REFERENCES

[1] G. D. Forney, "Coset codes—Part I: Introduction and geometrical classification," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1123–1151, Sept. 1998.

[2] G. D. Forney and L. Wei, "Multidimensional constellations—Part I: Introduction, figures of merit, and generalized cross constellations," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 941–958, Aug. 1989.

[3] A. R. Calderbank and L. H. Ozarow, "Nonequiprobable signaling on the Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 36, pp. 726–740, July 1990.

[4] G. D. Forney, "Trellis shaping," *IEEE Trans. Inform. Theory*, vol. 38, pp. 281–300, Mar. 1992.

[5] A. K. Khandani and P. Kabal, "Shaping multidimensional signal spaces—Part I: Optimum shaping, shell mapping," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1799–1808, Nov. 1993.

[6] J. N. Livingston, "Shaping using variable-size regions," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1347–1353, July 1992.

[7] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 39, pp. 913–929, May 1993.

[8] A. R. Calderbank and M. Klimesh, "Balanced codes and nonequiprobable signaling," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1119–1122, May 1992.

[9] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proc. IEEE*, vol. 85, pp. 265–298, Feb. 1997.

[10] G. P. Agrawal, *Fiber-Optic Communication Systems.* New York: Wiley, 1997.

[11] J. G. Proakis, *Digital Communications,* 3rd ed. New York: McGraw-Hill, 1995.

On the Inverse Windowed Fourier Transform

Laura Rebollo-Neira and Juan Fernandez-Rubio

Abstract— **The inversion problem concerning the windowed Fourier transform is considered. It is shown that, out of the infinite solutions that the problem admits, the windowed Fourier transform is the "optimal" solution according to a maximum-entropy selection criterion.**

Index Terms— **Gabor transform, inversion problems, maximum entropy, windowed Fourier transform.**

I. INTRODUCTION

The use of a generalized Fourier integral to convey simultaneous time and frequency information was first introduced by Gabor (1946). In [2], he defines a windowed Fourier integral, using a Gaussian window. Later, the window was generalized to any function in $L^2(R)$, the space of square integrable functions. The so generalized Gabor

transform is mostly referred to as the windowed Fourier transform (WFT).

Restricting the space of signals to $L^2(R)$, the WFT is a mapping from $L^2(R)$ to $L^2(R^2)$ which is not bijective. As a consequence, lack of uniqueness of the inverse problem must be expected. In this contribution, we focus on a statistical analysis of the inversion problem. First the problem is shown to admit an infinite number of solutions. We then work on the space of possible solutions adopting a statistical description as the essential tool. The possible solutions are considered as a stochastic process distributed according to a (to be determined) probability density. The desired solution is estimated as the mean value of the random process. Among all the probability densities capable of yielding admissible mean-value solutions we single out *one*, adopting the maximum-entropy principle (MEP). Finally, we show that, from the maximum-entropy (ME) probability density a mean-value solution is inferred which is identical to the WFT. Thereby the WFT is shown to be an "optimal" solution according to an ME selection criterion. This result also holds as a property within the *Frame Theory* [9].

II. THE WFT INVERSE PROBLEM

Definition: Let $f(x) \in L^2(R)$ be a given signal and $g(x) \in L^2(R)$ be any fixed function in $L^2(R)$. The WFT of $f(x)$ is a function $F(\omega, t) \in L^2(R^2)$ defined by

$$F(\omega, t) = \langle e^{i\omega x} g(x-t) \mid f(x) \rangle = \int_R e^{-i\omega x} g^*(x-t) f(x) \, dx \quad (1)$$

where $g^*(x)$ denotes the complex conjugate of $g(x)$.

The signal can be reconstructed from its WFT through the inversion formula [1], [3], [6]

$$f(x) = \frac{1}{C_g} \int_{R^2} e^{i\omega x} g(x-t) F(\omega, t) \, d\omega \, dt \quad (2)$$

where

$$C_g = \|g\|^2 = \int_R |g(x)|^2 \, dx.$$

Although the inversion formula (2) allows the recovery of a signal from its WFT, the inversion is not unique. Let us denote $\mathcal{W}$ to the image of the WFT, i.e.,

$$\mathcal{W} = \left\{ F(\omega, t); F(\omega, t) = \int_R e^{-i\omega x} g^*(x-t) f(x) \, dx; \right.$$
$$\left. \text{for some } f(x) \in L^2(R) \right\}. \quad (3)$$

$\mathcal{W}$ is only a closed subspace, not all of $L^2(R^2)$ (not every function $h(\omega, t) \in L^2(R^2)$ belongs to $\mathcal{W}$). The next theorem, whose proof is given in [6, p. 56], provides the necessary and sufficient condition for $h(\omega, t) \in \mathcal{W}$.

Theorem 1: A function $h(\omega, t)$ belongs to $\mathcal{W}$ if and only if it is square integrable and, in addition, satisfies

$$h(\omega', t') = \frac{1}{C_g} \int_{R^2} K(\omega', t', \omega, t) h(\omega, t) \, d\omega \, dt \quad (4)$$

where

$$K(\omega', t', \omega, t) = \langle e^{i\omega' x} g(x-t') \mid e^{i\omega x} g(x-t) \rangle$$
$$= \int_R e^{-i\omega' x} g^*(x-t') e^{i\omega x} g(x-t) \, dx. \quad (5)$$

The function $K(\omega', t', \omega, t)$ is called the *reproducing kernel* determined by the window g and (4) is called the associated *consistency condition*.

Theorem 2: All functions $h^\perp(\omega, t)$ belonging to $\mathcal{W}^\perp$ (the orthogonal complement of $\mathcal{W}$) satisfy

$$\int_{R^2} e^{i\omega x} g(x-t) h^\perp(\omega, t) \, d\omega \, dt = 0. \tag{6}$$

Proof: Multiplying the right-hand side of (6) by $e^{-i\omega' x} g^*(x-t')$ and integrating over x we have

$$\int_{R^2} \left(\int_R e^{i\omega x} g(x-t) e^{-i\omega' x} g^*(x-t') \, dx \right) h^\perp(\omega, t) \, d\omega \, dt. \tag{7}$$

From the definition of $\mathcal{W}$ it follows that

$$\int_R e^{-i\omega x} g^*(x-t) e^{i\omega' x} g(x-t') \, dx \in \mathcal{W}.$$

Consequently,

$$\int_{R^2} \left(\int_R e^{i\omega x} g(x-t) e^{-i\omega' x} g^*(x-t') \, dx \right) h^\perp(\omega, t) \, d\omega \, dt$$
$$= \int_{R^2} \left(\int_R e^{-i\omega x} g^*(x-t) e^{i\omega' x} g(x-t') \, dx \right)^* h^\perp(\omega, t) \, d\omega \, dt$$
$$= 0 \tag{8}$$

because $h^\perp(\omega, t) \in \mathcal{W}^\perp$ is orthogonal to every function in $\mathcal{W}$.

Notice that (8) can be recast in the form

$$\langle e^{i\omega' x} g(x-t') \mid F(x) \rangle = \int_R e^{-i\omega' x} g^*(x-t') F(x) \, dx = 0 \tag{9}$$

where

$$F(x) = \int_{R^2} e^{i\omega x} g(x-t) h^\perp(\omega, t) \, d\omega \, dt \tag{10}$$

and since

$$\text{span}\{e^{i\omega' x} g(x-t')\}_{(\omega', t') \in R^2}$$

is dense in $L^2(R)$

$$\langle e^{i\omega' x} g(x-t') \mid F(x) \rangle = 0, \qquad \forall (\omega', t')$$

implies $F(x) \equiv 0$, whereby the proof is completed. $\square$

The lack of uniqueness of the inverse WFT is an immediate consequence of Theorem 2. Indeed, besides $F(\omega, t)$, for any $h^\perp(\omega, t) \in \mathcal{W}^\perp$ the function $h(\omega, t) = F(\omega, t) + h^\perp(\omega, t)$ also reconstructs the same signal. The inversion formula (2) corresponds to the particular choice $h^\perp(\omega, t) = 0$ and obviously gives rise to a solution of the inverse problem which is "optimal" in a minimum norm (MN) sense. The MN requirement may be a reasonable criterion to be adopted in the case of some applications, but, *a priori*, certainly not in *all* of them. In this correspondence we address the problem of deciding on an appropriate estimate for the unknown solution $h(\omega, t)$ by recourse to a postulate originally conceived for the purpose of making decisions in indeterminate situations, namely, the MEP [4], [5]. In the next section we show that the WFT is also an "optimal" solution of the inverse problem according to a ME selection criterion, as it turns out to be the mean of the probability density that maximizes the entropy.

III. THE ME APPROACH

The problem we address now is that of inverting for h the equation

$$f(x) = \frac{1}{C_g} \int_{R^2} e^{i\omega x} g(x-t) h(\omega, t) \, d\omega \, dt. \tag{11}$$

We begin by splitting the above complex equation into real and imaginary parts so that it becomes

$$f^u(x) = \frac{1}{C_g} \int_{R^2} \left(g_{\omega,t}^u(x) h^u(\omega, t) - g_{\omega,t}^v(x) h^v(\omega, t) \right) d\omega \, dt \tag{12}$$

$$f^v(x) = \frac{1}{C_g} \int_{R^2} \left(g_{\omega,t}^v(x) h^u(\omega, t) + g_{\omega,t}^u(x) h^v(\omega, t) \right) d\omega \, dt \tag{13}$$

where $f^u(x)$, $f^v(x)$ are the real and imaginary parts of $f(x)$ whereas $h^u(\omega, t)$, $h^v(\omega, t)$ are the real and imaginary parts of $h(\omega, t)$ and $g_{\omega,t}^u(x)$, $g_{\omega,t}^v(x)$ are the real and imaginary parts of $e^{i\omega x} g(x-t)$, respectively.

As discussed in the previous section, there exist several functions $h(\omega, t)$ capable of satisfying (12) and (13). Our aim is that of selecting *one* of those solutions as "optimal" in an ME sense. In order to achieve such a goal, we regard the possible solutions as a stochastic process and estimate the desired solution as its mean value that we denote

$$\overline{h(\omega, t)} = \overline{h^u(\omega, t)} + i\overline{h^v(\omega, t)}; (\omega, t) \in R^2.$$

To deal with the stochastic process in a discrete way, we divide R^2 into squares of area $\Delta r = \frac{1}{M}$, centered at the points $r_j = (\omega_j, t_j)$ and take $\lim M \to \infty$. With this discretization, (12) and (13), which provide the constraints to be satisfied by the desired solution, are evaluated as

$$f^u(x) = \lim_{M \to \infty} \frac{1}{MC_g} \sum_{j=1}^M \left(g_{r_j}^u(x) \overline{h^u(r_j)} - g_{r_j}^v(x) \overline{h^v(r_j)} \right) \tag{14}$$

$$f^v(x) = \lim_{M \to \infty} \frac{1}{MC_g} \sum_{j=1}^M \left(g_{r_j}^v(x) \overline{h^u(r_j)} + g_{r_j}^u(x) \overline{h^v(r_j)} \right). \tag{15}$$

At a fixed point r_j, both $h^u(r_j)$ and $h^v(r_j)$ are now random variables. To simplify notation let us denote $\boldsymbol{h^u} = h^u(r_1), \cdots, h^v(r_M)$ and $\boldsymbol{h^v} = h^v(r_1), \cdots, h^u(r_M)$. Assuming that these $2M$ random variables are distributed according to a probability density $P(\boldsymbol{h^u}, \boldsymbol{h^v})$, the mean values $\overline{h^u(r_j)}$, $\overline{h^v(r_j)}$ involved in (14) and (15) are calculated as

$$\overline{h^u(r_j)} = \int_{R^{2M}} P(\boldsymbol{h^u}, \boldsymbol{h^v}) h^u(r_j) \, d\boldsymbol{h^u} \, d\boldsymbol{h^v}, \qquad j = 1, \cdots, M \tag{16}$$

$$\overline{h^v(r_j)} = \int_{R^{2M}} P(\boldsymbol{h^u}, \boldsymbol{h^v}) h^v(r_j) \, d\boldsymbol{h^u} \, d\boldsymbol{h^v}, \qquad j = 1, \cdots, M \tag{17}$$

where

$$d\boldsymbol{h^u} = dh^u(r_1), \cdots, dh^u(r_M)$$

and

$$d\boldsymbol{h^v} = dh^v(r_1), \cdots, dh^v(r_M).$$

Since $P(\boldsymbol{h^u}, \boldsymbol{h^v})$ is a probability density we must require it satisfies the constraint

$$\int_{R^{2M}} P(\boldsymbol{h^u}, \boldsymbol{h^v}) \, d\boldsymbol{h^u} \, d\boldsymbol{h^v} = 1. \tag{18}$$

In addition, we should set a constraint to ensure $\overline{h(\omega, t)} \in L^2(R^2)$. This is guaranteed under the requirement that $\|h\|^2$ be finite, which

also ensures that the variance of the probability density is finite. Consequently, we will set the additional constraint

$$\overline{\|h\|^2} = \lim_{M \to \infty} \frac{1}{M} \sum_{j=1}^{M} \int_{R^{2M}} P(\boldsymbol{h^u}, \boldsymbol{h^v})(h^u(r_j)^2 + h^v(r_j)^2) \, d\boldsymbol{h^u} \, d\boldsymbol{h^v}$$
$$= C \qquad (19)$$

where C is an *unknown* constant.

Constraints (14), (15), (18), and (19), to be satisfied by the probability density we are looking for, are not enough to determine it in a unique way. Among all the $P(\boldsymbol{h^u}, \boldsymbol{h^v})$ capable of fulfilling these constraints, we shall select *one* adopting the MEP. This criterion yields the probability density that, being consistent with the available data, is maximally noncommittal with respect to the lack of information (entropy) [4], [5].

The entropy, or uncertainty, associated with the probability density is given by the generalization of Shannon's measure [8] to continuous-type random variables [7], i.e.,

$$H(\boldsymbol{h^u}, \boldsymbol{h^v}) = -\int_{R^{2M}} P(\boldsymbol{h^u}, \boldsymbol{h^v}) \ln P(\boldsymbol{h^u}, \boldsymbol{h^v}) \, d\boldsymbol{h^u} \, d\boldsymbol{h^v}. \qquad (20)$$

Since we should take $\lim M \to \infty$ to represent our stochastic process, the appropriate measure to be used is the entropy rate $\bar{H}$, or entropy per degree of freedom, defined as [7]

$$\bar{H} = \lim_{M \to \infty} \frac{1}{2M} H(\boldsymbol{h^u}, \boldsymbol{h^v}). \qquad (21)$$

We then look for the probability density that maximizes $\bar{H}$ with constraints (14), (15), (18), and (19). In order to introduce the constraints (14) and (15) into the variational process, we divide the axis R into intervals of length $\Delta x = \frac{1}{N}$ centered at the points x_i and take $\lim_{N \to \infty}$, at the end. Assuming that $f^u(x)$ and $f^v(x)$ are continuous functions we incorporate each constraint (14), evaluated at $x = x_i$, through a Lagrange multiplier that we write $\lambda^u_{x_i} \Delta x$ and each constraint (15) through a Lagrange multiplier $\lambda^v_{x_i} \Delta x$. Constraints (18) and (19) are introduced through the Lagrange multipliers λ_0 and β, respectively. Thus the functional, S, to be maximized is cast

$$S = -\frac{1}{2M} \int_{R^{2M}} P(\boldsymbol{h^u}, \boldsymbol{h^v})$$
$$\cdot \left(\ln P(\boldsymbol{h^u}, \boldsymbol{h^v}) + 2\beta \sum_{j=1}^{M} (h^u(r_j)^2 + h^v(r_j)^2) \right) d\boldsymbol{h^u} \, d\boldsymbol{h^v}$$
$$- \lambda_0 \int_{R^{2M}} P(\boldsymbol{h^u}, \boldsymbol{h^v}) \, d\boldsymbol{h^u} \, d\boldsymbol{h^v}$$
$$- \frac{1}{N} \sum_{i=1}^{N} \lambda^u_{x_i} \frac{1}{MC_g} \sum_{j=1}^{M} \left(g^u_{r_j}(x_i) \overline{h^u(r_j)} - g^v_{r_j}(x_i) \overline{h^v(r_j)} \right)$$
$$- \frac{1}{N} \sum_{i=1}^{N} \lambda^v_{x_i} \frac{1}{MC_g} \sum_{j=1}^{M} \left(g^v_{r_j}(x_i) \overline{h^u(r_j)} + g^u_{r_j}(x_i) \overline{h^v(r_j)} \right) \qquad (22)$$

$\overline{h^u(r_j)}$ and $\overline{h^v(r_j)}$ are calculated as in (16) and (17).

From the condition $\frac{\delta S}{\delta P} = 0$ we obtain

$$P(\boldsymbol{h^u}, \boldsymbol{h^v}) = \exp{-(2M\lambda_0 + 1)}$$
$$\cdot \exp\left(-2 \sum_{j=1}^{M} (h^u(r_j)\gamma_1(r_j) + h^v(r_j)\gamma_2(r_j) \right.$$
$$\left. + \beta h^u(r_j)^2 + \beta h^v(r_j)^2 \right) \qquad (23)$$

where

$$\gamma_1(r_j) = \frac{1}{NC_g} \sum_{i=1}^{N} \left(\lambda^u_{x_i} g^u_{r_j}(x_i) + \lambda^v_{x_i} g^v_{r_j}(x_i) \right) \qquad (24)$$

and

$$\gamma_2(r_j) = \frac{1}{NC_g} \sum_{i=1}^{N} \left(\lambda^v_{x_i} g^u_{r_j}(x_i) - \lambda^u_{x_i} g^v_{r_j}(x_i) \right). \qquad (25)$$

Since the entropy (20) is a convex functional [7] it takes on its absolute maximum at $P(\boldsymbol{h^u}, \boldsymbol{h^v})$ given in (23).

The normalization constraint (18) entails

$$\exp(2M\lambda_0 + 1) = \int_{R^{2M}} \exp\left(-2 \sum_{j=1}^{M} (h^u(r_j)\gamma_1(r_j) + h^u(r_j)\gamma_2(r_j) \right.$$
$$\left. + \beta h^u(r_j)^2 + \beta h^v(r_j)^2 \right) d\boldsymbol{h^u} \, d\boldsymbol{h^v}$$
$$= \left(\frac{\pi}{2\beta} \right)^M \prod_{j=1}^{M} \exp\left(\frac{\gamma_1(r_j)^2}{2\beta} \right) \exp\left(\frac{\gamma_2(r_j)^2}{2\beta} \right). \qquad (26)$$

The remaining Lagrange multipliers, $\lambda^u_{x_i}, \lambda^v_{x_i}; i = 1, \cdots, N$, and β, should be obtained by using (23) in (14), (15), and (19) and solving the equations. However, as we shall see below, the functional form of $P(\boldsymbol{h^u}, \boldsymbol{h^v})$, given in (23), already provides the information that is needed to determine the mean value function $\overline{h(r_j)}$ that such probability density will predict. Indeed, by replacing (23) in (16) and (17) and performing the integrals we have

$$\overline{h^u(r_j)} = -\frac{\gamma_1(r_j)}{2\beta} = -\frac{1}{2\beta NC_g} \sum_{i=1}^{N} \left(\lambda^u_{x_i} g^u_{r_j}(x_i) + \lambda^v_{x_i} g^v_{r_j}(x_i) \right) \qquad (27)$$

$$\overline{h^v(r_j)} = -\frac{\gamma_2(r_j)}{2\beta} = -\frac{1}{2\beta NC_g} \sum_{i=1}^{N} \left(\lambda^v_{x_i} g^u_{r_j}(x_i) - \lambda^u_{x_i} g^v_{r_j}(x_i) \right). \qquad (28)$$

Taking now $\lim N \to \infty$, the above equations yield

$$\overline{h^u(r_j)} = -\frac{1}{2\beta C_g} \int_R \left(\lambda^u_x g^u_{r_j}(x) + \lambda^v_x g^v_{r_j}(x) \right) dx \qquad (29)$$

$$\overline{h^v(r_j)} = -\frac{1}{2\beta C_g} \int_R \left(\lambda^v_x g^u_{r_j}(x) - \lambda^u_x g^v_{r_j}(x) \right) dx \qquad (30)$$

or

$$\overline{h(r_j)} = \overline{h(\omega_j, t_j)} = \overline{h^u(\omega_j, t_j)} + i\overline{h^v(\omega_j, t_j)}$$
$$= \int_R e^{-i\omega_j x} g^*(x - t_j) w(x) \, dx \qquad (31)$$

with

$$w(x) = -\frac{1}{2\beta C_g} (\lambda^u_x + i\lambda^v_x).$$

From (31) and definition (3) we gather that

$$\overline{h(\omega_j, t_j)} \in \mathcal{W}, \qquad (\omega_j, t_j) \in R^2.$$

So that, by Theorem 1, we are in a position to reveal $\overline{h(\omega, t,)}$. In fact, by using $\overline{h(\omega, t)}$ in (11) and performing the inner product of both sides with $e^{i\omega' x} g(x - t')$ we have

$$\langle e^{i\omega' x} g(x - t') \mid f(x) \rangle$$
$$= F(\omega', t')$$
$$= \frac{1}{C_g} \int_{R^2} \langle e^{i\omega' x} g(x - t') \mid e^{i\omega x} g(x - t) \rangle \overline{h(\omega, t)} \, d\omega \, dt \qquad (32)$$

and, since $\overline{h(\omega, t)} \in \mathcal{W}$, by Theorem 1 the *consistency condition* (4) is verified. Hence, from (32) and Theorem 1 we conclude that

$\overline{h(\omega', t')} = F(\omega', t')$, which states the WFT as an optimal solution of the inverse problem according to an ME selection criterion. Since such a solution is also optimal in an MN sense, we are led to conclude that the MN requirement works by averaging functions in a maximally noncommittal way. In other words, we give here a new argument supporting the use of the MN solution for the inverse windowed Fourier transform problem, as it has been shown to be the "least biased" assignment one can make on the basis of the available information.

REFERENCES

[1] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.

[2] D. Gabor, "Theory of communications," *J. Inst. Elec. Eng.*, vol. 93, pp. 429–457, 1946.

[3] C. Heil and D. Walnut, "Continuous and discrete wavelet transforms," *SIAM Rev.*, vol. 31, pp. 628–666, 1989.

[4] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.*, vol. 106, pp. 620–630, 1957.

[5] ——, "Where do we stand maximum entropy?," in *The Maximum Entropy Formalism*, R. Levine and M. Tribus, Eds. Boston, MA: MIT Press, 1979.

[6] G. Kaiser, *A Friendly Guide to Wavelets*. Berlin, Germany: Birkhäuser, 1994.

[7] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1991.

[8] C. E. Shannon, "The mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.

[9] L. Rebollo-Neira, J. F. Rubio, and A. Plastino, "Frames: A maximum entropy statistical estimate of the inverse problem," *J. Math. Phys.*, vol. 38, no. 9, pp. 4863–4871, 1997.