interference terms obey rather simple construction rules, which can be derived by considering simplified signal models, based on either "atoms" or "chirp-like" structures.

Let us first suppose that we start with a time-frequency "atom" (or "logon" in Gabor's terminology [17]), i.e., a signal, such as a Gaussian, which is well localized in both time and frequency. Denoting by $x_0(t)$ such a signal and viewing it as a "mother" signal, we can then construct two "children" signals by shifting $x_0(t)$ in both time and frequency

$$x_0(t) \rightarrow \{x_j(t) \equiv x_0(t - t_j)e^{i2\pi f_j t}, j = 1, 2\}$$

and consider the sum $x(t) \equiv x_1(t) + x_2(t)$. Application of the definition (1) to this sum gives

$$W_x(t, f) = W_{x_0}(t - t_1, f - f_1) + W_{x_0}(t - t_2, f - f_2)$$
$$+ J_{x_1, x_2}(t - t_{12}, f - f_{12}), \quad (5)$$

with

$$J_{x_1, x_2}(t, f) = 2W_{x_0}(t, f) \cos(2\pi\{(f_1 - f_2)t - (t_1 - t_2)f\} + \varphi_{12})$$
$$(6)$$

and

$$t_{12} \equiv \frac{t_1 + t_2}{2}, \quad f_{12} \equiv \frac{f_1 + f_2}{2}, \quad \varphi_{12} \equiv 2\pi(f_1 - f_2)t_{12}.$$

The first two terms of the right-hand side of (5) correspond to the individual contributions of $x_1(t)$ and $x_2(t)$ (the "auto-terms"), and their form expresses the fact that the Wigner–Ville distribution is covariant with respect to shifts in time and frequency. The extra term (6) reveals the main features attached to an interference term (see Fig. 1):

1. it is located *midway* between the two interfering components;
2. it is *oscillatory*, with a frequency of oscillations in the time–frequency plane which increases as interfering components move apart;
3. it has a direction of oscillation *orthogonal* to the line joining the interfering components.

2.3. Localization

Although the interference construction was derived for only time-frequency atoms, it can formally be applied pointwise to more general signals [4, 16, 24]. This amounts to
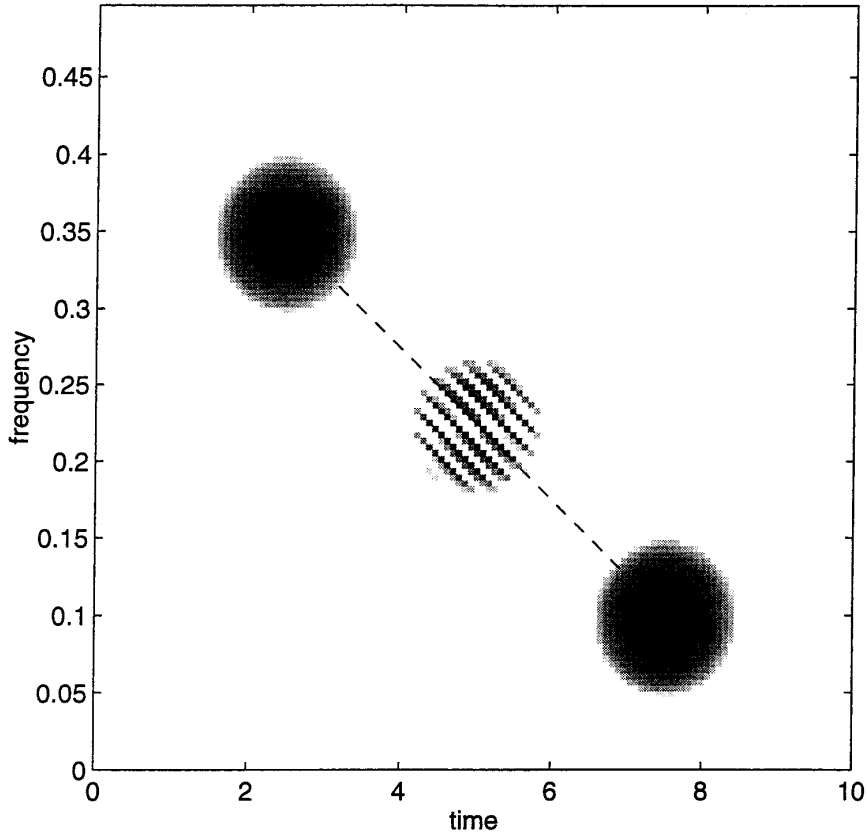
---

[2] Most of the results given in this Section 2 are detailed in [24] and discussed more briefly in [14, 16, 22]. For a general presentation of the Wigner–Ville distribution and its properties, see, e.g. [7, 14, or 23].

FIG. 1. Wigner–Ville distribution: interference term of two logons. The interference term is located midway between the two interfering logons and oscillates in a direction orthogonal to the chord joining them (dashed line). For all the figures, unless otherwise specified, dimensions of the time and frequency axes are respectively in seconds and normalized frequency units.

saying that any two points (t_1, f_1) and (t_2, f_2) located on some "components" (or subparts) of a signal will interfere and create a contribution at a third point (t_i, f_i) whose location in the time-frequency plane is determined by the arithmetic means

$$t_i = \frac{t_1 + t_2}{2} \tag{7}$$

and

$$f_i = \frac{f_1 + f_2}{2}. \tag{8}$$

This rule is particularly relevant in the case of amplitude and frequency modulated (AM-FM) signals of the type

$$X(f) = A_X(f) e^{i\Phi_X(f)}, \tag{9}$$

with

$$t_X(f) = -\frac{1}{2\pi} \dot{\Phi}_X(f) \tag{10}$$

being the corresponding group delay (a dot above a function stands for its usual derivative). The rule outlined above determines the region in time-frequency of all the interference terms generated by the Wigner–Ville distribution, a situation referred to as *inner interference* in the case of monocomponent signals[3] [22]. This region results from the interference of any two points $(t_X(f_1), f_1), (t_X(f_2), f_2)$ belonging to the group delay, and it can be defined as the locus of all the points (t_i, f_i) such that

$$t_i = \frac{t_X(f_1) + t_X(f_2)}{2}$$

and

$$f_i = \frac{f_1 + f_2}{2}.$$

[3] Loosely speaking, a monocomponent signal corresponds to a signal of the form (9) and such that its time–frequency distribution occupies a connected time–frequency domain around its modulation law (group delay $t_X(f)$, as defined in (10), or instantaneous frequency).

If we consider the specific quadratic phase $\Phi_X(f) = -2\pi(t_0 f + (\alpha/2)f^2)$, which corresponds to the linear group delay $t_X(f) = t_0 + \alpha f$, the above construction rules state that all of the interference points $(t_i, f_i) = (t_0 + \alpha(f_1 + f_2)/2, (f_1 + f_2)/2)$ lie exactly on the structure of the signal, i.e., on the support of the group delay in the time–frequency plane. This situation gives rise to a natural interpretation of the so-called localization property of the Wigner–Ville distribution on linear "chirps," according to which

$$X(f) = e^{-i2\pi(t_0 f + (\alpha/2)f^2)} \Rightarrow W_X(t, f) = \delta(t - t_X(f)).$$

Geometrically, localization of the Wigner–Ville distribution on linear "chirps" can be viewed as a by-product of interference in the sense that the locus of all the interference points coincides exactly with the signal structure itself.

A companion interpretation of the Wigner–Ville distribution can be given by reorganizing the terms in the arithmetic means (7) and (8), thus leading to

$$\begin{pmatrix} t_2 \\ f_2 \end{pmatrix} = 2 \begin{pmatrix} t_i \\ f_i \end{pmatrix} - \begin{pmatrix} t_1 \\ f_1 \end{pmatrix}. \qquad (11)$$

This last expression shows that significant contributions to the Wigner–Ville distribution exist at any points which are symmetric with respect to a third point which also contribute significantly to the Wigner–Ville distribution (and which plays therefore the role of a center of symmetry). Hence, the localized structure of the Wigner–Ville distribution on linear group delays ("chirps"), and only on them, can be viewed as resulting from the fact that straight lines are the only curves which are formed by all of their centers of symmetry.

2.4. Fine Structure of Interference Terms

Construction rules based on the arithmetic mean allow us to *localize* interference terms in the Wigner–Ville distribution but, in the case of AM–FM signals, more can be said about the *structure* of those terms.

Given the signal model $X(f) = A_X(f)e^{i\Phi_X(f)}$ and assuming that the envelope A_X is slowly varying compared to the variations of the phase Φ_X, approximations of the Wigner–Ville distribution can be obtained by invoking the principle of *stationary phase* [32, 33]. This principle states that, when integrating a function that oscillates quickly, the main contributions to the integration comes from the vicinity of the points at which the phase derivative is zero. In order to apply this principle to the evaluation of a Wigner–Ville distribution, the first step is to rewrite (2) as

$$W_X(t, f) = \int_{-\infty}^{\infty} L_X(\nu; f) e^{i\Psi_X(\nu; t, f)} d\nu,$$

with

$$L_X(\nu; f) \equiv A_X\left(f + \frac{\nu}{2}\right) A_X\left(f - \frac{\nu}{2}\right)$$

and

$$\Psi_X(\nu; t, f) \equiv \Phi_X\left(f + \frac{\nu}{2}\right) - \Phi_X\left(f - \frac{\nu}{2}\right) + 2\pi t\nu. \quad (12)$$

Assuming first that the phase $\Psi_X(\nu; t, f)$ has only a finite number of non-degenerate *stationary points* $\{\nu_n, n = 1, \ldots N\}$, characterized by the two simultaneous conditions

$$\frac{\partial \Psi_X(\nu_n; t, f)}{\partial \nu} = 0 \qquad (13)$$

and

$$\frac{\partial^2 \Psi_X(\nu_n; t, f)}{\partial \nu^2} \neq 0, \qquad (14)$$

we obtain the approximation

$$W_X(t, f) \approx \sum_{n=1}^{N} L_X(\nu_n; f) \left| \frac{\partial^2 \Psi_X(\nu_n; t, f)}{\partial \nu^2} \right|^{-1/2}$$
$$\times e^{i\Psi_X(\nu_n; t, f) + i(\pi/4)\text{sign}(\partial^2 \Psi_X(\nu_n; t, f)/\partial \nu^2)}. \quad (15)$$

From (13) and (9), the evaluation of the stationary points ν_n implies quite naturally that

$$t = \frac{1}{2}\left(t_X\left(f + \frac{\nu_n}{2}\right) + t_X\left(f - \frac{\nu_n}{2}\right)\right), \qquad (16)$$

which should be compared with (7) and (8). From (16), it appears that the stationary phase condition singles out those points which are midpoints of a chord joining any two points of the group delay [4], thus confirming that the Wigner–Ville distribution geometry is based upon the usual arithmetic mean. In particular, for a quadratic phase, (16) reduces to $t = t_0 + \alpha f$, thus recovering, from a new perspective, the perfect localization of the Wigner–Ville distribution on linear "chirps".

Moreover, it follows from the symmetric structure of (16) that the stationary points ν_n always occur in pairs (because, if ν_n is a solution of (16), then $-\nu_n$ is a solution as well). Hence, when only two points $(t_X(f_1), f_1)$ and $(t_X(f_2), f_2)$ of the group delay $t_X(f)$ interfere at[4] $(t, f) = (\frac{1}{2}(t_x(f_1) + t_x(f_2)), \frac{1}{2}(f_1 + f_2))$, the stationary points are determined by the trivial solution $\nu_n = \pm(f_1 - f_2)$. Assuming arbitrarily that $f_2 > f_1$, the stationary phase approximation of (15) simplifies to

---

[4] For consistency with (7) and (8), cross terms location should be referred to as (t_i, f_i). However, we will retain instead the notation (t, f) for a sake of simplicity, since the evaluation point is by construction a solution of (13).
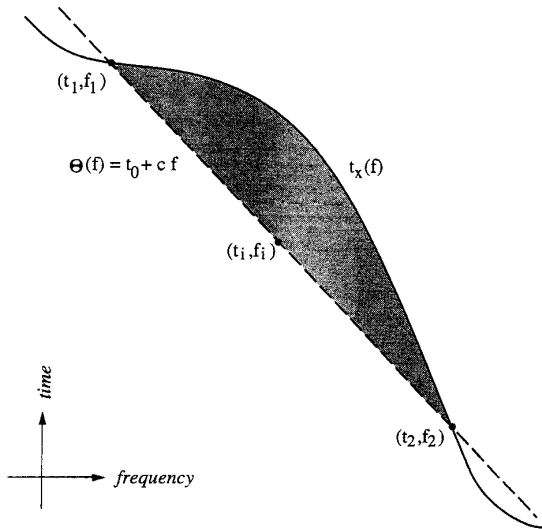
FIG. 2. Wigner–Ville distribution: area rule for the stationary phase approximation. In the case of a nonlinear frequency modulation, and within the stationary phase approximation, the local oscillations of the Wigner–Ville distribution are controlled by the (shaded) area bounded by the signal group delay (solid line) and the chord (dashed line) joining the two interfering points.

$$W_X(t,f) \approx 2 \frac{A_X(f_1)A_X(f_2)}{\sqrt{\frac{\pi}{2}|\dot{t}_X(f_1) - \dot{t}_X(f_2)|}}$$

$$\times \cos\left[2\pi \mathscr{A}_X(t,f) + \frac{\pi}{4}\text{sign}\{\dot{t}_X(f_1) - \dot{t}_X(f_2)\} \right] \quad (17)$$

where

$$\mathscr{A}_X(t,f) = \int_{f_1}^{f_2} t_X(\nu)d\nu - \int_{f_1}^{f_2} \theta(\nu)d\nu$$

corresponds to the area bounded by the signal group delay $t_X(f)$ and the straight line $\theta(f)$ joining $(t_X(f_1), f_1)$ to $(t_X(f_2), f_2)$ (see Fig. 2).

Except for linear chirps, stationary points define *regions* where the Wigner–Ville distribution has nonnegligible values. However, note that the stationary phase approximation is not valid when $\partial^2 \Psi_X(\nu_n; t, f)/\partial \nu^2 = 0$. This corresponds to when $\dot{t}_X(f_1) = \dot{t}_X(f_2)$, i.e., to the situation for which the slopes of the group delay at the interfering points are the same. Second-order singularities define, therefore, a subset of the stationary points set, which consists of time–frequency *lines,* referred to as *fold lines* (a term borrowed from the terminology of catastrophe theory [33]).

This situation is first encountered in the limit of null length chords ($\nu_n \to 0, t_1 \to t_2$ and $f_1 \to f_2$), thus defining the signal structure itself as a locus of second-order singular points. In this case, the stationary phase approximation

diverges and it becomes necessary to make use of refined methods, such as uniform approximations [10], if we want to end up with divergence-free approximations. However, if we are mainly interested in the local behavior of W_X in the vicinity of the curve $t_X(f)$, a less involved technique consists of expanding the frequency phase $\Psi_X(\nu; t, f)$ up to third order with respect to ν. Assuming a constant amplitude $A_X(f)$, we get, instead of (17), the following *transitional* approximation

$$W_X(t,f) \approx \frac{1}{\epsilon(f)}\text{Ai}\left(\frac{1}{\epsilon(f)}(t_X(f) - t) \right), \quad (18)$$

with

$$\epsilon(f) = \left(\frac{\ddot{t}_X(f)}{32\pi^2} \right)^{1/3}$$

and where $\text{Ai}(\cdot)$ stands for the Airy function, defined as [33]

$$\text{Ai}(y) \equiv \int_{-\infty}^{\infty} e^{i(yu+u^3/3)}du. \quad (19)$$

An example illustrating this behavior is given in Fig. 3.

Besides the signal structure itself, second-order singularities of Ψ_X may define "ghost" fold lines in the time–frequency plane. Specific points on such lines can themselves be the locus of a third-order singularity if $\partial^3 \Psi_X(\nu_n; t, f)/\partial \nu^3 = 0$. This simply corresponds to the condition $\ddot{t}_X(f_1) = -\ddot{t}_X(f_2)$, i.e., to the situation for which not only the slopes of the group delay at the interfering points are the same, but also the curvatures are equal in magnitude and opposite in sign. Generally, such singularities can only occur at isolated points which are *cusp points* of a fold line; an approximation of the distribution in the vicinity of a cusp point can be achieved in terms of Pearcey functions [33].

Finally, higher-order singularities can be observed at points which are the center of a perfect symmetry (or skew-symmetry) in the plane.

This classification exhausts the typical behaviors which can be observed in the case of the Wigner–Ville distribution.

3. TIME–FREQUENCY LOCALIZATION IN THE AFFINE CLASS

3.1. The Bertrands' Affine Class

The first step in constructing the affine class of the Bertrands [6] is to consider quadratic time–frequency representations $\Omega_X(t, f)$ that are covariant with respect to time shifts and dilations, i.e., for which the commutativity of the diagram
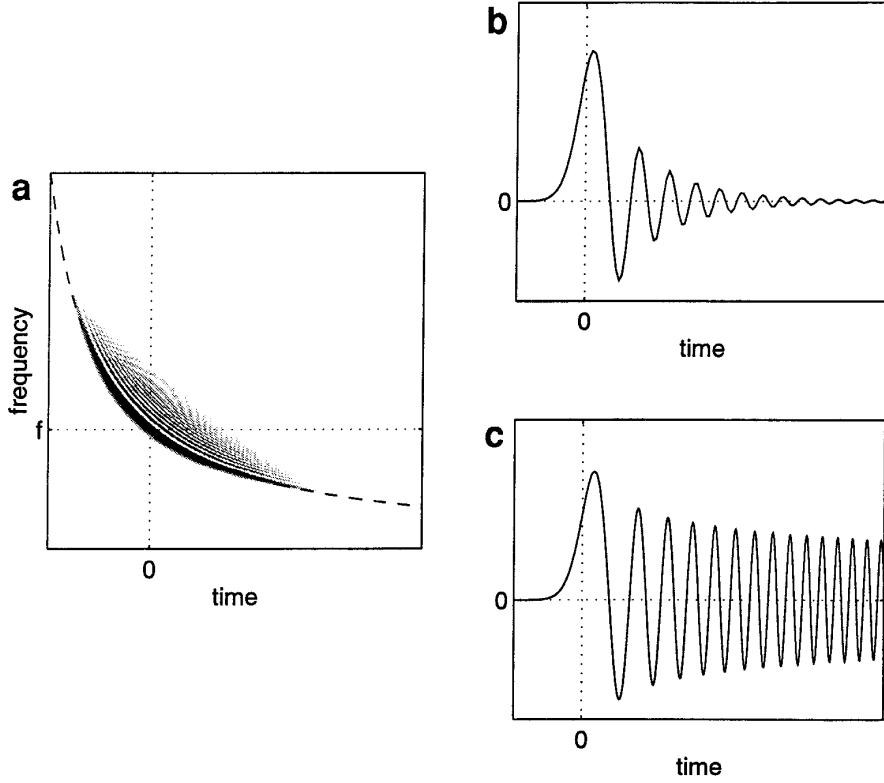
FIG. 3. Wigner–Ville distribution: transitional approximation. In the case of a nonlinear frequency modulation, the Wigner–Ville distribution possesses "inner" interference terms, which develop inside the concavity of the group delay. In the vicinity of the group delay, a transitional approximation allows us to describe these terms by an Airy function. (a) Wigner–Ville distribution of a signal with a hyperbolic group delay. (b) Section of the Wigner–Ville distribution at frequency f. (c) Airy function, renormalized so as to correspond to the theoretical prediction of (18).

$$
\begin{array}{ccc}
X(f) & \rightarrow & \Omega_X(t,f) \\
\downarrow & & \downarrow \\
a^{1/2}X(af)e^{-i2\pi t_0 f} & \rightarrow & \Omega_X(a^{-1}(t-t_0), af)
\end{array}
\tag{20}
$$

is guaranteed for any $a > 0$.

This covariance requirement gives rise to a large class of distributions, referred to as the *affine class* [6, 35]. Within this class, we can cite some well-known representations, such as the scalogram (squared modulus of a continuous wavelet transform) and the Wigner–Ville distribution, this latter distribution being also frequency-shift covariant.[5] Furthermore, we can find a subset of the affine class possessing a three parameter extended covariance. More precisely, it has been shown by the Bertrands [6] that the unique covariance properties, compatible with the affine covariance requirement (i.e., defining a three dimensional solvable Lie algebra containing the affine algebra), are associated with the commutativity of the diagrams

$$
\begin{array}{llll}
k \neq 0,1 & X(f) & \rightarrow & \Omega_X(t,f) \\
& \downarrow & & \downarrow \\
& e^{-i2\pi\alpha L_k f^k}X(f) & \rightarrow & \Omega_X(t - k\alpha L_k f^{k-1}, f) \\
k = 0 & X(f) & \rightarrow & \Omega_X(t,f) \\
& \downarrow & & \downarrow \\
& e^{-i2\pi\alpha(L_0 + \beta \log f)}X(f) & \rightarrow & \Omega_X(t - \beta\alpha f^{-1}, f) \\
k = 1 & X(f) & \rightarrow & \Omega_X(t,f) \\
& \downarrow & & \downarrow \\
& e^{-i2\pi\alpha f(L_1 + \log f)}X(f) & \rightarrow & \Omega_X(t - \alpha(1 + L_1 + \log f), f),
\end{array}
$$

$$
\tag{21}
$$

where L_k, L_0, L_1, β, and α are real constants.

It follows that the time–frequency representations that are covariant with respect to the three-parameter groups $\mathscr{G}_k$ associated with each of the diagrams in (21) can all be labeled by a real-valued parameter k and take on the following

---

[5] We must note, however, that frequency shifts may transform positive frequencies into negative ones, thus leading one to consider the whole real line of frequencies. In contrast, dilations force one to deal separately with the half real lines of positive and negative frequencies. In the rest of this paper, and unless otherwise specified, the discussion will always be restricted to positive frequencies only (analytic signals are a particular case for which the contributions at negative frequencies are set to zero).

form

$$\Omega_X^{(k)}(t,f) = \int_{-\infty}^{\infty} \mu_k(u)X(\lambda_k(u)f)X^*(\lambda_k(-u)f)e^{i2\pi ft\zeta_k(u)}du,$$
(22)

with

$$\begin{cases} \lambda_k(u) = \left(k\frac{e^{-u}-1}{e^{-ku}-1}\right)^{1/(k-1)}, & k \neq 0, 1, \\ \lambda_0(u) = \frac{u}{1-e^{-u}}, \\ \lambda_1(u) = \exp\left(1 + \frac{ue^{-u}}{e^{-u}-1}\right). \end{cases}$$
(23)

and

$$\zeta_k(u) = \lambda_k(u) - \lambda_k(-u).$$
(24)

The weight $\mu_k(u)$ is a real-valued function fixing some properties of the associated distribution $\Omega_X^{(k)}(t,f)$. Let us stress here two of these properties [6]:

1. *Unitarity.* The condition

$$\left|\int_{-\infty}^{\infty} X_1(f)X_2^*(f)df\right|^2 = \int_0^{\infty}\int_{-\infty}^{\infty} \Omega_{X_1}^{(k)}(t,f)\Omega_{X_2}^{(k)*}(t,f)dtdf$$

requires that

$$\mu_k(u) = \mu_k^U(u) \equiv \left(\lambda_k(u)\lambda_k(-u)\frac{d\zeta_k(u)}{du}\right)^{1/2}.$$
(25)

2. *Time localization.* The condition

$$X_{t_0}(f) = f^{-1/2}e^{-i2\pi t_0 f} \Rightarrow \Omega_{X_{t_0}}^{(k)}(t,f) = f^{-1}\delta(t-t_0)$$

requires simultaneously that

$$k \leqslant 0$$
$$\mu_k(u) = \mu_k^L(u) \equiv (\lambda_k(u)\lambda_k(-u))^{1/2}\frac{d\zeta_k(u)}{du}.$$
(26)

The unique function λ_k such that $\mu_k = \mu_k^U = \mu_k^L$, i.e., satisfying both conditions of unitarity and time localization, corresponds to $k = 0$, and results in the *unitary Bertrand distribution* [5, 6]

$$B_X(t,f) \equiv f\int_{-\infty}^{\infty}\frac{u}{2\sinh(u/2)}X\left(\frac{ue^{u/2}}{2\sinh(u/2)}f\right)$$

$$\times X^*\left(\frac{ue^{-u/2}}{2\sinh(u/2)}f\right)e^{i2\pi ftu}du, \quad (27)$$

which is associated with the parameterization given by (23).

3.2. The Affine Localized Class ($k \leqslant 0$)

Combining the time localization property of (26) with the covariances of (21) gives rise to distributions which are perfectly localized on specific nonlinear time–frequency curves. More precisely, we get for $k < 0$ [6]

$$X(f) = f^{-1/2}e^{-i2\pi t_0 f} \rightarrow \Omega_X^{(k)}(t,f) = f^{-1}\delta(t-t_0)$$
$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$
$$X^{(k)}(f) = f^{-1/2}e^{-i2\pi(t_0 f + \alpha L_k f^k)} \rightarrow \Omega_{X^{(k)}}^{(k)}(t,f) = f^{-1}\delta(t$$
$$- (t_0 + k\alpha L_k f^{k-1})),$$

and, for $k = 0$,

$$X(f) = f^{-1/2}e^{-i2\pi t_0 f} \rightarrow B_X(t,f)$$
$$= f^{-1}\delta(t-t_0)$$
$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$
$$X^{(0)}(f) = f^{-1/2}e^{-i2\pi(t_0 f + \alpha(L_0 + \beta\log f))} \rightarrow B_{X^{(0)}}(t,f) = f^{-1}\delta(t$$
$$- (t_0 + \beta\alpha f^{-1})).$$

Hence, for any power law group delay of the form $\theta^{(k)}(f) = t_0 + cf^{k-1}$, with $k \leqslant 0$ and $(t_0, c) \in \mathbb{R} \times \mathbb{R}$, the affine class provides a corresponding localized distribution $\Omega_X^{(k)}(t,f)$. The condition $\mu_k = \mu_k^L$ guarantees that the distribution $\Omega_X^{(k)}(t,f)$ behaves as a *Dirac* distribution along the trajectory $\theta^{(k)}(f)$, when it is computed on the signal

$$X^{(k)}(f) = f^{-1/2}e^{i\Phi^{(k)}(f)},$$

with

$$\Phi^{(k)}(f) \equiv -2\pi\int_{-\infty}^{f}\theta^{(k)}(\nu)d\nu$$
$$= \begin{cases} -2\pi(t_0 + \alpha(L_0 + \beta\log f)), & k = 0 \\ -2\pi(t_0 f + \alpha L_k f^k), & k < 0 \end{cases}.$$

For instance, the case $k = 0$ leads to the unitary Bertrand distribution (27), which localizes on *hyperbolae* on the time-frequency plane, whereas the case $k = -1$, together with the localization constraint, leads to the "active" form of the Unterberger distribution [6]

$$U_X^a(t,f) \equiv f\int_{-\infty}^{\infty}\left(1 + \frac{1}{\gamma^2}\right)X(\gamma f)X^*\left(\frac{f}{\gamma}\right)e^{i2\pi(\gamma - 1/\gamma)tf}d\gamma,$$
(28)

which localizes along *squared hyperbolae* $\theta^{(-1)}(f) = t_0 + \alpha L_{-1}f^{-2}$.

As for the Wigner–Ville distribution, perfect localization of this affine subclass on specific curves in the plane can be associated with particular geometries and modified means (in contrast with the arithmetic one, which holds for the Wigner–Ville distribution). The next section is devoted to this topic.
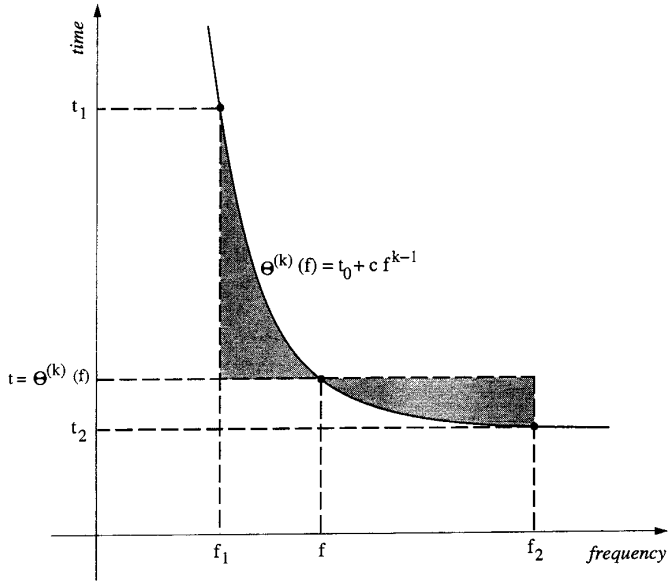
FIG. 4. Frequency mean. For an affine distribution indexed by k, the location of an interference point is fixed by the frequency mean (evaluated on the interval defined by the two interfering points) of the "matched" power-law group delay of degree $(k-1)$. This corresponds to the equality of the shaded areas.

3.3. Geometric Interpretation of Localization

Assuming that $k \leq 0$ and $\mu_k = \mu_k^L$, let us substitute the signal $X^{(k)}(f) = f^{-1/2}e^{i\Phi^{(k)}(f)}$ into (22). We get

$$\Omega_{X^{(k)}}^{(k)}(t,f) = \int_{-\infty}^{\infty} e^{i[\Phi^{(k)}(\lambda_k(u)f) - \Phi^{(k)}(\lambda_k(-u)f)]}e^{i2\pi ft\zeta_k(u)}d\zeta_k(u). \quad (29)$$

Given k, the "matched" group delay $\theta^{(k)}(f) = -(1/2\pi)\dot{\Phi}^{(k)}(f)$ is of the form $t_0 + cf^{k-1}$, and we know from the previous section that the distribution $\Omega_{X^{(k)}}^{(k)}(t,f)$ is such that

$$\Omega_{X^{(k)}}^{(k)}(t,f) = f^{-1}\delta\left(t + \frac{1}{2\pi}\dot{\Phi}^{(k)}(f)\right). \quad (30)$$

Therefore, identifying the result given by this last expression with the general form of (29) yields the equation

$$\frac{\Phi^{(k)}(\lambda_k(u)f) - \Phi^{(k)}(\lambda_k(-u)f))}{\lambda_k(u)f - \lambda_k(-u)f} = \dot{\Phi}^{(k)}(f). \quad (31)$$

This relation merely states that a nonzero value for $\Omega_{X^{(k)}}^{(k)}$ at a point $(t = \theta^{(k)}(f), f)$ results from the interaction between any two points $(t_1 = \theta^{(k)}(f_1), f_1 = \lambda_k(-u)f)$ and $(t_2 = \theta^{(k)}(f_2), f_2 = \lambda_k(u)f)$ such that the finite difference of the phase $\Phi^{(k)}$ between the interacting points equals the phase derivative at the interference point (t,f). Making use

of the relation between $\Phi^{(k)}(f)$ and $\theta^{(k)}(f)$, and assuming for convenience that $f_2 > f_1$, (31) becomes

$$t = \theta^{(k)}(f) = \frac{1}{f_2 - f_1}\int_{f_1}^{f_2}\theta^{(k)}(\nu)d\nu. \quad (32)$$

In words, the frequency location of the interference point is fixed by the frequency mean (evaluated on the interval defined by the two interfering points) of the "matched" group delay $\theta^{(k)}$. The corresponding time location follows directly from the one-to-one relation $t = \theta^{(k)}(f)$ (see Fig. 4).

It thus becomes simple to explain the perfect localization of $\Omega^{(k)}$ on power-law group delays $\theta^{(k)}(f)$ of degree $(k-1)$ as a by-product of interference geometry: this directly results from the fact that any two points $(\theta^{(k)}(\lambda_k(u)f), \lambda_k(u)f)$ and $(\theta^{(k)}(\lambda_k(-u)f), \lambda_k(-u)f)$ belonging to the "matched" group delay $\theta^{(k)}(f)$ interfere at a third point that also belongs to the same group delay.

This geometrical interpretation of localization generalizes the one previously discussed in the case of the Wigner–Ville distribution and can be formulated as follows:

A distribution is perfectly localized on a given time-frequency curve as long as this curve is the locus of all of its interference points.

3.4. Construction Rule for Interference Terms

Although, in the previous derivations, we have considered only monocomponent signals defined by a continuous power-law group delay, we can imagine that, as we did for the Wigner–Ville distribution, we can generalize the pointwise application of (31) and (32). For any two points (t_1, f_1) and (t_2, f_2), lying on an arbitrary signal structure, the continuity and the monotonic property of λ_k ensure that there always exist constants u and f such that

$$f_1 = \lambda_k(-u)f$$

$$f_2 = \lambda_k(u)f,$$

as well as a unique pair (t_0, c) such that $\theta^{(k)}(f) = t_0 + cf^{k-1}$, thus fitting a "matched" group delay to the two considered points. Then, from the pointwise application of (32) and from the general definition (23) of λ_k, we see that the interference between (t_1, f_1) and (t_2, f_2) will be located at the point $(t = \theta^{(k)}(f), f = \Theta^{(k)}(f_1, f_2))$, with

$$\Theta^{(k)}(f_1, f_2) \equiv \left(\frac{1}{k}\frac{f_2^k - f_1^k}{f_2 - f_1}\right)^{1/(k-1)}. \quad (33)$$

We note that, as with the definition (23) of λ_k, the specific cases $k = 0$ and $k = 1$ can be obtained by continuity. However, we can also consider them separately, and, in either case, we get the closed form expressions
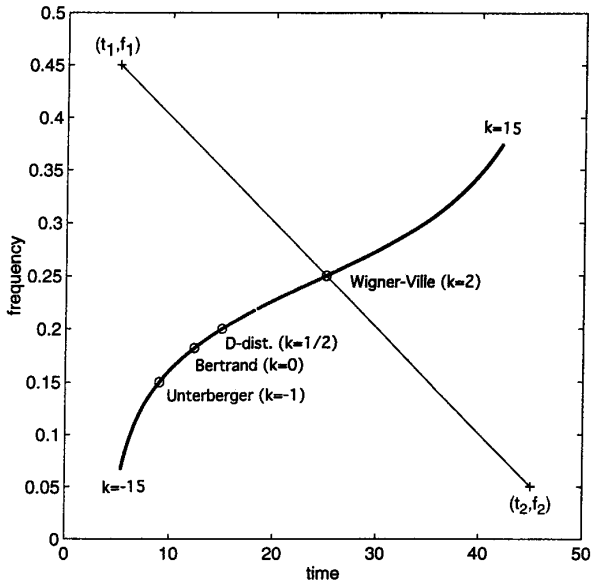
FIG. 5. Midpoint construction rule. Any two points of the time–frequency plane interfere at a third point according to (33) and (34). The graph shows the evolution of the midpoints constructed this way when k is varying.

$$\Theta^{(0)}(f_1, f_2) = \frac{f_2 - f_1}{\log(f_2/f_1)},$$

$$\Theta^{(1)}(f_1, f_2) = \exp\left[\frac{f_2(\log f_2 - 1) - f_1(\log f_1 - 1)}{f_2 - f_1}\right].$$

An analogous construction rule can be obtained for the time-location of the interference point, by inverting the equation $t = \theta^{(k)}(f)$. Introducing

$$\tau \equiv t - t_0 = c f^{k-1}$$

yields

$$f = \left(\frac{\tau}{c}\right)^{1/(k-1)},$$

and, substituting into (33), we obtain

$$\left(\frac{\tau}{c}\right)^{1/(k-1)} = \left(\frac{1}{k}\frac{\left(\frac{\tau_2}{c}\right)^{k/(k-1)} - \left(\frac{\tau_1}{c}\right)^{k/(k-1)}}{\left(\frac{\tau_2}{c}\right)^{1/(k-1)} - \left(\frac{\tau_1}{c}\right)^{1/(k-1)}}\right)^{1/(k-1)},$$

which reduces to

$$\tau^{1/(k-1)} = \Theta^{(k)}\left(\tau_1^{1/(k-1)}, \tau_2^{1/(k-1)}\right), \qquad (34)$$

It follows from (33) and (34) that the general construction rule of interference points is governed only by the function $\Theta^{(k)}$, whose behavior is plotted in Fig. 5. From the point of view of the localized affine class, only negatives k's make sense, although considering $\Theta^{(k)}$ on its own and allowing positive k's remains interesting. In particular, it seems quite natural to consider $\Theta^{(k)}$ as defining a kind of modified "mean", and it is remarkable that the usual (arithmetic) mean is precisely associated with the value $k = 2$, known to be connected to the Wigner–Ville distribution within the affine class [6] (see also the next section). Therefore, and before proceeding to a closer study of $\Theta^{(k)}$, it is instructive to say a few words about a possible generalization of interference constructions to distributions with positive k's.

3.5. Localization Properties Extended to the Case $k > 0$

Perfect localization of affine distributions along nonlinear curves on the time–frequency plane requires that the time localization property of (26) be satisfied. Whatever the phase spectrum $\Phi^{(k)}(f)$ in (29), *perfect* localization (in the sense of having a distribution with a Dirac-type behavior) is only achievable if the integration bounds are infinite. This means that the function $u \mapsto \zeta_k(u)$ must be in a one-to-one correspondence from $\mathbb{R}$ to $\mathbb{R}$ and this is true only for $k \leq 0$ [6].

If we now consider k to be positive, the function

$$\zeta_k(u) = 2\left(k\frac{\sinh(u/2)}{\sinh(ku/2)}\right)^{1/(k-1)}\sinh\frac{u}{2},$$

remains continuous and monotonous, but maps $\mathbb{R}$ to the interval $[-k^{1/(k-1)}, k^{1/(k-1)}]$. Therefore, this induces finite integration bounds in (29), which prevent the corresponding integral from behaving as a *Dirac* distribution. This limitation of the integration domain plays the role of an implied weighting function applied to the variable $\zeta_k(u)$. This results, in the time–frequency domain, in a convolution which spreads the signature of the distribution around the theoretical curve of the group delay.

For instance, the value $k = 2$ yields the function

$$\lambda_2(u) = \frac{2}{e^{-u} + 1} = \frac{e^{u/2}}{\cosh(u/2)} = \tanh\frac{u}{2} + 1,$$

and the difference

$$\zeta_2(u) = 4\frac{\sinh^2(u/2)}{\sinh u} = 2\tanh\frac{u}{2},$$

that is a one-to-one correspondence from $\mathbb{R}$ to the interval $[-2, 2]$. The corresponding time–frequency distribution

takes on the form

$$\Omega_X^{(2)}(t,k) = f \int_{-\infty}^{\infty} \mu_2(u) X\left(f\left(1 + \tanh\frac{u}{2}\right)\right)$$

$$\times X^*\left(f\left(1 - \tanh\frac{u}{2}\right)\right) e^{i4\pi t f \tanh u/2} du.$$

This expression can be simplified to

$$\Omega_X^{(2)}(t,f) = \int_{-2f}^{2f} \mu_2(u) X\left(f + \frac{\xi}{2}\right) X^*\left(f - \frac{\xi}{2}\right) e^{i2\pi\xi t}(fJ)d\xi,$$

by introducing the change of variable $\xi = 2f\tanh(u/2)$, whose Jacobian J is such that

$$(fJ)^{-1} = \frac{d\xi}{du}$$

$$= \frac{4f^2 - \xi^2}{4f^2}$$

$$= 1 - \tanh^2\frac{u}{2}$$

$$= \left[\left(1 + \tanh\frac{u}{2}\right)\left(1 - \tanh\frac{u}{2}\right)\right]^{1/2}$$

$$\times \left(1 - \tanh^2\frac{u}{2}\right)^{1/2}$$

$$= (\lambda_2(u)\lambda_2(-u))^{1/2}\left(\frac{d}{du}\tanh\frac{u}{2}\right)^{1/2}$$

$$= \left(\lambda_2(u)\lambda_2(-u)\frac{d\zeta_2(u)}{du}\right)^{1/2}.$$

Therefore, from the unitarity condition (25), we have that $(fJ)^{-1} = \mu_2^U$, and, forcing $\mu_2 = \mu_2^U$, the resulting distribution can be written as

$$\Omega_X^{(2)}(t,f) = \frac{1}{4}\int_{-\infty}^{\infty} X_a\left(f + \frac{\xi}{2}\right) X_a^*\left(f - \frac{\xi}{2}\right) e^{i2\pi\xi t}d\xi, \quad (35)$$

where $X_a(f)$ stands explicitly for the analytic signal associated to $X(f)$ (i.e., the signal such that $X_a(f) = 2X(f)$ for $f \geqslant 0$ and $X_a(f) \equiv 0$ for $f < 0$). We recognize in the above expression the definition of the Wigner–Ville distribution, restricted to analytic signals. It is clear that this version of the Wigner–Ville distribution does not possess the strict localization property, since the Wigner–Ville distribution has been shown to be Dirac with respect to "chirps," whose frequency modulation extends along the entire real line. Moreover, the weight μ_2 has been chosen so as to coincide with μ_2^U (unitarity) and not with μ_2^L (localization). Thereby, as for

the usual form (2) of the Wigner–Ville distribution, $\Omega^2(t,f)$ is a unitary time–frequency representation, but it cannot be expected to be perfectly localized on its specific group delay $\theta^{(2)}(f) = t_0 + cf$ for $f \geqslant 0$.

However, if we formally apply the pointwise construction rules of interferences given in (33) and (34), we get for $k = 2$

$$f = \frac{f_1 + f_2}{2}$$

$$\tau = \frac{\tau_1 + \tau_2}{2} \Leftrightarrow t = \frac{t_1 + t_2}{2},$$

which are the classical arithmetic means associated with the usual Wigner–Ville distribution. We can therefore expect that, apart from the spreading effect mentioned previously, distributions $\Omega^{(k)}$ will still exhibit an approximate localization on their "matched" group delay in the case of positive k's. This claim will be supported by numerical experiments in the next section and by stationary phase arguments in Section 4. Before proceeding with these points, we will first investigate further the properties of the function $\Theta^{(k)}$ controlling the interference geometry.

3.6. Generalized Means

Within the affine localized class, we have shown that interference geometry is governed by the function $\Theta^{(k)}$, defined as

$$\Theta^{(k)}(\omega_1, \omega_2) = \left(\frac{1}{k}\frac{\omega_2^k - \omega_1^k}{\omega_2 - \omega_1}\right)^{1/(k-1)}. \quad (36)$$

This function generalizes the notion of midpoint associated to the case $k = 2$ (arithmetic mean, Wigner–Ville distribution) and defines modified geometries. It is therefore tempting to consider $\Theta^{(k)}$ as defining a "generalized mean" which would allow one—as in (11)—to characterize all these modified geometries in terms of generalized symmetries. To make this point more precise, let us recall the definition of a generalized mean introduced by Kolmogorov [27] and Nagumo [29], as revisited by Aczél [1] (see also [21]). By definition, a two-dimensional function M defines a generalized mean if there exists a one-dimensional, continuous and monotonic increasing function g such that

$$M(\omega_1, \omega_2) = g^{-1}\left(\frac{g(\omega_1) + g(\omega_2)}{2}\right). \quad (37)$$

Given this definition, we can establish the following result:

The midpoint geometric rule (36) *corresponds to a generalized mean for only three values of the parameter k*: $k = -1$ (*geometric mean*), $k = \frac{1}{2}$ (*square root mean*) and $k = 2$ (*arithmetic mean*).

The proof of this claim is given in Appendix A. The two cases $k = -1$ and $k = 2$ have already been noted to correspond respectively to the "active" Unterberger distribution and to the Wigner–Ville distribution. The time–frequency representation corresponding to the value $k = \frac{1}{2}$ is the so-called D-distribution [14], and is associated with the function

$$\lambda_{1/2}(u) = \left(\frac{2}{e^{-u/2} + 1} \right)^2 = \left(1 + \tanh \frac{u}{4} \right)^2.$$

In this function, the change of variable $\gamma = 4 \tanh(u/4)$ yields

$$\zeta_{1/2}(u) = \gamma,$$

and

$$du = \left(1 - \left(\frac{\gamma}{4} \right)^2 \right)^{-1} d\gamma.$$

Thus, setting

$$\mu_{1/2}(u) = \lambda_{1/2}(u)\lambda_{1/2}(-u),$$

the D-distribution can be equivalently written as

$$D_X(t, f) = f \int_{-4}^{4} \left(1 - \left(\frac{\gamma}{4} \right)^2 \right) X \left(f \left(1 + \frac{\gamma}{4} \right)^2 \right)$$
$$\times X^* \left(f \left(1 - \frac{\gamma}{4} \right)^2 \right) e^{i2\pi f t \gamma} d\gamma. \quad (38)$$

It is interesting to note that the integration bounds, given by $[-k^{1/(k-1)}, k^{1/(k-1)}]$ with $k = \frac{1}{2}$, prohibit a perfect localization of the D-distribution along the theoretical group delay $\theta^{(1/2)}(f) = t_0 + cf^{-1/2}$. Similar to the case of $k = 2$, it is better to speak of an approximate localization around $\theta^{(k)}$.

The knowledge of the values of k for which (36) defines a generalized mean allows us to identify their exact type (this is detailed in Appendix A):

- $k = 2$ (arithmetic mean)

$$\Theta^{(2)}(\omega_1, \omega_2) = \frac{\omega_1 + \omega_2}{2}.$$

- $k = -1$ (geometric mean)

$$\Theta^{(-1)}(\omega_1, \omega_2) = e^{(1/2)(\log \omega_1 + \log \omega_2)} = \sqrt{\omega_1 \omega_2}.$$

- $k = \frac{1}{2}$ (square root mean)

$$\Theta^{(1/2)}(\omega_1, \omega_2) = \left(\frac{\sqrt{\omega_1} + \sqrt{\omega_2}}{2} \right)^2.$$

Let us point out that, although the condition (37), which defines an exact generalized mean, is only satisfied for $k = -1, k = \frac{1}{2}$ and $k = 2$, other values of k give rise to midpoints which are not that far from an exact Kolmogorov–Nagumo mean. One can be convinced of the accuracy of this approximation by evaluating the criterion derived by Aczél in [1], and which states that a function $M(\omega_1, \omega_2)$ defines a Kolmogorov–Nagumo mean if and only if it satisfies the property of *bisymmetry*

$$M(M(\omega_1, \omega_2), M(\omega_3, \omega_4)) = M(M(\omega_1, \omega_3), M(\omega_2, \omega_4)).$$

Figure 6 displays the behavior of the ratio

$$\rho_{\omega_1, \omega_2, \omega_3, \omega_4}(k) = \frac{\Theta^{(k)}(\Theta^{(k)}(\omega_1, \omega_2), \Theta^{(k)}(\omega_3, \omega_4))}{\Theta^{(k)}(\Theta^{(k)}(\omega_1, \omega_3), \Theta^{(k)}(\omega_2, \omega_4))},$$

and illustrates its weak deviation from unity for a wide range of values of the parameters, and especially when $-1 \leq k \leq 2$. This behavior is consistent with the approximation discussed in [18], and according to which (36) satisfies

$$\Theta^{(k)}(\omega_1, \omega_2) \approx \tilde{g}_k^{-1} \left(\frac{\tilde{g}_k(\omega_1) + \tilde{g}_k(\omega_2)}{2} \right)$$

with

$$\tilde{g}_k(\omega) = \omega^{(k+1)/3}.$$

3.7. An Illustration

The localization properties, which have been discussed so far from a theoretical perspective, are illustrated in Fig. 7. The example consists of three (time-limited) signals with a power-law group delay, the degree of which is fixed to $-2, -1$ and $-\frac{1}{2}$. These three signals, whose group delays can be viewed as "matched" group delays $\theta^{(k)}(f)$, associated respectively with $k(\text{sig.}) = -1, k(\text{sig.}) = 0$ and $k(\text{sig.}) = \frac{1}{2}$, are each analyzed by means of the three corresponding distributions, i.e., the Unterberger distribution ($k(\text{dis.}) = -1$), the unitary Bertrand distribution ($k(\text{dis.}) = 0$) and the D-distribution $k(\text{dis.}) = \frac{1}{2}$).

The numerical computations are in good agreement with the theoretical predictions. As expected, localization (up to a spreading due to time limitation) is achieved in the subfigures along the main diagonal, i.e., when the index of the analysis is matched to the index of the group delay. Off-diagonal plots reveal a further spreading of interference terms and, as predicted by the construction rule, this spreading develops on one side only of the signal group delay, depending on the sign of $\delta k = k(\text{sig.}) - k(\text{dis.})$: when $\delta k < 0 \, (\delta k > 0)$, the interference terms occur on the convex (concave) side of the signal group delay.
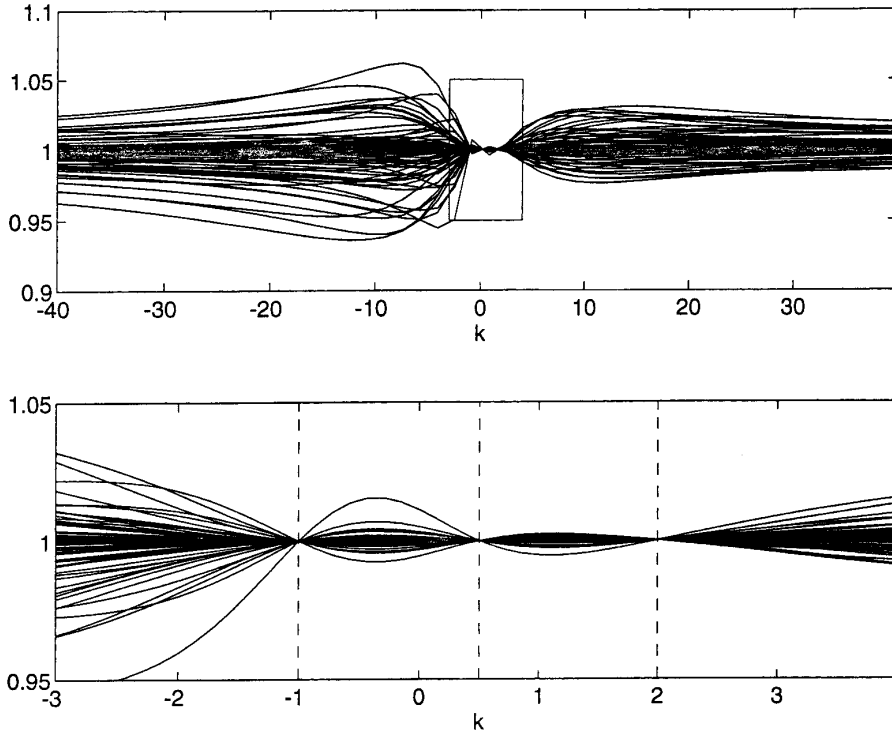
FIG. 6. Midpoints and generalized means. Not all the midpoints of Fig. 5 can be defined as generalized means. This is true only for $k = -1, k = \frac{1}{2}$ and $k = 2$. The graph illustrates this theoretical result by plotting the Aczel's test of bisymmetry $\rho_{\omega_1,\omega_2,\omega_3,\omega_4}(k)$ (see text) for a wide range of the parameters (100 trials with random values of $\omega_1, \omega_2, \omega_3$ and ω_4, uniformly distributed on the interval [0.1, 30]). The bottom diagram is an enlargement of the box in the top diagram.

Construction rules, as they are applied in Fig. 7, can only provide information about the *location* of interference terms, but not about their *structure*. The next section is devoted to this topic, in order to refine the theoretical predictions and, e.g., to explain the fringes observed in Fig. 7b.

4. INTERFERENCE TERMS AND FINE STRUCTURE OF AFFINE DISTRIBUTIONS

In the case of frequency modulated signals of the form (9), a more complete study of the local behavior of affine distributions can be undertaken using stationary phase-type approximations. This section is devoted to this study and will provide results generalizing those summarized in Section 2.4 (Wigner–Ville case). To complement the results we will give a specific, analytically tractable example, which will be used throughout this section. It consists of a two-chirp model (see Fig. 8), the group delay of which is given by

$$t_X(f) = \left| \frac{f - f_0}{\alpha} \right| U(f_0 - f)\text{sign}(t) = \begin{cases} t_1(f), \ t_X(f) \leqslant 0 \\ t_2(f), \ t_X(f) \geqslant 0, \end{cases} \quad (39)$$

where U is the unit step function and sign is the signum function.

Analytic calculations corresponding to this signal and its analysis by means of the Unterberger distribution ($k = -1$) are detailed in Appendix B.

4.1. Critical Manifold

As for the Wigner–Ville distribution (see Section 2.4), the key idea, for using stationary phase-type approximation, is, given the model (9), to rewrite any localized affine distribution (22) as an oscillatory integral:

$$\Omega_X^{(k)}(t, f) = \int_{-\infty}^{\infty} L_X^{(k)}(u; f) e^{i\Psi_X^{(k)}(u;t,f)} d\zeta_k(u), \quad (40)$$

with

$$L_X^{(k)}(u; f) \equiv f(\lambda_k(u)\lambda_k(-u))^{1/2} A_X(\lambda_k(u)f) A_X(\lambda_k(-u)f), \quad (41)$$

$$\Psi_X^{(k)}(u; t, f) \equiv \Phi_X(\lambda_k(u)f) - \Phi_X(\lambda_k(-u)f) + 2\pi t f \zeta_k(u). \quad (42)$$
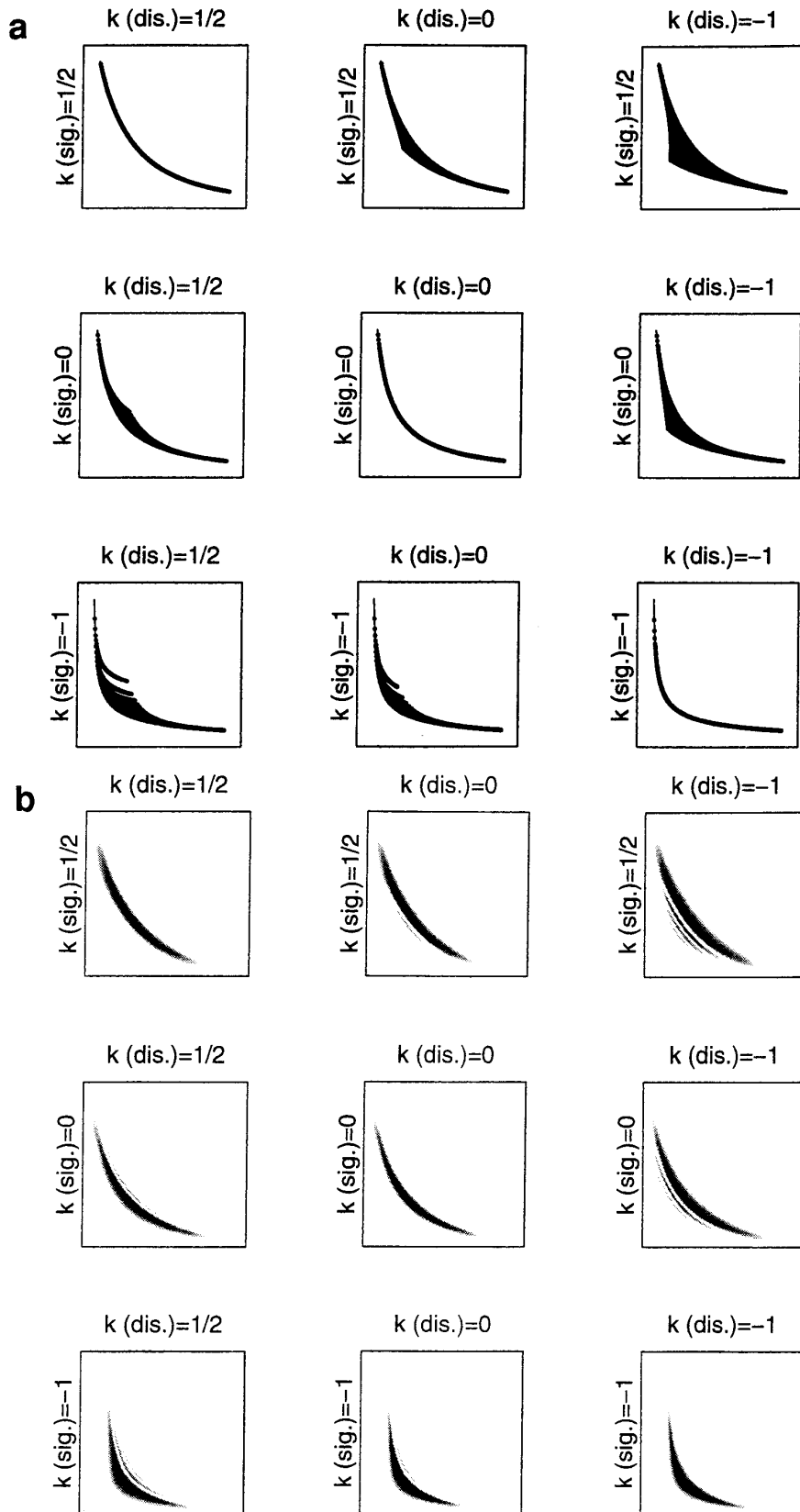
FIG. 7. Affine distributions: localization and interference. Power-law group delays of degree $(k-1)$ are "matched" to affine distributions indexed by k. When the index k(sig.)" of the signal group delay is different from the index k(dis.)" of the distribution, spurious inner interference terms appear. The graphs illustrate this result with 3 signals with different group delays analyzed by the 3 corresponding distributions. (a) Predictions for the locus of interference, from a pointwise application of the rules of Fig. 5. (b) Numerical computations of the distributions.
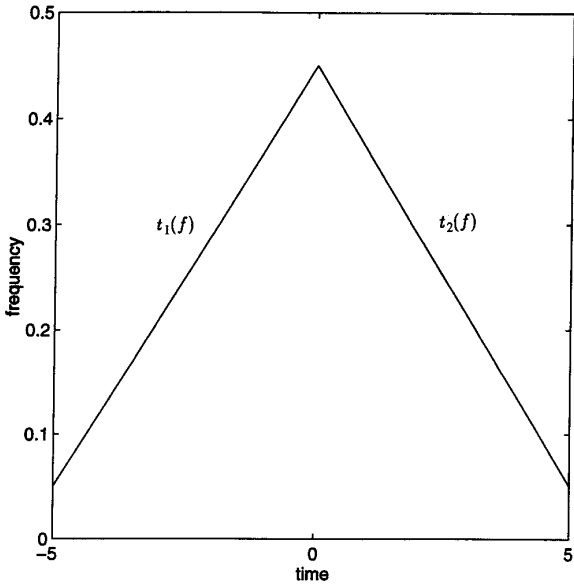
FIG. 8.   The two chirp signal model.

$$t = \frac{\dot{\lambda}_k(u)(t_0 + c\lambda_k^{k-1}(u)f^{k-1}) + \dot{\lambda}_k(-u)(t_0 + c\lambda_k^{k-1}(-u)f^{k-1})}{\dot{\lambda}_k(u) + t\lambda_k(-u)}$$

$$= t_0 + cf^{k-1}\frac{\dot{\lambda}_k(u)\lambda_k^{k-1}(u) + \dot{\lambda}_k(-u)\lambda_k^{k-1}(-u)}{\dot{\lambda}_k(u) + \dot{\lambda}_k(-u)}$$

$$= t_0 + cf^{k-1}\frac{1}{k}\frac{(d/du)(\lambda_k^k(u) - \lambda_k^k(-u))}{(d/du)(\lambda_k(u) - \lambda_k(-u))}.$$

Recalling that [6]

$$\lambda_k(-u) = e^{-u}\lambda_k(u), \tag{45}$$

we then have

$$
\begin{aligned}
\lambda_k^k(u) - \lambda_k^k(-u) &= \lambda_k(u)\lambda_k^{k-1}(u) - e^{-u}e^{-(k-1)u}\lambda_k(u)\lambda_k^{k-1}(u) \\
&= \lambda_k(u)\lambda_k^{k-1}(u)(1 - e^{-ku}) \\
&= \lambda_k(u)\left(k\frac{e^{-u}-1}{e^{-ku}-1}\right)(1 - e^{-ku}) \\
&= k\lambda_k(u)(1 - e^{-u}) \\
&= k(\lambda_k(u) - \lambda_k(-u)).
\end{aligned}
$$

This implies that

$$\frac{(d/du)(\lambda_k^k(u) - \lambda_k^k(-u))}{(d/du)(\lambda_k(u) - \lambda_k(-u))} = k,$$

and thus

$$t = t_0 + cf^{k-1} = \theta^{(k)}(f).$$

This is clearly consistent with the geometric interpretation of interference terms according to which interfering and interference points of a localized distribution $\Omega_{X^{(k)}}^{(k)}$ all belong to the same specific group delay $\theta^{(k)}(f)$.

Approximation. Given a finite number of stationary points $\{u_n, n = 1 \cdots N\}$, and provided that the non-degeneracy condition

$$\frac{\partial^2 \Psi_X^{(k)}(u_n; t, f)}{\partial u^2} \neq 0 \tag{46}$$

is satisfied, the distribution $\Omega_X^{(k)}$ can be given the following stationary phase approximation

As a consequence, and assuming as previously that the envelope A_X is slowly varying compared to the variations of the phase Φ_X, it can be justified (see, e.g., [33]) that the qualitative behavior of (40) entirely depends on the singularities (i.e., vanishing derivatives) of the *critical manifold* $\mathscr{C}$, defined as

$$\mathscr{C} = \left\{(u, t, f) \;\middle|\; \frac{\partial \Psi_X^{(k)}(u)}{\partial u} = 0\right\}. \tag{43}$$

4.2. Stationary Phase Regions

Locus. Generically, the critical manifold (43) defines time–frequency *regions* associated with u's that are stationary points of the oscillatory integral (40), and on which the value of the distribution is nonnegligible. Making use of both (42) and (43), the stationary phase condition reads

$$t = \frac{\dot{\lambda}_k(u)t_X(\lambda_k(u)f) + \dot{\lambda}_k(-u)t_X(\lambda_k(-u)f)}{\dot{\lambda}_k(u) + \dot{\lambda}_k(-u)}. \tag{44}$$

We claim that the time–frequency points which satisfy this condition can be exactly identified with the interference points obtained by a pointwise application of the construction rule determined previously. In other words,

The projection of the critical manifold onto the time-frequency plane corresponds to the locus of the interference terms.

In order to prove this claim, consider the "matched" group delay $\theta^{(k)}(f) = t_0 + cf^{k-1}$ passing through the interfering points $(t_X(\lambda_k(u)f), \lambda_k(u)f)$ and $(t_X(\lambda_k(-u)f), \lambda_k(-u)f)$. From (44), we get
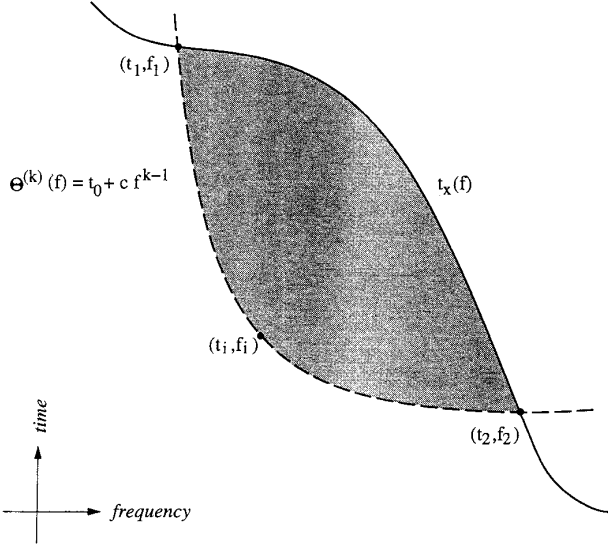
FIG. 9. Affine distributions: area rule for the stationary phase approximation. In the case of a nonlinear frequency modulation, and within the stationary phase approximation, the local oscillations of an affine distribution indexed by k are controlled by the area bounded by the signal group delay (solid line) and the "matched" power-law group delay of degree $(k - 1)$ (dashed line) joining the two interfering points.

$$\Omega_X^{(k)}(t, f) \approx \sum_{n=1}^{N} L_X^{(k)}(u_n; f) \left| \frac{\partial^2 \Psi_X^{(k)}(u_n; t, f)}{\partial u^2} \right|^{-1/2}$$

$$\times e^{i\Psi_X^{(k)}(u_n; t, f) + i(\pi/4)\text{sign}(\partial^2 \Psi_X^{(k)}(u_n; t, f)/\partial u^2)}. \quad (47)$$

It is easy to check from (44) that stationary points always occur in pairs. More precisely, if we consider two interfering points (t_1, f_1) and (t_2, f_2) which both belong to $t_X(f)$, we have

$$f_1 = \lambda_k(-u)f$$

$$f_2 = \lambda_k(u)f.$$

This corresponds to two distinct solutions for (44) and, hence, to two stationary points. Making use of (45), these stationary points are explicitly given by $u_1 = u_0, u_2 = -u_0$, with

$$u_0 = \log \frac{f_2}{f_1}.$$

In the case of a single contribution at a given point (t, f) (and assuming that $f_2 > f_1$), it then follows that the sum (47) reduces to

$$\Omega_X^{(k)}(t, f) \approx 2L_X^{(k)}(u_0; f) \left| \frac{\partial^2 \Psi_X^{(k)}(u_0; t, f)}{\partial u^2} \right|^{-1/2}$$

$$\times \cos\left(2\pi \mathscr{A}_X^{(k)}(t, f) + \frac{\pi}{4}\text{sign}\frac{\partial^2 \Psi_X^{(k)}(u_0; t, f)}{\partial u^2}\right), \quad (48)$$

with

$$\mathscr{A}_X^{(k)}(t, f) = \int_{f_1}^{f_2} t_X(\nu)d\nu - t(f_2 - f_1)$$

$$= \int_{f_1}^{f_2} t_X(\nu)d\nu - \int_{f_1}^{f_2} \theta^{(k)}(\nu)d\nu. \quad (49)$$

Oscillations. The above quantity $\mathscr{A}_X^{(k)}(t, f)$ is nothing but the area limited by the signal group delay $t_X(f)$ and the "matched" group delay $\theta^{(k)}(f)$, both passing through the two interfering points (t_1, f_1) and (t_2, f_2) (see Fig. 9). Clearly, (48) reveals an oscillatory structure which generalizes the Wigner–Ville case (17), for which the group delay was simply a linear law (see Fig. 2).

A simple method to quantify further the fine structure of the oscillations in a neighborhood of an interference point (t, f) studies $\mathscr{A}_X^{(k)}(t + \delta t, f + \delta f)$ as a function of δt and δf (both supposed to be small). A first order expansion of the phase (at the vicinity of a stationary point u_0) leads to

$$2\pi \mathscr{A}_X^{(k)}(t + \delta t, f + \delta f) = \Phi_X((f + \delta f)\lambda_k(u_0))$$

$$- \Phi_X((f + \delta f)\lambda_k(-u_0)) + 2\pi\zeta_k(u_0)(f + \delta f)(t + \delta t)$$

$$\approx \Phi_X(\lambda_k(u_0)f) + \delta f \lambda_k(u_0)\dot\Phi_X(\lambda_k(u_0)f)$$

$$- \Phi_X(\lambda_k(-u_0)f) - \delta f \lambda_k(-u_0)\dot\Phi_X(\lambda_k(-u_0)f)$$

$$+ 2\pi\zeta_k(u_0)[ft + f\delta t + t\delta f + \delta t\delta f].$$

Neglecting the second order term $\delta t \delta f$ in the bracket, we get

$$\mathscr{A}_X^{(k)}(t + \delta t, f + \delta f) \approx \mathscr{A}_X^{(k)}(t, f) + \delta\mathscr{A}_X^{(k)}(\delta t, \delta f),$$

with

$$\delta\mathscr{A}_X^{(k)}(\delta t, \delta f) = \delta t(f_2 - f_1)$$

$$+ \frac{1}{f}[f_2(t - t_2) - f_1(t - t_1)]\delta f. \quad (50)$$

Taking into account that the three points $(t, f), (t_1, f_1)$ and (t_2, f_2) are on the same power-law group delay $\theta^{(k)}$ of degree $(k - 1)$ and that $f = \Theta^{(k)}(f_2, f_1)$ (cf. (33)), we get

$$\delta\mathscr{A}_X^{(k)}(\delta t, \delta f)$$

$$= \delta t(f_2 - f_1) + \frac{1}{f}[f_2(cf^{k-1} - cf_2^{k-1})$$

$$- f_1(cf^{k-1} - cf_1^{k-1})]\delta f$$

$$= \delta t(f_2 - f_1) + \frac{c}{f}[f^{k-1}(f_2 - f_1) - (f_2^k - f_1^k)]\delta f$$
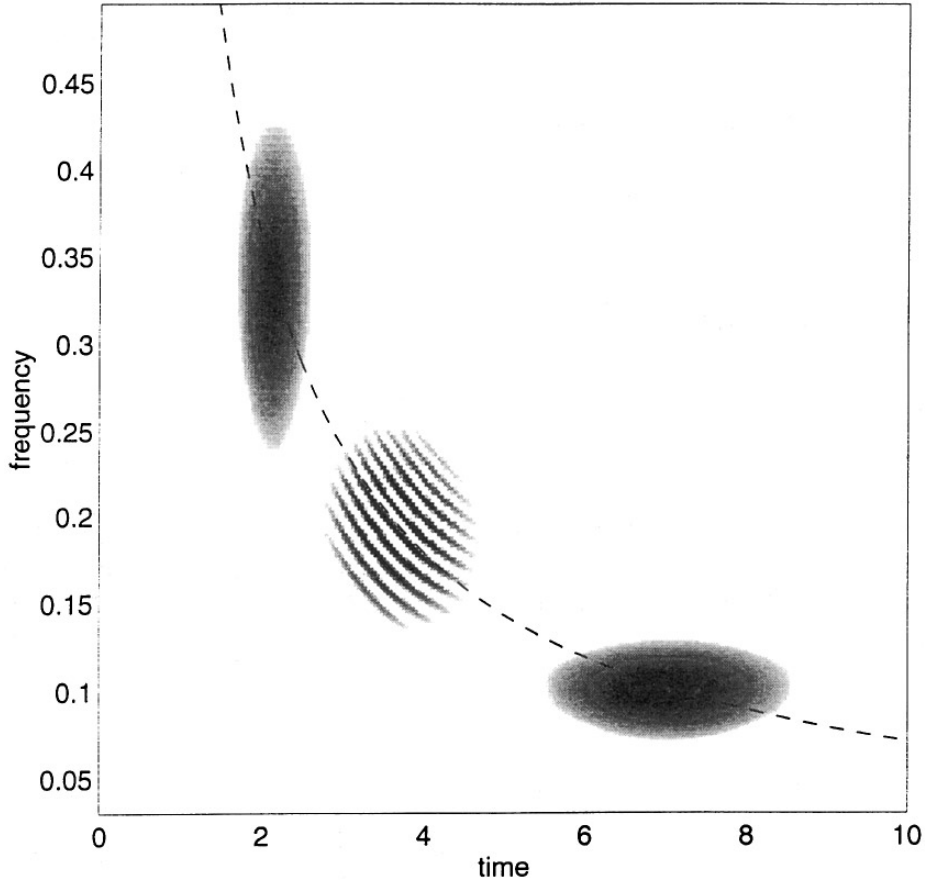
FIG. 10. Bertrand distribution: interference term of two logons. The interference term is located as explained in Fig. 5 and oscillates in a direction orthogonal to the "matched" group delay joining the two interfering components (dashed curve).

$$= \delta t(f_2 - f_1) + \frac{c}{f}(f_2 - f_1)\left[f^{k-1} - k\left(\frac{1}{k}\frac{f_2^k - f_1^k}{f_2 - f_1}\right)\right]\delta f$$

$$= (f_2 - f_1)[\delta t - c(k-1)f^{k-2}\delta f]$$

$$= (f_2 - f_1)(\delta t - \dot{\theta}^{(k)}(f)\delta f).$$

This result reveals some interesting features of the local oscillations of affine distributions (see Fig. 10).

1. *Frequency of oscillations.* As for the Wigner–Ville distribution (see (6)), interference terms of affine distributions oscillate locally at a frequency which increases as the frequency distance $(f_2 - f_1)$ between the interfering terms is increased. Moreover, the frequency of the oscillations is also increased as the local slope $\dot{\theta}^{(k)}(f)$ is increased. This contrasts with the case of the Wigner–Ville distribution, for which $\theta^{(2)}(f)$ is linear and so $\dot{\theta}^{(2)}(f)$ is constant.

2. *Direction of oscillations.* Isocontours of $\Omega_X^{(k)}(t+\delta t, f + \delta f)$ are determined by the equiphase relation

$$\delta \mathcal{A}_X^{(k)}(\delta t, \delta f) = C,$$

where C is some constant. Again, assuming that δt and δf are both small, this is equivalent to the equation

$$\delta t = C + \dot{\theta}^{(k)}(f)\delta f,$$

which defines a set of time–frequency lines which are all parallel to the direction given by the local slope $\dot{\theta}^{(k)}(f)$. Therefore, interference terms at a point (t, f) oscillate in a direction locally orthogonal to the specific power-law group delay $\theta^{(k)}$ attached to the corresponding interfering points. Again, this generalizes in a natural way the Wigner–Ville case.

4.3. Fold Lines

It is clear from (48) that the stationary phase approximation breaks down when the second derivative of the phase $\Psi_X^{(k)}$ vanishes. We will now consider this situation, i.e., we consider stationary points for which we have simultaneously

$$\frac{\partial \Psi_X^{(k)}(u; t, f)}{\partial u} = \frac{\partial^2 \Psi_X^{(k)}(u; t, f)}{\partial u^2} = 0$$
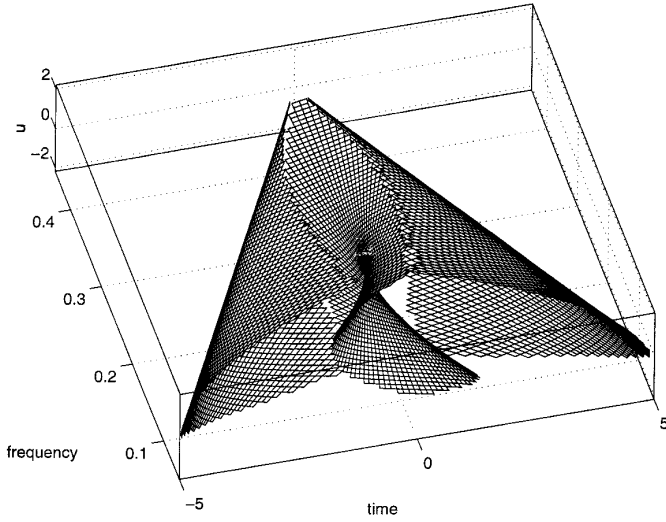
FIG. 11. Unterberger distribution of the two chirp signal: critical manifold. It consists of 3 different parts, associated respectively with autoterms (left and right parts) and cross-terms (middle part). The projection of the critical manifold onto the time–frequency plane ($u = 0$) defines stationary phase regions, with singularities in the case of vertical tangents.

and

$$\frac{\partial^3 \Psi_X^{(k)}(u;t,f)}{\partial u^3} \neq 0.$$

These points lie beneath the critical manifold $\mathscr{C}$ and, more precisely, they correspond to the locus of coalescence of two stationary points (i.e., to the vertical tangents of $\mathscr{C}$ along the u-axis). In the time-frequency plane, these points single out curves within the stationary phase domain, above which the critical manifold is folded.

Locus. Evaluating the second derivative of $\Psi_X^{(k)}(u;t,f)$ and equating the result to zero yields

$$t = \frac{\begin{aligned}&\ddot{\lambda}_k(u)t_X(\lambda_k(u)f) - \ddot{\lambda}_k(-u)t_X(\lambda_k(-u)f) \\ &\quad + f(\dot{\lambda}_k^2(u)\dot{t}_X(\lambda_k(u)f) - \dot{\lambda}_k^2(-u)\dot{t}_X(\lambda_k(-u)f))\end{aligned}}{\ddot{\lambda}_k(u) - \ddot{\lambda}_k(-u)}.$$

Because the u's we are interested in are also stationary points, (44) must apply here as well. Therefore, equating both expressions for t leads to

$$(t_X(f_1) - t_X(f_2))\frac{\ddot{\lambda}_k(u)\dot{\lambda}_k(-u) + \ddot{\lambda}_k(-u)\dot{\lambda}_k(u)}{\dot{\lambda}_k(u) + \dot{\lambda}_k(-u)}$$

$$= f(\dot{\lambda}_k^2(u)\dot{t}_X(f_2) - \dot{\lambda}_k^2(-u)\dot{t}_X(f_1)),$$

where, as previously, we have posed for convenience $f_1 = \lambda_k(-u)f$ and $f_2 = \lambda_k(u)f$.

From the geometric construction rule of interference terms, we know that there exists a power-law group delay

$\theta^{(k)}$, of degree $(k - 1)$, such that

$$t_X(f_1) = \theta^{(k)}(f_1) = t_0 + cf_1^{k-1} = t_0 + c\lambda_k^{k-1}(-u)f^{k-1}$$

$$t_X(f_2) = \theta^{(k)}(f_2) = t_0 + cf_2^{k-1} = t_0 + c\lambda_k^{k-1}(u)f^{k-1}$$

and

$$\dot{\theta}^{(k)}(f_1) = c(k-1)\lambda_k^{k-2}(-u)f^{k-2}$$

$$\dot{\theta}^{(k)}(f_2) = c(k-1)\lambda_k^{k-2}(u)f^{k-2}.$$

This allows us to write

$$t_X(f_1) - t_X(f_2) = \frac{1}{k-1}f(\lambda_k(-u)\dot{\theta}^{(k)}(f_1) - \lambda_k(u)\dot{\theta}^{(k)}(f_2)).$$

Finally, the first order singularity of the critical manifold defines fold lines, which are given by the set of time–frequency points (t, f) such that

$$\dot{\lambda}_k^2(u)\dot{t}_X(f_2) - \lambda_k(u)\Lambda_k(u)\dot{\theta}^{(k)}(f_2)$$

$$= \dot{\lambda}_k^2(-u)\dot{t}_X(f_1) - \lambda_k(-u)\Lambda_k(-u)\dot{\theta}^{(k)}(f_1), \quad (51)$$

with

$$\Lambda_k(u) = \frac{1}{1-k}\frac{\ddot{\lambda}_k(u)\dot{\lambda}_k(-u) + \ddot{\lambda}_k(-u)\dot{\lambda}_k(u)}{\dot{\lambda}_k(u) + \dot{\lambda}_k(-u)}.$$

This characterization is somewhat difficult to interpret, since it involves nontrivial weighted differences between the slope of the signal group delay t_X and the slope of the specific power-law group delay $\theta^{(k)}$ attached to the considered distribution. Nevertheless, (51) can be viewed as generalizing the way fold lines are created in the Wigner–Ville distribution. In this latter case, we have $k = 2$ and $\lambda_2(u) = 1 + \tanh u/2$, which results in $\Lambda_2(u) = 0$, thus reducing condition (51) to the well-known simpler condition of parallelism [16, 24]

$$\dot{t}_X(f_2) = \dot{t}_X(f_1).$$

More generally, note that, for any k, (51) always admits the trivial solution $u = 0$, for which $f_1 = f_2$. As for the Wigner–Ville distribution, this situation corresponds to the limit of a null length "chord" between the two interfering points, thus making of the signal group delay $t_X(f)$ itself a fold line. This behavior is illustrated in Fig. 11, which shows that the individual group delays of the two-chirp model are indeed associated with vertical tangents of the critical manifold. The signal structure appears therefore as the locus of coalescence of two interfering points, one above and one below the $u = 0$ plane, i.e., the time–frequency plane. (Let us recall that, due to the divergence of the stationary phase approximation, the distribution is expected to exhibit large amplitudes at these points.)
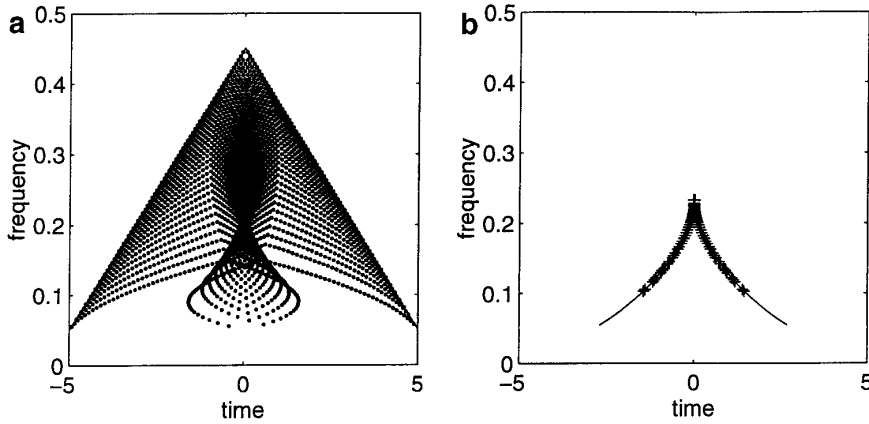
FIG. 12. Unterberger distribution of the two chirp signal: interference regions and fold lines. (a) Locus of interference points, resulting from the pointwise application of the construction rule of Fig. 5, with $k = -1$. (b) "Ghost" fold line as derived from the general construction rule (51) with $k = -1$ (crosses) and from the theoretical cuspoid (52) (line).

However, within the chosen model, it is clear from Fig. 11 that the signal structure is not the unique locus of coalescence for stationary points. For instance, between the two individual group delays, the critical manifold is twisted, with the consequence that there exist vertical tangents whose projections onto the time-frequency plane define fold lines which are distinct from those associated with the signal structure. These "ghost" fold lines arise from non-trivial solutions of (51) and can be predicted numerically (see Fig. 12). As shown in Appendix B, they can also be predicted analytically when analyzing the two-chirp signal with the active Unterberger distribution. The result is a cuspoïd, whose equation is given by

$$f = \frac{1}{2}(f_0^{2/3} - (\alpha t)^{2/3})^{3/2}. \qquad (52)$$

Figure 12 displays this theoretical prediction, together with points singled out by applying the general rule (51). It displays also the locus of interference points, as they result from a pointwise application of the geometric construction rule. Actually, this interference layer corresponds to the projection of the critical manifold onto the time–frequency plane, and the fold lines are the projected images of the folding of this manifold (i.e., of the vertical tangents of $\mathscr{C}$).

The numerical computation of the Unterberger distribution for the two-chirp signal is given in Fig. 13. It shows an excellent agreement with the theoretical predictions concerning the loci of singularities. It also reveals a fine structure for the interference terms, which will be explained further in the following.

Approximation. As for the Wigner–Ville case, the stationary phase approximation (47) breaks down beneath the singularities of $\mathscr{C}$. This drawback can be overcome by using a *uniform* approximation (see, e.g. [10]), but we will not follow this general approach here. We will rather restrict

ourselves to establishing a *transitional* approximation, in the neighborhood of an arbitrary group delay $t_X(f)$, thus generalizing (18) to arbitrary k's.

As mentioned previously, fold lines result form two coalescing stationary points ($\pm u_0 \to 0$) or, equivalently, from the interference between two neighboring time–frequency points that both belong to the signal structure ($t_X(f_1) \to t_X(f_2)$). Since the interference construction is governed by the group delay $\theta^{(k)}(f) = t_0 + cf^{k-1}$ passing through the two interfering points, this specific group delay becomes, in the limit, tangent to the signal group delay $t_X(f)$ at the interference point denoted by[6] $(\theta^{(k)}(f_0), f_0)$.

In order to establish a transitional approximation (i.e., to characterize a local behavior in the vicinity of the signal structure), we can therefore approximate the signal group delay by a power-law group delay that locally has an identical slope but a different curvature. Introducing the difference ϕ existing between $\theta^{(k)}$ and t_X in the vicinity of a given frequency f_0, we are led to the following system of equations

$$t_X(f) = \theta^{(k)}(f) + \phi(f)$$
$$\dot{t}_X(f) = \dot{\theta}^{(k)}(f) + \dot{\phi}(f)$$
$$\ddot{t}_X(f) = \ddot{\theta}^{(k)}(f) + \ddot{\phi}(f),$$

with the initial conditions

$$t_X(f_0) = \theta^{(k)}(f_0); \quad \phi(f_0) = 0$$
$$\dot{t}_X(f_0) = \dot{\theta}^{(k)}(f_0); \quad \dot{\phi}(f_0) = 0$$
$$\ddot{t}_X(f_0) = \ddot{\theta}^{(k)}(f_0) + \beta; \quad \ddot{\phi}(f_0) = \beta, \beta \in \mathbb{R}.$$

[6] This requires that the signal group delay is continuous and differentiable, which is assumed here.
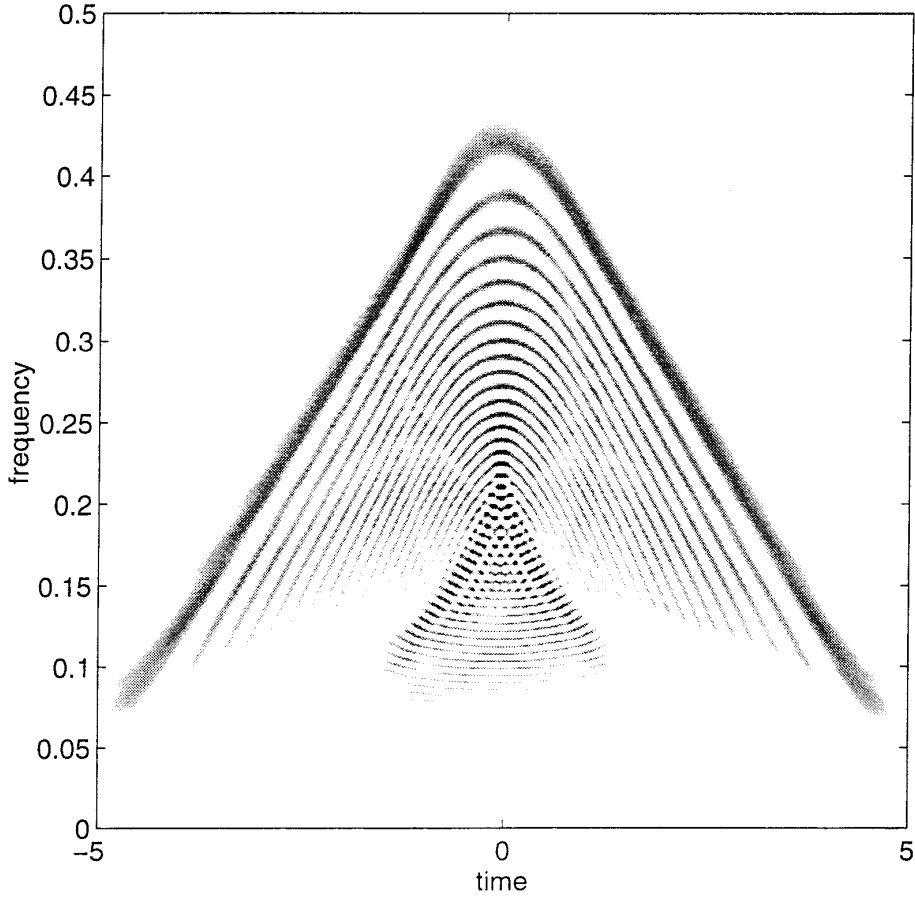
FIG. 13. Unterberger distribution of two chirp signal: numerical computation. Only the positive values of the distribution are plotted using a logarithmic dynamic scale.

The solution reads

$$\phi(f) = \frac{\beta}{2}(f - f_0)^2$$

and the companion approximation for the frequency phase of the signal is simply

$$\Phi_X(f) = -2\pi\left(t_0 f + \frac{c}{k}f^k\right) - 2\pi\frac{\beta}{6}(f - f_0)^3.$$

As a consequence,

$$\Psi_X^{(k)}(u;t,f) = \Phi_X(\lambda_k(u)f) - \Phi_X(\lambda_k(-u)f) + 2\pi t f \zeta_k(u)$$

$$= -2\pi\left(t_0\lambda_k(u)f + \frac{c}{k}\lambda_k^k(u)f^k\right) - 2\pi\frac{\beta}{6}(\lambda_k(u)f$$

$$- f_0)^3 + 2\pi\left(t_0\lambda_k(-u)f + \frac{c}{k}\lambda_k^k(-u)f^k\right)$$

$$+ 2\pi\frac{\beta}{6}(\lambda_k(-u)f - f_0)^3 + 2\pi t f \zeta_k(u)$$

$$= -2\pi f \zeta_k(u)(t_0 - t + cf^{k-1})$$

$$- 2\pi\frac{\beta}{6}\underbrace{[(\lambda_k(u)f - f_0)^3 - (\lambda_k(-u)f - f_0)^3]}_{\mathscr{H}(u;f,f_0)}.$$

Expanding $\mathscr{H}(u;f,f_0)$ and reorganizing the different terms in order to factor $\zeta_k(u)$, we get

$$\mathscr{H}(u;f,f_0)$$

$$= f^3(\lambda_k^3(u) - \lambda_k^3(-u)) - 3f^2 f_0(\lambda_k^2(u) - \lambda_k^2(-u))$$

$$+ 3f f_0^2 \zeta_k(u)$$

$$= f\zeta_k(u)[f^2(\lambda_k^2(u) + \lambda_k^2(-u) + \lambda_k(u)\lambda_k(-u))$$

$$- 3f f_0(\lambda_k(u) + \lambda_k(-u)) + 3f_0^3]$$

$$= f\zeta_k(u)[f^2\zeta_k^2(u) + 3(f^2\lambda_k(u)\lambda_k(-u) - f f_0(\lambda_k(u)$$

$$+ \lambda_k(-u)) + f_0^2)]$$

$$= f^3 \zeta_k^3(u) \left[1 + 3 \frac{(f\lambda_k(u) - f_0)(f\lambda_k(-u) - f_0)}{f^2 \zeta_k^2(u)} \right].$$

Since we are interested in the behavior of $\Psi_X^{(k)}(u; t, f)$ when u tends toward zero and f tends toward f_0, we can expand to first order the term of $\mathcal{H}$ into brackets in the vicinity of $u = 0$, which leads to

$$[\cdots]_{f=f_0} \underset{u \to 0}{\approx} 1 + 3 \frac{(\lambda_k(0) + u\dot{\lambda}_k(0) - 1)(\lambda_k(0) - u\dot{\lambda}_k(0) - 1)}{(\lambda_k(0) + u\dot{\lambda}_k(0) - \lambda_k(0) + u\dot{\lambda}_k(0))^2}$$

$$= 1 + 3 \frac{-u^2 \dot{\lambda}_k^2(0)}{4u^2 \dot{\lambda}_k^2(0)}$$

$$= \frac{1}{4}.$$

Therefore, $\Psi_X^{(k)}(u; t, f)$ can be rewritten, in the vicinity of $t_X(f)$, as

$$\Psi_X^{(k)}(u; t, f) \approx -2\pi(t_X(f) - t)f\zeta_k(u) - 2\pi \frac{\beta}{24} (f\zeta_k(u))^3$$

$$= -(t_X(f) - t) \left(\frac{32\pi^2}{\ddot{t}_X(f) - \ddot{\theta}^{(k)}(f)} \right)^{1/3} w - \frac{1}{3}w^3,$$

where we have used the changes of variables $f\zeta_k(u) = w(4/\pi\beta)^{1/3}$ and $\beta = \ddot{t}_X(f) - \ddot{\theta}^{(k)}(f)$.

Finally, the behavior of the distribution $\Omega_X^{(k)}$ can be modelled, in the neighborhood of an arbitrary signal group delay t_X, by the following form

$$\Omega_X^{(k)}(t, f) \approx \frac{1}{2\pi f} \frac{1}{\varepsilon^{(k)}(f)} \mathrm{Ai}\left(\frac{1}{\epsilon^{(k)}(f)} (t_X(f) - t) \right), \quad (53)$$

with

$$\varepsilon^{(k)}(f) = \left(\frac{\ddot{t}_X(f) - \ddot{\theta}^{(k)}(f)}{32\pi^2} \right)^{1/3}, \quad (54)$$

and where Ai stands for the Airy function (19).

This result brings to several interesting features:

1. The location and the structure of the oscillations are essentially controlled by a difference between group delay curvatures. This means in particular that the oscillations take place inside the concavity of the group delay when $\ddot{t}_X(f) - \ddot{\theta}^{(k)}(f) \geqslant 0$ and outside for the opposite case, thus justifying the behaviors reported in Fig. 7.

2. Since the argument of the Airy function is dependent on the difference of curvatures, the same oscillating structure can arise (up to a sign) from different situations. For instance, analogous interference terms will be generated by

either a unitary Bertrand distribution ($k = 0$) calculated on a linear chirp, or a Wigner–Ville distribution ($k = 2$) applied to a hyperbolic frequency modulation. The same kind of behavior is qualitatively observed in Fig. 7.

3. If we formally consider the case of the Wigner–Ville distribution ($k = 2$), the specific group delay $\ddot{\theta}^{(2)}(f)$ is a straight line and, hence, $\dddot{\theta}^{(2)}(f) = 0$, which is clearly in agreement with the result of (18).

4. If we consider the two-chirp model (39), we have $\ddot{t}_1(f) = 0$ (where $t_1(f)$ stands for the group delay of the up-going chirp) and, together with the condition $\dot{t}_1(f) = 1/\alpha$, it is easy to show that $\dddot{\theta}^{(-1)}(f) = 3/\alpha f$, thus recovering exactly for $\epsilon^{(-1)}(f)$ the result (71) given in Appendix B. The Airy fringes are clearly visible in the Unterberger distribution plotted in Fig. 13 and detailed in Fig. 14.

5. In the limiting case of a perfect match between the signal and the distribution, we get $\epsilon^{(k)}(f) = 0$ and, hence, we recover the localization property of (30) because of the identity

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon} \mathrm{Ai}\left(\frac{t}{\epsilon} \right) = 2\pi\delta(t).$$

4.4. Cusp Points

Fold lines are formed by a one-dimensional subset of the time–frequency plane that lies beneath the critical manifold and that corresponds to the coalescence of two stationary points, i.e., to the vanishing of the first and second derivatives of the phase. A further vanishing of the third derivative will now single out isolated points belonging to fold lines. These points are referred to as *cusp points* [33], and defined by the three simultaneous conditions

$$\frac{\partial \Psi_X^{(k)}(u; t, f)}{\partial u} = \frac{\partial^2 \Psi_X^{(k)}(u; t, f)}{\partial u^2} = \frac{\partial^3 \Psi_X^{(k)}(u; t, f)}{\partial u^3} = 0. \quad (55)$$

Locus. The explicit evaluation of the requirement (55) is tedious but straightforward. Since it leads to extremely cumbersome equations which are hardly interpretable, our analysis will not be reproduced here in its most general form. Rather, we will focus on the specific two-chirp example.

In the previous section, we identified (for this example, and in the case of the Unterberger distribution) the existence of a "ghost" fold line, and it is clear from Figs. 12 and 13 that this extra line exhibits a singularity, in the form of a cusp point. Considering the corresponding critical manifold, it turns out that the locus of this singularity coincides with the projection, onto the time–frequency plane, of the point where two pairs of stationary points coalesce.

This claim can be made explicit by focusing on the case $k = -1$ and overconstraining the conditions of existence of the extra fold line (i.e., (72) and (73)) by imposing the
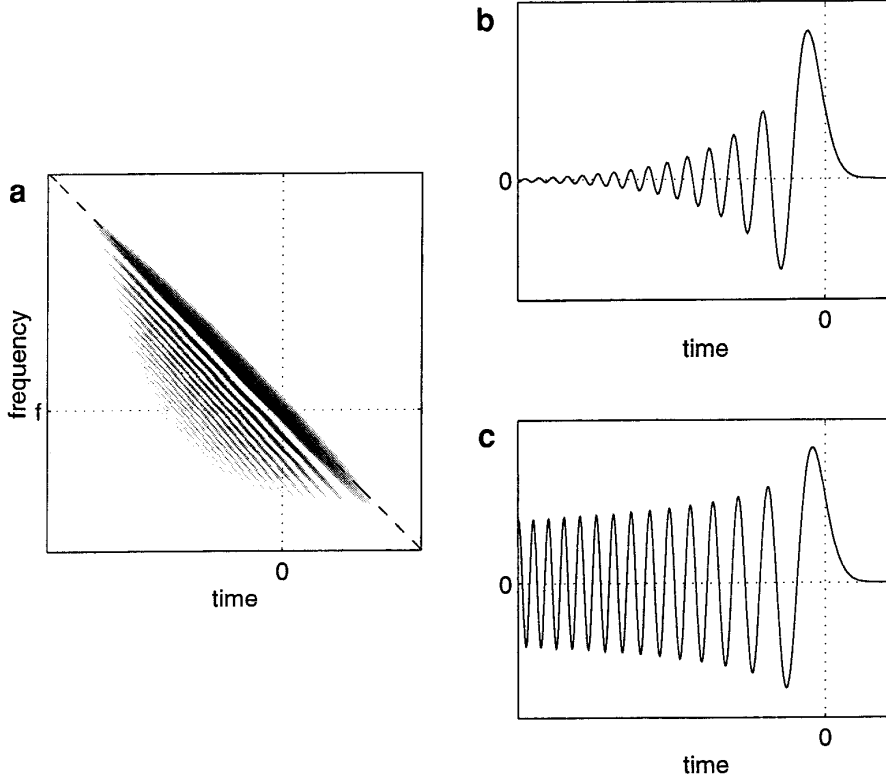
FIG. 14. Unterberger distribution of two chirp signal: transitional approximation. In the vicinity of the group delay, a transitional approximation allows us to describe the interference terms by an Airy function. Therefore, only one half of the two chirp signal described in Fig. 8, is sufficient to characterize the fine structure of interference at its vicinity. (a) Unterberger distribution of the second half (down-going chirp) of the two chirp signal. (b) Section of the Unterberger distribution at frequency f. (c) Airy function, renormalized so as to correspond to the theoretical prediction of (71).

third derivative of $\Psi_{12}^{(-1)}(u; t, f)$ to be zero as well. The corresponding system of equations reads

$$f \sinh u = f_0 \sinh \frac{u}{2} + \alpha t \cosh \frac{u}{2}$$

$$2f \cosh u = f_0 \cosh \frac{u}{2} + \alpha t \sinh \frac{u}{2}$$

$$4f \sinh u = f_0 \sinh \frac{u}{2} + \alpha t \cosh \frac{u}{2}.$$

One can easily check that assuming $u \neq 0$ yields no solution, whereas the trivial solution $u = 0$ leads to the unique cusp point defined as

$$(t_c^{(-1)}, f_c^{(-1)}) = \left(0, \frac{f_0}{2}\right).$$

Figures 12 and 13 show the excellent agreement between this theoretical result and the actual position of the apex of the cuspoid.

This result can be extended to the more general case of an arbitrary k. The general system one has to solve takes the form

$$f[\dot\lambda_k(u)\lambda_k(u) - \dot\lambda_k(-u)\lambda_k(-u)]$$
$$= f_0[\dot\lambda_k(u) - \dot\lambda_k(-u)] + \alpha t[\dot\lambda_k(u) + \dot\lambda_k(-u)]$$

$$f[\ddot\lambda_k(u)\lambda_k(u) + \dot\lambda_k^2(u) + \ddot\lambda_k(-u)\lambda_k(-u) + \dot\lambda_k^2(-u)]$$
$$= f_0[\ddot\lambda_k(u) + \ddot\lambda_k(-u)] + \alpha t[\ddot\lambda_k(u) - \ddot\lambda_k(-u)]$$

$$f[(d^3\lambda_k(u)/du^3)\lambda_k(u) + 3\ddot\lambda_k(u)\dot\lambda_k(u)$$
$$- 3\ddot\lambda_k(-u)\dot\lambda_k(-u) - (d^3\lambda_k(-u)/du^3)\lambda_k(-u)]$$
$$= f_0[d^3\lambda_k(u)/du^3 - d^3\lambda_k(-u)/du^3]$$
$$+ \alpha t[d^3\lambda_k(u)/du^3 + d^3\lambda_k(-u)/du^3].$$

and always admits the trivial solution[7]

$$u = 0 \Rightarrow t_c^{(k)} = 0,$$

so that the frequency position of the cusp point can be derived from the second equation, leading to

$$f_c^{(k)} = \left(1 + \frac{\dot\lambda_k^2(0)}{\ddot\lambda_k(0)}\right)^{-1} f_0.$$

---

[7] In contrast with the special case $k = -1$, the unicity of this solution has not been proved explicitly. Nevertheless, experiments indicate that the unique singularity of the cuspojd systematically occurs for $u = 0$ and is located at $t = 0$.

A straightforward calculation shows that

$$\frac{\dot{\lambda}_k^2(0)}{\ddot{\lambda}_k(0)} = \frac{3}{2-k},$$

from which we deduce that the frequency location of the cusp point (the apex of the "ghost" fold line) is given by the general expression

$$f_c^{(k)} = \frac{2-k}{5-k} f_0. \tag{56}$$

Naturally, this form simply reduces to $f = f_0/2$ when the particular value $k = -1$ of the example is considered.

Approximation. The conditions under which the stationary phase approximation breaks down beneath the first order singularities of $\mathscr{C}$ remain effective at cusp points. Moreover, in this case, they extend to second-order singularities and it is the Airy approximation itself which is no longer valid. Uniform approximation theory [10, 33] tells us that, in the vicinity of cusp points, the Airy function describing the oscillatory structure of $\Omega_X^{(k)}$ has to be replaced by a *Pearcey* function, defined by the "cusp canonical integral"

$$\mathrm{Pe}(\xi, \eta) = \int_{-\infty}^{\infty} e^{i(u^4/4 + \xi(u^2/2) + \eta u)} du. \tag{57}$$

It is shown in Appendix B how a transitional approximation brings out such a Pearcey structure by reducing (57) to a "canonical integral" [33]. The result is that, in the neighborhood of the cusp point $(t_c^{(-1)} = 0, f_c^{(-1)} = f_0/2)$, the Unterberger distribution can be approximated by

$$\Omega_X^{(-1)}(t, f) \approx \frac{\kappa^{(-1)}(f)}{\pi f} \mathrm{Re} \left\{ e^{i(2\pi/\alpha)(f-f_0)^2} \right.$$
$$\left. \times \mathrm{Pe}\left(\frac{[\kappa^{(-1)}(f)]^2}{2\pi\alpha f} \left(f - \frac{f_0}{2} \right), \kappa^{(-1)}(f)t \right) \right\}, \tag{58}$$

with

$$\kappa^{(-1)}(f) \equiv 2 \left(\frac{8\alpha\pi^3 f^3}{f_0} \right)^{1/4}. \tag{59}$$

We can verify the consistency of this approximation by comparing the "canonical Pearcey cuspöjd" (or "caustic"), defined by [10]

$$4\xi^3 = 27\eta^2, \tag{60}$$

with the local structure of the "ghost" fold line (52) in the vicinity of its apex. In the case of (58), the canonical cuspöjd (60) becomes

$$f = \frac{f_0}{2} \left[1 - \frac{3}{2} \left(\frac{\alpha t}{f_0} \right)^{2/3} \right],$$

which is precisely the local approximation of the extra fold line (52) when t is small. The excellent agreement between these theoretical predictions and the numerical computation is illustrated in Fig. 15, which compares a Pearcey function with the local behavior of the Unterberger distribution in the vicinity of its cusp point.

The approach of Appendix B for the Unterberger distribution can be reproduced, step by step, in the more general case of arbitrary k. Expanding the phase
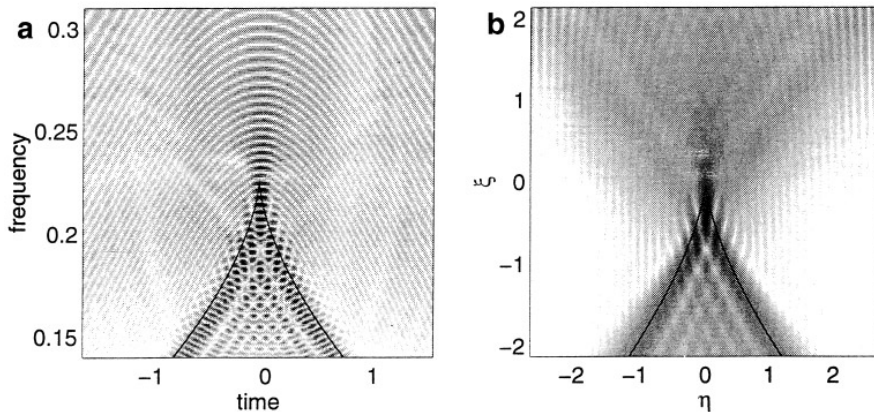

FIG. 15. Unterberger distribution of the two chirp signal: detail of the cusp. (a) Enlargement of the central part of Fig. 13 (in this case, the magnitude of the distribution is plotted using a linear dynamic scale). (b) Magnitude of the Pearcey function, renormalized so as to correspond to the theoretical prediction of (58). In both cases, the caustic (60) has been superimposed (solid line). One clearly observes an excellent agreement between the two diagrams, up to the modulation induced in (a) by the additional quadratic phase appearing in (58).

$$\Psi_{12}^{(k)}(u;t,f) = -\frac{\pi}{\alpha}[(\lambda_k(u)f - f_0)^2$$

$$+ (\lambda_k(-u)f - f_0)^2 - 2\alpha tf\zeta_k(u)],$$

in the vicinity of $u = 0$ (solution of the cusp location) up to fourth order yields

$$\Psi_{12}^{(k)}(u;t,f) \approx -\frac{2\pi}{\alpha}(f - f_0)^2 + 4\pi tf\dot{\lambda}_k(0)u$$

$$- \frac{2\pi f}{\alpha}(\ddot{\lambda}_k(0) + \dot{\lambda}_k^2(0))(f - f_c^{(k)})u^2 + \frac{\pi ff_0}{2\alpha}Au^4,$$

with

$$A \equiv \frac{\dot{\lambda}_k^2(0)d^4\lambda_k(0)/du^4 - 3\ddot{\lambda}_k^2(0) - 4\dot{\lambda}_k(0)\ddot{\lambda}_k(0)d^3\lambda_k(0)/du^3}{3(\ddot{\lambda}_k(0) + \dot{\lambda}_k^2(0))}$$

a constant. The change of variables

$$\frac{\pi ff_0}{2\alpha}|A|u^4 = \frac{1}{4}y^4 \Leftrightarrow u = \left(\frac{\alpha}{2\pi ff_0|A|}\right)^{1/4}y,$$

leads to

$$\Psi_{12}^{(k)}(u(y);t,f) \approx -\frac{2\pi}{\alpha}(f - f_0)^2$$

$$+ 2t\dot{\lambda}_k(0)\left(\frac{8\alpha\pi^3 f^3}{f_0|A|}\right)^{1/4}y - 4(\ddot{\lambda}_k(0) + \dot{\lambda}_k^2(0))$$

$$\times (f - f_c^{(k)})\left(\frac{\pi f}{2\alpha f_0|A|}\right)^{1/2}\frac{y^2}{2} + \text{sign}(A)\frac{y^4}{4}.$$

Therefore, in the vicinity of its cusp points, the local behavior of any distribution $\Omega_X^{(k)}$ can be approximated in terms of a Pearcey function by

$$\Omega_X^{(k)}(t,f) \approx \frac{1}{\dot{\lambda}_k(0)}\frac{\kappa^{(k)}(f)}{2\pi f}\text{Re}\left\{e^{i(2\pi/\alpha)(f-f_0)^2}\right.$$

$$\left.\times \text{Pe}\left(\frac{[\kappa^{(k)}(f)]^2}{2\pi\alpha^{(k)}f}(f - f_c^{(k)}), \kappa^{(k)}(f)t\right)\right\}, \quad (61)$$

with

$$\kappa^{(k)}(f) \equiv 2\dot{\lambda}_k(0)\left(\frac{8\alpha\pi^3 f^3}{f_0|A|}\right)^{1/4} \quad (62)$$

and

$$\frac{1}{\alpha^{(k)}} \equiv \frac{1}{2\alpha}\left(1 + \frac{\ddot{\lambda}_k(0)}{\dot{\lambda}_k^2(0)}\right).$$

The Unterberger specific case is such that $A = -\frac{1}{16}$, thus reducing (62) to (59) and (61) to (58).

4.5. Higher-order Singularities and Symmetries

The previous classification of singularities in terms of stationary phase *regions,* fold *lines* and cusp *points* almost exhaust the typical behaviors which can be observed in affine distributions. However, it must be noted that degenerate situations can occur at points associated with higher-order singularities.

In the case of the Wigner–Ville distribution, it is known [16, 24] that degeneracies are related to the existence of
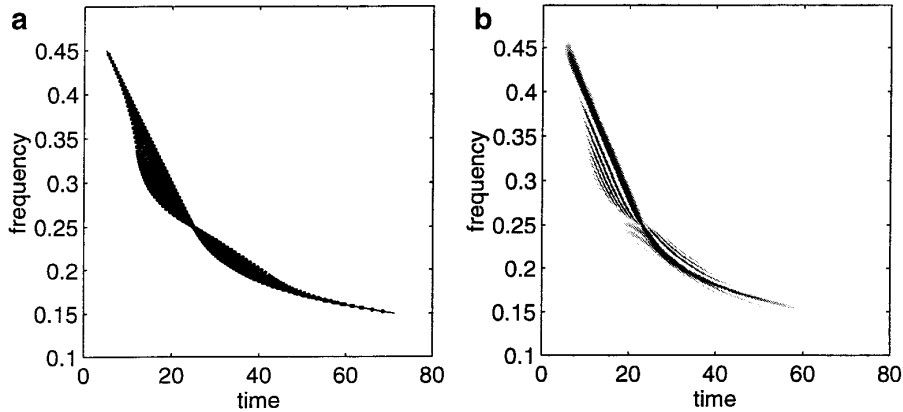

FIG. 16. Symmetry. Higher-order singularities can be observed in affine distributions when the signal structure presents "symmetries" in the sense of the geometry. The diagrams present an example of such a situation in the case $k = -1$. (a) Signal model and locus of interference. (b) Numerical computation of the corresponding Unterberger distribution. Both diagrams demonstrate the existence of one point of accumulation which is "center" of an infinite number of "power-law chords," thus leading to a large amplitude at this point.
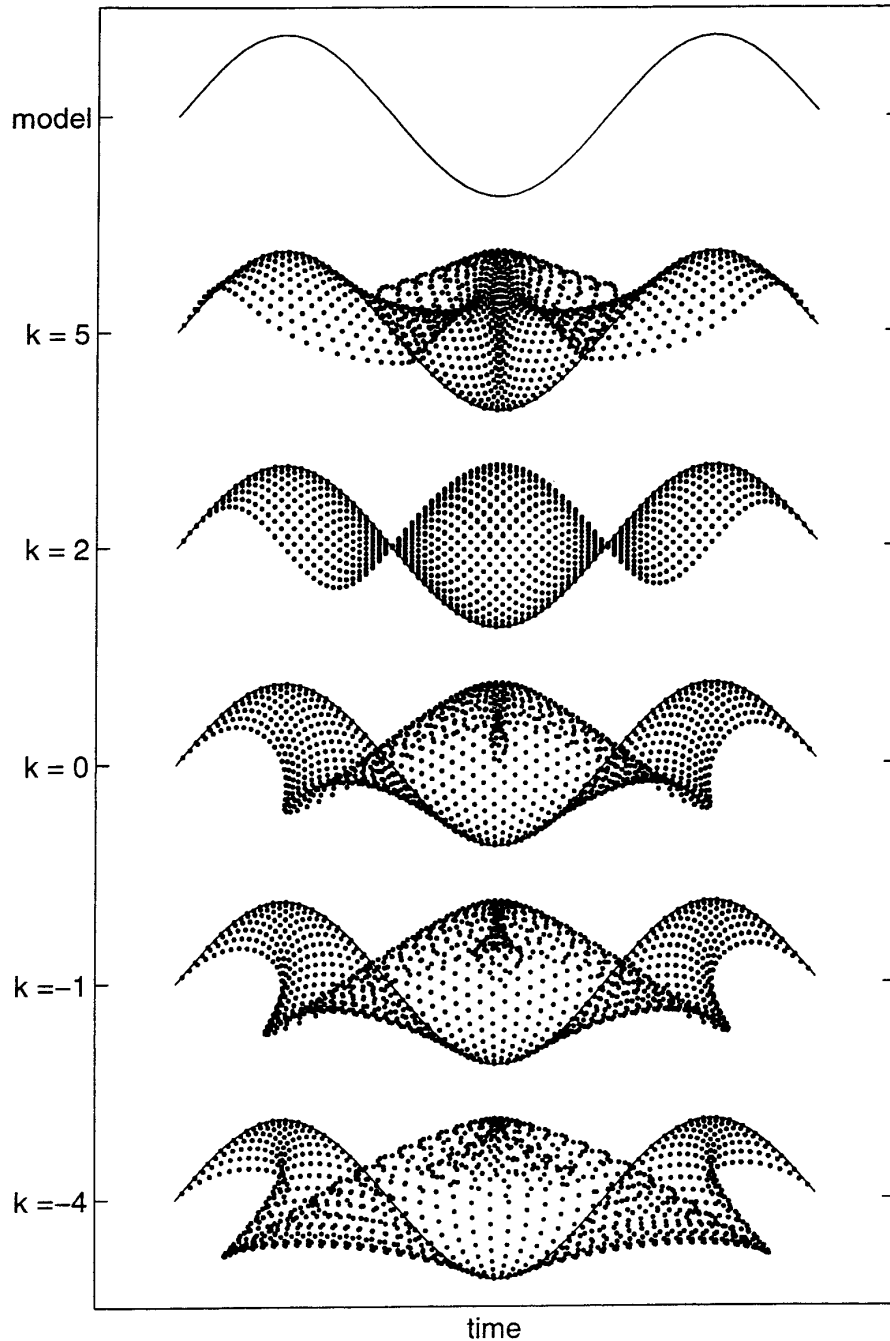
FIG. 17. Interference diagrams for a sinusoidal frequency modulation. In the case of any nonlinear frequency modulation, the pointwise application of the construction rules of Fig. 5 can be used as a way of predicting the structure of the corresponding affine distribution. This is illustrated here with the example of a sinusoidal frequency modulation (whose model is given at the top of the diagram) with various values of k.

a (skew-)symmetry in the time–frequency structure of the analyzed signal, because this results in a singular point (the center of symmetry) which is the midpoint of an infinite number of chords. In the case of affine distributions, a similar situation is likely to be observed, provided that the concept of "symmetry" is suitably modified. This amounts to considering time–frequency structures for which a given

time–frequency point can be viewed as the "midpoint," in the sense of the rule (36), of an infinite number of "power-law chords" $\theta^{(k)}(f)$.

The construction of examples of this type can be easily achieved when the rule $\Theta^{(k)}$ also defines a Kolmogorov–Nagumo generalized mean (37). In this case, given a point ω considered as a "center" and assuming that there exist
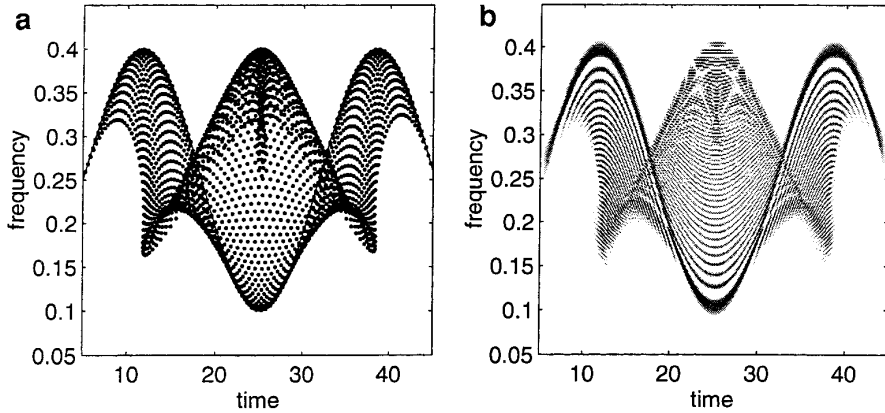
FIG. 18.   Bertrand distribution for a sinusoidal frequency modulation. The effectiveness of the predictions of Fig. 16 can be checked by a comparison with the result of a numerical computation, as illustrated here in the case of the Bertrand distribution ($k = 0$). (a) Prediction. (b) Numerical computation.

two points ω_1 and ω_2 such that $\omega = \Theta^{(k)}(\omega_1, \omega_2)$, we readily obtain

$$\omega_1 = g^{-1}[2g(\omega) - g(\omega_2)],$$

where g stands for the nonlinear function attached to the generalized mean. This defines a point transformation which is involutive and which can be considered as a generalized central symmetry.[8] If $k = 2$ (Wigner–Ville), we end up with the usual *central symmetry*

$$\omega_1 = 2\omega - \omega_2,$$

whereas, if $k = -1$ (Unterberger), the resulting operation takes the form of an *inversion:*

$$\omega_1 = \frac{\omega^2}{\omega_2}.$$

More precisely, the symmetry rule (11), which applies to the Wigner–Ville distribution, can be generalized in the following way: given a time–frequency point (t, f), the symmetric image of any point (t_1, f_1) is given by the point (t_2, f_2) such that

$$t_2 = t + (t_1 - t)\frac{f_2^{k-1} - f^{k-1}}{f_1^{k-1} - f^{k-1}},$$

with

$$f_2 = g^{-1}[2g(f) - g(f_1)].$$

Symmetric time–frequency structures can therefore be constructed by applying the above rule to time–frequency curves, thus making of the chosen center of symmetry a

[8] This interpretation is closely related to the fact that symmetry can be used as a constructive argument for defining time–frequency distributions. In particular, it is shown in [34] that the Wigner–Ville distribution can be expressed as the expectation value of a parity operator whereas, in [20], the authors use a frequency inversion operator for defining a distribution which turns out to coincide with the Unterberger distribution.

point of accumulation for an infinite number of "power-law chords." An example of such a degenerate situation is given in Fig. 16. (Of course, localized distributions are another example of degeneracy, since their "matched" group delay is the locus of all of its centers of symmetry.)

5. CONCLUSION

From the point of view of their geometry, affine distributions have been shown to obey construction rules that generalize most of the results known in the case of the Wigner–Ville distribution. These rules (which however are analytically much more intricate) have been identified by the use of both approximations and specific models, but it is clear that their applicability remains qualitatively valid in more general situations. For example, Fig. 17 provides predicted interference diagrams, as derived from the rules established in this paper, in the case of a chirping signal with a sinusoidal frequency modulation. As an illustration of the accuracy of the prediction, Fig. 18 compares theory with the numerical computation of the corresponding distribution (in this case, the unitary Bertrand distribution). One clearly observes a good agreement between both diagrams, thus making the geometrical approach an ingredient which is believed to be of key importance for a better understanding of the fine structure of an affine distribution of an analyzed signal.

A. MIDPOINTS AND GENERALIZED MEANS

Given the definition (37), a result due to de Finetti [12] states that a two-dimensional function $M(\omega_1, \omega_2)$ is a (Kolmogorov–Nagumo) generalized mean if and only if

$$\frac{\partial M(\omega_1, \omega_2)/\partial \omega_1}{\partial M(\omega_1, \omega_2)/\partial \omega_2} = \frac{\varphi(\omega_1)}{\varphi(\omega_2)}, \tag{63}$$

where

$$\varphi(\omega) = \frac{dg(\omega)}{d\omega}. \tag{64}$$

For the specific rule given in (36), we can verify that

$$\frac{\partial \Theta^{(k)}(\omega_1, \omega_2/\partial \omega_1}{\partial \Theta^{(k)},(\omega_1, \omega_2)/\partial \omega_2} = \frac{h(\omega_1, \omega_2)}{h(\omega_2, \omega_1)},$$

where

$$h(\omega_1, \omega_2) = (k-1)\omega_1^k - k\omega_1^{k-1}\omega_2 + \omega_2^k.$$

Assuming that $\omega_1 = \alpha\omega_2 (\alpha > 0)$, we get

$$\frac{\partial \Theta^{(k)}(\alpha\omega_2, \omega_2)/\partial \omega_1}{\partial \Theta^{(k)}(\omega_2, \alpha\omega_2)/\partial \omega_2} = \psi(\alpha),$$

with

$$\psi(\alpha) = \frac{(k-1)\alpha^k - k\alpha^{k-1} + 1}{(k-1) - k\alpha + \alpha^k}. \tag{65}$$

This implies that the ratio involved in (63) must satisfy

$$\varphi(\alpha\omega_2) = \psi(\alpha)\varphi(\omega_2), \tag{66}$$

and therefore

$$\varphi(\alpha\beta\omega_2) = \psi(\alpha)\varphi(\beta\omega_2)$$

$$= \psi(\alpha)\psi(\beta)\varphi(\omega_2)$$

$$\varphi(\alpha\beta\omega_2) = \psi(\alpha\beta)\varphi(\omega_2).$$

It follows that ψ must verify the equation

$$\psi(\alpha\beta) = \psi(\alpha)\psi(\beta),$$

the (continuous) solution of which is of the form [2]

$$\psi(\alpha) = \alpha^c, \quad c \in \mathbb{R}. \tag{67}$$

Replacing this result in (65), the problem reduces to finding the k's which are solution of the equation

$$(k-1)\alpha^k - k\alpha^{k-1} - (k-1)\alpha^c + k\alpha^{c+1} - \alpha^{k+c} + 1 = 0. \tag{68}$$

This equation must hold for any value of α, imposing the joint choice of k and c to balance the constant 1. The trivial solutions $k = 0$ and $k = 1$ do not match since both give rise

to $\varphi(\omega) = 0$ and, from (64), to the solution $g(\omega) = $ constant. This latter is not consistent with the required monotony of g, imposed to ensure its reversibility. Finally, the only possible values for c are $c = 0, c = -1$ and $c = -k$. Let us consider each of these cases:

• $c = 0$.

This value leads to the equation

$$(k-2)\alpha^k - k\alpha^{k-1} + k\alpha = 0,$$

and thus, to the solution $k = 2$.

• $c = -1$.

In this case, we have

$$(k-1)\left(\alpha^k - \frac{1}{\alpha}\right) - (k+1)\alpha^{k-1} = 0,$$

with the solution $k = -1$.

• $c = -k$.

Here, the condition reads

$$k\alpha^{1-k} - (1-k)\alpha^k = k\alpha^{-(1-k)} - (1-k)\alpha^{-k},$$

which imposes the equality $k = 1 - k$, and thus the solution $k = \frac{1}{2}$ (associated with $c = -\frac{1}{2}$).

The knowledge of the values of k for which (36) defines a Kolmogorov–Nagumo mean allows us to identify the exact type of these generalized means, i.e., the function $g(\omega)$. From (66) and (67), one can write

$$\varphi(\alpha\omega) = \alpha^c\varphi(\omega).$$

Differentiating this equation with respect to α leads to

$$\frac{d\varphi}{\varphi} = c\frac{d\omega}{\omega} \Rightarrow \varphi(\omega) = a\omega^c$$

whose solution, using the result of (64), is given by

$$g_c(\omega) = A\omega^{c+1} + B, c \neq -1$$

$$g_{-1}(\omega) = A\log\omega,$$

where A and B are real constants.

We end up therefore with the following Kolmogorov–Nagumo generalized means:

• $c = 0, k = 2$ (arithmetic mean)

$$g_0(\omega) = \omega \Rightarrow \Theta^{(2)}(\omega_1, \omega_2) = \frac{\omega_1 + \omega_2}{2}.$$

• $c = -1, k = -1$ (geometric mean)

$$g_{-1}(\omega) = \log\omega \Rightarrow \Theta^{(-1)}(\omega_1, \omega_2) = e^{(1/2)(\log\omega_1 + \log\omega_2)} = \sqrt{\omega_1\omega_2}.$$

• $c = -\frac{1}{2}, k = \frac{1}{2}$ (square root mean)

$$g_{-1/2}(\omega) = \sqrt{\omega} \Rightarrow \Theta^{(1/2)}(\omega_1, \omega_2) = \left(\frac{\sqrt{\omega_1} + \sqrt{\omega_2}}{2}\right)^2$$

B. SINGULARITIES OF THE UNTERBERGER DISTRIBUTION IN THE CASE OF A TWO CHIRP SIGNAL

B.1. Signal Model

The group delay of the two chirp signal model is defined by (39), thus corresponding to the phase response

$$\Phi_X(f) = -2\pi \int_{-\infty}^{f} t_X(\nu)d\nu$$

$$= \begin{cases} \Phi_1(f) = -\frac{\pi}{\alpha}(f - f_0)^2, \ t_X(f) \leq 0 \\ \Phi_2(f) = +\frac{\pi}{\alpha}(f - f_0)^2, \ t_X(f) \geq 0 \end{cases}.$$

B.2. Singularities Attached to Auto-Terms

Auto-Term Phase. Given the model (39), two phase terms are associated with the "auto terms," i.e., interference terms between points belonging to the same group delay. Symmetry arguments allow one to study only one of these phase terms, and we will therefore consider only

$$\Psi_{11}^{(k)}(u; t, f) \equiv \Phi_1(\lambda_k(u)f) - \Phi_1(\lambda_k(-u)f) + 2\pi f t \zeta_k(u)$$

$$= -\frac{\pi}{\alpha}[(\lambda_k(u)f - f_0)^2 - (\lambda_k(-u)f - f_0)^2 - 2\alpha t f \zeta_k(u)].$$

Critical Manifold. The critical manifold $\mathscr{C}_1$, which is characterized by the condition $\partial\Psi_{11}^{(k)}(u; t, f)/\partial u = 0$, is defined by

$$f(\dot{\lambda}_k(u)\lambda_k(u) + \dot{\lambda}_k(-u)\lambda_k(-u))$$

$$- (f_0 + \alpha t)(\dot{\lambda}_k(u) + \dot{\lambda}_k(-u)) = 0.$$

In the special case of the Unterberger distribution, for which $k = -1$, λ_k is such that

$$\lambda_{-1}(u) = e^{u/2} \Rightarrow \dot{\lambda}_{-1}(u) = \frac{1}{2}e^{u/2},$$

so that the points that lie beneath $\mathscr{C}_1$ are those for which

$$f \cosh u = (f_0 + \alpha t) \cosh \frac{u}{2}. \tag{69}$$

Fold Lines and Local Approximation. Fold lines are defined by the points (t, f) which verify (69), with a further vanishing of the second derivative of the phase $\Psi_{11}^{(k)}$. This second condition leads to

$$f \sinh u = \frac{1}{2}(f_0 + \alpha t) \sinh \frac{u}{2} \tag{70}$$

which, together with (69), imposes

$$\tanh u = \frac{1}{2} \tanh \frac{u}{2}.$$

The only solution of this equation is $u = 0$ and, making use again of (69), it corresponds to

$$t = \frac{f - f_0}{\alpha} = t_1(f),$$

which indicates that the group delay is a fold line of the distribution in the time–frequency plane. In the vicinity of this fold line, a transitional approximation can be derived by expanding the phase $\Psi_{11}^{(-1)}$ up to third order with respect to u. Starting from

$$\Psi_{11}^{(-1)}(u; t, f) = -\frac{2\pi f}{\alpha}\left[f \sinh u - 2(f_0 + \alpha t)\sinh\frac{u}{2}\right]$$

and using the approximation

$$\sinh u \approx u + \frac{u^3}{6},$$

we get

$$\Psi_{11}^{(-1)}(u; t, f) \approx -\frac{2\pi f}{\alpha}(f - (f_0 + \alpha t))u$$

$$- \frac{1}{3}\frac{2\pi f}{2\alpha}\left(f - \frac{1}{4}(f_0 + \alpha t)\right)u^3.$$

This can be put in the form

$$\Psi_{11}^{(-1)}(u; t, f) \approx yv + \frac{1}{3}v^3,$$

by introducing the change of variables

$$v \equiv -\left(\frac{\pi f}{\alpha}\left(f - \frac{1}{4}(f_0 + \alpha t)\right)\right)^{1/3}u$$

and

$$y \equiv \frac{2\pi f}{\alpha}(f - (f_0 + \alpha t))\left(\frac{\pi f}{\alpha}\left(f - \frac{1}{4}(f_0 + \alpha t)\right)\right)^{-1/3}.$$

Because we are interested in the local behavior of the Unterberger distribution $U^a = \Omega^{(-1)}$ in the vicinity of the fold line $f = f_0 + \alpha t$, we can further consider that

$$f - \frac{1}{4}(f_0 + \alpha t) \approx \frac{3}{4}f$$

and, hence, that

$$y \approx \left(\frac{32\pi^2}{3/\alpha f} \right)^{1/3} (t_1(f) - t).$$

Using this result and the definition (19) of the Airy function, this corresponds finally to the approximation

$$U_1^a(t, f) \approx \frac{1}{2\pi f} \frac{1}{\epsilon^{(-1)}(f)} \mathrm{Ai} \left(\frac{1}{\epsilon^{(-1)}(f)} (t_1(f) - t) \right),$$

with

$$\epsilon^{(-1)}(f) = \left(\frac{3/\alpha f}{32\pi^2} \right)^{1/3}. \qquad (71)$$

B.3. Singularities Attached to Cross-Terms

Cross-Term Phase. The phase of (47) which is associated with the cross term reads[9]

$$\Psi_{12}^{(k)}(u; t, f) \equiv \Phi_1(\lambda_k(u)f) - \Phi_2(\lambda_k(-u)f) + 2\pi f t \zeta_k(u)$$

$$= -\frac{\pi}{\alpha} [(\lambda_k(u)f - f_0)^2 + (\lambda_k(-u)f - f_0)^2 - 2\alpha t f \zeta_k(u)].$$

Critical Manifold. The corresponding *critical manifold* $\mathscr{C}_{12}$, characterized by the condition $\partial \Psi_{12}^{(k)}(u; t, f)/\partial u = 0$, is therefore defined by

$$f(\dot{\lambda}_k(u)\lambda_k(u) - \dot{\lambda}_k(-u)\lambda_k(-u)) - f_0(\dot{\lambda}_k(u)$$

$$- \dot{\lambda}_k(-u)) - \alpha t(\dot{\lambda}_k(u) + \dot{\lambda}_k(-u)) = 0.$$

The special case of this general equation, associated with the value $k = -1$ of the Unterberger distribution, reads

$$f \sinh u = f_0 \sinh \frac{u}{2} + \alpha t \cosh \frac{u}{2} \qquad (72)$$

or, equivalently, when $u \neq 0$,

$$f = \frac{1}{2} \left[\frac{f_0}{\cosh \frac{u}{2}} + \frac{\alpha t}{\sinh \frac{u}{2}} \right].$$

"Ghost" Fold Lines. Second-order singularities of the phase $\Psi_{12}^{(k)}$ are characterized by the relation $\partial^2 \Psi_{12}^{(k)}(u; t, f)/\partial u^2 = 0$, leading to

$$f \cosh u = \frac{1}{2} \left(f_0 \cosh \frac{u}{2} + \alpha t \sinh \frac{u}{2} \right) \qquad (73)$$

[9] In fact, one also has to consider the phase $\Psi_{21}^{(k)}(u; f, t) \equiv \Phi_2(\lambda_k(u)f) - \Phi_1(\lambda_k(-u)f) + 2\pi f t \zeta_k(u)$. This leads to similar results, whose derivation is not reproduced here.

when $k = -1$. Imposing simultaneously (72) and (73), we get, after identification of both values of f and the change of variable $\gamma = \tanh u/2$,

$$\frac{1 + \gamma^2}{1 - \gamma^2} \left(f_0 \sqrt{1 - \gamma^2} + \alpha t \frac{\sqrt{1 - \gamma^2}}{\gamma} \right)$$

$$= f_0 \frac{1}{\sqrt{1 - \gamma^2}} + \alpha t \frac{\gamma}{\sqrt{1 - \gamma^2}},$$

which reduces to

$$(1 + \gamma^2) \left(f_0 + \frac{\alpha t}{\gamma} \right) = f_0 + \alpha t.$$

The solution is

$$\gamma = - \left(\frac{\alpha t}{f_0} \right)^{1/3}$$

and, substituting this result in (72), we finally get

$$f = \frac{1}{2} \sqrt{1 - \gamma^2} \left(f_0 + \frac{\alpha t}{\gamma} \right)$$

$$= \frac{1}{2} f_0 \sqrt{1 - \left(\frac{\alpha t}{f_0} \right)^{2/3}} \left(1 - \left(\frac{\alpha t}{f_0} \right)^{2/3} \right)$$

$$= \frac{1}{2} f_0 \left(1 - \left(\frac{\alpha t}{f_0} \right)^{2/3} \right)^{3/2}$$

$$= \frac{1}{2} (f_0^{2/3} - (\alpha t)^{2/3})^{3/2}.$$

Cusp Points and Local Approximation. The cusp point can be identified as the apex of the cuspoïd (52). As before, instead of using a uniform approximation in the vicinity of this point, we can restrict ourselves to a transitional approximation, which amounts to approximating the phase $\Psi_{12}^{(-1)}(u; t, f)$ by a suitable fourth-order polynomial in order to reduce (57) to a "canonical integral" [33]. A natural way to do so is to expand the phase term (with respect to u) in the vicinity of $u = 0$, which is the solution associated to the cusp point. Starting from

$$\Psi_{12}^{(-1)}(u; t, f) = -\frac{\pi}{\alpha} [(e^{u/2}f - f_0)^2$$

$$+ (e^{-u/2}f - f_0)^2 - 2\alpha t f (e^{u/2} - e^{-u/2}) = -\frac{2\pi}{\alpha}$$

$$\times \left[f^2 \cosh u + f_0^2 - 2f f_0 \cosh \frac{u}{2} - 2\alpha t f \sinh \frac{u}{2} \right]$$

and expanding this expression so as to take into account all the necessary terms up the fourth order, we get the following approximation

$$\Psi_{12}^{(-1)}(u;t,f)$$

$$\approx -\frac{2\pi}{\alpha}\left[f^2\left(1+\frac{u^2}{2}+\frac{u^4}{24}\right)+f_0^2-2ff_0\right.$$

$$\left.\times\left(1+\frac{1}{2}\frac{u^2}{4}+\frac{1}{16}\frac{u^4}{24}\right)-2\alpha tf\left(\frac{u}{2}+\frac{1}{6}\frac{u^3}{8}\right)\right]$$

$$=-\frac{2\pi}{\alpha}\left[(f-f_0)^2-\alpha tfu+\frac{u^2}{2}f\left(f-\frac{f_0}{2}\right)\right.$$

$$\left.-\frac{u^3}{24}\alpha tf+\frac{u^4}{24}f\left(f-\frac{f_0}{8}\right)\right]$$

$$=-\frac{2\pi}{\alpha}(f-f_0)^2+2\pi tf\underbrace{\left(1+\frac{u^2}{24}\right)}_{\epsilon_1}u-\frac{\pi}{\alpha}f$$

$$\times\left(f-\frac{f_0}{2}\right)\underbrace{\left(1+\frac{u^2}{12}\right)}_{\epsilon_2}u^2-\frac{\pi}{32\alpha}ff_0u^4.$$

Since we are merely interested in a local approximation in the vicinity of the cusp point $(t_c^{(-1)}=0, f_c^{(-1)}=f_0/2)$, we can reasonably neglect the higher-order terms ϵ_1 and ϵ_2. Within this approximation, and setting the change of variables

$$\frac{\pi}{32\alpha}ff_0u^4=\frac{y^4}{4}\Leftrightarrow u=\left(\frac{8\alpha}{\pi ff_0}\right)^{1/4}y,$$

the polynomial phase approximation becomes

$$\Psi_{12}^{(-1)}(u(y);t,f)\approx -\frac{2\pi}{\alpha}(f-f_0)^2+2t\left(\frac{8\alpha\pi^3 f^3}{f_0}\right)^{1/4}$$

$$\times y-4\left(f-\frac{f_0}{2}\right)\left(\frac{2\pi f}{\alpha f_0}\right)^{1/2}\frac{y^2}{2}-\frac{y^4}{4},$$

with a similar result for the companion phase $\Psi_{21}^{(-1)}$.

Summing the two contributions due to $\Psi_{12}^{(-1)}$ and $\Psi_{21}^{(-1)}$, and using the fact that $\mathrm{Pe}(\xi,-\eta)=\mathrm{Pe}(\xi,\eta)$, the Unterberger distribution can finally be approximated, in the neighbor-

hood of the cusp point $(t_c^{(-1)}=0, f_c^{(-1)}=f_0/2)$, by

$$U_X^a(t,f)\approx\frac{\kappa^{(-1)}(f)}{\pi f}\mathrm{Re}\left\{e^{i(2\pi/\alpha)(f-f_0)^2}\right.$$

$$\left.\times\mathrm{Pe}\left(\frac{[\kappa^{(-1)}(f)]^2}{2\pi\alpha f}\left(f-\frac{f_0}{2}\right),\kappa^{(-1)}(f)t\right)\right\},$$

with

$$\kappa^{(-1)}(f)\equiv 2\left(\frac{8\alpha\pi^3 f^3}{f_0}\right)^{1/4}.$$

ACKNOWLEDGMENTS

Note added in proof. After completion of the manuscript, we discovered that the midpoint function (33–36) exactly coincides with a specific form of (weighted) generalized mean proposed in [36]—without any connection to the present context—as a way to generalize the usual logarithmic mean.

REFERENCES

1. J. Aczél, On mean values, *Bull. Amer. Math. Soc.* **54** (1948), 393–400.

2. J. Aczél and Z. Daróczy, "On Measures of Information and their Characterizations," Academic Press, New York, 1975.

3. R. G. Baraniuk and D. L. Jones, Warped wavelet bases—Unitary equivalence and signal processing, "IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP-93," pp. III.320–III.323, Minneapolis, 1993.

4. M. V. Berry, Semiclassical mechanics in phase space—A study of Wigner's function, *Phil. Trans. Roy. Soc. A* **287** (1977), 237–271.

5. J. Bertrand and P. Bertrand, Représentations temps-fréquence des signaux, *C. R. Ac. Sc. Paris* **299** (1984), 635–638.

6. J. Bertrand and P. Bertrand, A class of affine Wigner distributions with extended covariance properties, *J. Math. Phys.* **33**, No. 7 (1992), 2515–2527.

7. T. A. C. M. Claasen and W. F. G. Mecklenbräuker, The Wigner distribution—A tool for time–frequency signal analysis, Parts I–III, *Philips J. Res.* **35** (1980), 217–250, 276–300, 372–389.

8. L. Cohen, Generalized phase-space distribution functions, *J. Math. Phys.* **7**, No. 5 (1966), 781–786.

9. L. Cohen, Time–frequency distributions—A review, *Proc. IEEE* **77**, No. 7 (1989), 941–981.

10. J. N. L. Connor, Practical methods for the uniform asymptotic evaluation of oscillating integrals with several coalescing saddle points, *in* "Asymptotic and Computational Analysis" (R. Wong, Ed.), pp. 137–173, Dekker, New York, 1990.

11. G. Courbebaisse, Transformée bilinéaire temps-échelle des signaux asymptotiques d'énergie finie, "14ème Colloque GRETSI sur le Traitement du Signal et des Images," pp. 41–44, Juan-les-Pins, 1993.

12. B. de Finetti, Sul concetto di media, *Giornale Ist. Ital. Attuari* **2** (1931), 369–396.

13. P. Flandrin, Some features of time-frequency representations of multicomponent signals, "IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP-84," pp. 41.B.4.1–41.B.4.4, San Diego, 1984.

14. P. Flandrin, "Temps-Fréquence," Hermès, Paris, 1993.

15. P. Flandrin and P. Gonçalvès, From wavelets to time-scale energy distributions, *in* "Recent Advances in Wavelet Analysis" (L. L. Schumaker and G. Webb, Eds.), pp. 309–334, Academic Press, San Diego, 1994.

16. P. Flandrin and F. Hlawatsch, Signal representation geometry and catastrophes in the time–frequency plane, *in* "Mathematics in Signal Processing" (T. S. Durrani *et al.,* Eds.), pp. 3–14, Clarendon Press, Oxford, 1987.

17. D. Gabor, Theory of communication, *J. IEE* **93,** No. III (1946), 429–457.

18. P. Gonçalvès, "Représentations temps-fréquence et temps-échelle bilinéaires—Synthèse et contributions," Thèse de Doctorat, INP Grenoble, 1993.

19. P. Gonçalvès and P. Flandrin, Sur la géométrie et la localisation des distributions affines, "14ème Colloque GRETSI sur le Traitement du Signal et des Images," pp. 355–358, Juan-les-Pins, 1993.

20. A. Grossmann and B. Escudié, Une représentation bilinéaire en temps et échelle des signaux d'énergie finie, "13ème Colloque GRETSI sur le Traitement du Signal et des Images," pp. 33–36, Juan-les-Pins, 1991.

21. G. H. Hardy, J. E. Littlewood, and G. Polya, "Inequalities," Cambridge Univ. Press, Cambridge, 1934.

22. F. Hlawatsch, Interference terms in the Wigner distribution, "Int. Conf. on Digital Signal Processing," pp. 363–367, Florence, 1984.

23. F. Hlawatsch and G. F. Boudreaux-Bartels, Linear and quadratic time–frequency signal representations, *IEEE Signal Proc. Mag.* **9,** No. 2 (1992), 21–27.

24. 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.), to appear.

25. 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, "27th Asilomar Conf. on Signals, Syst. and Comp.," Pacific Grove, 1993.

26. F. Hlawatsch and R. Urbanke, Bilinear time–frequency representations of signals—The shift-scale invariant class, *IEEE Trans. Signal Proc.* **42,** No. 2 (1994), 357–366.

27. A. Kolmogorov, Sur la notion de moyenne, *Atti R. Accad. Naz. Lincei* **6,** No. 12 (1930), 388–391.

28. S. Kadambe and G. F. Boudreaux-Bartels, A comparison of the existence of "cross-terms" in the Wigner distribution and the squared magnitude of the wavelet transform and the short-time Fourier transform, *IEEE Trans. Signal Proc.* **40,** No. 10 (1992), 2498–2517.

29. M. Nagumo, Uber eine Klasse der Mittelwerte, *Jap. J. Math.* **VI,** No. 1 (1930), 71.

30. J.-Ph. Ovarlez, "La transformation de Mellin—Un outil pour l'analyse des signaux à large bande," Thèse de Doctorat, Univ. Paris 6, 1992

31. 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 member, *IEEE Trans. Signal Proc.* **41,** No. 12 (1993), 3425–3444.

32. A. Papoulis, "Signal Analysis," McGraw-Hill, New York, 1977.

33. T. Poston and I. Stewart, "Catastrophe Theory and its Applications," Pitman, 1978.

34. A. Royer, Wigner function as expectation value of a parity operator, *Phys. Rev. A* **15,** No. 2 (1977), 449–450.

35. O. Rioul and P. Flandrin, Time-scale energy distributions—A general class extending wavelet transforms, *IEEE Trans. Signal Proc.* **40,** No. 7 (1992), 1746–1757.

36. K. B. Stolarsky, Generalizations of the logarithmic mean, *Math. Mag.* **48** (1975), 87–92.