Separability, positivity, and minimum uncertainty in time–frequency energy distributions

Patrick Flandrin[a)]
Ecole Normale Supérieure de Lyon, Laboratoire de Physique (URA 1325 CNRS), 46 allée d'Italie, 69364 Lyon Cedex 07, France

(Received 24 June 1997; accepted for publication 4 February 1998)

Gaussian signals play a very special role in classical time–frequency analysis because they are solutions of apparently unrelated problems such as minimum uncertainty, and positivity and separability of Wigner–Ville distributions. We investigate here some of the logical connections which exist between these different features, and we discuss some examples and counterexamples of their extension to more general joint distributions within Cohen's class and the affine class. © *1998 American Institute of Physics.* [S0022-2488(98)00508-8]

I. INTRODUCTION

In most cases, physical signals are defined as functions of time. Until the somewhat recent past, signal analysis has been mainly concerned with such a temporal description of signals, or with a dual description in the frequency domain, obtained from a classical Fourier transform. Much information can, of course, be gained from both descriptions, but it is also clear that time-only and frequency-only representations, being in some sense orthogonal to each other, are not best suited for displaying mixed information about time and frequency. In fact, there are many situations in which a Fourier representation, although mathematically correct, is not able to capture in a direct way time-dependent spectral features which may be intuitively expected. This is so, for instance, with music signals, for which our everyday experience suggests that the ''frequency of tones is continuously changing.'' In this case, Fourier analysis is clearly not well adapted and intuition rather calls for a joint time–frequency description, revealing not only the different frequencies occurring in a piece of music, but also their time of occurrence, their duration,..., i.e., the kind of information which is indeed coded on a musical score.

Although interpretation may be different, the search for a joint time–frequency description of signals has much to share with the problem of finding joint distributions of position and momentum in Quantum Mechanics. It is therefore quite natural that most of the tools which have been developed in either domain have indeed found applications in both. This is especially the case for all the Wigner-based distributions which have been proposed and extensively studied since the pioneering works of Wigner in quantum mechanics[1] and Ville in signal theory.[2] One can, however, remark that, in quantum mechanics, joint distributions $P_\psi(q,p)$ are mostly used as a computational tool, allowing us to write

$$(\psi,\mathbf{G}\psi)=\int\int_{-\infty}^{+\infty} g(q,p)P_\psi(q,p)dq\,dp,$$

where $\mathbf{G}$ stands for the operator associated with a given classical quantity $g(q,p)$ in the sense of some (nonunique) correspondence rule. This nonunicity—which is due to the fact that two dual variables connected by a Fourier transform relationship are associated with elementary operators which do not commute—carries over to the definition of $P_\psi(q,p)$, and the situation is similar for time and frequency. In signal theory, however, a joint distribution is basically considered as a (quasi-) density function of its variables, and the main issue is much more to get a readable time–frequency ''picture'' of a signal, which has therefore to be interpreted *per se* and not only through inner products with test functions.

[a)]Electronic mail: flandrin@physique.ens-lyon.fr

An extensive body of literature has been devoted to time–frequency analysis with Wigner-type distributions, from the point of view of both theory and interpretation (for a survey, see, e.g., Ref. 3 or 4), and here we would like to focus further on some specific questions related to both aspects. More precisely, if we consider Gaussian signals of the form

$$g(t)=Ce^{-\alpha t^2}, \tag{1}$$

with $C\in\mathbf{C}$ and $\alpha>0$, we know that, besides the fact that their Fourier transform is also Gaussian, namely that

$$G(f)\equiv\int_{-\infty}^{+\infty} g(t)e^{-i2\pi ft}\,dt=C\sqrt{\frac{\pi}{\alpha}}\,e^{-\pi^2 f^2/\alpha},$$

such signals happen to play a very special role in classical time–frequency analysis. (Throughout the paper, we will adopt the convention of using a lower case symbol for representing a signal in the time domain, and the corresponding upper case symbol for denoting its Fourier transform, in the frequency domain.) This is in fact so for at least three different reasons.

(1) Minimum uncertainty. Gaussians (1) are the only minimizers for the time–frequency uncertainty relation,[5]

$$\Delta t_x \Delta f_x \geq \frac{1}{4\pi}, \tag{2}$$

where

$$\Delta t_x^2=\frac{1}{E_x}\int_{-\infty}^{+\infty} t^2|x(t)|^2\,dt, \tag{3}$$

$$\Delta f_x^2=\frac{1}{E_x}\int_{-\infty}^{+\infty} f^2|X(f)|^2\,df, \tag{4}$$

and

$$E_x=\int_{-\infty}^{+\infty}|x(t)|^2 dt.$$

(2) Positivity. Gaussians (1) are—up to linear and quadratic phase terms—the only signals for which the Wigner–Ville distribution

$$W_x(t,f)\equiv\int_{-\infty}^{+\infty} x\left(t+\frac{\tau}{2}\right)\overline{x\left(t-\frac{\tau}{2}\right)}e^{-i2\pi f\tau}\,d\tau=\int_{-\infty}^{+\infty} X\left(f+\frac{\nu}{2}\right)\overline{X\left(f-\frac{\nu}{2}\right)}e^{i2\pi\nu t}\,d\nu, \tag{5}$$

is everywhere non-negative (Hudson's theorem;[6] cf. Proposition 1 below).

(3) Separability. The Wigner–Ville distribution (5) of Gaussians (1) is separable, namely,

$$W_g(t,f)=|C|^2\sqrt{\frac{2\pi}{\alpha}}\,e^{-2\alpha t^2}e^{-(2\pi^2/\alpha)f^2}. \tag{6}$$

These three properties are important ones for attaching to Gaussian signals a specific status and for offering a nice physical interpretation of their time–frequency content by means of the Wigner–Ville distribution.

(1) Minimum uncertainty is often invoked for considering Gaussian signals as time–frequency ''atoms'' and for using them as elementary building blocks (the so-called ''Gabor logons'' [5]) of linear signal decompositions.

(2) Although it is known that no time–frequency distribution—be it positive or not—can be interpreted pointwise,[7] positivity improves the readability of a time–frequency diagram.

(3) Separability provides further insight in the probability picture of time–frequency distributions by corresponding to a notion of ''independence,'' according to which a logon is a state whose time and frequency behaviors are decoupled.[8]

Given these facts, an interesting question is to know whether the three mentioned properties (minimum uncertainty, positivity, and separability) have logical reasons to be related or, on the contrary, if their simultaneous occurrence is only fortuitous. Moreover, since many other joint distributions do exist besides the Wigner–Ville distribution, a further question which is worth investigating is how the results pertaining to the pair ''Gaussian/Wigner–Ville'' can be extended to other situations, especially in the important case of affine (scale-based) distributions.

The paper is devoted to these questions and is organized as follows. In Sec. II, the situation of the Wigner–Ville distribution is considered in some detail by reviewing and/or establishing a number of its specific properties connected in a direct way with the considered problem. In Sec. III we address the question of a possible extension of such results to the more general ''Cohen's class,'' whereas in Sec. IV we are concerned with similar questions within the ''affine class'' framework. Finally, some conclusions are drawn, together with the possibility of getting further extensions in more general situations.

II. WIGNER–VILLE DISTRIBUTION

In this section, we will review and/or establish basic properties—related to positivity, separability, and minimum uncertainty—which hold in the case of the Wigner–Ville distribution. As far as positivity is concerned, it has to be noted that many other important results exist (see, e.g., Refs. 9–12), even with extensions beyond the Wigner–Ville case. The purpose of this paper is not to review all of them, and only those results which are connected in some way with the two other issues of separability and minimum uncertainty will be considered here.

A. Positivity

Proposition 1 (Hudson's theorem[6]): The Wigner–Ville distribution is positive for signals of the type

$$g_{\alpha,\beta,\gamma}(t) \equiv e^{-(\alpha t^2+\beta t+\gamma)}, \tag{7}$$

with $(\alpha,\beta,\gamma)\in \mathbf{C}^3$ *and* $Re\{\alpha\}>0$, *and only for them.*

Proof: The fact that (generalized) Gaussian signals of the form (7) have a positive Wigner–Ville distribution follows from a direct calculation, according to which

$$W_{g_{\alpha,\beta,\gamma}}(t,f) = \sqrt{\frac{2\pi}{\mathrm{Re}\{\alpha\}}}\exp\left(-2\,\mathrm{Re}\{\alpha t^2+\beta t+\gamma\} - \frac{2\pi^2}{\mathrm{Re}\{\alpha\}}\,[f-f_i(t)]^2\right),$$

with

$$f_i(t) \equiv \frac{1}{2\pi}\,\mathrm{Im}\{2\alpha t+\beta\}.$$

The original proof of the converse was first stated by Hudson in Ref. 6. It consists of introducing—according to the definition (7)—the test signal $g_{1,z,\gamma}(t)$, for which the corresponding Wigner–Ville distribution is everywhere non-negative. Making use of Moyal's formula,[4] we have

$$\left|\int_{-\infty}^{+\infty} x(t)\overline{g_{1,z,\gamma}(t)}\,dt\right|^2 = \int\int_{-\infty}^{+\infty} W_x(t,f)\,W_{g_{1,z,\gamma}}(t,f)\,dt\,df,$$

and this quantity is guaranteed to be everywhere positive for those signals $x(t)$ whose Wigner–Ville distribution is non-negative. It follows that the function

$$F(z) = e^{\gamma}\int_{-\infty}^{+\infty} x(t)\overline{g_{1,z,\gamma}(t)}\,dt \tag{8}$$

is analytic, has no zeros, and furthermore satisfies

$$|F(z)|^2 \leqslant \sqrt{\frac{\pi}{2}}\, E_x \, \exp\left(\frac{1}{2}\, (\mathrm{Re}\{z\})^2\right),$$

as can be checked by applying the Cauchy–Schwarz inequality.

As a consequence, $F(z)$ is an entire function of order at most 2, without zeros. Hence, by Hadamard's factorization theorem, $F(z)$ is necessarily the exponential of some quadratic form in z. Letting $z = i2\pi f$ and using the fact that Fourier transforms of Gaussians are still Gaussians, we can deduce from (8) that $x(t)$ is itself the exponential of some quadratic form in t and the proof is complete.

Two remarks.

(1) In order to have finite energy signals—i.e., $x(t) \in L^2(\mathbf{R})$—we have assumed that $\mathrm{Re}\{\alpha\} > 0$ and this, in turn, guarantees that the Wigner–Ville distribution is bounded since it is easy to show that $|W_x(t,f)| \leqslant 2E_x$. A degeneracy can, however, be observed whenever $\mathrm{Re}\{\alpha\} = \mathrm{Re}\{\beta\} = 0$, a situation for which the Wigner–Ville distribution reads as

$$W_{g_{\alpha,\beta,\gamma}}(t,f) = \exp(-2\,\mathrm{Re}\{\gamma\})\, \delta(f - f_i(t)).$$

(2) Minimum uncertainty implies positivity of the Wigner–Ville distribution, but the class of signals (7) which have a positive Wigner–Ville distribution is larger than the class (1) of signals with minimum uncertainty. In fact, the former can be seen as a modulation of the latter by means of a "chirp" wave form whose phase is quadratic in time and, hence, whose instantaneous frequency is linear. Precisely, we get

$$\Delta t_{g_{\alpha,\beta,\gamma}} \Delta f_{g_{\alpha,\beta,\gamma}} = \frac{1}{4\pi}\left(1 + \left(\frac{\mathrm{Im}\{\alpha\}}{\mathrm{Re}\{\alpha\}}\right)^2\right)^{1/2} \geqslant \frac{1}{4\pi},$$

with equality if and only if $\mathrm{Im}\{\alpha\} = 0$, i.e., if the Gaussian is not modulated in frequency, thus reducing to (1).

B. Separability

Proposition 2: If a Wigner–Ville distribution is separable, it is positive and factorizable as

$$W_x(t,f) = \frac{|x(t)|^2 |X(f)|^2}{E_x}. \tag{9}$$

Proof: Starting from the definition (5), it is well known (and easy to check) that a Wigner–Ville distribution always has correct marginals, i.e., that

$$\int_{-\infty}^{+\infty} W_x(t,f)\, dt = |X(f)|^2, \tag{10}$$

$$\int_{-\infty}^{+\infty} W_x(t,f)\, df = |x(t)|^2. \tag{11}$$

Assuming that separability holds, there exist two functions $p(t)$ and $Q(f)$ such that $W_x(t,f) = p(t)Q(f)$. Using (11) and (10), we get $P(0)Q(f) = |X(f)|^2 \geqslant 0$ and $p(t)q(0) = |x(t)|^2 \geqslant 0$. It follows that $P(0)$ and $Q(f)$ on one hand, and $p(t)$ and $q(0)$ on the other hand, are necessarily of the same sign. One more integration leads to $P(0)q(0) = E_x > 0$, and this concludes the proof.

Remark: We know from Moyal's formula that, for any finite energy signal $x(t)$, we have

$$\int \int_{-\infty}^{+\infty} W_x^2(t,f)\, dt\, df = E_x^2.$$

Applying this result to the factorized distribution (9), we are led to the conclusion that signals with a separable Wigner–Ville distribution must necessarily be such that

$$\left(\int_{-\infty}^{+\infty}|x(t)|^4\,dt\right)^{1/4}\left(\int_{-\infty}^{+\infty}|X(f)|^4\,df\right)^{1/4}=E_x,$$

a property which holds for Gaussians. Gaussians are therefore solutions for separability. In fact, it can be shown that they are the only solutions, as claimed in the following.

Proposition 3: The Wigner–Ville distribution is separable for Gaussian signals only.

Proof: This result can be viewed as a Corollary to Propositions 2 and 1 (''separability⇒positivity⇒Gaussians''), but it can also be derived in a direct way as follows. Assuming that separability holds for a Wigner–Ville distribution, we can start from the factorization of (9) and, after taking an inverse Fourier transform over frequency, we get

$$\frac{1}{|x(t)|^2}\,x\left(t+\frac{\tau}{2}\right)\overline{x\left(t-\frac{\tau}{2}\right)}=\gamma_x(\tau), \tag{12}$$

with

$$\gamma_x(\tau)\equiv\frac{1}{E_x}\int_{-\infty}^{+\infty}x\left(t+\frac{\tau}{2}\right)\overline{x\left(t-\frac{\tau}{2}\right)}dt.$$

This defines a functional equation whose first implication is that its left-hand side must be a function of τ only. Writing $x(t)=|x(t)|\exp\{i\arg x(t)\}\equiv\exp\{\varphi(t)+i\psi(t)\}$, we end up with two necessary conditions upon a modulus term and a phase term, respectively. The phase condition (of independence upon the t variable) reads as

$$\frac{\partial}{\partial t}\left[\psi\left(t+\frac{\tau}{2}\right)-\psi\left(t-\frac{\tau}{2}\right)\right]=0,$$

for any t and τ, with the consequence that the phase must necessarily be of the form $\psi(t)=at+b$, with $(a,b)\in\mathbf{R}^2$. After a change of variables, the companion condition on the modulus leads, for any v and w, to

$$\dot{\varphi}\left(\frac{v+w}{2}\right)=\frac{\dot{\varphi}(v)+\dot{\varphi}(w)}{2},$$

where the overdot indicates the (time) derivative.

We recognize in this relation *Jensen's functional equation,*[13] whose most general (continuous) solution in $\dot{\varphi}(t)$ is given by linear functions of the form $\dot{\varphi}(t)=a't+b'$ with $(a',b')\in\mathbf{R}^2$. It follows that $\varphi(t)=(a'/2)t^2+b't+c'$, with $c'\in\mathbf{R}$ and the requirement that $a'\leqslant 0$ in order to guarantee that $x(t)\in L^2(\mathbf{R})$.

Gaussian signals therefore appear as candidates for defining separable Wigner–Ville distributions. In order to be admissible solutions, we have to further check that the left-hand side of (12) is not only independent of t, but also proportional to the deterministic correlation function $\gamma_x(\tau)$, which an elementary calculation proves to be true.

C. Minimum uncertainty

Proposition 4 (Ref. 14): Given any $T>0$, the Wigner–Ville distribution satisfies the time–frequency uncertainty relation,

$$\Sigma_{tf}^2(W_x)\equiv\frac{1}{E_x}\int\int_{-\infty}^{+\infty}\left(\frac{t^2}{T^2}+T^2f^2\right)W_x(t,f)\,dt\,df\geqslant\frac{1}{2\pi}, \tag{13}$$

with equality if and only if $x(t)$ is a Gaussian signal.

Proof: This follows directly from the marginal properties (10)–(11) of the Wigner–Ville distribution and from the definitions (3)–(4), since we can write

$$\Sigma_{tf}^2(W_x)=\frac{1}{E_x}\left[\frac{1}{T^2}\int_{-\infty}^{+\infty}t^2|x(t)|^2dt+T^2\int_{-\infty}^{+\infty}f^2|X(f)|^2df\right]$$
$$=\frac{\Delta t_x^2}{T^2}+T^2\Delta f_x^2=\left(\frac{\Delta t_x}{T}-T\Delta f_x\right)^2+2\Delta t_x\Delta f_x\geqslant 2\Delta t_x\Delta f_x, \tag{14}$$

with equality if the arbitrary duration T is ''matched'' to the signal $x(t)$ according to $T=\sqrt{\Delta t_x/\Delta f_x}$. In this case, we know from (2) that the right-hand side of (14) is lower bounded too, with equality in the only case of Gaussians, whence the result.

Remark: In the above derivation, it has been implicitly assumed that the mean time and the mean frequency of the analyzed signal are both zero, i.e., that

$$\int_{-\infty}^{+\infty}t|x(t)|^2\,dt=\int_{-\infty}^{+\infty}f|X(f)|^2\,df=0.$$

As for the standard Gabor–Heisenberg uncertainty relation (2), this of course does not reduce the generality of the result, since the Wigner–Ville distribution being covariant with respect to shifts in time and frequency, uncertainty relations pertaining to signals with nonzero mean time and/or mean frequency can be established *mutatis mutandis* by introducing suitable shifts.

Because a Wigner–Ville distribution attains generally negative values, the interpretation of (13) as a measure of time–frequency energy spread can be questioned. A companion result can, however, be obtained for *squared* Wigner–Ville distributions, thus reinforcing the prominent role of Gaussians. This is given by the following.

Proposition 5 (Ref. 15): Given any $T>0$, the squared Wigner–Ville distribution satisfies the time–frequency uncertainty relation

$$\Sigma_{tf}^2(W_x^2)=\frac{1}{E_x^2}\int\int_{-\infty}^{+\infty}\left(\frac{t^2}{T^2}+T^2f^2\right)W_x^2(t,f)\,dt\,df\geqslant\frac{1}{4\pi},$$

with equality if and only if $x(t)$ is a Gaussian signal.

Proof: The proof proceeds as previously, except that we have to evaluate second-order moments with respect to $W_x^2(t,f)$. It is easy to show that

$$\int_{-\infty}^{+\infty}W_x^2(t,f)\,df=\int_{-\infty}^{+\infty}\left|x\left(t+\frac{\tau}{2}\right)\right|^2\left|x\left(t-\frac{\tau}{2}\right)\right|^2d\tau,$$

and, therefore,

$$\frac{1}{E_x^2}\int\int_{-\infty}^{+\infty}t^2W_x^2(t,f)\,dt\,df=\frac{1}{E_x^2}\int\int_{-\infty}^{+\infty}\left(\frac{v+w}{2}\right)^2|x(v)|^2|x(w)|^2dv\,dw$$
$$=\frac{1}{4E_x^2}\left[2E_x\int_{-\infty}^{+\infty}v^2|x(v)|^2dv+2\left(\int_{-\infty}^{+\infty}v|x(v)|^2dv\right)^2\right]\geqslant\frac{\Delta t_x^2}{2},$$

with equality for signals whose mean time is zero. It can be proved in a similar way that

$$\frac{1}{E_x^2}\int\int_{-\infty}^{+\infty}f^2W_x^2(t,f)\,dt\,df\geqslant\frac{\Delta f_x^2}{2},$$

with equality for signals whose mean frequency is zero.

Given these two properties, the proof is finished by reorganizing terms exactly as in the nonsquared case.

Remark: The fact that both the Wigner–Ville distribution $W_x(t,f)$ and its square $W_x^2(t,f)$ have a minimum time–frequency spread which is attained for Gaussian signals suggests a natural link between minimum uncertainty and positivity. For simplicity, let us assume that $x(t)$ is of unit energy so that both $W_x(t,f)$ and $W_x^2(t,f)$ integrate to 1 in the time–frequency plane. In fact, as far as the spread related to $W_x(t,f)$ only is concerned, a decrease in the spread measure could *a priori* be obtained not only by concentrating positive values around the origin, but also by accepting negative values far from the origin. This, however, turns out to be in contradiction with the spread inequality related to $W_x^2(t,f)$ since a distribution with negative values would create ''holes''—in between positive contributions—when squared, thus increasing the spread measure as compared to a positive distribution.

It is possible to find many other ways of formulating uncertainty relations by means of the Wigner–Ville distribution, among which we can cite the following results.

Proposition 6: Any signal $x(t)$ of finite energy E_x satisfies the ''uncertainty relation''

$$\alpha_x \equiv \frac{1}{E_x}\int_{-\infty}^{+\infty} t^2 |\dot{x}(t)|^2 dt \geqslant \frac{1}{4},$$

with

$$\lim_{\gamma\to 0_+} \alpha_{x_\gamma} = \tfrac{1}{4},$$

for signals of the form

$$x_\gamma(t) \equiv C \exp(-\beta|t|^\gamma), \tag{15}$$

with $C\in\mathbf{C}$ and $(\beta,\gamma)\in\mathbf{R}_+^2$.

Proof: For convenience, let us introduce the quantity

$$J \equiv \int_{-\infty}^{+\infty} t\dot{x}(t)\overline{x(t)}dt,$$

whose imaginary part is sometimes referred to as the ''covariance'' of the signal.[3] By applying the Cauchy–Schwarz inequality, we get $|J|^2 \leqslant \alpha_x E_x^2$, whereas an integration by parts shows that $\mathrm{Re}\{J\} = -E_x/2$, and the claimed inequality follows from the fact that we always have $|J|^2 \geqslant (\mathrm{Re}\{J\})^2$.

From the way the lower bound has been derived, it is not clear whether it can be effectively attained. However, it can be shown by a direct calculation that signals of the form (15) are such that $\alpha_{x_\gamma} = (\gamma+1)/4$, with the consequence that $\lim_{\gamma\to 0_+}\alpha_{x_\gamma} = 1/4$, thus completing the proof.

Three remarks.

(1) Although the lower bound $\alpha_x = \frac{1}{4}$ is asymptotically sharp for signals $x_\gamma(t)$ when $\gamma\to 0_+$, it has to be remarked that the bandwidth Δf_{x_γ} and, hence, the usual uncertainty measure remain finite only if $\gamma > \frac{1}{2}$.

(2) The ''alpha moment'' α_x (a quantity which has been introduced and studied by Titlebaum; see Ref. 16) is such that

$$\alpha_x = \frac{1}{E_x}\int_{-\infty}^{+\infty} t^2 |\dot{x}(t)|^2 dt = \frac{1}{E_x}\int_{-\infty}^{+\infty} f^2 |\dot{X}(f)|^2 df.$$

It follows that any property of α_x which holds for a given signal $x(t)$ also holds for its Fourier transform $X(f)$. In particular, the lower bound for α_x is also attained for signals whose spectrum reads as $X(f) = C \exp(-\beta|f|^\gamma)$, in the limit $\gamma\to 0_+$, with the consequence that the corresponding signals have a finite mean-square duration—and, hence, a finite usual uncertainty measure—only if $\gamma > \frac{1}{2}$.

(3) The uncertainty measure given by the ''alpha moment'' α_x is not minimum for Gaussian signals. In fact, expanding $x(t)\in L^2(\mathbf{R})$ as $x(t) = \sum_{n=0}^{+\infty} x_n \psi_n(t)$ on the orthonormal basis of Hermite functions $\{\psi_n(t); n\in\mathbf{N}\}$ (defined as in Ref. 17), we get

$$E_x\alpha_x = \sum_{n=0}^{+\infty}\left(\left[\frac{3}{4} + \frac{n(n+1)}{2}\right]|x_n|^2 - \frac{\sqrt{(n+1)(n+2)(n+3)(n+4)}}{2}\,\mathrm{Re}\{x_n x_{n+4}\}\right).$$

This proves that $\alpha_{\psi_n}=3/4+n(n+1)/2$ and thus that the Gaussian (associated to $n=0$) may be considered as a minimizer for α_x, but in the restricted class of Hermite functions.

Proposition 7: For any signal $x(t)$ of finite energy E_x and of mean time and mean frequency identically zero, the Wigner–Ville distribution satisfies the uncertainty relation

$$\sigma_{tf}^2(W_x^2)\equiv\frac{1}{E_x^2}\int\int_{-\infty}^{+\infty}t^2f^2W_x^2(t,f)\,dt\,df\geqslant\frac{1}{32\pi^2}\left(\alpha_x-\frac{1}{2}+4\pi^2\Delta t_x^2\Delta f_x^2\right),$$

with equality for signals whose ''covariance'',

$$\mathrm{cov}_{tf}(x)\equiv\mathrm{Im}\left(\int_{-\infty}^{+\infty}t\dot{x}(t)\overline{x(t)}\,dt\right), \tag{16}$$

is zero.

Proof: A way of proving this result is to make use of the so-called ''inner interference formula,''[15] according to which

$$W_x^2(t,f)=\int\int_{-\infty}^{+\infty}W_x\left(t+\frac{\tau}{2},\ f+\frac{\nu}{2}\right)W_x\left(t-\frac{\tau}{2},\ f-\frac{\nu}{2}\right)d\tau\,d\nu.$$

By making a change of variable and reorganizing terms, this leads to

$$\sigma_{tf}^2(W_x^2)=\frac{1}{8E_x^2}\int\int\int\int_{-\infty}^{+\infty}K(v,w,\xi,\zeta)W_x(v,\xi)W_x(w,\zeta)\,dv\,dw\,d\xi\,d\zeta,$$

with

$$K(v,w,\xi,\zeta)=v^2\xi^2+v^2\zeta^2+2v^2\xi\zeta+2vw\xi^2+4vw\xi\zeta.$$

The contribution of the first term of $K(v,w,\xi,\zeta)$ to the integral reduces to

$$\int\int\int\int_{-\infty}^{+\infty}v^2\xi^2W_x(v,\xi)W_x(w,\zeta)\,dv\,dw\,d\xi\,d\zeta=E_x^2\sigma_{tf}^2(W_x),$$

and it can be shown[4] that

$$\sigma_{tf}^2(W_x)=\frac{1}{4\pi^2}\left(\alpha_x-\frac{1}{2}\right).$$

The second term contributes as

$$\int\int\int\int_{-\infty}^{+\infty}v^2\zeta^2W_x(v,\xi)W_x(w,\zeta)\,dv\,dw\,d\xi\,d\zeta$$

$$=\int\int_{-\infty}^{+\infty}v^2\zeta^2|x(v)|^2|X(\zeta)|^2\,dv\,d\zeta=E_x^2\Delta t_x^2\Delta f_x^2,$$

whereas the third and fourth ones provide a null contribution because of the assumptions of mean time and mean frequency equal to zero.

Finally, the fifth term is such that

$$\int\int\int\int_{-\infty}^{+\infty}vw\xi\zeta\,W_x(v,\xi)W_x(w,\zeta)\,dv\,dw\,d\xi\,d\zeta=\frac{1}{4\pi^2}\,\mathrm{cov}_{tf}^2(x)\geqslant 0,$$

with (following Cohen[3]), the ''covariance'' of $x(t)$ defined by (16).

Remark: If $x(t)$ is chosen to be $x_\gamma(t)$ as defined in (15), we get explicitly

$$\sigma_{tf}^2(W_{x_\gamma}^2)=\frac{1}{128\pi^2}\left[(\gamma-1)+\gamma^2\,\frac{\Gamma(3/\gamma)\Gamma(2-1/\gamma)}{\Gamma^2(1/\gamma)}\right],$$

and a numerical estimate of the γ_0 minimizing the spread measure gives $\gamma_0\approx 1.33$, with $\sigma_{tf}^2(W_{x_{\gamma_0}}^2)\approx 0.77/64\pi^2$. This is less than in the Gaussian case ($\gamma=2$) for which $\sigma_{tf}^2(W_{x_2}^2)=1/64\pi^2$.

III. COHEN'S CLASS

Some results pertaining to Wigner–Ville distributions can be generalized to members of *Cohen's class*, the general class of quadratic time–frequency energy distributions which are covariant to time and frequency shifts. By definition, Cohen's class can be written as[4]

$$C_x(t,f;\varphi)=\int\int_{-\infty}^{+\infty}\varphi(\xi,\tau)A_x(\xi,\tau)e^{-i2\pi(t\xi+f\tau)}d\xi\, d\tau, \tag{17}$$

where $\varphi(\xi,\tau)$ is some arbitrary kernel function and $A_x(\xi,\tau)$ is the so-called ''ambiguity function,'' obtained as a two-dimensional Fourier transform of the Wigner–Ville distribution and defined explicitly by

$$A_x(\xi,\tau)=\int_{-\infty}^{+\infty}x\left(t+\frac{\tau}{2}\right)\overline{x\left(t-\frac{\tau}{2}\right)}e^{i2\pi\xi t}\,dt.$$

Within this more general framework, Gaussian signals still appear to be central with respect to separability properties, as illustrated by the following claims.

Proposition 8: Smoothed pseudo-Wigner–Ville distributions—characterized by a separable kernel—are the only members of Cohen's class which are separable when applied to Gaussian signals.

Proof: Assuming that the analyzed signal $x(t)$ is a Gaussian, we know that its Wigner–Ville distribution $W_x(t,f)$ is separable, a property which carries over to its two-dimensional Fourier transform $A_x(\xi,\tau)$. Using (17), we get

$$\varphi(\xi,\tau)=\frac{1}{A_x(\xi,\tau)}\int\int_{-\infty}^{+\infty}C_x(t,f;\varphi)e^{i2\pi(t\xi+f\tau)}dt\, df,$$

and, if we require $C_x(t,f;\varphi)$ to be separable, it follows that the kernel function $\varphi(\xi,\tau)$ must itself be separable, a situation corresponding to distributions referred to as ''smoothed pseudo-Wigner–Ville distributions.''[4]

Proposition 9: Smoothed pseudo-Wigner–Ville distributions—characterized by a separable kernel—are separable for Gaussian signals only.

Proof: The proof is very similar to the previous one since, according to (17), we can write

$$W_x(t,f)=\int\int_{-\infty}^{+\infty}\left(\frac{1}{\varphi(\xi,\tau)}\int\int_{-\infty}^{+\infty}C_x(t',f';\varphi)e^{i2\pi(t'\xi+f'\tau)}dt'\, df'\right)e^{-i2\pi(t\xi+f\tau)}d\xi\, d\tau.$$

Assuming therefore that both the kernel function $\varphi(\xi,\tau)$ and the associated smoothed Wigner–Ville distribution $C_x(t,f;\varphi)$ are separable, we get that the Wigner–Ville distribution itself is separable, which in turn implies that $x(t)$ is necessarily a Gaussian.

Proposition 10: When applied to Gaussian signals, spectrograms are separable if and only if their short-time window is Gaussian.

Proof: Assuming that the analyzed signal $x(t)$ is a Gaussian, we know that both its Wigner–Ville distribution $W_x(t,f)$ and its ambiguity function $A_x(\xi,\tau)$ are separable. If we now consider a spectrogram,

$$S_x^{(h)}(t,f)\equiv\left|\int_{-\infty}^{+\infty}x(s)\overline{h(s-t)}e^{-i2\pi fs}\,ds\right|^2=\left|\int_{-\infty}^{+\infty}X(\xi)\overline{H(\xi-f)}e^{i2\pi\xi t}\,d\xi\right|^2$$

having a short-time window $h(t)$ which is Gaussian, we get that the associated kernel function[4] $\varphi(\xi,\tau)=\overline{A_h(\xi,\tau)}$ is separable, and it follows that the spectrogram—as an inverse Fourier transform of $\varphi(\xi,\tau)A_x(\xi,\tau)$—is itself separable.

Conversely, if we suppose that a spectrogram is separable when applied to a Gaussian signal $g(t)$, we get from (17) [with $\varphi(\xi,\tau)=\overline{A_h(\xi,\tau)}$] that

$$\overline{A_h(\xi,\tau)}=\frac{1}{A_g(\xi,\tau)}\int\int_{-\infty}^{+\infty} S_g^{(h)}(t',f')e^{i2\pi(t'\xi+f'\tau)}dt'\ df'$$

is separable, with the consequence that $W_h(t,f)$ itself—as an inverse Fourier transform of $A_x(\xi,\tau)$—is separable and, hence, that the window $h(t)$ must be Gaussian.

More generally, we can remark that imposing the requirement of separability to spectrograms leads to specific constraints which have to be jointly satisfied by the analyzed signal $x(t)$ and the analysis window $h(t)$. Starting from the definition of a spectrogram of window $h(t)$ and assuming that this spectrogram is separable according to $S_x^{(h)}(t,f)=p(t)Q(f)$, we get

$$p(t)q(0)=\int_{-\infty}^{+\infty} S_x^{(h)}(t,f)df=\int_{-\infty}^{+\infty}|x(s)|^2|h(s-t)|^2ds$$

and

$$P(0)Q(f)=\int_{-\infty}^{+\infty} S_x^{(h)}(t,f)dt=\int_{-\infty}^{+\infty}|X(\xi)|^2|H(\xi-f)|^2d\xi.$$

Since $P(0)q(0)=E_xE_h$, separable spectrograms happen to be necessarily of the form

$$S_x^{(h)}(t,f)=\int\int_{-\infty}^{+\infty}\frac{|x(s)|^2|X(\xi)|^2}{E_x}\frac{|h(s-t)|^2|H(\xi-f)|^2}{E_h}\ ds\ d\xi,$$

a form which has to be compared to the (always valid) expression

$$S_x^{(h)}(t,f)=\int\int_{-\infty}^{+\infty}W_x(s,\xi)W_h(s-t,\xi-f)ds\ d\xi,$$

or even to

$$S_x^{(h)}(t,f)=\int\int_{-\infty}^{+\infty}C_x(s,\xi;\varphi)\overline{C_h(s-t,\xi-f:\varphi)}ds\ d\xi,$$

with the requirement that the parametrization function $\varphi(\xi,\tau)$ be such that $|\varphi(\xi,\tau)|=1$ (see, e.g., Ref. 4).

Reasoning along the same lines in the more general case of any distribution within Cohen's class, we can end up with the result that a separable distribution is necessarily such that

$$\begin{aligned}C_x(t,f;\varphi)&=\int\int_{-\infty}^{+\infty}W_x(s,\xi)\Pi(s-t,\xi-f)ds\ d\xi\\&=\int\int_{-\infty}^{+\infty}\frac{|x(s)|^2|X(\xi)|^2}{E_x}\frac{F(s-t,0)\psi(0,\xi-f)}{\varphi(0,0)}\ ds\ d\xi,\end{aligned}\qquad(18)$$

where $\Pi(t,f)$ stands for the two-dimensional Fourier transform of $\varphi(\xi,\tau)$, and $F(t,\tau)$ and $\psi(\xi,\nu)$ for one-dimensional partial Fourier transforms of the same quantity, over frequency and time, respectively. This leads to the following.

Proposition 11: If a Cohen's class distribution has correct marginals and is separable, it is positive and factorizable as

$$C_x(t,f;\varphi)=\frac{|x(t)|^2|X(f)|^2}{E_x}. \tag{19}$$

Proof: This can be proved either directly—as for Proposition 2—or by using (18), since we know[4] that having correct marginals implies that $\varphi(\xi,0)=\varphi(0,\tau)=1$.

Remark: Given *any* signal $x(t)$, one could imagine to formally obtain a distribution factorizable as in (19) by choosing for the parametrization function the (signal-dependent) quantity $\varphi(\xi,\tau)\equiv E_x\gamma_X(\xi)\gamma_x(\tau)/A_x(\xi,\tau)$, with the γ's defined as in Proposition 3. In the general case, this would unfortunately correspond to pathological φ's, with singularities at those points where the ambiguity function vanishes while the individual time and frequency correlation functions do not. Regular solutions can be guaranteed by imposing to the ambiguity function to be everywhere positive but, as mentioned in Ref. 18, imposing positivity of the ambiguity function is equivalent to imposing positivity of the Wigner–Ville distribution, with Gaussians as only solutions.

Proposition 12: The Wigner–Ville distribution is the only member of Cohen's class which both has correct marginals and is separable when applied to Gaussian signals.

Proof: By Fourier transformation, (19) becomes $\varphi(\xi,\tau)A_x(\xi,\tau)=E_x\gamma_X(\xi)\gamma_x(\tau)$. Assuming that $x(t)$ is a Gaussian signal, this simplifies to $\varphi(\xi,\tau)=1$, thus defining the Wigner–Ville distribution as the only solution.

Proposition 13 (Ref. 4): For any signal $x(t)$ whose mean time and mean frequency are zero, and any $T>0$, distributions of Cohen's class satisfy the time–frequency uncertainty relation

$$\Sigma_{tf}^2(C_x)\geqslant\frac{1}{2\pi}-\frac{1}{4\pi^2}\left[\frac{1}{T^2}\frac{\partial^2\varphi}{\partial\xi^2}(0,0)+T^2\frac{\partial^2\varphi}{\partial\tau^2}(0,0)\right],$$

with equality if and only if $x(t)$ is a Gaussian signal.

Proof: The proof is straightforward and follows directly from the fact that $C_x(t,f;\varphi)$ has for Fourier transform $\varphi(\xi,\tau)A_x(\xi,\tau)$. Defining

$$\Sigma_t^2(C_x)\equiv\frac{1}{E_x}\int_{-\infty}^{+\infty}t^2\left(\int_{-\infty}^{+\infty}C_x(t,f;\varphi)\,df\right)dt,$$

we get

$$\begin{aligned}\Sigma_t^2(C_x)&=\frac{1}{E_x}\int_{-\infty}^{+\infty}|x(s)|^2\left(\int_{-\infty}^{+\infty}t^2F(s-t,0)\,dt\right)ds\\&=\frac{1}{E_x}\int_{-\infty}^{+\infty}|x(s)|^2\left(\int_{-\infty}^{+\infty}(s^2-2s\theta+\theta^2)F(\theta,0)\,d\theta\right)ds=\varphi(0,0)\Delta t_x^2-\frac{1}{4\pi^2}\frac{\partial^2\varphi}{\partial\xi^2}(0,0),\end{aligned}$$

and, similarly,

$$\Sigma_f^2(C_x)\equiv\frac{1}{E_x}\int_{-\infty}^{+\infty}f^2\left(\int_{-\infty}^{+\infty}C_x(t,f;\varphi)\,dt\right)df=\varphi(0,0)\Delta f_x^2-\frac{1}{4\pi^2}\frac{\partial^2\varphi}{\partial\tau^2}(0,0).$$

Assuming that $C_x(t,f;\varphi)$ is a valid energy distribution [i.e., that $\varphi(0,0)=1$], we are led to

$$\Sigma_{tf}^2(C_x)=\frac{\Sigma_t^2(C_x)}{T^2}+T^2\Sigma_f^2(C_x)=\Sigma_{tf}^2(W_x)-\frac{1}{4\pi^2}\left[\frac{1}{T^2}\frac{\partial^2\varphi}{\partial\xi^2}(0,0)+T^2\frac{\partial^2\varphi}{\partial\tau^2}(0,0)\right],$$

and therefore to the claim, from the result established previously for $\Sigma_{tf}^2(W_x)$.

Proposition 14: For any signal $x(t)$ whose mean time and mean frequency are zero, and any $T>0$, spectrograms with a window $h(t)$ satisfy the time–frequency uncertainty relation,

$$\Sigma_{tf}^2(S_x^{(h)})=\Sigma_{tf}^2(W_x)+\Sigma_{tf}^2(W_h)\geqslant\frac{1}{\pi},$$

with equality, if and only if $x(t)$ and $h(t)$ are matched Gaussian signals.

Proof: This can be viewed as a Corollary to the previous Proposition since, in the case of spectrograms, we get

$$-\frac{1}{4\pi^2}\left[\frac{1}{T^2}\frac{\partial^2\varphi}{\partial\xi^2}(0,0)+T^2\frac{\partial^2\varphi}{\partial\tau^2}(0,0)\right]=\frac{1}{T^2}\frac{\partial^2\overline{A_h}}{\partial\xi^2}(0,0)+T^2\frac{\partial^2\overline{A_h}}{\partial\tau^2}(0,0),$$

$$=\Sigma_{tf}^2(W_h)$$

and hence

$$\Sigma_{tf}^2(S_x^{(h)})=\Sigma_{tf}^2(W_x)+\Sigma_{tf}^2(W_h).$$

The total time–frequency spread being the sum of two independent positive quantities, it is minimized when each of these quantities is itself minimized, which is obtained when both $x(t)$ and $h(t)$ are Gaussian signals (see Proposition 4). In order to ensure that the two individual minimizations are compatible, the same arbitrary T has to be chosen for both $\Sigma_{tf}^2(W_x)$ and $\Sigma_{tf}^2(W_h)$. This leads to

$$T=\sqrt{\frac{\Delta t_x}{\Delta f_x}}=\sqrt{\frac{\Delta t_h}{\Delta f_h}},$$

and, hence, to $\Delta t_h=\Delta t_x$ and $\Delta f_h=\Delta f_x$, since $x(t)$ and $h(t)$ being Gaussians, $\Delta t_x\Delta f_x=\Delta t_h\Delta f_h=1/4\pi$.

Remark: Although the derivation is different, this result is equivalent to the one given by Janssen in Ref. 19.

Proposition 15: For any signal $x(t)$ whose mean time and mean frequency are zero, and any $T>0$, spectrograms with a real-valued and even window $h(t)$ are such that

$$\sigma_{tf}^2(S_x^{(h)})=\frac{1}{4\pi^2}\,[(\alpha_x+\alpha_h-1)+4\pi^2(\Delta t_x^2\Delta f_h^2+\Delta t_h^2\Delta f_x^2)].$$

Proof: Let us introduce the time–frequency spread measure,

$$\sigma_{tf}^2(C_x)\equiv\frac{1}{E_x}\int\int_{-\infty}^{+\infty}t^2f^2C_x(t,f;\varphi)dt\;df, \tag{20}$$

where $C_x(t,f;\varphi)$ stands for any distribution within Cohen's class. This spread measure (20) can be equivalently expressed as

$$\sigma_{tf}^2(C_x)=\frac{1}{16\pi^4E_x}\frac{\partial^4}{\partial\xi^2\partial\tau^2}\left(\varphi(\xi,\tau)A_x(\xi,\tau)\right)\Big|_{(0,0)},$$

with

$$\frac{\partial^4}{\partial\xi^2\,\partial\tau^2}(\varphi A_x)=\frac{\partial^4\varphi}{\partial\xi^2\,\partial\tau^2}A_x+\frac{\partial^2\varphi}{\partial\tau^2}\frac{\partial^2A_x}{\partial\xi^2}+\frac{\partial^2\varphi}{\partial\xi^2}\frac{\partial^2A_x}{\partial\tau^2}+f\,\frac{\partial^4A_x}{\partial\xi^2\,\partial\tau^2}$$
$$+2\left(\frac{\partial^3\varphi}{\partial\xi\,\partial\tau^2}\frac{\partial A_x}{\partial\xi}+\frac{\partial^3\varphi}{\partial\xi^2\,\partial\tau}\frac{\partial A_x}{\partial\tau}+\frac{\partial\varphi}{\partial\tau}\frac{\partial^3A_x}{\partial\xi^2\,\partial\tau}+\frac{\partial\varphi}{\partial\xi}\frac{\partial^3A_x}{\partial\tau\,\partial\tau^2}\right).$$

This holds for any distribution within Cohen's class and, in particular, for spectrograms $S_x^{(h)}(t,f)$ with a window $h(t)$. Given such a window, (normalized) spectrograms are known to correspond to the kernel function $\varphi(\xi,\tau)=\overline{A_h(\xi,\tau)}/E_h$. Assuming, furthermore, that the window is real valued and even, we can show that

$$\frac{\partial A_h}{\partial\xi}(0,0)=\frac{\partial A_h}{\partial\tau}(0,0)=0$$

and

$$\frac{\partial^3 A_h}{\partial \xi^2\, \partial \tau}\,(0,0)=\frac{\partial^3 A_h}{\partial \xi\, \partial \tau^2}\,(0,0)=0,$$

whereas, for any signal $x(t)$,

$$\frac{\partial^2 A_x}{\partial \xi^2}\,(0,0)=4\pi^2 E_x \Delta t_x^2,$$

$$\frac{\partial^2 A_x}{\partial \tau^2}\,(0,0)=4\pi^2 E_x \Delta f_x^2,$$

and

$$\frac{\partial^4 A_x}{\partial \xi^2\, \partial \tau^2}\,(0,0)=4\pi^2 E_x\left(\alpha_x-\frac{1}{2}\right).$$

It thus follows that, in the case of spectrograms, the overall spread measure (20) reads as announced.

Remark: Positivity of spectrograms guarantees that $\sigma_{tf}^2(S_x^{(h)})\geqslant 0$. In the special case where $h(t)=x(t)$ (matched spectrograms), we get $2\alpha_x-1+4\pi^2(2\Delta t_x^2\Delta f_x^2)\geqslant 0$ and therefore $\alpha_x\geqslant 1/2-4\pi^2\Delta t_x^2\Delta f_x^2\geqslant 1/4$, as already proved in Proposition 6.

IV. AFFINE TIME–FREQUENCY DISTRIBUTIONS

The Wigner–Ville distribution and, more generally, all members of Cohen's class, all belong to a family of time–frequency distributions which are covariant with respect to shifts in time and frequency. Moreover, assuming that an energy distribution depends quadratically on the signal, it can be shown that covariance with respect to shifts in time and frequency is sufficient to reduce the class of admissible solutions to Cohen's class. Following this line, it becomes therefore possible to generate other classes of distributions by modifying in a suitable manner the covariance requirement. Among the various solutions which have been derived this way, a prominent example is given by the class of the so-called *affine* time–frequency distributions, whose construction relies basically on the consideration of dilations in place of frequency shifts.

According to Ref. 20, the general class of affine time–frequency distributions $P_X^{(k)}(t,f)$ which are covariant with respect to each solvable three-parameter extension of the affine group is defined on $\mathbf{R}\times\mathbf{R}_+$ and can be parametrized by

$$P_X^{(k)}(t,f)=f^{2(r+1)-q}\int_{-\infty}^{+\infty}\mu_k(u)X(\lambda_k(u)f)\overline{X(\lambda_k(-u)f)}e^{-i2\pi tf\zeta_k(u)}\,du, \tag{21}$$

with $\zeta_k(u)=\lambda_k(u)-\lambda_k(-u)$, and where r and q are real-valued free parameters. In this expression, $X(f)$ stands as before for the Fourier spectrum of the analyzed signal, but only positive frequencies are considered. The parametrization function $\lambda_k(u)$ is given explicitly by

$$\lambda_k(u)=\left(k\,\frac{e^{-u}-1}{e^{-ku}-1}\right)^{1/(k-1)}, \tag{22}$$

with $k\in\mathbf{R}$, whereas $\mu_k(u)$ is a weight function whose value can be made dependent on $\lambda_k(u)$ so as to guarantee specific requirements for $P_X^{(k)}(t,f)$.[20]

From (21), it is apparent that affine distributions depend naturally on the product variable tf, a quantity which can be interpreted as a Mellin variable. To justify this point, let us recall that a definition of the Mellin transform $\underline{X}(s)$ of $X(f)$ can be given by[21]

$$\underline{X}(s)\equiv\int_0^{+\infty}X(f)f^{i2\pi s+r}\,df,$$

thus corresponding to an expansion of the spectrum $X(f)$ on ''hyperbolic chirps,''

$$C_0(f) \equiv f^r e^{-i2\pi s \log f}, \quad f \geqslant 0,$$

whose group delay is

$$t_{C_0}(f) = -\frac{1}{2\pi}\frac{d}{df}(-2\pi s \log f) = \frac{s}{f}.$$

In this picture, the Mellin variable s appears therefore as a *hyperbolic chirp rate* related to time and frequency by a relation of the type $s = tf$.

A companion interpretation amounts to consider the Mellin variable s as *a scale* variable.[22,3] This comes from the fact that, choosing, e.g., Weyl's correspondence rule, the operator $\mathscr{S}$ associated to $s = tf$ can be written as

$$\mathscr{S} = \frac{\mathscr{T}\mathscr{F} + \mathscr{F}\mathscr{T}}{2},$$

where $\mathscr{T}$ and $\mathscr{F}$ stand for the time and frequency operators, defined, respectively, by $(\mathscr{T}x)(t) = tx(t)$ and $(\mathscr{F}x)(t) = (-i/2\pi)\dot{x}(t)$. It thus follows that, for any $\lambda > 0$,

$$(e^{i2\pi\lambda S}X)(f) = e^{-\lambda/2}X(e^{-\lambda}f),$$

which allows us to interpret $\mathscr{S}$ as the infinitesimal generator of a scaling operator (whose eigenfunctions are precisely hyperbolic chirps).

Whatever the interpretation, and while $P_X^{(k)}(t,f)$ is usually considered as a time–frequency distribution, it can equally well be reparametrized as

$$\tilde{P}_X^{(k)}(s,f) \equiv P_X^{(k)}\left(\frac{s}{f}, f\right), \tag{23}$$

so that it becomes a function of (Mellin) scale and (Fourier) frequency variables.

A. Minimum uncertainty

Because the operators attached to scale and frequency do not commute, namely

$$[\mathscr{S},\mathscr{F}] \equiv \mathscr{S}\mathscr{F} - \mathscr{F}\mathscr{S} = \frac{i}{2\pi}\mathscr{F},$$

the corresponding variables obey a specific form of uncertainty relation. Precisely, let $X(f)$ be an analytic signal such that

$$E_X \equiv \int_0^{+\infty} |X(f)|^2 f^{2r+1} df < +\infty,$$

$$\frac{1}{E_X}\int_{-\infty}^{+\infty} s|\underline{X}(s)|^2 ds = s_0,$$

and

$$\frac{1}{E_X}\int_0^{+\infty} f|X(f)|^2 df = f_0.$$

If we denote by Δs_X and Δf_X its mean-square deviations in scale and frequency, defined, respectively, by

$$\Delta s_X^2 \equiv \frac{1}{E_X} \int_{-\infty}^{+\infty} (s-s_0)^2 |\underline{X}(s)|^2 \, ds \tag{24}$$

and

$$\Delta f_X^2 \equiv \frac{1}{E_X} \int_0^{+\infty} (f-f_0)^2 |X(f)|^2 f^{2r+1} \, df, \tag{25}$$

we can prove the following.[3,21–23]

Proposition 16: Scale and frequency satisfy the uncertainty relation

$$\Delta s_X \Delta f_X \geqslant \frac{f_0}{4\pi}, \tag{26}$$

with equality if and only if $X(f)$ is a ''Klauder wavelet'' of the form

$$X(f) = C f^{\alpha} e^{-\beta f} f^{i\gamma} U(f), \tag{27}$$

with $C \in \mathbf{C}$, $(\alpha,\beta,\gamma) \in \mathbf{R}_+ \times \mathbf{R}_+ \times \mathbf{R}$ and $U(\cdot)$ the unit step function.

Proof: In order to get simplified notations, let us introduce the quantities $\mathscr{S}_0 \equiv \mathscr{S} - s_0 \mathscr{I}$ and $\mathscr{F}_0 \equiv \mathscr{F} - f_0 \mathscr{I}$, with $\mathscr{I}$ the identity operator. It follows from the unitarity of the Mellin transform that

$$\Delta s_X^2 = \int_0^{+\infty} |(S_0 X)(f)|^2 f^{2r+1} \, df,$$

so that we can write, by using the Cauchy–Schwarz inequality,

$$E_X^2 \Delta s_X^2 \Delta f_X^2 = \int_0^{+\infty} |(S_0 X)(f)|^2 f^{2r+1} \, df \int_0^{+\infty} |(\mathscr{F}_0 X)(f)|^2 f^{2r+1} \, df$$

$$\geqslant \left| \int_0^{+\infty} \overline{(S_0 X)(f)} (\mathscr{F}_0 X)(f) f^{2r+1} \, df \right|^2 = \left| \int_0^{+\infty} \overline{X(f)} (S_0 \mathscr{F}_0 X)(f) f^{2r+1} \, df \right|^2.$$

Defining the anticommutator of any two operators $\mathscr{A}$ and $\mathscr{B}$ as $[\mathscr{A},\mathscr{B}]_+ \equiv \mathscr{A}\mathscr{B} + \mathscr{B}\mathscr{A}$ and remarking that $[\mathscr{S}_0,\mathscr{F}_0] = [\mathscr{S},\mathscr{F}]$, we get, in the specific case of scale and frequency, that

$$\mathscr{S}_0 \mathscr{F}_0 = \tfrac{1}{2}([\mathscr{S}_0,\mathscr{F}_0] + [\mathscr{S}_0,\mathscr{F}_0]_+) = \tfrac{1}{2}\left(\frac{i}{2\pi} \mathscr{F} + [\mathscr{S}_0,\mathscr{F}_0]_+ \right),$$

and it follows that

$$\Delta s_X^2 \Delta f_X^2 \geqslant \frac{f_0^2}{16\pi^2},$$

with equality if

$$[\mathscr{S}_0,\mathscr{F}_0]_+ = 0, \tag{28}$$

and for those signals $X(f)$ which satisfy, for some $\lambda \in \mathbf{C}$,

$$(\mathscr{S}_0 X)(f) = \lambda (\mathscr{F}_0 X)(f). \tag{29}$$

Plugging (29) into (28) yields that λ is necessarily purely imaginary (namely that $\lambda^2 = -\Delta s_X^2 / \Delta f_X^2$). By developing explicitly Eq. (29) which characterizes the signals with minimum uncertainty in scale and frequency, we end up with the differential equation

$$2 f \dot{X}(f) = [4\pi \, \mathrm{Im}\{\lambda\}(f-f_0) - 1 - i4\pi s_0] X(f),$$

whose solution is given by

$$X(f)=K\ \exp(2\pi\ \mathrm{Im}\{\lambda\}f-(2\pi\ \mathrm{Im}\{\lambda\}f_0+\tfrac{1}{2}+i2\pi s_0)\log f),$$

with $K\in\mathbf{C}$ and $\mathrm{Im}\{\lambda\}<0$, so that the signal is of finite energy. This solution, which is only defined for positive frequencies, is of the form (27), and this concludes the proof.

In order to get an analog of Proposition 4 in the case of Mellin–Fourier variables, it is natural to consider the *unitary* Bertrand distribution $B_X(t,f)\equiv P_X^{(0)}(t,f)$, which, in many respects, is the counterpart of the Wigner–Ville distribution in the Mellin–Fourier domain. By definition, the unitary Bertrand distribution (also referred to as the *tomographic* distribution[20]) is obtained from (21) by setting $k=0$ in (22), and it reads explicitly as

$$B_X(t,f)=f^{2(r+1)-q}\int_{-\infty}^{+\infty}\left(\frac{u/2}{\sinh(u/2)}\right)^{2(r+1)}X\left(\frac{(u/2)e^{u/2}f}{\sinh(u/2)}\right)\overline{X\left(\frac{(u/2)e^{-u/2}f}{\sinh(u/2)}\right)}e^{i2\pi ftu}\,du. \quad (30)$$

With this definition, we are led to the following.

Proposition 17: Given any $S>0$, *the unitary Bertrand distribution satisfies the Mellin-frequency uncertainty relation,*

$$\Sigma_{sf}^2(\tilde{B}_X)\equiv\frac{1}{E_X}\int_{-\infty}^{+\infty}\int_0^{+\infty}\left[\left(\frac{s-s_0}{S}\right)^2+S^2\left(\frac{f}{f_0}-1\right)^2\right]\tilde{B}_X(s,f)f^{q-1}ds\,df\geq\frac{1}{2\pi},$$

with equality, if and only if $X(f)$ *is a Klauder wavelet.*

Proof: The proof relies on marginal properties of the unitary Bertrand distribution, according to which

$$\int_{-\infty}^{+\infty}\tilde{B}_X(s,f)ds=f^{2(r+1)-q}|X(f)|^2$$

and

$$\int_0^{+\infty}\tilde{B}_X(s,f)f^{q-1}\,df=|\underline{X}(s)|^2.$$

Using the definitions (24) and (25), it follows directly from these relations that

$$\Sigma_{sf}^2(\tilde{B}_X)=\frac{\Delta s_X^2}{S^2}+S^2\,\frac{\Delta f_X^2}{f_0^2}=\left(\frac{\Delta s_X}{S}-S\,\frac{\Delta f_X}{f_0}\right)^2+2\,\frac{\Delta s_X\Delta f_X}{f_0}\geq 2\,\frac{\Delta s_X\Delta f_X}{f_0},$$

with equality if S is ''matched'' to the signal according to $S=\sqrt{f_0\Delta s_X/\Delta f_X}$. In this case, we know from (26) that the product $\Delta s_X\Delta f_X$ is lower bounded too, with equality in the only case of Klauder wavelets, whence the result.

B. Separability

The question we are interested in here can be stated as follows: is it possible to get affine distributions (23) which are separable in their Mellin and Fourier variables for some signal $X(f)$, i.e., such that they can be put in the form

$$\tilde{P}_X^{(k)}(s,f)=p(s)Q(f), \quad (31)$$

where $p(s)$ and $Q(f)$ are functions to be determined?

A first answer to this question can be looked after in the spirit of Proposition 3, which has been established thanks to the correct marginal properties of the Wigner–Ville distribution, in both time and frequency. In the specific case of the unitary Bertrand distribution (30), we have, in fact,[20] the two marginal properties

$$\int_{-\infty}^{+\infty} B_X(t,f)\,dt = f^{2r+1-q}|X(f)|^2 \tag{32}$$

and

$$\int_{-\infty}^{+\infty}\int_0^{+\infty} B_X(t,f)\,\delta\!\left(t-\frac{s}{f}\right) f^{q-1}\,dt\,df = |\underline{X}(s)|^2. \tag{33}$$

Within this new context, and considering first the frequency marginal property (32), it is easy to show that the separability condition (31) leads to

$$\begin{aligned}|X(f)|^2 &= f^{q-2r-1}\int_{-\infty}^{+\infty} B_X(t,f)\,dt\\ &= f^{q-2(r+1)}\int_{-\infty}^{+\infty} \tilde{B}(s,f)\,\frac{ds}{f} = f^{q-2(r+1)}\int_{-\infty}^{+\infty} p(s)Q(f)\,\frac{ds}{f} = f^{q-2(r+1)}P(0)\,\frac{Q(f)}{f} \geqslant 0,\end{aligned}$$

whereas, by using the Mellin marginal (33), we obtain, in a similar way,

$$\begin{aligned}|\underline{X}(s)|^2 &= \int_{-\infty}^{+\infty}\int_0^{+\infty} B_X(t,f)\,\delta\!\left(t-\frac{s}{f}\right) f^{q-1}dt\,df\\ &= \int_0^{+\infty} \tilde{B}_X(s,f) f^{q-1}\,df = p(s)\int_0^{+\infty} Q(f) f^{q-1}\,df \geqslant 0.\end{aligned}$$

Combining these two marginals, we get

$$P(0)\int_0^{+\infty} Q(f) f^{q-1}\,df = E_X \geqslant 0,$$

with the consequence that

$$\tilde{B}_X(s,f) = \frac{|\underline{X}(s)|^2 f^{2(r+1)-q}|X(f)|^2}{E_X} \geqslant 0.$$

We could therefore think of having proved that, as for Wigner–Ville, separability implies positivity. The situation is, however, different since, unfortunately, the assumption of separability can be proved not to be admissible for unitary Bertrand distributions. This is evidenced by the following.

Proposition 18: The only affine distributions which are compatible with separability are characterized by (21) and (22) with $k=-1$, 1/2, *or* 2.

Proof: Assuming that separability holds, we can take the Fourier transform of (31) with respect to s, and we get

$$\begin{aligned}P(\sigma)Q(f) &= \int_{-\infty}^{+\infty} \tilde{P}_X^{(k)}(s,f)e^{-i2\pi\sigma s}\,ds\\ &= f^{2(r+1)-q}\int_{-\infty}^{+\infty} \mu_k(u)X(\lambda_k(u)f)\overline{X(\lambda_k(-u)f)}\,\delta(\sigma-\zeta_k(u))du\\ &= f^{2(r+1)-q}\mu_k(\zeta_k^{-1}(\sigma))w(\sigma)X(\lambda_k(\zeta_k^{-1}(\sigma))f)\overline{X(\lambda_k(-\zeta_k^{-1}(\sigma))f)},\end{aligned}$$

with $w(\sigma)=(d/d\sigma)\zeta_k^{-1}(\sigma)$. Setting $\zeta_k^{-1}(\sigma)=u$, we end up with

$$P(\zeta_k(u))Q(f) = f^{2(r+1)-q}\mu_k(u)w(\zeta_k(u))X(\lambda_k(u)f)\overline{X(\lambda_k(-u)f)},$$

a relation which implies that

$$X(\lambda_k(u)f)\overline{X(\lambda_k(-u)f)}=\frac{P(\zeta_k(u))}{\mu_k(u)w(\zeta_k(u))}\,\frac{Q(f)}{f^{2(r+1)-q}}\equiv\tilde{P}(u)\tilde{Q}(f). \tag{34}$$

We know from Ref. 20 that $\lambda_k(0)=1$. Imposing $u=0$ in (34), we get therefore $|X(f)|^2=\tilde{P}(0)\tilde{Q}(f)$ and, hence,

$$\frac{X(\lambda_k(u)f)\overline{X(\lambda_k(-u)f)}}{|X(f)|^2}=\frac{\tilde{P}(u)}{\tilde{P}(0)}. \tag{35}$$

This states that the left-hand side of (35) is necessarily a function of u only. In order to find solutions (and conditions for their existence) to the functional equation (35), we can factor $X(f)$ in a (non-negative) modulus term and a phase term according to $X(f)=|X(f)|\exp\{i\arg X(f)\}\equiv\exp\{\Phi(f)+i\Psi(f)\}$, thus leading to

$$\frac{X(\lambda_k(u)f)\overline{X(\lambda_k(-u)f)}}{|X(f)|^2}=e^{\{\Phi(\lambda_k(u)f)+\Phi(\lambda_k(-u)f)-2\Phi(f)]+i[\Psi(\lambda_k(u)f)-\Psi(\lambda_k(-u)f)\}}.$$

Within this formulation, imposing (35) to hold requires that both the real and imaginary parts of the above exponential do not depend upon f. The condition on the phase $\Psi(f)$ is

$$\frac{\partial}{\partial f}[\Psi(\lambda_k(u)f)-\Psi(\lambda_k(-u)f)]=0,$$

or, equivalently [by letting $\lambda_k(u)f=\xi$ and $\lambda_k(-u)f=\zeta$], $\xi\dot{\Psi}(\xi)=\zeta\dot{\Psi}(\zeta)$ for any $(\xi,\zeta)\in\mathbf{R}_+^2$. This does not impose any restriction on the admissible distributions but constrains the phase to be of the form

$$\Psi(f)=a\log f+b, \tag{36}$$

with $(a,b)\in\mathbf{R}^2$. Concerning the modulus, the condition can be written as

$$\frac{\partial}{\partial f}[\Phi(\lambda_k(u)f)+\Phi(\lambda_k(-u)f)-2\Phi(f)]=0,$$

or, equivalently,

$$\dot{\Phi}(f)=\frac{\lambda_k(u)\dot{\Phi}(\lambda_k(u)f)+\lambda_k(-u)\dot{\Phi}(\lambda_k(-u)f)}{2}.$$

This expression can be rewritten by introducing, as before, $\lambda_k(u)f=\xi$ and $\lambda_k(-u)f=\zeta$, thus defining f as the *Stolarsky's generalized mean* of ξ and ζ,[24] whose expression reads explicitly as $f=\Theta(\xi,\zeta)$, with

$$\Theta(\xi,\zeta)=\left(\frac{1}{k}\,\frac{\xi^k-\zeta^k}{\xi-\zeta}\right)^{1/(k-1)}. \tag{37}$$

What we get is $\Theta(\xi,\zeta)\dot{\Phi}(\Theta(\xi,\zeta))=(\xi\dot{\Phi}(\xi)+\zeta\dot{\Phi}(\zeta))/2$, and finally, if we let

$$m(f)\equiv f\dot{\Phi}(f), \tag{38}$$

we end up with $m(\Theta(\xi,\zeta))=(m(\xi)+m(\zeta))/2$ or, equivalently,

$$\Theta(\xi,\zeta)=m^{-1}\left(\frac{m(\xi)+m(\zeta)}{2}\right).$$

We recognize therefore that the Stolarsky's generalized mean (37) must necessarily be of the form of a *quasiarithmetic generalized mean* (in the sense of Kolmogorov and Nagumo), a situation we know to be true only if $k=-1$, 1/2, or 2, as proved in (Ref. 25, Appendix A). This concludes the proof.

It appears therefore that the unitary Bertrand distribution $(k=0)$ is excluded from the class of affine distributions which may be separable in their Fourier and Mellin variables. The three cases mentioned in the Proposition can, however, be studied in further detail by specifying the corresponding generalized means and the associated distributions, and by identifying the signals which guarantee separability.

1. Case k=2

The generalized mean (37) reduces in this case to the arithmetic mean[25]

$$\Theta(\xi,\zeta)=\frac{\xi+\zeta}{2}\Leftrightarrow m(f)=af+b,$$

and, according to (38), $\dot{\Phi}(f)=a+b/f$, which leads to $\Phi(f)=af+b\log f+c$. Together with (36), the associated class of signals is therefore of the form

$$X_2(f)=Cf^{\alpha}e^{-\beta f}f^{i\gamma}U(f),$$

with $C\in\mathbf{C}$, $(\alpha,\beta,\gamma)\in\mathbf{R}^3$, and $U(\cdot)$ the unit step function, thus corresponding to the family of the ''Klauder wavelets'' (27) when α and β are both positive.

Assuming that $X(f)$ is analytic, i.e., vanishes for negative frequencies, it is known[20,25] that the case $(k=2,\ q=0,\ r=-1/2)$ corresponds to the Wigner–Ville distribution,

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

for a suitable choice of the weighting function $\mu_2(u)$, namely $\mu_2(u)=\cosh^{-2}(u/2)$. In this case, a direct calculation shows that

$$\tilde{W}_{X_2}(s,f)=2|C|^2f^{2\alpha+1}e^{-2\beta f}\int_{-1}^{+1}(1-w^2)^{\alpha}\left(\frac{1+w}{1-w}\right)^{i\gamma}e^{i4\pi sw}\,dw\;U(f).$$

One can remark that, in the real case where $\gamma=0$, this result can be given a closed form expression which reads as

$$\tilde{W}_{X_2}(s,f)=2|C|^2\sqrt{\pi}\Gamma(\alpha+1)f^{2\alpha+1}e^{-2\beta f}\,\frac{J_{\alpha+1/2}(4\pi|s|)}{(2\pi|s|)^{\alpha+1/2}}\,U(f),$$

where $J_\nu(\cdot)$ stands for the Bessel function of the first kind.[26]

Remark: Whereas in this special case, $\tilde{W}_{X_2}(s,f)$ is separable and is of the form $\tilde{W}_{X_2}(s,f)=k(s)f|X(f)|^2$, it can be checked that $k(s)$ is neither proportional to the Mellin density $|\underline{X}(s)|^2$ nor positive.

2. Case k=−1

The generalized mean (37) reduces in this case to the geometric mean,[25]

$$\Theta(\xi,\zeta)=\sqrt{\xi\zeta}\Leftrightarrow m(f)=a\log f+b,$$

and, according to (38), $\dot{\Phi}(f)=a(\log f)/f+b/f$, which leads to $\Phi(f)=(a/2)\log^2 f+b\log f+c$. Together with (36), the associated class of signals reads as

$$X_{-1}(f)=Cf^{\alpha}e^{-\beta\log^2 f}f^{i\gamma}U(f), \tag{39}$$

and it generalizes a family of wave forms defined by Altes[27] (the special case of the above expression with $\alpha=0$ and $\beta>0$).

It is known[20,25] that the case ($k=-1$, $q=0$, $r=-1/2$) corresponds to the so-called *Unterberger distribution*, which, in its ''active form'' [i.e., when setting $\mu(u)=\cosh(u/2)^{20}$], reads as

$$\tilde{U}_X(s,f)=f\int_{-\infty}^{+\infty}\left(1+\frac{1}{\gamma^2}\right)X(\gamma f)\overline{X\left(\frac{f}{\gamma}\right)}e^{i2\pi(\gamma-1/\gamma)s}\,d\gamma. \tag{40}$$

A direct calculation shows that

$$\tilde{U}_{X_{-1}}(s,f)=|C|^2f^{2\alpha+1}e^{-2\beta(\log^2 f+\gamma^2/4)}\int_{-\infty}^{+\infty}e^{-2\beta(a\sinh w-i(\gamma/2))^2}e^{i4\pi sw}\,dw\;U(f).$$

Remark: The family of wave forms (39) has been introduced by Altes in a context of active sonar, such as encountered in bat echolocation. They are, in fact, solutions to the problem of ''Doppler tolerance,'' which consists in estimating at best (no bias and the highest signal-to-noise ratio) a time delay in the presence of some unknown Doppler shift.

3. Case k=1/2

The generalized mean (37) reduces in this case to the square-root mean[25]

$$\Theta(\xi,\zeta)=\left(\frac{\sqrt{\xi}+\sqrt{\zeta}}{2}\right)^2\Leftrightarrow m(f)=a\sqrt{f}+b,$$

and, according to (38), $\dot{\Phi}(f)=a/\sqrt{f}+b/f$, which leads to $\Phi(f)=2a\sqrt{f}+b\log f+c$. Together with (36), the associated class of signals reads as

$$X_{1/2}(f)=Cf^{\alpha}e^{-\beta\sqrt{f}}f^{i\gamma}U(f).$$

It is known[20,25] that the case ($k=1/2$, $q=0$, $r=-1/2$) comprises the so-called ''D distribution:''

$$\tilde{D}_X(s,f)=f\int_{-4}^{+4}\left[1-\left(\frac{\gamma}{4}\right)^2\right]X\left(\left(1+\frac{\gamma}{4}\right)^2f\right)\overline{X\left(\left(1-\frac{\gamma}{4}\right)^2f\right)}e^{i2\pi s\gamma}d\gamma,$$

and a direct calculation shows that

$$\tilde{D}_{X_{1/2}}(s,f)=4|C|^2f^{2\alpha+1}e^{-2\beta\sqrt{f}}\int_{-1}^{+1}(1-w^2)^{\alpha}\left(\frac{1+w}{1-w}\right)^{i\gamma}e^{i8\pi sw}\,dw\;U(f).$$

If $\gamma=0$, this result can be simplified to

$$\tilde{D}_{X_{1/2}}(s,f)=4|C|^2\sqrt{\pi}\Gamma(\alpha+1)f^{2\alpha+1}e^{-2\beta\sqrt{f}}\,\frac{J_{\alpha+1/2}(8\pi|s|)}{(4\pi|s|)^{\alpha+1/2}}\,U(f).$$

C. Positivity

It has already been mentioned that Wigner–Ville distributions attain generally negative values, but that they are everywhere positive when applied to a Gaussian signal and, more generally, to the exponential of a quadratic polynomial.[6] It is our purpose in this section to point out related results in the case of affine time–frequency distributions.

1. Unitary Bertrand distributions

We will first show that, in the case of the (unitary) Bertrand distribution (30), ''Klauder wavelets'' play a role which has much to share with the one played by Gaussians in the Wigner–Ville case. To this end, it is convenient to generalize and reparametrize the Klauder wavelet (27) as

$$X_2(f)\equiv Cf^{-(r+1)}e^{2\pi\lambda(f_0\log f-f)}e^{-i2\pi(\beta_0\log f+\xi f)}U(f), \tag{41}$$

with $C\in\mathbf{C}$, and r, λ, f_0, β_0, and ξ real-valued parameters such that $\lambda\geqslant 0$ and $f_0>0$. This reformulation allows a simple physical interpretation of the parameters since the mean frequency f_m and the Q factor defined by $Q\equiv f_m/\Delta f$ are given, respectively, by $f_m=f_0$ and $Q^2=4\pi\lambda$. Moreover, the phase term is written so as to correspond to a modulation with a hyperbolic group delay $t_{X_2}(f)=\xi+\beta_0/f$.

Given these notations, we have the following.

Proposition 19: The unitary Bertrand distribution of the Klauder wavelet (41) is everywhere non-negative.

Proof: The proof relies on properties of characteristic functions (which are Fourier transforms of non-negative functions), and it proceeds as follows. We first start from the definition (41) of the Klauder wavelet and we plug it into the definition (30) of the Bertrand distribution. After some manipulations, the result can be expressed as

$$B_{X_2}(t,f)=|C|^2 f^{\alpha-q}\int_{-\infty}^{+\infty}M_{\alpha,\beta}(u)e^{-i2\pi\zeta u}\,du\;U(f), \tag{42}$$

with

$$M_{\alpha,\beta}(u)\equiv\left(\frac{u/2}{\sinh(u/2)}\right)^{\alpha}\exp\{-\beta u\coth(u/2)\} \tag{43}$$

and

$$\alpha\equiv 4\pi\lambda f_0;\quad \beta\equiv 2\pi\lambda f;\quad \zeta\equiv\beta_0-(t-\xi)f. \tag{44}$$

In order to prove that (42) corresponds to a non-negative quantity, it is then sufficient to show that $M_{\alpha,\beta}(u)$ is—up to a positive constant—a characteristic function, a property which will obviously hold if each of the two factors of (43) is itself proportional to a characteristic function.[28] Using first the fact that[29]

$$\sinh x=x\prod_{k=1}^{\infty}\left(1+\frac{x^2}{k^2\pi^2}\right),$$

we can write

$$\left(\frac{u/2}{\sinh(u/2)}\right)^{\alpha}=\prod_{k=1}^{\infty}\left(1+i\,\frac{u}{2k\pi}\right)^{-\alpha}\left(1-i\,\frac{u}{2k\pi}\right)^{-\alpha}.$$

This is clearly a characteristic function since, α being positive, each of the factors of this infinite product is the characteristic function of a Gamma distribution.[28] Using then the fact that[29]

$$\coth\pi x=\frac{1}{\pi x}+\frac{2x}{\pi}\sum_{k=1}^{\infty}\frac{1}{x^2+k^2},$$

we can write.

$$\exp\{-\beta u\coth(u/2)\}=e^{-2\beta}\prod_{k=1}^{\infty}\exp\left\{4\beta\left(\frac{1}{1+(u/2k\pi)^2}-1\right)\right\}. \tag{45}$$

Since $(1+(u/2k\pi)^2)^{-1}$ is a characteristic function (namely the characteristic function of a Laplace distribution) and $\beta\geqslant 0$, it suffices to make use of a theorem by Lukacs[28]—stating that, if $g(u)$ is a characteristic function and γ a positive real number, $\exp\{\gamma(g(u)-1)\}$ is itself a characteristic function—for guaranteeing that each of the factors of (45) and, hence, (45) itself is a characteristic function. The two factors of (43) each being a characteristic function, the Fourier transform of their product is everywhere non-negative and the proof is complete.

Two remarks.

(1) Assuming as we did that $\lambda > 0$, we have $X_2(f) \in L^2(\mathbf{R}_+, f^{2r+1}\, df)$ and $B_{X_2}(t,f)$ is a bounded square-integrable function. A degeneracy is, however, observed in the case of the ''hyperbolic chirps'' $C_0(f)$ corresponding to $X_2(f)$ with $\lambda = 0$, and for which we know[20] that

$$B_{C_0}(t,f) = |C|^2 f^{-q-1} \delta\left(t - \left(\xi + \frac{\beta_0}{f}\right)\right) U(f),$$

a result which also follows directly from (42), (43), and (44), since we have then $\alpha = \beta = 0$, and hence $M_{\alpha,\beta} = 1$.

(2) According to the definition (27), the nonzero frequency f_0 can be interpreted as the central frequency of the (envelope of the) Klauder wavelet. Within the narrow band assumption $Q^2 = 4\pi\lambda \gg 1$, we get the approximation

$$X_2(f) \sim C' \exp\left\{ -\frac{\pi}{f_0}(\lambda - i\beta_0)\delta f^2 - i2\pi\left(\xi + \frac{\beta_0}{f_0}\right)\delta f \right\},$$

where $\delta f \equiv f - f_0$ and C' is some suitable (complex-valued) constant. This means that narrow band Klauder wavelets reduce to Gaussians, in clear accordance with the fact that the Bertrand distribution reduces to the Wigner–Ville one for narrow band signals,[20] and that Wigner–Ville distributions are non-negative for Gaussians.

2. Unterberger distributions

A result similar to the one of the previous Proposition can be obtained in the case of the active Unterberger distribution (40), but it requires replacing the Klauder wavelet by the signal

$$X_{-1}(f) \equiv C f^{-(r+1)} e^{-2\pi\lambda(\alpha_0/f + f)} e^{-i2\pi(\xi f - \alpha/f)} U(f),$$

with $\alpha_0 > 0$. More precisely, we can prove the following.

Proposition 20: The active Unterberger distribution of $X_{-1}(f)$ is everywhere non-negative.

Proof: The proof follows from a direct calculation according to which

$$U_{X_{-1}}(t,f) = 2|C|^2 f^{-(2r+1)} \int_{-\infty}^{+\infty} e^{-4\pi\lambda(\alpha_0/f + f)\sqrt{1+(w/2)^2}} e^{i4\pi(tf - \xi f - \alpha/f)w}\, dw\ U(f),$$

a result which can be equivalently expressed as (see, e.g., Ref. 30)

$$U_{X_{-1}}(t,f) = 4|C|^2 f^{-(2r+1)} G_\beta\left(4\pi\left(tf - \xi f - \frac{\alpha}{f}\right)\right) U(f),$$

with $\beta \equiv 4\pi\lambda(f + \alpha_0/f)$ and

$$G_\beta(y) \equiv \frac{2\beta}{\sqrt{\beta^2 + y^2}} K_1(\beta^2 + y^2), \tag{46}$$

$K_1(\cdot)$ being a modified Bessel function of the first kind.[26] Positivity of the distribution is therefore guaranteed by the simple fact that modified Bessel functions $K_\nu(z)$ are known to be positive if $\nu > -1$ and $z > 0$.[26]

Unterberger distributions also exist under a ''passive'' form,[20] whose definition reads as

$$V_X(t,f) = f \int_{-\infty}^{+\infty} X(\gamma f) \overline{X\left(\frac{f}{\gamma}\right)} e^{i2\pi(\gamma - 1/\gamma)tf} \frac{d\gamma}{\gamma}.$$

''Active'' Unterberger distributions are known to be perfectly localized for signals whose group delay is defined by squared hyperbolas of the form $\alpha + \gamma/f^2$, but they are also known not to be unitary (in fact, it can be shown[20] that unitary Bertrand distributions are the only distributions guaranteeing both localization and unitarity). ''Passive'' Unterberger distributions share a lot of properties with ''active'' ones, at the notable exception of localization. Both forms are, in fact, ''dual'' in the sense that, whereas none has simultaneously properties similar to the unitary Bertrand distribution, they can be combined so as to satisfy the Moyal-type formula

$$\int_{-\infty}^{+\infty}\int_{0}^{+\infty} U_X(t,f)V_Y(t,f)\,dt\,\frac{df}{f}=\left|\int_{0}^{+\infty} X(f)\overline{Y(f)}\,\frac{df}{f}\right|^2.$$

With the above definition for $V_X(t,f)$, the following can be proved.

Proposition 21: The passive Unterberger distribution of $X_{-1}(f)$ is everywhere non-negative.

Proof: This can be checked by a direct calculation, according to which

$$V_X(t,f)=2|C|^2 f^{-(2r+1)}K_0\left(\sqrt{\beta^2+4\pi^2\left(tf-\xi f-\frac{\alpha}{f}\right)^2}\right)U(f),$$

which concludes the proof exactly as in the previous Proposition.

Remark: The closed form expression given above can be found almost exactly in Ref. 31, where the authors introduce a time-scale energy distribution whose definition has essentially the form of a passive Unterberger distribution.

In the case of the Wigner–Ville distribution, Hudson's theorem guarantees that Gaussian signals are the only ones ensuring positivity, whereas the question is left open to prove that Klauder wavelets would be the only signals with a (unitary) Bertrand distribution everywhere positive. In the case of the (active) Unterberger distribution, a class of signals ensuring positivity has been evidenced in Proposition 20, but it can be proved that this solution is not unique. Two counterexamples to unicity for positive Unterberger distributions are provided by the two following claims.

Proposition 22: The active Unterberger distribution of

$$Y_{-1}(f)\equiv Cf^{-(r+1)}e^{-2\pi\lambda f^2}e^{-i2\pi(\xi f-\alpha/f)}U(f),$$

is everywhere non-negative.

Proof: The active Unterberger distribution (40) admits the (time–frequency) equivalent form

$$U_X(t,f)=2f\int_{-\infty}^{+\infty}\cosh u\,X(e^u f)\overline{X(e^{-u}f)}e^{i4\pi tf\sinh u}\,du\,U(f).$$

Plugging the definition of $Y_{-1}(f)$ into this expression yields

$$U_{Y_{-1}}(t,f)=\frac{|C|^2}{\sqrt{2\lambda}}\,f^{-(2r+1)}e^{-4\pi\lambda f^2}e^{(-2\pi/\lambda)(t-\xi-\alpha/f^2)^2}U(f),$$

whence the result.

Remark: When applied to $Y_{-1}(f)$ with $\alpha=0$ (a constant group delay), the active Unterberger distribution is not only positive, but also separable in time and frequency.

Proposition 23: The active Unterberger distribution of

$$Z_{-1}(f)\equiv Cf^{-(r+1)}e^{-2\pi\lambda f}e^{-i2\pi(\xi f-\alpha/f)}U(f)$$

is everywhere non-negative.

Proof: The proof follows from a direct calculation, according to which

$$U_{Z_{-1}}(t,f)=2|C|^2 f^{-(2r+1)}G_{4\pi\lambda f}\left(4\pi\left(tf-\xi f-\frac{\alpha}{f}\right)\right)U(f), \tag{47}$$

with $G(\cdot)$ defined as in (46), and positivity stemming from the usual assumption $\lambda>0$.

V. CONCLUDING REMARKS

In this paper, a number of results have been given which—although still preliminary in many respects and far from exhausting the subject—are believed to clarify some of the links which exist between positivity, separability, and minimum uncertainty in time–frequency energy distributions. In fact, whereas in the case of the Wigner–Ville distribution, these features may admit order relations of the type ''separability⇒positivity⇒minimum uncertainty,'' it has been shown that the situation of other distributions is more intricate, with possible incompatibilities between the three different properties. Affine distributions (based on frequency and on a Mellin variable, in place of the usual time variable) have especially been considered in some detail, with two main consequences which make them depart from the Wigner–Ville case.

(1) In comparison with the known fact that minimizers of the classical time–frequency uncertainty—Gaussians with a linear phase—have a positive and separable Wigner–Ville distribution, signals with minimum frequency-scale uncertainty—namely, Klauder wavelets—do possess a positive (unitary) Bertrand distribution, but this latter is not separable.
(2) Whereas positivity is an exception in the Wigner–Ville case and can be observed with Gaussian signals only, it turns out that this situation of unicity is no longer true in the affine case, different classes of signals having been evidenced to lead to positive distributions.

As already mentioned, many questions are still left open, such as, e.g., the question of unicity for the positivity of *unitary* Bertrand distributions. From another perspective, it can be finally mentioned that other types of extensions (e.g., to the hyperbolic class[32] and, more generally, to ''warped'' classes[33]) could have been considered, but have not, since the corresponding results can be readily anticipated from those given here. In fact, such extensions being related to usual classes (Cohen of affine) by a warping operation, all the results obtained in one case can be transformed, *mutatis mutandis*, to get the corresponding results in the other case. For instance, the Altes-Marinovic distribution,[34]

$$Q_X(s,f)=f\int_{-\infty}^{+\infty} X(e^{u/2}f)\overline{X(e^{-u/2}f)}e^{i2\pi su}\,du,$$

which is known to belong to (and be a central member of) the hyperbolic class,[32] can be expressed as well in terms of the usual Wigner distribution as

$$Q_X(s,f)=W_{\tilde{X}}\left(\frac{s}{f_0},f_0\log\frac{f}{f_0}\right),$$

with $\tilde{X}(f)=e^{f/2f_0}X(e^{f/f_0})$ and $f_0>0$. It follows that any result pertaining to the Wigner–Ville distribution can be transposed to the Altes–Marinovic distribution, provided that it is applied to signals warped the suitable way. In particular, an analog of Hudson's theorem can be given as (no new proof is really necessary)

Proposition 24: The Altes–Marinovic distribution is positive for ''Altes signals'' of the type (39), and only for them.

ACKNOWLEDGMENTS

Thanks to Richard G. Baraniuk for many useful discussions and comments, and to the anonymous reviewers for their careful reading of the manuscript.

[1] E. P. Wigner, ''On the quantum correction for thermodynamic equilibrium,'' Phys. Rev. **40**, 749–759 (1932).
[2] J. Ville, ''Théorie et applications de la notion de signal analytique,'' Câbles Transm. **2ème A**, 61–74 (1948).
[3] L. Cohen, *Time–Frequency Analysis* (Prentice–Hall, Englewood Cliffs, 1995).
[4] P. Flandrin, *Temps–Fréquence* (Hermès, Paris, 1993).
[5] D. Gabor, ''Theory of communication,'' J. Inst. Electr. Eng. **93**, 429–457 (1946).
[6] R. L. Hudson, ''When is the Wigner quasi-probability density non-negative?,'' Rep. Math. Phys. **6**, 249–252 (1974).
[7] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, ''On the time–frequency discrimination of energy distributions—Can they look sharper than Heisenberg?,'' *IEEE-ICASSP'84*, San Diego, CA, 1984, pp. 41B7.1–41B7.4.
[8] K. Urbanik, ''Joint probability distributions of observables in quantum mechanics,'' Stud. Math. **XXI**, 117–133 (1961).
[9] A. J. E. M. Janssen, ''Positivity of weighted Wigner distributions,'' SIAM J. Math. Anal. **12**, 752–758 (1981).

[10] A. J. E. M. Janssen, ''Positivity properties of phase-plane distribution functions,'' J. Math. Phys. **25**, 2240–2252 (1984).
[11] A. J. E. M. Janssen, ''Bilinear phase-plane distribution functions and positivity,'' J. Math. Phys. **26**, 1986–1994 (1985).
[12] A. J. E. M. Janssen, ''Positivity of time–frequency distribution functions,'' Signal Process. **14**, 243–252 (1988).
[13] J. Aczél, *Lectures on Functional Equations and Their Applications* (Academic, New York, 1966).
[14] T. A. C. M. Claasen and W. F. G. Mecklenbräuker, ''The Wigner distribution—A tool for time–frequency signal analysis. Part I: Continuous-time signals,'' Philips J. Res. **35**, 217–250 (1980).
[15] A. J. E. M. Janssen, ''On the locus and spread of pseudo-density functions in the time–frequency plane,'' Philips J. Res. **37**, 79–110 (1982).
[16] R. A. Altes, ''Methods of wide-band signal design for radar and sonar systems,'' Ph.D. dissertation, University of Rochester, Rochester, NY, 1970.
[17] N. J. De Bruijn, ''Uncertainty principles in Fourier analysis,'' in *Inequalities*, edited by O. Shisha (Academic, New York, 1967), pp. 55–71.
[18] E. H. Lieb, ''Integral bounds for radar ambiguity functions and Wigner distributions,'' J. Math. Phys. **31**, 594–599 (1990).
[19] A. J. E. M. Janssen, ''Optimality property of the Gaussian window spectrogram and an application to lattice expansions,'' IEEE Trans. Signal Process. **SP-39**, 202–204 (1991).
[20] J. Bertrand and P. Bertrand, ''A class of affine Wigner distributions with extended covariance properties,'' J. Math. Phys. **33**, 2515–2527 (1992).
[21] J. Ph. Ovarlez, ''La transformation de Mellin—Un outil pour l'analyse des signaux à large bande,'' Thèse de doctorat, Univ. Paris 6, 1992.
[22] L. Cohen, ''The scale representation,'' IEEE Trans. Signal Process. **41**, 3275–3292 (1993).
[23] J. R. Klauder, ''Path integrals for affine variables,'' in *Functional Integration: Theory and Applications*, edited by J.-P. Antoine and E. Tirapegui (Plenum, New York, 1980), pp. 101–119.
[24] K. B. Stolarsky, ''Generalizations of the logarithmic mean,'' Math. Mag. **48**, 87–92 (1975).
[25] P. Flandrin and P. Gonçalvès, ''Geometry of affine time–frequency distributions,'' Appl. Comp. Harm. Anal. **3**, 10–39 (1996).
[26] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables* (Dover, New York, 1972).
[27] R. A. Altes, ''Sonar for generalized target description and its similarity to animal echolocation systems,'' J. Acoust. Soc. Am. **59**, 97–105 (1976).
[28] E. Lukacs, *Characteristic Functions* (Griffin, London, 1970).
[29] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products* (Academic, New York, 1980).
[30] F. Oberhettinger, *Fourier Transforms of Distributions and Their Inverses—A Collection of Tables* (Academic, New York, 1973).
[31] B. Escudié and A. Grossmann, ''Une représentation bilinéaire en temps et échelle,'' 13ème Coll. GRETSI, Juan-les-Pins, France, 1991, pp. 33–36.
[32] A. Papandreou, F. Hlawatsch, and G. F. Boudreaux-Bartels, ''The hyperbolic class of quadratic time–frequency representations. Part I: Constant-Q warping, the hyperbolic peradigm, properties and members,'' IEEE Trans. Signal Process. **SP-41**, 3425–3444 (1993).
[33] R. G. Baraniuk and D. L. Jones, ''Unitary equivalence: A new twist on signal processing,'' IEEE Trans. Signal Process. **SP-43**, 2269–2282 (1995).
[34] R. A. Altes, ''Wideband, proportional-bandwidth Wigner–Ville analysis,'' IEEE Trans. Acoust., Speech, Signal Process. **ASSP-38**, 1005–1012 (1990).