DECOMPOSITION OF SIGNALS AND INFORMATION THEORY

A. Plastino*, L. Rebollo Neira**, F. Zyserman*

Centro de Física Teórica - Universidad Nacional de La Plata
C.C. 67, 1900 La Plata, Argentina
* Member of CONICET
** Member of CICPBA

ABSTRACT

We discuss a method that is able to determine, with the help of the Maximum Entropy Principle, the detailed nature of a complex signal that can be conceived as a linear superposition of independent, elementary signals. The physical relevance of this decomposition is discussed.

1. INTRODUCTION

Techniques for the decomposition of signals are of physical interest because they can be used to study systems that react, after being impinged upon by an external probe, by producing an output signal that is composed by a linear superposition of independent, "elementary" signals. The analysis of the nature of this superposition allows one to gain revealing insights concerning the structure of the system.

Signals are preconcerted signs that convey information. We envisage a situation in which a well-known probe, (for example, electromagnetic radiation) impinges upon a physical system, interacts with it and is afterward analyzed by an appropriate detection procedure.

Condensed Matter Theories, Vol. 7, Edited by A.N. Proto
and J.L. Aliaga, Plenum Press, New York, 1992

Shannon's[1] vectorial representation of signals independizes the pertinent considerations from the specific details characterizing the detection procedure. In a previous work[2], we have found it very convenient to proceed as follows: to any signal f a vector ket $|f\rangle$ is attributed and measurements performed upon f are described by linear functionals $\mathcal{L}_i$ that map $|f\rangle$ upon the set of the real numbers.

The process of decomposition of signals is thus applied with the idea of studying systems that react with the input probe producing a response signal that is a superposition of independent signals. "Statistical weights" appear as a coefficient in this superposition and they contain information about the (statistical) nature of the physical system. Illustrations of these ideas are given in Ref.2. The Maximum Entropy Principle (MEP)[3] is there employed and shown to provide one with a powerful algorithm that allows for successfully tackling this type of problems.

However, whenever recourse to the MEP was made, we have tacitly assumed in our previous work[2], that a <u>positive-definite</u> quantity, given as the exponential of a suitable linear form, is the protagonist of the concomitant algorithm. Finding it is the final goal to be achieved, that will provide the information one is searching for.

The purpose of the present effort is that of overcoming this restriction, at least with reference to the problem outlined in the first paragraphs above.

2.FORMALISM

We shall assume that our response signal $|f\rangle$, to be analyzed by a convenient detection procedure, belongs to a vector subspace U_M (of a suitable vector space) that is spanned by a basis $|n\rangle$ ($n = 1,\ldots,M$). Thus $|f\rangle$ acquires the form

$$|f\rangle = \sum_{n=1}^{M} C_n \; |n\rangle \tag{2.1}$$

The decomposition of $|f\rangle$ is the procedure that allows one to find out the coefficients C_n . To this end, $|f\rangle$ is to be subject of a <u>finite</u> number, N, of independent measurements that will allow for a numerical representation of the response signal[2], in the form of a set of numbers f_i ($i = 1, \ldots,N$), where

$$f_i = \mathcal{L}_i \, |f\rangle \qquad i = 1,\dots,N \tag{2.2}$$

which under the assumption of linearity can be rewritten as

$$f_i = \sum_{n=1}^{M} C_n \, \ell_{ni} \tag{2.3}$$

if ℓ_{ni} stands for the measurement, represented by $\mathcal{L}_i$ [2], performed upon the "elementary" output signal $|n\rangle$

$$\ell_{ni} = \mathcal{L}_i \, |n\rangle \qquad n = 1,\dots,M \tag{2.4}$$

The new idea to be discussed here is that of allowing for <u>non-positive definite</u> (i.e., negative) coefficients C_n. This is achieved by writting them down as

$$C_n = p_n - B \qquad n = 1,\dots,M \tag{2.5}$$

where the p_n <u>are</u> positive-definite figures and B is an unknown constant, to be self-consistently determined from the data (measurements upon the response signal) according to the algorithm to be developed below. With (2.5), (2.3) acquires the form

$$f_i = \sum_{n=1}^{M} \{p_n \, \ell_{ni} - B \, \ell_{ni}\} \qquad i = 1,\dots,N \tag{2.6}$$

Select one of these N equations, say the r-th one, so as to fix B

$$B = [\, \sum_{n=1}^{M} p_n \, \ell_{nr} - f_r \,] \,/\, [\, \sum_{n=1}^{M} \ell_{nr} \,] \tag{2.7}$$

and introduce the two definitions

$$F_i = f_i \sum_{n=1}^{M} \ell_{nr} - f_r \sum_{n=1}^{M} \ell_{ni} \tag{2.8}$$

$$O_{ni} = \ell_{ni} \sum_{n=1}^{M} \ell_{nr} - \ell_{nr} \sum_{n=1}^{M} \ell_{ni} \tag{2.9}$$

It becomes apparent that the system (2.6) can be recast as

$$F_i = \sum_{n=1}^{M} p_n \, O_{ni} \qquad i \neq r;\ i = 1,\dots,N \tag{2.10}$$

The set of positive numbers $\{p_n\}$ can be thought of as representing a non-normalized probability distribution, whose informational entropy is[4]

$$S = -\sum_{n=1}^{M} p_n \ln p_n + \sum_{n=1}^{M} p_n - 1 \tag{2.11}$$

Each equation of the system (2.10) tells us that the datum F_i is proportional to the mean value of a random variable whose values are given by the O_{ni} $(n = 1,\ldots,M)$ and "weighted" by the p_n.

The idea is now to solve the system (2.10) by recourse to the MEP. An iterative procedure will be followed in which an "optimal conjecture" is successively improved according to the MEP.

We start with a zeroth-order guess, in which a set $p_n^{(0)}$ is obtained by requiring that the form (2.11) be maximized (that is, $p_n = 1$, independent of n). This approximation provides our zeroth-order estimate for B, to be called $B^{(0)}$

$$B^{(0)} = 1 - f_r \Big/ \sum_{n=1}^{M} \ell_{nr} \tag{2.12}$$

with which we can <u>predict</u> for the result of the remaining measurements the values

$$f_i^{(0)} = \sum_{n=1}^{M} p_n^{(0)} \ell_{ni} - B^{(0)} \sum_{n=1}^{M} \ell_{ni} \qquad i \neq r \;;\; i = 1,\ldots,N \tag{2.13}$$

The quality of this conjecture can be measured by defining the "predictive error" ε_i (for the i-th measurement)

$$\varepsilon_i = |f_i - f_i^{(0)}| \,/\, |f_i| \qquad i = 1,\ldots,N \tag{2.14}$$

In order to improve upon the zeroth-order guess and construct a first-order estimate we select, among the members of the set $\{\varepsilon_i\}$, its largest one, to be called ε_{k1}. The first-order weights $p_n^{(1)}$ are chosen so as to maximize S subject to the constraint

$$F_{k1} = \sum_{n=1}^{M} p_n^{(1)} O_{nk1} \tag{2.15}$$

which is tantamount to enforce the fulfillment of the k1-th equation in the system (2.10). According to Jaynes' MaxEnt approach this leads to[3]

$$p_n^{(1)} = \exp(-\lambda_1 O_{nk1}) \tag{2.16}$$

where the Lagrange multiplier λ_1 is constructively obtained by solving (2.15) for it.

With the $p_n^{(1)}$ we can build up the "predictions" $f_i^{(1)}$ (i = 1,...,N) and the concomitant (new) set of ε_i. After selection of the largest element of this new $\{\varepsilon_i\}$-set, let us call it ε_{k2}, we obtain the $p_n^{(2)}$ by maximizing S subject to two constraints, namely, the fulfillment of the equations in the set (2.10) corresponding to both i = k1 and i = k2.

In general, the J-th order estimate is

$$p_n^{(J)} = \exp \{ - \sum_{i=1}^{J} \lambda_i \, O_{nki} \} \qquad n = 1,\dots,M \qquad (2.17)$$

where the Lagrange multipliers λ_i (i = 1,...,J) are obtained by solving the J equations

$$F_{ki} = \sum_{n=1}^{M} p_n^{(J)} \, O_{nki} \qquad i = 1,\dots,J \qquad (2.18)$$

The iterative process is to be ended when we reach the situation

$$\varepsilon_i \leq \Delta f_i \qquad i = 1,\dots,N \qquad (2.19)$$

where Δf_i are the errors (of whatever origin) that characterize the experimental data f_i. Let us suppose that this happy circumstance occurs when we reach the L-th iteration. Our final results will be

$$B^{(L)} = [\sum_{n=1}^{M} p_n^{(L)} \, \ell_{nr} - f_r] \, / \, [\sum_{n=1}^{M} \ell_{nr}] \qquad (2.20)$$

$$C_n^{(L)} = p_n^{(L)} - B^{(L)} \qquad n = 1,\dots,M \qquad (2.21)$$

and they allow for the <u>prediction</u> of any subsequent measurement performed upon $|f\rangle$, f_{N+1}, f_{N+2},..., f_{N+k},... These predictions read

$$f_{N+k}^{(L)} = \sum_{n=1}^{M} C_n^{(L)} \, \ell_{n(N+k)} \qquad k = 1,2,3,\dots \qquad (2.22)$$

3.A NUMERICAL ILLUSTRATION

As an application of the formalism expounded in the preceding paragraphs we shall consider that situation in which the measurements performed upon the signal $|f\rangle$ are obtained as a function of some appropriate parameter t_i (i = 1,...,N). This is a common occurrence indeed. We assume, of course, that we deal with N independent measurements, so that each value t_i can be regarded as defining an (orthogonal) direction $|t_i\rangle$ in an appropriate N-

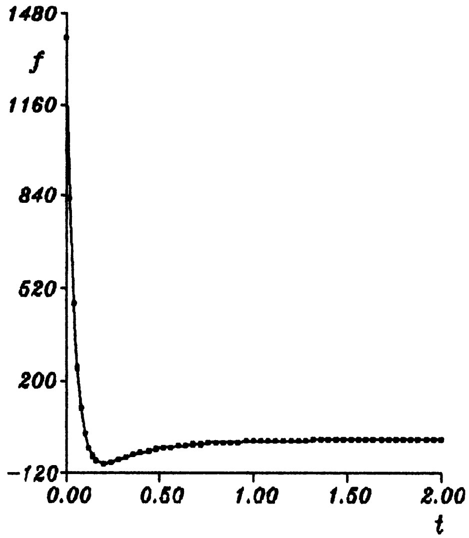

Fig.1 The continuous curve constitutes the theoretical predictions of the data (squares).

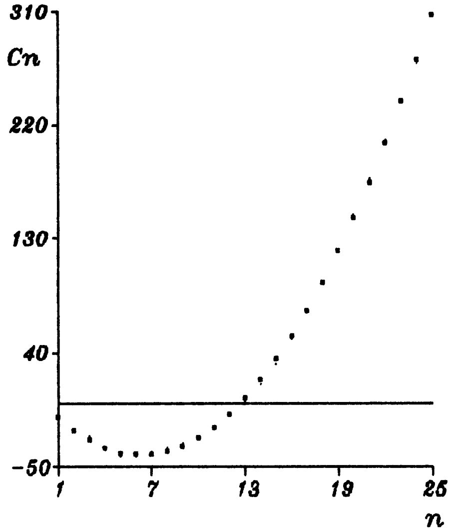

Fig.2 The exact Cn coefficients (squares) are compared with the results of the seventh order iterative process (points).

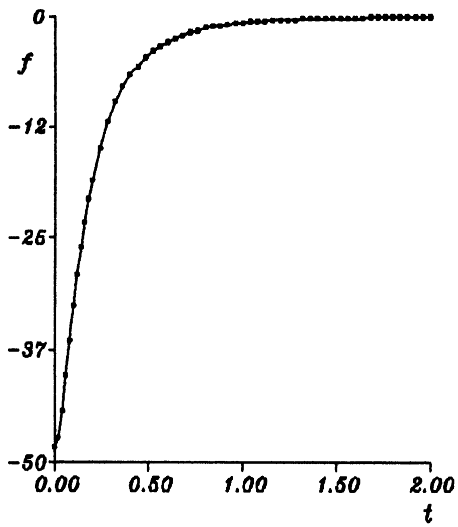

Fig.3 The curve is the theoretical prediction, the squares the numerical representation of the signal.

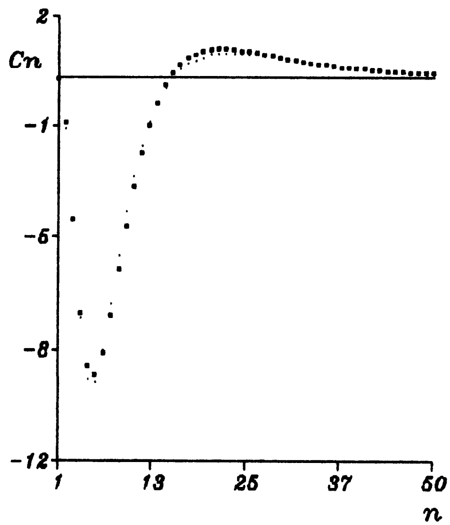

Fig.4 The exact Cn coefficients (squares) are compared with the results of the fifth order iterative process (points).

dimensional space. The figures f_i that result after performing each measurements can now be regarded as "projections" of $|f\rangle$ upon the "direction" defined by t_i

$$f_i = \langle t_i|f\rangle \qquad i = 1,\dots,N \tag{3.1}$$

and, in an analogous fashion, we set

$$\ell_{ni} = \langle t_i|n\rangle \qquad n = 1,\dots,M \tag{3.2}$$

We shall consider two situations. In both of them we take

$$\langle t_i|n\rangle = \exp(-nt_i) \tag{3.3}$$

and we simulate the data in two different manners. Those of Fig.1 (squares) arise from the expression

$$f_i = \sum_{n=1}^{M} n(n - 12.5)\exp(-nt_i) \tag{3.4}$$

with $t_i = (j - 1)0.02$ ($j = 1,\dots,100$). The iterative process of section 2 converges after 7 iterations for an error $\Delta f_i = 0.01$ for all i. The coefficients $C_n^{(7)}$ are displayed in Fig.2 (points), while the exact ones are represented by squares. Coming back to Fig.1, the continuous curve represents the predictions afforded by the $C_n^{(7)}$. In a second example we simulate the data (squares) of Fig.3 by means of a set of coefficients C_n

$$C_n = (n - 15.6)\exp\{-(\ln n - \ln 7)^2/(\ln 1.8)^2\} \tag{3.5}$$

Convergence, after 5 iterations and an error equal to that of Example 1 yields the points of Fig.4 which are to be compared to the exact ones (squares). The continuous curve in Fig.3 represents the predictions calculated by recourse to the $C_n^{(5)}$.

4.CONCLUSION

We have presented a method that is able to conveniently decompose a complex signal as a linear superposition of independent, elementary signals. With the help of the Maximum Entropy Principle, we have devised a practical algorithm that allows for the determination of the coefficients of that linear superposition.

The method can be of relevance for the study of systems that, upon interacting with well known probes, react producing complex output signals of the type described above.

Application to two simple examples allows one to appreciate the fact that this algorithm produces excellent results.

REFERENCES

1 - C.E. Shannon, Proc. IRE 37 (1949) 1

2 - A. Plastino, L. Rebollo Neira, A.G. Alvarez, Phys. Rev. A40 (1989) 1644

3 - E.T. Jaynes, "The Maximum Entropy Formalism" edited by R.D. Levine and M. Tribus (MIT, Boston, 1979)

4 - R. Balian, Y. Alhassid, H. Reinhardt, Phys. Rep. Vol 131 (1986)