

Frames: a Maximum Entropy Statistical Estimate of the Inverse Problem

L. Rebollo-Neira and J. Fernandez-Rubio

Departament de Teoria del Senyal i Comunicacions, Escola Tecnica Superior d'Enginyers de Telecomunicació, Campus Nord, UPC, Edifici D-4, c/. Gran Capita s/n. 08034, Barcelona, Spain

A. Plastino

Departamento de Física, Universidad Nacional de La Plata C.C. 727, 1900 La Plata, Argentina.

Abstract

A Maximum Entropy statistical treatment of an inverse problem concerning *frame theory* is presented. The problem arises from the fact that a *frame* is an overcomplete set of vectors that defines a mapping with *no unique* inverse. Although any vector in the concomitant space can be expressed as linear combination of frame elements, the coefficients of the expansion are *not unique*. *Frame theory* guarantees the existence of a set of coefficients which is “optimal” in a Minimum Norm sense. We show here that these coefficients are also “optimal” from a Maximum Entropy viewpoint.

1 Introduction

Frames were introduced by Duffin and Schaeffer within the context of non-harmonic Fourier series [1], where most of the theory was developed (a complete review is given in [2]). The interest in *frame theory* has received great impetus since that mathematical structure was adopted to study *coherent states*, among which one may cite *Weyl-Heisenberg coherent states* [3, 4, 5, 6], that are the result of translations and modulations of a single function, and *affine coherent states*, called *wavelets*, that arise as translations and dilations of a single function [4, 5, 6, 7, 8]. Typically, a *frame* is an over-complete set of vectors that, in spite of not being linearly independent, can nonetheless be used to express any vector as a linear combination of them.

The *frame condition* ensures that the inverse mapping does exist and that an appropriate set of coefficients can be obtained by means of the *reciprocal frame*. However, due to the lack of linear independence of the frame elements such a set of coefficients is *not unique*. The lack of uniqueness poses a problem that has to be surmounted if one expects the coefficients to be endowed with some relevant physical information. Now, if one wishes to recognize a particular set of coefficients as “optimal”, an appropriate decision criterion has to be adopted. It is well known that the *reciprocal frame* provides a set of coefficients which is “optimal” in a Minimum Norm (MN) sense [2, 4]. 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 paper we tackle the inverse problem from a statistical point of view and show that the *reciprocal frame* provides one with a set of coefficients that is also “optimal” in a Maximum Entropy (ME) sense.

The early *frame theory* was devised with the discrete case in mind, but an interesting generalization, recently proposed [8, 9, 10, 11], allows for the inclusion of *continuous* cases as part of the same general structure. This generalization includes continuous transforms, such as the Windowed Fourier Transform (WFT) or the Continuous Wavelet Transform (CWT), as special instances of a more general framework. Here we adopt the generalized structure and develop our statistical description of the inverse problem within the *generalized frame* definition.

We shall i) regard each admissible solution of the inverse problem as a stochastic process (random function) distributed according to a suitable probability density (to be determined) and ii) estimate the desired solution as the mean value of such a random function. Then, among all the probability densities capable of yielding admissible solutions we shall single out *one*, adopting the Maximum Entropy Postulate (MEP). Finally, we will show that, from the ME probability density, a mean value function is inferred that is provided by the *reciprocal frame*, being therefore identical with the MN solution.

The paper is organized as follows: In Section 2 the *generalized frame* definition is given and some properties to be used are summarized. In Section 3 the proposed Maximum Entropy statistical treatment of the inverse problem is developed. The WFT and the CWT are given as examples of the general formalism. Some conclusions are drawn in Section 4.

2 Background on Frames

Let \mathcal{H} be the Hilbert space of possible functions (on the real line R) to be analyzed and M a set of labels $M = \{m \in M\}$. Adopting Dirac's notation [12], we represent a vector $f \in \mathcal{H}$ as $|f\rangle$ and its dual as $\langle f|$. The identity operator in \mathcal{H} is then expressed in the fashion

$$\hat{I}_{\mathcal{H}} = \int_R |t\rangle\langle t| dt. \quad (1)$$

Let μ be a measure on M and let us denote as $L^2(\mu)$ the Hilbert space in which the identity operator reads

$$\hat{I}_{L^2(\mu)} = \int_M |m\rangle\langle m| d\mu(m). \quad (2)$$

For all $|f_1\rangle$ and $|f_2\rangle \in \mathcal{H}$ the functional representation of \mathcal{H} can be introduced by inserting (1) in $\langle f_1|f_2\rangle$, i.e.,

$$\langle f_1|f_2\rangle = \int_R \langle f_1|t\rangle\langle t|f_2\rangle dt \quad (3)$$

with $f_2(t) = \langle t|f_2\rangle$ and $f_1^*(t) = \langle f_1|t\rangle = \langle t|f_1\rangle^*$, where $f^*(t)$ indicates the complex conjugate of $f(t)$. In the same way, for $|g_1\rangle$ and $|g_2\rangle \in L^2(\mu)$, the functional representation of $L^2(\mu)$ is given by operator (2)

$$\langle g_1|g_2\rangle = \int_M \langle g_1|m\rangle\langle m|g_2\rangle d\mu(m). \quad (4)$$

Now we are in a position to give the definition of *generalized frame* [10]:

Definition: A family of vectors $|h_m\rangle \in \mathcal{H}$; $m \in M$ is called a *generalized frame* (henceforth to be referred as simply a *frame*) if, for every $|f\rangle \in \mathcal{H}$,

a) The function $\tilde{f}(m) = \langle m|\tilde{f}\rangle = \langle h_m|f\rangle$ is measurable.

b) There exists a pair of constants $0 < A \leq B < \infty$ such that

$$A\langle f|f\rangle_{\mathcal{H}} \leq \langle \tilde{f}|\tilde{f}\rangle_{L^2(\mu)} \leq B\langle f|f\rangle_{\mathcal{H}}. \quad (5)$$

The constants A and B are called the *frame bounds* and (5) the *frame condition*. The latter implies that $|\tilde{f}\rangle \in L^2(\mu)$ whenever $|f\rangle \in \mathcal{H}$, then the mapping $\hat{T} : \mathcal{H} \mapsto L^2(\mu)$ defines an operator, called the *frame operator*,

$$\hat{T} = \int_M |m\rangle\langle h_m| d\mu(m). \quad (6)$$

The adjoint operator $\hat{T}^\dagger : L^2(\mu) \mapsto \mathcal{H}$ is

$$\hat{T}^\dagger = \int_M |h_m\rangle\langle m| d\mu(m). \quad (7)$$

The *frame condition* can be expressed in terms of the operator $\hat{G} = \hat{T}^\dagger\hat{T} : \mathcal{H} \mapsto \mathcal{H}$ as

$$A\hat{I}_{\mathcal{H}} \leq \hat{G} \leq B\hat{I}_{\mathcal{H}}. \quad (8)$$

From (6) and (7) \hat{G} is explicitly given by

$$\hat{G} = \int_M |h_m\rangle\langle h_m| d\mu(m). \quad (9)$$

The inequality (8) entails that \hat{G} has a bounded inverse \hat{G}^{-1} . In fact, \hat{G}^{-1} satisfies [2, 4, 10]

$$B^{-1}\hat{I}_{\mathcal{H}} \leq \hat{G}^{-1} \leq A^{-1}\hat{I}_{\mathcal{H}}. \quad (10)$$

Assuming that \hat{G}^{-1} is known explicitly, the *reciprocal frame* $\{|h^m\rangle ; m \in M\}$, is calculated as $|h^m\rangle = \hat{G}^{-1}|h_m\rangle ; m \in M$. Thus, from (9) we obtain the following resolution of the unity operator in \mathcal{H}

$$\hat{I} = \int_M |h^m\rangle\langle h_m| d\mu(m) = \int_M |h_m\rangle\langle h^m| d\mu(m). \quad (11)$$

The family $\{|h^m\rangle; m \in M\}$ turns out to be a *frame* as well, with *frame bounds* B^{-1} and A^{-1} [2, 4, 10]. The associated *frame operator* $\hat{S} : \mathcal{H} \mapsto L^2(\mu)$ is here

$$\hat{S} = \int_M |m\rangle\langle h^m| d\mu(m) = \hat{T}\hat{G}^{-1} \quad (12)$$

and its adjoint $\hat{S}^\dagger : L^2(\mu) \mapsto \mathcal{H}$

$$\hat{S}^\dagger = \int_M |h^m\rangle\langle m| d\mu(m) = \hat{G}^{-1}\hat{T}^\dagger. \quad (13)$$

The *reciprocal frame* of $\{|h^m\rangle; m \in M\}$ happens to be, again, the original *frame* [2, 4, 10].

When the *frame bounds* are equal the *frame* is called a *tight* one and the *reciprocal frame* satisfies $|h^m\rangle = \frac{|h_m\rangle}{A}$. For the case $A = 1$ the *frame* is *self-reciprocal*.

Let \mathcal{F} be the range of the operator $\hat{T} : \mathcal{H} \mapsto L^2(\mu)$, i.e., the subspace

$$\mathcal{F} = \text{Ran}(\hat{T}) = \{|\tilde{f}\rangle; \hat{T}|f\rangle = |\tilde{f}\rangle; |f\rangle \in \mathcal{H}\}. \quad (14)$$

One also has $\text{Ran}(\hat{S}) = \text{Ran}(\hat{T})$. The operator $\hat{S}^\dagger = \hat{G}^{-1}\hat{T}^\dagger : L^2(\mu) \mapsto \mathcal{H}$ provides the reconstruction of $|f\rangle \in \mathcal{H}$ from $|\tilde{f}\rangle \in \mathcal{F}$ as $|f\rangle = \hat{S}^\dagger|\tilde{f}\rangle$. In fact, for $|\tilde{f}\rangle \in \mathcal{F}$, $|\tilde{f}\rangle = \hat{T}|f\rangle$ and we have

$$\hat{S}^\dagger|\tilde{f}\rangle = \int_M |h^m\rangle\langle m|\hat{T}|f\rangle d\mu(m) = \int_M |h^m\rangle\langle h_m|f\rangle d\mu(m) = \int_M |h_m\rangle\langle h^m|f\rangle d\mu(m) = |f\rangle. \quad (15)$$

Notice that \mathcal{F} is just a closed *subspace*, not *all* of $L^2(\mu)$ (not every $|g\rangle \in L^2(\mu)$ can be expressed as $|g\rangle = \hat{T}|f\rangle$). The orthogonal projection operator \hat{P} from $L^2(\mu)$ onto \mathcal{F} is $\hat{P} = \hat{T}\hat{S}^\dagger = \hat{S}\hat{T}^\dagger$ [4, 10], which, explicitly, adopts the appearance

$$\hat{P} = \int_M |m'\rangle\langle h_{m'}|h^m\rangle\langle m| d\mu(m)d\mu(m') = \int_M |m'\rangle\langle h^{m'}|h_m\rangle\langle m| d\mu(m)d\mu(m'). \quad (16)$$

Since $|\tilde{g}\rangle \in \mathcal{F}$ if and only if $|\tilde{g}\rangle = \hat{P}|\tilde{g}\rangle$, it follows that, $|\tilde{g}\rangle \in \mathcal{F}$ *if and only if*:

$$|\tilde{g}\rangle = \int_M |m'\rangle\langle h_{m'}|h^m\rangle\langle m|\tilde{g}\rangle d\mu(m)d\mu(m') = \int_M |m'\rangle\langle h^{m'}|h_m\rangle\langle m|\tilde{g}\rangle d\mu(m)d\mu(m'). \quad (17)$$

or

$$\tilde{g}(m') = \langle m' | \tilde{g} \rangle = \int_M \langle h_{m'} | h^m \rangle \langle m | \tilde{g} \rangle d\mu(m) = \int_M \langle h^{m'} | h_m \rangle \langle m | \tilde{g} \rangle d\mu(m). \quad (18)$$

The above *reproducing kernel* equation provides the necessary and sufficient condition for a vector $|\tilde{g}\rangle \in L^2(\mu)$ to belong to \mathcal{F} .

Although $|f\rangle \in \mathcal{H}$ can be reconstructed as in (15), the corresponding expansion is not unique. Indeed, all $|g\rangle \in L^2(\mu)$ can be written as $|g\rangle = |\tilde{g}\rangle + |\tilde{g}\rangle^\perp$, with $|\tilde{g}\rangle \in \mathcal{F}$ and $|\tilde{g}\rangle^\perp \in \mathcal{F}^\perp$, the orthogonal complement of \mathcal{F} in $L^2(\mu)$, and since $\hat{S}^\dagger |\tilde{g}\rangle^\perp = 0$ we have : $\hat{S}^\dagger |g\rangle = \hat{S}^\dagger (|\tilde{g}\rangle + |\tilde{g}\rangle^\perp) = \hat{S}^\dagger |\tilde{g}\rangle = |f\rangle$. This implies that the inversion problem for determining $\langle m | g \rangle$ from the equation

$$|f\rangle = \int_M |h^m\rangle \langle m | g \rangle d\mu(m) \quad (19)$$

has *no unique* solution. The most general solution is of the form $\langle m | g \rangle = \langle h_m | f \rangle + \langle m | \tilde{g} \rangle^\perp$, with $|\tilde{g}\rangle^\perp \in \mathcal{F}^\perp$. Therefore, by setting $|\tilde{g}\rangle^\perp = 0$ one obtains the MN solution. Equivalently, changing $|h^m\rangle$ to $|h_m\rangle$ in (19), the MN solution for the corresponding inverse problem is $\langle m | g \rangle = \langle h^m | f \rangle$.

In what follows we show that $\langle m | g \rangle = \langle h^m | f \rangle$ is also an “optimal” solution in a ME sense.

3 ME statistical estimate of the inverse problem

The problem we address now is that of inverting the equation

$$\langle t | f \rangle = \int_M \langle t | h_m \rangle \langle m | g \rangle d\mu(m). \quad (20)$$

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

$$f^u(t) = \int_M (h_m^u(t) g^u(m) - h_m^v(t) g^v(m)) d\mu(m) \quad (21)$$

$$f^v(t) = \int_M (h_m^v(t) g^u(m) + h_m^u(t) g^v(m)) d\mu(m), \quad (22)$$

where $f^u(t), f^v(t)$ are the real and imaginary parts of $\langle t | f \rangle$ while $h_m^u(t), h_m^v(t)$ are the real and imaginary parts of $\langle t | h_m \rangle$ and $g^u(m), g^v(m)$ are the real and imaginary parts of $\langle m | g \rangle$. As

discussed in the previous section, there exist several functions $\langle m|g \rangle$ capable of satisfying (21) and (22). Our aim is that of selecting ONE of those solutions as “optimal” in a ME sense.

The inversion problem is then transformed into one of *statistical inference*. The essential step in this respect is to regard each admissible solution $\langle m|g \rangle$ as a random function, distributed according to a (to be determined) probability density . This probability density represents our *ignorance* vis-a-vis the fact that there is not a unique solution. Within this statistical framework, we estimate the desired solution as the mean value of the random function $\langle m|g \rangle$ and denote it as $\overline{\langle m|g \rangle} = \overline{g^u(m)} + i\overline{g^v(m)}$. Let $\{A_j\}$ be the measurable set that allows one to calculate (21) and (22) as:

$$f^u(t) = \lim_{K \rightarrow \infty} \sum_{j=1}^K (h_{m_j}^u(t) \overline{g^u(m_j)} - h_{m_j}^v(t) \overline{g^v(m_j)}) \mu(A_j) \quad (23)$$

$$f^v(t) = \lim_{K \rightarrow \infty} \sum_{j=1}^K (h_{m_j}^v(t) \overline{g^u(m_j)} + h_{m_j}^u(t) \overline{g^v(m_j)}) \mu(A_j). \quad (24)$$

At the fixed points m_j ; $j = 1, \dots, K$, both $g^u(m_j)$ and $g^v(m_j)$ are random variables. To simplify the notation let us introduce $\mathbf{g}^u = g^u(m_1), \dots, g^u(m_K)$ and $\mathbf{g}^v = g^v(m_1), \dots, g^v(m_K)$.

Assuming that these $2K$ random variables are distributed according to a probability density $P(\mathbf{g}^u, \mathbf{g}^v)$, the mean values $\overline{g^u(m_j)}, \overline{g^v(m_j)}$ involved in (23) and (24) are calculated as:

$$\overline{g^u(m_j)} = \int_{-\infty}^{\infty} P(\mathbf{g}^u, \mathbf{g}^v) g^u(m_j) d\mathbf{g}^u d\mathbf{g}^v \quad ; \quad j = 1, \dots, K \quad (25)$$

$$\overline{g^v(m_j)} = \int_{-\infty}^{\infty} P(\mathbf{g}^u, \mathbf{g}^v) g^v(m_j) d\mathbf{g}^u d\mathbf{g}^v \quad ; \quad j = 1, \dots, K, \quad (26)$$

where $d\mathbf{g}^u = dg^u(m_1), \dots, dg^u(m_K)$ and $d\mathbf{g}^v = dg^v(m_1), \dots, dg^v(m_K)$. We shall make sure

$\overline{\langle m|g \rangle} \in L^2(\mu)$ through the more stringent requirement that $\overline{\|g\|^2}$ be finite. This also ensures the finitude of the variance of the probability density. Consequently, we have to deal with the constraint

$$\overline{\|g\|^2} = \int_{-\infty}^{\infty} \lim_{K \rightarrow \infty} \sum_{j=1}^K \int_{-\infty}^{\infty} P(\mathbf{g}^u, \mathbf{g}^v) (g^u(m_j)^2 + g^v(m_j)^2) \mu(A_j) d\mathbf{g}^u d\mathbf{g}^v = C \quad (27)$$

where C is an *unknown* constant. As $P(\mathbf{g}^u, \mathbf{g}^v)$ is normalized to unity, it satisfies

$$\int_{-\infty}^{\infty} P(\mathbf{g}^u, \mathbf{g}^v) d\mathbf{g}^u d\mathbf{g}^v = 1. \quad (28)$$

We face now the problem of determining $P(\mathbf{g}^u, \mathbf{g}^v)$ satisfying (23), (24), (27) and (28). Among all the probability densities capable of fulfilling these constraints, we select *one* adopting the MEP. This criterion yields the probability density that, being consistent with the available data, is maximally noncommittal with respect to our lack of information [13, 14].

The entropy, or uncertainty, associated with $P(\mathbf{g}^u, \mathbf{g}^v)$ is given by Shannon's measure

$$H(\mathbf{g}^u, \mathbf{g}^v) = - \int_{-\infty}^{\infty} P(\mathbf{g}^u, \mathbf{g}^v) \ln P(\mathbf{g}^u, \mathbf{g}^v) d\mathbf{g}^u d\mathbf{g}^v. \quad (29)$$

In order to proceed take $\lim K \rightarrow \infty$, which entails that, here, the appropriate measure is the entropy rate \overline{H} (entropy per degree of freedom), defined as [15]

$$\overline{H} = \lim_{K \rightarrow \infty} \frac{1}{2K} H(\mathbf{g}^u, \mathbf{g}^v). \quad (30)$$

We look then for the probability density that maximizes \overline{H} with constraints (23), (24), (27), (28). In order to introduce the constraints (23) and (24) into the variational process, we divide the axis R into intervals of length $\Delta t = \frac{1}{N}$ centered at the points t_i and take $\lim N \rightarrow \infty$ at the end of the calculation. We incorporate each constraint (23) evaluated at $t = t_i$ via a Lagrange multiplier that we write $\lambda_{t_i}^u \Delta t$ and each constraint (24) through a Lagrange multiplier $\lambda_{t_i}^v \Delta t$. Constraints (27) and (28) are introduced through the Lagrange multipliers β and λ_0 respectively. Thus the functional, S , to be maximized adopts the appearance

$$\begin{aligned} S &= - \int_{-\infty}^{\infty} P(\mathbf{g}^u, \mathbf{g}^v) \left(\frac{\ln P(\mathbf{g}^u, \mathbf{g}^v)}{2K} + \beta \sum_{j=1}^K (g^u(m_j)^2 + g^v(m_j)^2) \mu(A_j) \right) d\mathbf{g}^u d\mathbf{g}^v \\ &- \lambda_0 \int_{-\infty}^{\infty} P(\mathbf{g}^u, \mathbf{g}^v) d\mathbf{g}^u d\mathbf{g}^v - \frac{1}{N} \sum_{i=1}^N \lambda_{t_i}^u \sum_{j=1}^K (h_{m_j}^u(t_i) \overline{g^u(m_j)} - h_{m_j}^v(t_i) \overline{g^v(m_j)}) \mu(A_j) \\ &- \frac{1}{N} \sum_{i=1}^N \lambda_{t_i}^v \sum_{j=1}^K (h_{m_j}^v(t_i) \overline{g^u(m_j)} + h_{m_j}^u(t_i) \overline{g^v(m_j)}) \mu(A_j), \end{aligned} \quad (31)$$

and, from the condition $\frac{\delta S}{\delta P} = 0$, we obtain

$$P(\mathbf{g}^u, \mathbf{g}^v) = \exp -(2K \lambda_0 + 1) \exp(-2K \sum_{j=1}^K (g^u(m_j) \gamma_1(m_j) + g^v(m_j) \gamma_2(m_j) + \beta g^u(m_j)^2 + \beta g^v(m_j)^2) \mu(A_j)) \quad (32)$$

where

$$\gamma_1(m_j) = \frac{1}{N} \sum_{i=1}^N \lambda_{t_i}^u h_{m_j}^u + \lambda_{t_i}^v h_{m_j}^v \quad (33)$$

$$\gamma_2(m_j) = \frac{1}{N} \sum_{i=1}^N \lambda_{t_i}^v h_{m_j}^u - \lambda_{t_i}^u h_{m_j}^v, \quad (34)$$

while the normalization constraint (28) entails

$$\exp(2K \lambda_0 + 1) = \int_{-\infty}^{\infty} \exp(-2K \sum_{j=1}^K (g^u(m_j) \gamma_1(m_j) + g^v(m_j) \gamma_2(m_j) + \beta g^u(m_j)^2 + \beta g^v(m_j)^2) \mu(A_j)) d\mathbf{g}^u d\mathbf{g}^v. \quad (35)$$

Obviously, we are led to

$$\exp(2K \lambda_0 + 1) = \prod_{j=1}^K \frac{\pi}{2K \beta \mu(A_j)} \exp\left(\frac{K \gamma_1(m_j)^2 \mu(A_j)}{2\beta}\right) \exp\left(\frac{K \gamma_2(m_j)^2 \mu(A_j)}{2\beta}\right), \quad (36)$$

so that, by replacing (32) into (25) and (26) and calculating the pertinent integrals we have

$$\overline{g^u(m_j)} = -\frac{\gamma_1(m_j)}{2\beta} = -\frac{1}{2\beta N} \sum_{i=1}^N \lambda_{t_i}^u h_{m_j}^u(t_i) + \lambda_{t_i}^v h_{m_j}^v(t_i) \quad (37)$$

$$\overline{g^v(m_j)} = -\frac{\gamma_2(m_j)}{2\beta} = -\frac{1}{2\beta N} \sum_{i=1}^N \lambda_{t_i}^v h_{m_j}^u(t_i) - \lambda_{t_i}^u h_{m_j}^v(t_i). \quad (38)$$

Passing now to the limit $\lim N \rightarrow \infty$, the above equations yield

$$\overline{g^u(m_j)} = -\frac{1}{2\beta} \int_R (\lambda_t^u h_{m_j}^u(t) + \lambda_t^v h_{m_j}^v(t)) dt, \quad (39)$$

$$\overline{g^v(m_j)} = -\frac{1}{2\beta} \int_R (\lambda_t^v h_{m_j}^u(t) - \lambda_t^u h_{m_j}^v(t)) dt, \quad (40)$$

or

$$\overline{\langle m_j | g \rangle} = \overline{g^u(m_j)} + i \overline{g^v(m_j)} = \int_R \xi(t) h_{m_j}^*(t) dt = \langle h_{m_j} | \xi \rangle, \quad (41)$$

with $\xi(t) = -\frac{1}{2\beta}(\lambda_t^u + i\lambda_t^v)$.

From equation (41) we gather that $\overline{\langle m_j | g \rangle} \in \mathcal{F}$. Using $\overline{\langle m | g \rangle}$ in (20) and performing the inner product operation on both sides with $\langle h^{m'} |$ we have

$$\langle h^{m'} | f \rangle = \int_M \langle h^{m'} | h_m \rangle \overline{\langle m | g \rangle} d\mu(m), \quad (42)$$

and, since $\overline{\langle m | g \rangle} \in \mathcal{F}$, the *reproducing kernel* equation (18) is verified. Hence, $\overline{\langle m | g \rangle} = \langle h^m | f \rangle$ and we conclude that the statistical estimate $\overline{\langle m | g \rangle} = \langle h^m | f \rangle$ is an “optimal” solution in a ME sense.

3.1 Some special cases: The WFT, the CWT and discrete frames

The frame formulation proposed in [8, 9, 10, 11], and adopted here in order to develop the present statistical treatment of the inverse problem, allows one to derive the WFT and CWT as special cases of the same structure. In addition, the classical discrete frame formulation [1, 2, 4] also appears as a particular case of the generalized theory.

For the WFT, $M = R^2$ is the set of all the continuous parameters $m = (\omega, b)$ and $d\mu(m) = d\omega db$.

In this case, the frame elements are $|h_m\rangle = \frac{|w_{\omega,b}\rangle}{\|w\|}$ with

$$\langle t | w_{\omega,b} \rangle = w(t - b) \exp(i\omega t), \quad (43)$$

with $w(t)$ any function in \mathcal{H} . The inverse problem for the WFT involves then the inversion of the equation

$$\langle t | f \rangle = \frac{1}{\|w\|} \int_{R^2} w(t - b) \exp(i\omega t) \langle \omega, b | g \rangle d\omega db. \quad (44)$$

As the *frame* is *self-reciprocal*: $|w_{\omega,b}\rangle = \frac{|w_{\omega,b}\rangle}{\|w\|}$ and the ME estimate for $\langle \omega, b | g \rangle$ gives the WFT of $|f\rangle$, i.e.,

$$\overline{\langle \omega, b | g \rangle} = \frac{1}{\|w\|} \langle w_{\omega,b} | f \rangle = \frac{1}{\|w\|} \int_R w(t - b)^* \exp(-i\omega t) f(t) dt. \quad (45)$$

For the CWT, $M = R^2$, $m = (a, b)$ and $d\mu(m) = \frac{dad b}{C_\psi a^2}$. The frame elements are the wavelets

$$\langle t | \psi_{a,b} \rangle = a^{-1/2} \psi_{a,b} \left(\frac{t-b}{a} \right), \quad (46)$$

with $\psi(t)$ a mother wavelet satisfying the admissibility condition

$$C_\psi = \int_R \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \quad (47)$$

where $\hat{\psi}(\omega)$ is the Fourier Transform of $\psi(t)$.

The equation to be inverted in this case is

$$\langle t | f \rangle = \frac{1}{C_\psi} \int_{R^2} a^{-1/2} \psi_{a,b} \left(\frac{t-b}{a} \right) \langle a, b | g \rangle \frac{dad b}{a^2}. \quad (48)$$

The *reciprocal frame* is also trivial, as $|\psi^{a,b}\rangle = |\psi_{a,b}\rangle$, and the ME estimate of the inverse problem is the CWT of $|f\rangle$, i.e.,

$$\overline{\langle a, b | g \rangle} = \langle \psi_{a,b} | f \rangle = \int_R a^{-1/2} \psi^* \left(\frac{t-b}{a} \right) f(t) dt. \quad (49)$$

When $M = Z^n$ and μ is the counting measure ($\mu(A) = \text{number of elements in } A$), $L^2(\mu) = l^2(Z^n)$ and the *generalized theory* reduces to the classical discrete one. The discrete version of both the WFT and CWT, for the sampling density required to give rise to a *frame* [4, 5, 6], involves *reciprocal frames* which are of no trivial character and have to be calculated by recourse to iterative algorithms [4, 16].

4 Conclusions

A statistical treatment of the *frame* inverse problem has been presented. The problem has been transformed into a problem of *statistical inference* by considering the set of admissible solutions as a random function and adopting the MEP as a decision criterion to select the probability density that, being consistent with the data, is less committal with respect to our

lack of information. The statistical treatment presented here leads one to conclude that the *reciprocal frame* gives rise to a solution that, in addition to being “optimal” in a MN sense is also “optimal” from a ME viewpoint.

As special cases, the WFT and CWT have been obtained from the concomitant inverse problems as “optimal conjectures”, derived according to MEP strictures.

Acknowledgments

L. Rebollo-Neira is a member of CICPBA (Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Argentina). This work has been supported by Ministerio de Educación y Ciencia, Spain (Program: “Estancias Temporales de Científicos y Tecnólogos Extranjeros en España” and Grant TIC 96-0500-C10-01). A. Plastino thanks the Argentine national Research Council for financial support.

References

- [1] R. J. Duffin, A. C. Shaffer, “A Class of Nonharmonic Fourier Series,” *Trans. Amer. Math. Soc.* **72**, 341-366 (1952).
- [2] R. M. Young, *An introduction to Nonharmonic Fourier Series* (Academic Press, New York, 1980).
- [3] J. R. Klauder, S. K. Skagerstam, *Coherent States* (World Scientific, Singapore, 1985).
- [4] I. Daubechies, *Ten Lectures on Wavelets* (CBMS-NSF, SIAM, Philadelphia, 1992).
- [5] I. Daubechies, “The Wavelets Transform, Time Frequency Localization and Signal Analysis,” *IEEE Trans. Inform. Theory* **36**, 961-1005 (1990).

- [6] C. Heil, D. Walnut, “Continuous and discrete wavelet transforms,” *SIAM Rev.* **31**, 628-666 (1989).
- [7] S. T. Ali, J. P. Antoine, J. P. Gazeau, “Square integrability of group representation on homogeneous spaces. II. Coherent and quasi-coherent states. The case of the Poincaré group,” *Ann. Inst. Henri Poincaré* **55**, 857-890 (1991).
- [8] S. T. Ali, J. P. Antoine, J. P. Gazeau, “Continuous Frames in Hilbert Space,” *Annals of Physics* **222**, 1-37 (1993).
- [9] G. Kaiser, *Quantum Physics, Relativity and Complex Spacetime: Towards a New Synthesis* (North-Holland, Amsterdam, 1990).
- [10] G. Kaiser, *A Friendly Guide to Wavelets* (Birkhäuser, Berlin, 1994).
- [11] S. T. Ali, J. P. Antoine, J. P. Gazeau, “Square integrability of group representation on homogeneous spaces. I. Reproducing triples and frames,” *Ann. Inst. Henri Poincaré* **55**, 829-855 (1991).
- [12] P. A. M. Dirac, *The principles of Quantum Mechanics* (Oxford University Press, 1958).
- [13] E. T. Jaynes, “Information Theory and Statistical Mechanics,” *Phys. Rev.* **106**, 620-630 (1957).
- [14] E. T. Jaynes, “Where do we Stand Maximum Entropy?,” in *The Maximum Entropy Formalism*, Ed. R. Levine and M. Tribus (MIT, Boston, 1979).
- [15] A. Papoulis, *Probability, Random Variables and Stochastic Processes* (McGraw-Hill, Inc., New York, 1991).

- [16] K. Grochening, "Acceleration of the Frame Algorithm," IEEE Transactions on Signal Processing **41,12**, 3331-3340 (1993).