

UNIVERSITAT POLITÈCNICA DE CATALUNYA

DEPARTAMENT DE TEORIA DEL SENYAL I  
COMUNICACIONS



BIBLIOTECA RECTOR GABRIEL FERRATE  
Campus Nord

TESI DOCTORAL

TRANSFORMADA DE FOURIER EN PROCESSAMENT  
NO LINEAL DEL SENYAL

AUTORA: ALBA M<sup>a</sup> PAGÈS ZAMORA  
DIRECTOR: MIGUEL ÀNGEL LAGUNAS HERNÁNDEZ  
Barcelona, Octubre de 1996

## Agraïments

La llista de gent a la que m'agradaria agrair el seu granet de sorra en aquesta etapa de doctoranta al GPS és molt extensa. Així que, per a no allargar-me i amb la intenció de que ningú no s'ofengui, només citaré els que han posat una pedra en lloc d'un granet de sorra.

En primer lloc, vull agrair als meus pares l'encoratjament i suport que vaig rebre per part d'ells a l'inici de la tesi, fet que amb el temps es dilueix però que al començament va molt bé ja que no hi ha gaire gent que t'entengui. En Marc també va dir de la seva, tot i que a ell crec que se li ha de reconèixer la paciència que ha demostrat, sovint espectacular com podeu endevinar els que mínimament em coneixeu, i els ànims que m'ha ofert al llarg dels quatre anys que ha durat tot això.

D'en Miguel Ángel, a part de l'amistat que s'ha creat, se li pot donar el títol de que "sense ell això no hagués estat possible". Ja ho sé que és el que formalment es diu, però és que en aquest cas és del tot cert.

També vull esmentar els que han anat circulant per la sala D5-117 i els "ex-ocupes" de la biblioteca perquè gràcies a tots ells les encantadores sobretauls de cada dia han ajudat a pair millor (vosaltres ja m'enteneu). A part d'ells donar les gràcies als soferts polidors de la tesi (Pepe, Javier i Ana), a les actuals companyes de despatx perquè gràcies a elles tinc despatx (Ana i Txell), a la "coconut" que falta esmentar pels moments en comú (Montse), a en Gregori per la complicitat amistosa que m'ha ofert els darrers mesos i, en general, a tothom del grup de comunicacions i del grup GPS perquè són gent fenomenal.



## Abstract

The present work deals with the topic of modeling non linear systems (NLSs), being particularly focused on the chances of the trigonometric or Fourier series in this context. The model presented is based on the Fourier series development of the input/output relation that characterises the NLS under modeling. In fact, a truncated Fourier series is used in order to cope with a feasible model. Thus, the so-called Fourier model can be classified within the set of non-recursive and function series based models, whose paradigmatic model in non linear signal processing is the Volterra model.

This is the main reason to have devoted the second chapter to the Volterra/Wiener model, embracing aspects such as: inherent constraints of the Volterra series (which functions can be developed with polynomial and which not), computational load of the model in terms of number of multiplications per iteration (considering a discrete time implementation) and design capabilities (specially with adaptive methods).

In chapter 3, these three points are discussed for the Fourier model and also used to establish a comparison between both, the Fourier and Volterra model. The main results with respect to the first two aspects are that the Fourier model can deal with non-continuous functions (whereas the Volterra model does not) and that the computational load of both models are of the same order of magnitude, although the Fourier model is slightly larger.

Nevertheless, the most interesting results arise in the design stage of the Fourier model. As it happens in the Volterra systems, the Fourier model also depends linearly on its coefficients, allowing the use of design criteria already developed in the linear signal processing field. Thus, both the Wiener criterion of minimum mean square error (MMSE) and the least squares (LS) criterion have a unique solution, being also possible to apply their adaptive solutions (basically, the NLMS gradient method and the RLS recursive one are presented). In the off-line solution, the matrix resulting in the Fourier model has some symmetric properties (Toeplitz and hermitian symmetry in the memoryless case, and block-Toeplitz and hermitian symmetry in case of dealing with NLS with memory) that allow the use of efficient matrix inversion methods. Furthermore, these properties together with the fact that the complex exponential functions involved in the Fourier model have equal unity power, provide robustness to a gradient adaptive design.

This point is crucial to understand the chances of the Fourier model in adaptive design of NLS models in front of the Volterra model. The Volterra model has convergence problems when gradient adaptive methods are used for the design, due to the high eigenvalue spread of the autocorrelation matrix that emerges in the off-line design. This behaviour forces to use

recursive adaptive solutions that, apart from showing instability problems, also increase the computational cost. Since the Fourier model shows good performance in gradient adaptive solutions and it also has the same computational load of the Volterra model, it represents an alternative choice to adaptive Volterra modeling.

In order to make evident the potential and flexibility of the Fourier model, different versions of the general model are provided in case of dealing with even or odd NLSs. These simplified versions of the Fourier model involved less computational cost without decreasing the performance.

Looking back to the design of the Fourier model, it would be interesting to study in which cases the autocorrelation matrix that appears in the MMSE design becomes diagonal, avoiding in that case the inversion of the matrix. Two possible solutions emerge. The first one consists in driving the Fourier model with a process which has a uniform distribution and, the second one, becomes the orthogonalisation (in the defined norm) of the complex exponential functions involved in the Fourier model.

This last choice has been developed, applying the Gram-Schmidt procedure to the set of complex exponential of the memoryless Fourier model. It is interesting to remark that this procedure leads to a lattice filter, named Fourier-Lattice filter, which basically consists in a non-direct implementation of the memoryless Fourier model that shares some interesting properties with the lattice implementation of an FIR filter. This development, carried out at the end of the 3<sup>rd</sup> chapter, opens the possibility of a future work looking for a non-direct implementation of the Fourier model with memory as if it is already done with the so-called multilinear filters in the linear signal processing field.

With respect to the use of a uniform distributed input signal in order to have a diagonal autocorrelation matrix in the MMSE design, it is evident that it is not always possible to choose the kind of distribution that the input to a given NLS has. Nevertheless, whenever the identification of a NLS is concerned, the use of uniform distributed signals to design the Fourier model will simplify the solution, both in an off-line design or in an adaptive one. In any case, in chapter 4 it is shown how a previous “whitening” of the probability density function can be considered in order to make the learning of the Fourier model easier.

The behaviour of the Fourier model in front of a uniform distributed input signal is symptomatic of a strong relation carried by the Fourier model between the power spectrum density (PSD) of a process and the probability density function (PDF) of a random variable. In fact, a uniform distribution plays the same role in the Fourier model as if it was a white noise (with “uniform” PSD) supplying a MA model since the matrices of the MMSE design of both models would become diagonal. That is, the Fourier model deals with the PDF of the input signal as if it was a PSD of the input signal of an FIR filter.

The formal discussion of this parallelism between the Fourier model and an MA model, as well as, between the PDF of the input signal to the Fourier model and the PSD of the input signal to the MA model is included in chapter 4. This relationship offers the possibility of using spectral estimate methods for providing a PDF estimate when used in the frame of Fourier modeling. Both parametric and non-parametric spectral estimate methods are formulated to provide a PDF estimation, being included both the univariate case as well as the multidimensional one. Probably, the estimation of the joint PDF of a multivariate random variable is the main contribution of the PDF-PSD relation, together with a vector-matrix notation of the PDF estimation useful for evaluating the computational cost of the involved estimation.

Finally, chapter 5 gathers all the simulation results that have been organised in two different topics: NLS modeling (memoryless and with memory) and PDF estimation.



# Índex

<b>Capítol 1</b>	<b>Introducció</b>	<b>1</b>
1.1	Models de sistemes no lineals.....	2
1.2	Visió general de la tesi .....	3
1.3	Referències.....	4
<b>Capítol 2</b>	<b>Sèries de Volterra/Wiener per a modelar sistemes no lineals</b>	<b>5</b>
2.1	Sistemes de Volterra .....	6
2.1.1	Sèrie de Taylor.....	6
2.1.2	Introducció històrica al model de Volterra.....	8
2.1.3	Sèries de Volterra.....	9
2.1.4	Arquitectura per a implementar el model de Volterra.....	11
2.1.5	Limitacions del model de Volterra.....	13
2.2	Motivació del model de Wiener.....	14
2.3	El model de Wiener per a sistemes no lineals .....	17
2.3.1	Els G-funcionals: una base ortonormal de sèries funcionals. ....	17
2.3.2	Descomposició en funcions ortonormals del nucli de Wiener.....	19
2.3.3	El treball originari de Wiener.....	20
2.3.4	Arquitectura del model de Wiener .....	21
2.3.5	Validació del model de Wiener.....	27
2.4	Disseny adaptatiu.....	28
2.4.1	Mètodes de gradient: l'algorisme LMS .....	29
2.4.1.1	L'algorisme LMS Normalitzat (NLMS) .....	31
2.4.1.2	L'algorisme NLMS amb $\alpha(m)$ .....	31
2.4.2	Mètodes recursius: l'algorisme RLS .....	32
2.4.3	Estructures en gelosia o "lattice".....	34
2.5	Conclusions.....	35
2.6	Referències.....	36
<b>Capítol 3</b>	<b>Model de Fourier de sistemes no lineals basat en la sèrie trigonomètrica</b>	<b>39</b>
3.1	Model de Fourier per a sistemes no lineals sense memòria.....	40
3.1.1	Sèries de Fourier. Principals propietats.....	41
3.1.2	Model de Fourier sense memòria.....	45



3.1.3	Arquitectura per implementar el model de Fourier.....	48
3.1.4	Disseny del model de Fourier.....	51
3.1.4.1	Disseny de mínims quadrats del model de Fourier sense memòria.....	56
3.1.4.2	Disseny adaptatiu del model de Fourier.....	59
3.1.5	Perspectiva general del model de Fourier per SNLs sense memòria.....	61
3.2	Model de Fourier per sistemes no lineals amb memòria.....	62
3.2.1	Sèrie de Fourier multidimensional.....	62
3.2.2	Model de Fourier amb memòria.....	64
3.2.3	Arquitectura per implementar el model de Fourier amb memòria..	65
3.2.4	Disseny del model de Fourier amb memòria.....	70
3.2.4.1	Disseny de mínims quadrats del model de Fourier amb memòria.....	73
3.2.4.1	Disseny adaptatiu del model de Fourier amb memòria..	74
3.2.5	Perspectiva general del model de Fourier per SNLs sense memòria.....	74
3.3	Disseny en gelosia o “lattice”.....	74
3.3.1	Implementació lattice del model de Fourier sense memòria.....	75
3.3.1.1	Disseny en bloc del model de Fourier-Lattice sense memòria.....	85
3.3.1.2	Disseny adaptatiu del model de Fourier-Lattice sense memòria.....	87
3.4	Valoracions del model de Fourier.....	88
3.5	Referències.....	89

**Capítol 4 Model de Fourier en l'estimació de funcions densitat de probabilitat mitjançant mètodes d'estimació espectral 95**

4.1	Una perspectiva nova sobre el model de Fourier.....	96
4.2	El model de Fourier i la caracterització estadística dels senyals.....	98
4.2.1	Blanqueig de la funció densitat de probabilitat.....	100
4.3	Mètodes d'estimació espectral per a estimació de FDP. Cas unidimensional.....	102
4.3.1	Mètodes clàssics d'estimació espectral: Periodograma i mètode de Blackman-Tukey.....	107

4.3.2	Mètodes d'Estimació Espectral basats en filtratge: mètode de Míxima Variància.....	110
4.3.3	Mètodes paramètrics d'estimació espectral. ....	115
4.3.3.1	Mètode de Predicció Lineal.....	116
4.3.3.2	Mètode de Màxima Entropia.....	120
4.4	Mètodes d'estimació espectral per a estimació de FDP conjunta. Cas multivariable .....	122
4.4.1	Mètodes clàssics.....	125
4.4.2	Mètodes paramètrics .....	127
4.4.3	Mètode de Míxima Variància.....	129
4.5	Tècniques basades en DVS aplicades a variables aleatòries discretes. ....	131
4.6	Conclusions.....	132
4.7	Referències.....	133
<b>Capítol 5</b>	<b>Avuallació del comportament experimental</b>	<b>139</b>
5.1	Modelatge de SNLs amb el model de Fourier.....	140
5.1.1	Identificació de SNLs sense memòria mitjançant el model de Fourier sense memòria. ....	140
5.1.1.1	Disseny de mínims quadrats en la identificació d'un SNL sense memòria.....	142
5.1.1.2	Disseny adaptatiu en la identificació d'un SNL sense memòria.....	145
5.1.2	Identificació de SNLs amb memòria mitjançant el model de Fourier amb memòria.....	150
5.1.2.1	Disseny de mínims quadrats en la identificació d'un SNL amb memòria .....	152
5.1.2.2	Disseny adaptatiu en la identificació d'un SNL amb memòria.....	154
5.2	"Blanqueig" de la funció densitat de probabilitat .....	157
5.3	Estimació de la funció densitat de probabilitat. ....	161
5.3.1	Estimació de FDP. Cas univariable. ....	162
5.3.2	Estimació de FDP. Cas multivariable .....	166
5.4	Conclusions.....	171
5.5	Referències.....	172
<b>Capítol 6</b>	<b>Conclusions i futures línies de recerca</b>	<b>173</b>



# Capítol 1

## Introducció

Nombroses aplicacions en el camp de processament del senyal mostren signes de característiques no lineals. Aquestes aplicacions s'emmarquen en camps tant diversos com sonar, HF radar, processament d'imatge, control, econometria, comunicacions, biologia i disseny d'altaveus, entre d'altres. Per altra banda, i pel que fa a la caracterització estadística del senyal, poden aparèixer en alguns casos la propietat de no estacionarietat del senyal així com diversitat en el tipus de distribució del senyal d'entrada.

En enfrontar-nos amb un sistema que té un comportament no lineal, la primera eina que es fa necessària en el seu estudi (síntesi, identificació, compensació,...) és la d'un model matemàtic que sigui capaç de reproduir amb fidelitat el citat comportament. En aquest sentit és rellevant destacar l'elevat nombre de diversos models existents, que han anat sorgint en funció de les mancances d'uns i altres, tot adequant-se a diferents entorns o problemàtiques. Pel que fa a l'estacionarietat i posterior adequació del model al tipus de senyal, l'opció d'un aprenentatge adaptatiu sembla ésser la més utilitzada.

Aquest primer capítol pretén ésser una introducció molt descriptiva dels models existents de sistemes no lineals, cosa que serà d'utilitat per a emmarcar la present tesi. A continuació, es descriurà detalladament cadascun dels capítols que segueixen.

## 1.1 Models de sistemes no lineals

Tal i com Billings proposa en el segon capítol del llibre [Bil84] dedicat a la identificació de sistemes no lineals, es poden distingir 3 tipus de desenvolupaments per a la modelització de sistemes no lineals. En primer lloc, s'hauria de citar les equacions diferencials [Vid78] (o equacions en diferències per a sistemes discrets) que bàsicament sorgeixen en la interpretació de fenòmens físico-químics i que, sempre que és possible de determinar-les, són capaces de caracteritzar un determinat fenomen amb un nombre de paràmetres relativament baix ja que les seves equacions inclouen valors actuals i passats, no tan sols del senyal d'entrada sino que també del de sortida. Un cas particular d'aquest grup de sistemes no lineals és el model bilineal que consisteix en un model ARMA que, a més a més, inclou termes de segon ordre que barrejen la diversitat temporal de l'entrada amb la de la sortida [Sub81]. Aquesta propietat, per contra, provoca problemes d'instabilitat i a part, no sempre és possible una formulació tancada del comportament del sistema en forma d'equacions diferencials.

Per tal de suplir aquesta manca de generalitat dels models basats en equacions en diferències, un segon grup de models de sistemes no lineals es troba en els models desenvolupats a partir de sèries de funcions, com és el cas del model de Volterra o el de Wiener [Sch89,Wie58]. En aquests casos no existeix cap realimentació, és a dir, l'actual sortida només depèn del valor actual i valors anteriors del senyal d'entrada. Aquest fet proporciona estabilitat al model però fa que alhora sigui necessari un elevat nombre de paràmetres, alguns cops desorbitat, per a caracteritzar certs sistemes.

En una altra línia, i per tal de disposar de models amb un nombre reduït de paràmetres que permeti el disseny del mateix, han sorgit d'altres sistemes que consisteixen en la connexió en sèrie de sistemes lineals amb sistemes no lineals no dinàmics (és a dir, sense memòria). Així, si el primer bloc consisteix en un sistema no dinàmic i no lineal seguit d'un altre sistema lineal, l'arquitectura resultant s'en diu de *Hammerstein* [Cas83, Bil84], mentre que si la col·locació és la inversa, es coneix pel nom de model de Wiener, per la seva relació amb la sèrie funcional de Wiener [Wie58,Sch89]. Alguns cops també s'opta per la connexió en paral·lel de sistemes com els citats anteriorment, i llur aprenentatge es pot assolir mitjançant mètodes basats en funcions d'estadística d'ordre superior [Ben90].

Per altra banda, es troben els models basats en xarxes neuronals (veure per exemple [Hec90]) com poden ésser les xarxes basades en funcions de base radial que van tenir un grau d'acceptació elevat en l'equalització de canals de comunicació ([Che91, Cid94]).

En aquest entorn de modelatge de sistemes no lineals es situa l'objectiu principal de la present tesi que consisteix en estudiar les possibilitats d'un model basat en la sèrie funcional trigonomètrica o també coneguda, en el món del processament del senyal, com a *sèrie de*

*Fourier*. Els esforços s'han encaminat a presentar-la tot establint una comparació amb el model de Volterra, tant a nivell d'arquitectura com de disseny, bàsicament adaptatiu ja que és l'entorn en el que el model de Volterra s'ha desenvolupat considerablement en relació als altres models. Així, a continuació es presenta un resum del contingut dels següents capítols, cosa que ajudarà a establir en quins termes s'ha enfocat el treball.

## 1.2 Visió general de la tesi

En vistes de que el model de Fourier per a sistemes no lineals consisteix en un model basat en sèrie de funcions, el capítol 2 de la present tesi s'ha dedicat exclusivament a presentar el model de Volterra i el desenvolupat per Wiener arrel del primer, com a models paradigmàtics d'aquest grup de models no lineals. Es presta més atenció al model de Volterra, tot estudiant-lo en termes de generalitat, complexitat i disseny (tant en bloc com adaptatiu). Pel que fa al model de Wiener, aquest es presenta com el model pioner basat en les sèries de Volterra que va intentar evitar els defectes bàsicament de disseny que el model de Volterra presenta. En una segona part del capítol 2 s'ha fet un recull de la bibliografia relacionada amb el model de Volterra per tal d'emmarcar les línies bàsiques de recerca en les que actualment es debat aquest model.

El 3<sup>er</sup> capítol ja s'endinsa en l'aportació d'aquesta tesi, presentant el model de sistemes no lineals basat en la sèrie de Fourier (o trigonomètrica). La discussió del model s'estableix en els mateixos termes que pel cas del model de Volterra per tal de facilitar la tasca de comparació.

En un altre sentit, el capítol 4 enfoca el model de Fourier des d'una altra perspectiva com a conseqüència del paral·lelisme que sorgeix entre el model de Fourier i un model MA en termes de la funció densitat de probabilitat del primer i de la densitat espectral de potència del segon. Aquesta relació ens condueix, tal i com es mostra en el citat capítol, a la possibilitat d'utilitzar mètodes d'estimació espectrals però aplicats a l'estimació de funcions densitat de probabilitat.

A continuació, el 5<sup>é</sup> capítol conté totes les simulacions que s'han considerat oportunes. En elles s'ha volgut deixar constància de quin és el comportament del model de Fourier, tant pel que fa a modelatge de sistemes no lineals en dissenys adaptatius i en bloc, com pel que fa a l'estimació de funcions densitat de probabilitat d'una o múltiples variables aleatòries.

Finalment, el capítol 6 recull les conclusions a les que s'ha arribat en aquesta tesi i també es proposen futures vies de recerca.

### 1.3 Referències

- [Ben90] Bendat J.S., *Nonlinear System Analysis & Identification*, John Wiley & Sons, Inc., 1990.
- [Bil84] Billings S.A., *Nonlinear Systems Design*, Capítol 2, Peter Peregrinus Ltd., London, U.K., 1984.
- [Cas83] Casar-Corredera J.R., *Aplicación de las Estructuras No Lineales de Wiener y Hammerstein al Modelado y Tratamiento de Señales*, Tesi Doctoral, Universitat Politècnica de Madrid, 1983.
- [Cid94] Cid-Sueiro J., *Redes Neuronales para Igualación y Detección en Comunicaciones Digitales*, Tesi Doctoral, Universitat Politècnica de Madrid, 1994.
- [Che91] Chen S., Gibson G.J., Cowan C.F.N., "Reconstruction of Binary Signals Using an Adaptive Radial Basis Function Equalizer", *Signal Processing*, Vol. 20, No. 2, pp.77-93, 1991.
- [Hec90] Hecht-Nielsen R., *Neurocomputing*, Addison-Wesley Publishing Company, USA, 1990.
- [Sch89] Schetzen M., *The Volterra/Wiener Theories of Nonlinear Systems*, Robert E. Krieger Publishing Company, Florida, 1989. (1<sup>a</sup> edició 1980).
- [Sub81] Subba Rao T., "On the Theory of Bilinear Time Series Models", *J. Royal Statistics Society*, Sèrie B., Vol.43, 1981.
- [Vid78] Vidyasagar M., *Nonlinear Systems Analysis*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1978.
- [Wie58] Wiener N., *Nonlinear problems in random theory*, The Technology Press, M.I.T. and John Wiley & Sons, Inc., New York, 1958.

## Capítol 2

### Sèries de Volterra/Wiener per a modelar sistemes no lineals

Aquest capítol està dedicat a la presentació dels anomenats models de Volterra [Vol13] i de Wiener [Wie58] que van ser desenvolupats respectivament al principi i a mitjan segle XX per a modelar sistemes no lineals.

El capítol s'ha estructurat seguint un ordre cronològic, és a dir, primer es presenta el model de Volterra sense memòria i després amb memòria, tot incidint en les mancances que pateix i que posteriorment varen originar el treball de Wiener. La valoració d'aquests models, que servirà per comparar-los entre ells i envers el model que es proposa al capítol 3, es basarà en diversos aspectes. Un d'ells serà "l'abast" dels diferents models, és a dir, quin tipus de no linealitats són capaços d'aproximar amb un error arbitràriament baix. En aquest sentit, serà rellevant l'estudi de la convergència de la sèrie en la que es basa el model. A l'altra banda, i en constant contraposició, estarà la complexitat del model que es mesura mitjançant el nombre de coeficients que inclou i s'expressa en funció del "ordre" i "memòria" del model. Un darrer aspecte que també es ressaltarà és la incidència que l'estadística del senyal d'entrada té sobre el model, tant a nivell estructural com de disseny.

Al llarg del capítol es podrà comprovar que tots dos models, el de Wiener i el de Volterra, són models paramètrics que depenen linealment dels coeficients. Aquesta important



propietat permet un disseny dels models mitjançant mètodes adaptatius, i és en aquest camp on han proliferat les seves aplicacions (especialment el model de Volterra). Per aquesta raó, s'ha considerat interessant incloure un darrer apartat dedicat a mostrar les peculiaritats que els algorismes adaptatius més coneguts presenten en ésser aplicats a models no lineals (de fet, només es farà referència al model de Volterra ja que és el més estès).

## 2.1 Sistemes de Volterra

El model proposat per Volterra sorgeix com una extensió de la caracterització dels sistemes lineals invariants amb el temps (SLIs) mitjançant la resposta impulsional  $h(n)$  (Eq.2.1).

$$y(n) = \sum_{m=-\infty}^{+\infty} h(m) \cdot x(n-m) \quad (2.1)$$

Per començar, però, ens centrarem en el model de Volterra que serveix per modelar sistemes no lineals (SNLs) sense memòria, o també anomenats SNLs estàtics. Aquest primer pas ajuda a la posterior consecució i avaluació del model de Volterra més general que contempla memòria.

### 2.1.1 Sèrie de Taylor

Tot sistema lineal sense memòria es pot caracteritzar per una relació entrada/sortida que només depèn del valor actual de l'entrada. Llavors, la representació del senyal de sortida,  $y(n)$ , en funció de l'entrada,  $x(n)$ , proporciona una corba  $g(\cdot)$  que, sota determinades condicions, pot expressar-se com una sèrie de potències o polinomis en termes de  $x$ .

$$y(n) = g[x(n)] = \sum_{p=0}^{+\infty} c_p \cdot x^p(n) \quad ; c_p = \frac{1}{p!} \left. \frac{\partial^p g(x)}{\partial x^p} \right|_{x=0} \quad (2.2)$$

Aquesta representació és el desenvolupament en sèrie de Taylor de la funció  $g(\cdot)$  i constitueix l'origen del model de Volterra ja que, tal i com es veurà al següent apartat, la sèrie de Volterra es pot interpretar com la generalització de la sèrie de Taylor en el cas d'aproximar funcions entrada/sortida de SNLs amb memòria. Aquesta relació entre ambdues sèries permet intuir les limitacions que presentarà la sèrie de Volterra, tot extrapolant les de la sèrie de Taylor a un situació multivariable. En aquest sentit, sembla interessant presentar una reflexió sobre la

sèrie de Taylor de manera que, tot i fer-se des d'una perspectiva poc estricta, permeti discutir les propietats que ofereix.

Una d'aquestes propietats, molt important en els models de SNLs mitjançant sèries de funcions, és la convergència. En aquest sentit, cal destacar en detriment del model que no totes les funcions admeten un desenvolupament en sèrie de Taylor que convergeixi en tot el rang d'entrada  $x \in ]-\infty, +\infty[$  i, en conseqüència, no sempre la igualtat anterior (Eq.2.2) és correcte (per exemple, qualsevol funció que presenti una discontinuïtat no admet un desenvolupament en sèrie de Taylor [Sch81]). És més, fins i tot en el cas de que la sèrie de Taylor convergeixi en tot el rang dels reals, poden aparèixer problemes de convergència en el model. En efecte, en voler aconseguir un model realitzable a partir de la sèrie proposada (Eq.2.2), s'haurà de treballar amb una versió d'ordre finit  $P < +\infty$ , resultant-ne una aproximació de la funció  $g(\cdot)$  que s'indica per  $\hat{g}(\cdot)$ .

$$y(n) \equiv \hat{g}[x(n)] = \sum_{p=0}^P c_p \cdot x^p(n) \quad ; 0 < P < +\infty \quad (2.3)$$

Aquesta versió truncada de la sèrie de Taylor (Eq.2.3) només serà convergent dins d'un rang determinat de l'entrada (fora del rang divergirà tendint a  $\pm\infty$ ). Aquest fet és clarament un inconvenient de les sèries en potències ja que gran part dels sistemes no lineals trobats a la pràctica presenten un efecte de saturació que, estrictament, no podrà ser modelat en tot el rang per una sèrie truncada com l'anterior.

A la pràctica, es predetermina un rang limitat de l'entrada dins del qual, suposant que la sèrie de Taylor original convergeixi, el model proposat (Eq.2.3) serà vàlid. L'error comès per l'aproximació es calcula com la diferència entre la funció real,  $g(\cdot)$ , i la proporcionada pel model,  $\hat{g}(\cdot)$ .

$$E_P(x) = g(x) - \sum_{p=0}^P c_p \cdot x^p \quad (2.4)$$

El fet de que la sèrie de Taylor convergeixi assegura que, per tot valor real de  $x$ , l'error  $E_P(x)$  tendeix a 0 a mesura que  $P \rightarrow \infty$ . Llavors, qualsevol intent de disminuir l'error comès pel model implicarà un augment de l'ordre de l'aproximació  $P$  i per tant un increment de la complexitat del model (és a dir, més paràmetres  $c_p$ ). Sorgeix així, un compromís entre complexitat del model i exactitud assolida que serà present al llarg de tota la tesi.

A part del problema de la convergència, un darrer aspecte que també serà present a Volterra és la dificultat per a determinar els paràmetres del model fent servir la sèrie de Taylor. Per una banda està l'ordre de l'aproximació  $P$  que es decideix a cada aplicació en base al mètode

de prova i error, o senzillament es limita per problemes tecnològics o d'implementació (per exemple, disposar d'un nombre limitat d'operacions per segon en aplicacions en temps real). De totes maneres, el principal problema radica en el disseny dels coeficients  $\{c_p \forall p=0, \dots, P\}$ . En aquest punt, i segons la definició de la sèrie de Taylor (Eq.2.2), es necessita de  $g(x)$  per determinar-los però en una situació real aquesta funció es desconeix [Vid78].

Tot i que el que s'entén actualment per model de Volterra guarda poca, fins i tot molt poca, relació amb el treball originari de Taylor, sembla interessant retrocedir en el temps fins a aquest punt ja que, tal i com s'anirà plasmant al llarg del capítol, els problemes de convergència i, sobretot, de disseny també són presents en el model de Volterra, com si es tractés d'una reminiscència del desenvolupament originari basat en sèrie de Taylor.

### 2.1.2 Introducció històrica al model de Volterra

La sèrie de Volterra sorgeix del treball que Vito Volterra va desenvolupar al començament de l'anàlisi funcional [Vol13] amb la finalitat de generalitzar la sèrie de Taylor al cas multivariable. Amb la finalitat de modelar SNLs amb memòria, la relació entrada/sortida  $g(\cdot)$  és en efecte una funció multivariable ja que no només depèn del valor de l'entrada en l'instant de temps present, sino que també depèn de valors de l'entrada previs i, en el cas més general, dels posteriors. El resultat, formulat per a sistemes discrets, es pot expressar com una sèrie funcional en termes de potències de la diversitat temporal de l'entrada,

$$\begin{aligned}
 y(n) = & h_0 + \sum_{m_1=-\infty}^{+\infty} h_1(m_1) \cdot x(n-m_1) + \sum_{m_1=-\infty}^{+\infty} \sum_{m_2=-\infty}^{+\infty} h_2(m_1, m_2) \cdot x(n-m_1)x(n-m_2) + \dots \\
 & + \sum_{m_1=-\infty}^{+\infty} \sum_{m_2=-\infty}^{+\infty} \dots \sum_{m_p=-\infty}^{+\infty} h_p(m_1, m_2, \dots, m_p) \cdot x(n-m_1)x(n-m_2) \dots x(n-m_p) + \dots
 \end{aligned}
 \tag{2.5}$$

on les funcions  $\{h_p(m_1, m_2, \dots, m_p) \ p > 1\}$  van ser originàriament definides per Volterra en termes de la derivada  $p$ -èssima de la funció multivariable  $g(\cdot)$  que es vol aproximar. D'aquí sorgeix el fet de que sovint es fa referència a la sèrie de Volterra com la sèrie de Taylor amb memòria i, en aquest sentit, és interessant destacar que si les funcions  $\{h_p(m_1, m_2, \dots, m_p) \ p > 1\}$  de la sèrie de Volterra es particularitzessin com a productes de deltes de Kronecker centrades en l'origen, és a dir  $\{\delta(m_1)\delta(m_2), \dots, \delta(m_p) \ p > 1\}$ , en resultaria una sèrie polinòmica com la de Taylor [Rug81]. En conseqüència, es pot entendre que quan es fa referència tant a sèries com sistemes de Volterra amb memòria, també s'inclou el cas sense memòria.

Com a conseqüència d'aquesta relació, és força coherent pensar que les limitacions de mesura dels coeficients  $c_p$  i de convergència de la sèrie de Taylor també estaran presents en la

sèrie de Volterra [Sch89]. Aquest fet va motivar diversos estudis que permeten una redefinició de les funcions  $\{h_p(m_1, m_2, \dots, m_p) \ p > 1\}$  per tal de facilitar-ne la seva determinació i també, per augmentar el nombre de funcions susceptibles d'ésser aproximades per a una sèrie funcional com l'anterior (Eq.2.5).

Així, en el cas sense memòria de la sèrie de Taylor, el teorema de Weierstrass assegura l'existència d'un polinomi d'ordre finit  $P$  del tipus

$$\hat{g}(x) = \sum_{p=0}^P c_p \cdot x^p \quad (2.6)$$

(on ara  $c_p$  no respon als coeficients de Taylor de l'expressió (Eq.2.2)), capaç d'aproximar amb un error absolut arbitràriament baix qualsevol funció real i contínua  $g(x)$  dins d'un interval tancat  $[a, b]$  de l'eix de coordenades. De la mateixa manera, el treball de Stone-Weierstrass [Rug81] representa la versió amb memòria del treball de Weierstrass, permetent una expressió polinòmica del tipus,

$$\begin{aligned} \hat{g}(x) = & h_0 + \sum_{m_1=-\infty}^{+\infty} h_1(m_1) \cdot x(n - m_1) + \sum_{m_1=-\infty}^{+\infty} \sum_{m_2=-\infty}^{+\infty} h_2(m_1, m_2) \cdot x(n - m_1)x(n - m_2) + \\ & \dots + \sum_{m_1=-\infty}^{+\infty} \sum_{m_2=-\infty}^{+\infty} \dots \sum_{m_p=-\infty}^{+\infty} h_p(m_1, m_2, \dots, m_p) \cdot x(n - m_1)x(n - m_2) \dots x(n - m_p) \end{aligned} \quad (2.7)$$

per a aproximar amb un error arbitràriament baix tota funció multivariable  $g(\cdot)$  real i contínua dins d'un recinte compacte. Aquests teoremes d'existència varen encoratjar a seguir treballant amb la sèrie funcional (Eq.2.5) i desenvolupar mètodes diferents dels de Taylor o Volterra que permetessin la determinació dels paràmetres que contenen ( $\{c_p \ p > 1\}$  pel cas sense memòria i  $\{h_p(m_1, m_2, \dots, m_p) \ p > 1\}$  pel cas amb memòria).

Per concloure aquesta breu introducció històrica del model de Volterra és important remarcar que, malgrat que en general les funcions  $\{h_p(m_1, m_2, \dots, m_p) \ p > 1\}$  seran diferents de les proposades inicialment per Volterra, les sèries funcionals que segueixen una formulació igual a la presentada anteriorment, (Eq.2.5) i els models que d'ella s'en deriven, continuen portant la denominació de Volterra, fent honor a aquest matemàtic de principis del segle XX.

### 2.1.3 Series de Volterra

La sèrie funcional de Volterra (Eq.2.8) ens proporciona el fonament matemàtic per a modelar un SNL invariant temporalment i amb memòria a partir d'una aproximació de la relació

entrada/sortida del SNL,  $g(\cdot)$ , basada en potències de la diversitat temporal del senyal d'entrada.

$$\begin{aligned}
 y(n) = & h_0 + \sum_{m_1=-\infty}^{+\infty} h_1(m_1) \cdot x(n-m_1) + \sum_{m_1=-\infty}^{+\infty} \sum_{m_2=-\infty}^{+\infty} h_2(m_1, m_2) \cdot x(n-m_1)x(n-m_2) + \dots \\
 & + \sum_{m_1=-\infty}^{+\infty} \sum_{m_2=-\infty}^{+\infty} \dots \sum_{m_p=-\infty}^{+\infty} h_p(m_1, m_2, \dots, m_p) \cdot x(n-m_1)x(n-m_2) \dots x(n-m_p) + \dots
 \end{aligned} \tag{2.8}$$

De la mateixa manera que un SLI està caracteritzat per la resposta impulsional  $h(n)$  (Eq.2.1), el SNL invariant ho està pel conjunt de funcions  $\{h_0, h_p(m_1, m_2, \dots, m_p) \ p > 1\}$  que es coneixen pel nom de *nuclis de Volterra* ("Volterra's kernels" en anglès). És interessant notar que, a diferència d'un SLI, la resposta impulsional de la sèrie de Volterra (és a dir, la sortida que s'obté quan l'entrada és la funció delta de Kronecker  $x(n) = \delta(n)$ ) no caracteritza completament al sistema. En efecte, s'obtidria informació de tots els nuclis  $h_p(m_1, m_2, \dots, m_p)$  però no avaluats en qualsevol punt de l'hiperespai de dimensió  $p$ , sino solament en els punts  $(m_1, m_2, \dots, m_p)$  que complissin  $\{m_i = n \ \forall i = 1, \dots, p\}$ .

Generalment, a l'igual que en SLIs, es suposen tot un seguit de propietats sobre el SNL que condicionen els nuclis [Sch89]. Exemples de les propietats més comunes són la causalitat temporal del sistema (Eq.2.9a), o bé la simetria dels nuclis (Eq.2.9b) ja que es pot demostrar que no es perd generalitat. A part, tot i que no s'entrarà en detall, també hi ha condicions sobre els nuclis per garantir-ne l'estabilitat .

$$h_p(m_1, m_2, \dots, m_p) = 0 \quad ; \forall m_i < 0 \quad ; \forall i = 1, \dots, p \quad ; \forall p \geq 1 \tag{2.9a}$$

$$h_p(m_i) = h_p(m_j) \quad ; \forall m_i, m_j \equiv \left\{ \begin{array}{l} \text{vectors amb qualsevol de les } p! \\ \text{permutacions de } (m_1, m_2, \dots, m_p) \end{array} \right\} \tag{2.9b}$$

Una altra forma més compacta d'expressar la sèrie de Volterra és com la suma dels anomenats *operadors de Volterra d'ordre  $p$* ,  $\underline{H}_p[x(n)]$ ,

$$y(n) = h_0 + \sum_{p=1}^{+\infty} \underline{H}_p[x(n)] \tag{2.10}$$

de manera que l'operador d'ordre  $p$

$$\underline{H}_p[x] = \sum_{m_1=-\infty}^{+\infty} \sum_{m_2=-\infty}^{+\infty} \dots \sum_{m_p=-\infty}^{+\infty} h_p(m_1, m_2, \dots, m_p) x(n-m_1)x(n-m_2) \dots x(n-m_p) \tag{2.11}$$

és una funció que depèn del senyal d'entrada i del nucli de Volterra d'ordre  $p$  (Eq.2.11). Es diu que són operadors *homogenis* ja que inclouen un sol nucli de Volterra [Sch89].

#### 2.1.4 Arquitectura per a implementar el model de Volterra

A partir de la sèrie de Volterra (Eq.2.8) i assumint que aquesta convergeix, es construeix una versió truncada (Eq.2.12) per tal de que el model es pugui implementar. El truncament s'efectua en dos sentits. Primerament, es limita l'ordre màxim dels nuclis fins a un enter  $P < +\infty$  i, en segon lloc, es suposa que tots els nuclis  $h_p(m_1, m_2, \dots, m_p)$  prenen un valor despreciable per qualsevol  $m_i \forall i=1, \dots, p$  que superi a  $(N-1)$ .

$$\begin{aligned} \hat{g}(x) = & h_0 + \sum_{m_1=0}^{N-1} h_1(m_1) \cdot x(n-m_1) + \sum_{m_1=0}^{N-1} \sum_{m_2=0}^{N-1} h_2(m_1, m_2) \cdot x(n-m_1)x(n-m_2) + \\ & \dots + \sum_{m_1=0}^{N-1} \sum_{m_2=0}^{N-1} \dots \sum_{m_N=0}^{N-1} h_p(m_1, m_2, \dots, m_p) \cdot x(n-m_1)x(n-m_2) \dots x(n-m_p) \end{aligned} \quad (2.12)$$

S'en diu, llavors, que es tracta d'un sistema de Volterra d'ordre  $P$  i memòria  $N$ , ja que només es veuen implicats en el model els  $N$  valors  $\{x(n), x(n-1), \dots, x(n-N+1)\}$  obtinguts de la diversitat temporal de l'entrada, agrupats en potències de fins a  $P$  termes.

La complexitat del model es reflecteix en el nombre de coeficients que inclou, essent aquest de l'ordre de  $O(N^P)$  que correspon a la suma dels punts que contenen  $(P+1)$  hipercubs, cadascun de dimensió  $p$  (per  $p=0, \dots, P$ ), i amb  $N$  punts per aresta. De totes maneres, sota la hipòtesi de que els nuclis de Volterra són simètrics (Eq.2.9b), el darrer model (Eq.2.12) inclou un nombre de coeficients menor (Eq.2.13) que és igual a totes les possibles combinacions diferents de  $N$  termes agafats des de 0 en 0 (corresponent a  $h_0$ ) fins de  $P$  en  $P$ .

$$\text{Nombre de coeficients} = \sum_{p=0}^P \binom{N+p-1}{p} = \binom{P+N}{N} = \frac{(P+N)!}{N! P!} \quad ; \quad \binom{n}{k} = \frac{n!}{n!(n-k)!} \quad (2.13)$$

En general, l'ordre  $P$  s'escull força menor a  $N$ , de manera que el nombre de coeficients d'un model de Volterra( $P, N$ ) esdevé de l'ordre de  $O(N^P)$ . D'aquí el fet de que es digui que el nombre de coeficients d'un model de Volterra té un creixement exponencial.

La implementació pràctica del model es pot interpretar com un cas particular d'una configuració més general de models paramètrics [Fej94b] que clarament respon al que comunament es coneix pel model de Wiener [Wie57]. Aquesta estructura (Fig.2.1) consta de dues etapes: la primera és de sortida múltiple i s'anomena LD (de l'anglès "Linear-Dynamic"), i

la segona és d'entrada múltiple i es diu NMR (també de l'anglès "Nonlinear-Memoryless-Readout").

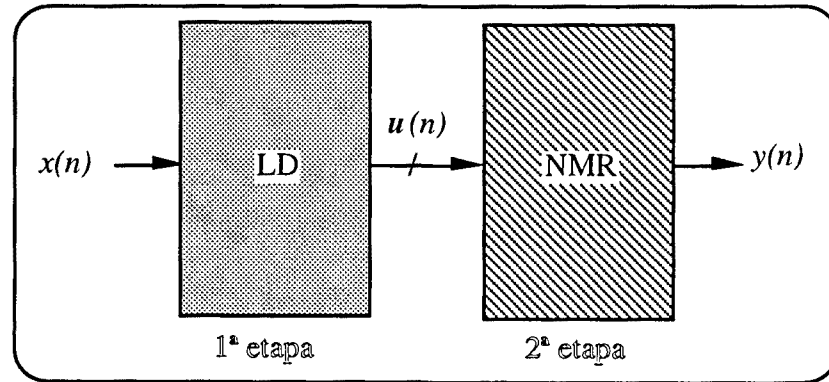


Figura 2.1 Configuració de dues etapes per models no lineals paramètrics ("Linear-Dynamic/Nonlinear-Memoryless-Readout").

En general, la part LD és l'encarregada de generar una diversitat temporal a partir del senyal d'entrada i sempre mitjançant operadors lineals. El resultat es recull en el vector  $u(n)$ , que tot seguit s'introdueix en el mòdul NMR on se li aplica una transformació no lineal sense memòria. Aquesta segona etapa conté els paràmetres del model que es combinen linealment amb les funcions no lineals obtingudes a partir de  $u(n)$ , i d'aquí el nom anglès "Readout", que traduït literalment vol dir *lectura de sortida*.

En aquest punt, tot i que s'incidirà més endavant, cal destacar que en un treball previ [Pag95] es va estudiar un determinat model de sistemes no lineals que respon a una arquitectura diferent de l'anterior però que també consta de dues etapes seguint les pautes del teorema de Kolmogorov.

Respecte de la implementació del sistema de Volterra( $P,N$ ) mitjançant la configuració general anterior, es poden dissenyar les dues etapes tal i com es recull a la figura (2.2). La primera part consisteix en una cadena de ( $N-1$ ) retards que generen un vector de sortida  $u(n)$  que, en aquest cas, coincideix amb el vector  $x(n)$  format per la diversitat temporal de l'entrada (l'apòstrof denota vector transposat).

$$u(n) = x(n) \equiv [x(n) \ x(n-1) \ \dots \ x(n-N+1)]' \quad (2.14)$$

L'etapa NMR es subdivideix ara en dos mòduls (Fig.2.2). El primer d'ells genera els monomis multivariables que inclou el model de Volterra( $N,P$ ) (Eq.2.12), de manera que a la sortida s'obté el vector  $z(x,n)$  que agrupa les combinacions diferents de productes entre components del vector  $x(n)$ .

$$z(x,n) = [ 1 \quad x(n) \quad x(n-1) \dots x(n-N+1); \\ x^2(n) \quad x(n)x(n-1) \dots x(n)x(n-N+1) \quad \dots \quad x^2(n-N+1); \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots ; \\ x^P(n) \quad x^{P-1}(n)x(n-1) \dots x^{P-1}(n)x(n-N+1) \quad \dots \quad x^P(n-N+1)]' \quad (2.15)$$

El segon mòdul combina linealment els diferents monomis junt amb els coeficients corresponents als nuclis de Volterra, que s'agrupen dins del vector  $w$ .

$$w = [ h_0 \quad h_1(0) \quad h_1(1) \dots h_1(N-1); \\ h_2(0,0) \quad h_2(0,1) \dots h_2(0,N-1) \quad \dots \quad h_2(N-1,N-1); \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots ; \\ h_P(0,0,\dots,0) \quad h_P(0,\dots,0,1) \dots h_P(0,\dots,0,N-1) \quad \dots \quad h_P(N-1,N-1,\dots,N-1)]' \quad (2.16)$$

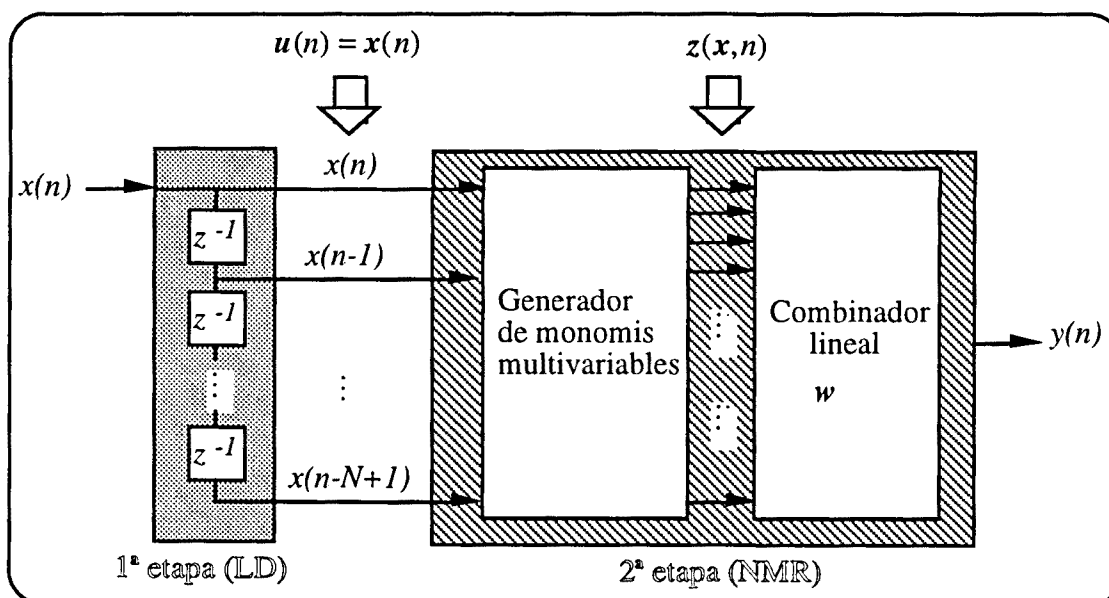


Figura 2.2 Implementació de la sèrie truncada de Volterra a partir de la configuració en dues etapes.

És interessant puntualitzar que un model sense memòria com el de la sèrie polinòmica truncada (Eq.2.6) també es pot implementar seguint l'arquitectura anterior (Fig.2.2); solament caldria eliminar la dinàmica temporal introduïda a la primera etapa de manera que  $u(n)=[x(n)]$ . Això reforça el fet de que el model de Volterra(P,N) de l'expressió (2.12) es pot interpretar com un sistema que serveix tant per modelar SNLs amb memòria ( $N>1$ ) com sense ( $N=1$ ).

### 2.1.5 Limitacions del model de Volterra

Els punts febles del model de Volterra amb memòria esdevenen la convergència de la sèrie i la dificultat de mesurar els nuclis. En relació al primer aspecte, ja s'ha recalcat que els



teoremes de Weierstrass i Stone-Weierstrass asseguruen que el model polinòmic (Eq.2.12) és vàlid sempre i quan l'entrada pertanyi a un recinte compacte. Aquesta premissa condiona en el sentit de que un model polinòmic serà vàlid per a modelar SNLs si el senyal d'entrada pertany a un rang limitat (des d'un punt de vista purista això no seria vàlid, però no es pretèn entrar en detalls matemàtics [Rug81]). Aquesta característica, que ja s'havia observat en la sèrie de Taylor, és fàcil de verificar donat que en el moment en que es delimiti un ordre  $P$  finit pel model de Volterra, s'està forçant a que la resposta del mateix divergeixi a mesura que  $x$  tendeix a  $\pm\infty$ .

Respecte de l'obtenció dels nuclis de Volterra,  $h_p(m_1, m_2, \dots, m_p)$ , ja s'ha comentat que es tracta d'una tasca difícil (especialment quan  $P > 2$ ) i això queda palès en els mètodes existents que es basen en la correlació creuada [Sch89, Koh85]. Intuitivament, val a dir que el problema radica en el fet de que els nuclis del model estan entrelaçats entre ells, dificultant-ne la determinació de cadascun per separat. Aquest problema va ser solucionat per Wiener que, tal i com es veurà en les properes seccions, en primer lloc va proposar un nou criteri d'estimació i, per altra banda, va introduir canvis en el propi model polinòmic de manera que aquest criteri d'estimació es pugués utilitzar de forma eficient en la determinació dels nuclis del model.

## 2.2 Motivació del model de Wiener

Una de les grans contribucions del treball de Wiener al processament del senyal va ser el proposar un nou criteri d'estimació de senyals basant-se en la minimització del valor esperat de l'error quadràtic (MMSE, de l'anglès "Minimum Mean Square Error"). Així, a partir d'unes dades obtingudes d'un senyal estacionari  $x(n)$  agrupades en el vector  $\mathbf{x}(n)$  i un senyal de referència  $d(n)$ , va considerar estimadors  $\hat{d}(n)$  d'aquesta referència del tipus,

$$d(n) \equiv \hat{d}(n) = \sum_{q=1}^{N_c} w_q^* \cdot z_q(\mathbf{x}, n) \equiv \mathbf{w}^H \cdot \mathbf{z}(\mathbf{x}, n) \quad (2.17)$$

que depenen d'un vector de pesos  $\mathbf{w}^H$  (on el superíndex "H" indica una transformació hermítica) que es combina linealment amb un grup de funcions predeterminades aplicades sobre les dades,  $\mathbf{z}(\mathbf{x}, n)$ . Per a aquests models que depenen linealment dels paràmetres (en aquest cas els coeficients  $\mathbf{w}^H$ ), Wiener va proposar un criteri de disseny que minimitzés una funció de cost denotada per  $\xi_w$  i definida com l'error quadràtic mig.

$$\xi_w \equiv E[|e(n)|^2] = E\left[|d(n) - \hat{d}(n)|^2\right] \xrightarrow{\xi_w |_{MIN}} \mathbf{w}_{opt} \quad (2.18)$$

L'error quadràtic mig es pot expressar vectorialment (Eq.2.19) en funció del vector de pesos  $w$ , la matriu d'autocorrelació de les funcions  $R_{zz}$ , la correlació creuada entre el vector de funcions i la referència  $r_{zd}$ , i la potència de la referència  $r_{dd}$ .

$$\begin{aligned}\xi_w &= E[d(n)^2] - w^H E[z(x,n)d^*(n)] - E[d(n)z^H(x,n)]w + w^H E[z(x,n)z^H(x,n)]w \\ &= w^H \cdot R_{zz} \cdot w - w^H \cdot r_{zd} - r_{zd}^H \cdot w + r_{dd}\end{aligned}\quad (2.19)$$

Es tracta doncs d'una forma quadràtica en termes del vector error de pesos [Hay91], cosa que ens assegura la unicitat de la solució per a  $w_{opt}$ , sempre i quan la matriu  $R_{zz}$  sigui definida positiva. Sota aquesta premissa, la funció de cost respon a una forma convexa amb un sol mínim que es pot obtenir igualant a zero el gradient de  $\xi_w$  respecte dels coeficients.

$$\xi_w \Big|_{MINIM} \Rightarrow \nabla_{w^H} \xi_w = R_{zz} \cdot w - r_{zd} \equiv 0 \quad (2.20)$$

El vector de pesos solució d'aquesta equació,  $w_{opt}$ , és òptim en el sentit de que minimitza l'error quadràtic mig de l'estimador, constituint el que es coneix per la solució de Wiener (Eq.2.21) que depèn de la inversa de la matriu d'autocorrelació de les funcions  $z(x,n)$  i la correlació creuada entre dites funcions i el senyal de referència.

$$w_{opt} = E[z(x,n) \cdot z^H(x,n)]^{-1} \cdot E[z(x,n) \cdot d^*(n)] = R_{zz}^{-1} \cdot r_{zd} \quad (2.21)$$

És important notar que la premissa de que  $R_{zz}$  sigui definida positiva assegura la seva no singularitat i, en conseqüència, l'existència de la solució de Wiener.

Posteriorment, Kalman va generalitzar aquesta solució pel cas en que el senyal  $x(n)$ , a partir del qual es conforma l'estimador, no fos estacionari [And79]. De tota manera, nosaltres ens centrarem en situacions estacionàries o bé quasi-estacionàries. Per ambdues situacions existeixen algorismes adaptatius [Hay91] en els que el vector de pesos és modificat a cada iteració de manera que convergeixi cap a la solució òptima. En cas d'un escenari quasi-estacionari, el disseny adaptatiu serà vàlid sempre i quan la velocitat de convergència d'aquests algorismes sigui suficientment alta com per seguir les variacions lentes de  $w_{opt}$ .

Respecte de la solució de Wiener per blocs (Eq.2.21) (generalment anomenada *off-line*), el càlcul de  $w_{opt}$  implica la inversió de la matriu de correlació de les funcions que consisteix en una matriu quadrada  $N_c \times N_c$ , essent  $N_c$  el nombre de coeficients del model. En el cas de que la solució existeixi, és a dir, que la matriu  $R_{zz}$  no sigui singular, el cost que comporta la inversió d'aquesta matriu per mètodes directes és de l'ordre de  $O(N_c^3)$  operacions (divisions+multiplicacions), però sota determinades condicions es pot reduir a  $O(N_c^2)$  ordenant el vector de funcions (Eq.2.15) de manera que la matriu d'autocorrelació sigui Toeplitz per

blocs [Mat88]. Donat que el model de Volterra( $P,N$ ) inclou un nombre de coeficients  $O(N^P)$ , la complexitat del càlcul de la solució de Wiener seria de  $O(N^{2P})$ , tot això en el supòsit de que la matriu no sigui singular. Per aquesta raó, el ritme de creixement del nombre de coeficients és un factor decisiu en els models no lineals, essent molt valuosa qualsevol reducció del mateix.

En un intent d'evitar problemes amb la inversió de la matriu de correlació i disminuir el cost computacional, Wiener va proposar l'ús de funcions  $z(\mathbf{x},n)$  que formessin un conjunt complet i que, a més a més, fossin ortogonals entre elles (Eq.2.22).

$$E[z_l(\mathbf{x},n) \cdot z_m^*(\mathbf{x},n)] = \begin{cases} \neq 0 & m = l \\ 0 & m \neq l \end{cases} \quad (2.22)$$

D'aquesta manera, l'obtenció de cadascun dels paràmetres de l'estimador és independent dels altres tal i com es mostra a continuació (Eq.2.23) ja que, si es compleix la condició (Eq.2.22), la matriu d'autocorrelació esdevé una matriu diagonal.

$$w_{l,\text{òptim}} = \frac{E[z_l(\mathbf{x},n) \cdot d^*(n)]}{E[|z_l(\mathbf{x},n)|^2]} \quad ; \forall l = 1, \dots, N_c \quad (2.23)$$

El model de Volterra( $P,N$ ) (Eq.2.12) es pot interpretar com un model lineal en termes dels elements dels seus nuclis, permetent una representació tal i com la proposada anteriorment (Eq.2.17). Així, el vector de funcions  $z(\mathbf{x},n)$  estaria format per tots els monomis multivariables que es poden obtenir a partir del vector de dades  $\mathbf{x}(n)=[x(n) \ x(n-1) \ \dots \ x(n-N+1)]'$  fins a un ordre de  $P$  (Eq.2.15) i el vector de pesos correspondria als coeficients dels nuclis de Volterra associats a cadascun d'aquests monomis (Eq.2.16). La correlació, en general no nul·la, entre els diferents monomis fa que els coeficients del model de Volterra no es puguin determinar per separat un dels altres, obligant a una inversió de la matriu de correlació que serà costosa en termes computacionals i que, en general, comporta problemes numèrics tant en un disseny per blocs com en un disseny adaptatiu (Apt.2.4).

El model de Wiener per a SNLs sorgeix com a conseqüència de les limitacions del model de Volterra en aquestes qüestions de disseny, i es basa en el teorema de Stone-Weierstrass que prova la complitud del conjunt format pels operadors de Volterra  $\{H_p[x(n)]\}$  (Eq.2.11). El treball de Wiener en el camp de processament no lineal del senyal va consistir, bàsicament, en idear un nou model de SNLs en termes d'uns operadors no homogenis, és a dir, que depenen de més d'un nucli de Volterra, anomenats *G-funcionals*. Aquests operadors es construeixen a partir dels operadors de Volterra de manera que siguin ortonormals entre ells, tot permetent la determinació de cadascun dels paràmetres del model per separat. Com a conseqüència de que els operadors de Volterra depenen del senyal d'entrada, els

G-funcionals només són ortogonals quan el senyal d'entrada  $x(n)$  és blanc i gaussià; d'aquí que Wiener utilitzés la lletra "G" per a anomenar-los.

## 2.3 El model de Wiener per a sistemes no lineals

A l'apartat anterior s'ha vist com el criteri d'estimació proposat per Wiener és especialment "atractiu" en disposar d'un conjunt complet de funcions ortogonals (Eq.2.22) a partir de les quals construir l'estimador (Eq.2.17), ja que llavors el càlcul de cadascun dels coeficients del model es pot efectuar per separat (Eq.2.23). El treball de Wiener en el camp de processament no lineal [Wie57] es pot interpretar com el desenvolupament i verificació pràctica (aplicat concretament al camp de la bioenginyeria) d'un model de SNLs invariants que aprofita aquestes propietats.

### 2.3.1 Els G-funcionals: una base ortonormal de sèries funcionals.

El problema de determinació dels nuclis  $\{h_p(m_1, m_2, \dots, m_p) \mid p > 1\}$  plantejat en el model de Volterra (Eq.2.12) va motivar a Wiener a desenvolupar un model de SNLs amb memòria,

$$y(n) = \sum_{p=0}^{\infty} G_p[k_p; x(n)] \quad ; k_p = [k_{0(p)} \ k_{1(p)} \ \dots \ k_{p(p)}]' \quad (2.24)$$

com a suma dels anomenats G-funcionals,  $\{G_p[k_p; x(n)] \mid \forall p=0, \dots, \infty\}$ , que són uns operadors no homogenis que depenen tant del senyal d'entrada,  $x(n)$  (que es suposa real), com d'un conjunt de funcions també reals que pel G-funcional d'ordre  $p$  s'agrupen en el vector  $k_p$ .

$$G_p[k_p; x(n)] = k_{0(p)} + \sum_{m=1}^p \left\{ \sum_{i_1=-\infty}^{+\infty} \dots \sum_{i_m=-\infty}^{+\infty} k_{m(p)}(i_1, \dots, i_m) \cdot x(n-i_1) \dots x(n-i_m) \right\} \quad (2.25)$$

A part de no ser homogenis, la particularitat que diferencia els G-funcionals dels operadors de Volterra (Eq.2.11) és que constitueixen un conjunt d'operadors ortonormals en la norma definida a continuació.

$$E\{G_p[k_p; x(n)] \cdot G_q[k_q; x(n)]\} = \begin{cases} 0 & p \neq q \\ 1 & p = q \end{cases} \quad (2.26)$$

Amparant-se en el teorema de Stone-Weierstrass que recolza la complitud dels operadors homogenis de Volterra, Wiener va calcular els G-funcionals mitjançant el procés

d'ortogonalització de Gram-Schmidt aplicat sobre aquests operadors, entenent el producte escalar com la correlació creuada de dos processos (Eq.2.26). Amb aquest objectiu, el G-funcional d'ordre  $p$ ,  $G_p[k_p; x(n)]$ , es defineix com un funcional no homogeni (Eq.2.25) que ha de ser ortogonal a qualsevol operador homogeni d'ordre menor (Eq.2.27.a). Un cop s'ha trobat un funcional que compleixi les condicions anteriors, es normalitza de manera que la seva norma sigui la unitat (Eq.2.27.b).

$$E\{G_p[k_p; x(n)] \cdot H_q[h_q; x(n)]\} = 0 \quad \forall 0 \leq q < p; \forall h_q(m_1, \dots, m_q) \quad (2.27.a)$$

$$E\{[G_p[k_p; x(n)]]^2\} = 1 \quad (2.27.b)$$

Iterant el procés des de  $p=0$  fins a l'ordre desitjat del model (a la pràctica el model de Wiener (Eq.2.24) es trunca fins a un ordre finit  $P$ ), i sota la hipòtesi de que el senyal d'entrada és gaussià i blanc, s'arriba a la següent expressió tancada pels G-funcionals [Sch89].

$$G_p[k_p; x(n)] = \sum_{m=1}^{\lfloor \frac{p}{2} \rfloor} \frac{(-1)^m p! \sigma^{2m}}{(p-2m)! m! 2^m} \cdot f[k_{p(p)}; x(n); m] \quad (2.28)$$

Cal assenyalar que el límit superior del sumatori,  $\lfloor p/2 \rfloor$ , representa l'enter després d'arrodonir per defecte la divisió  $p/2$ , i que la funció  $f(\cdot)$  està definida de la següent manera.

$$f[k_{p(p)}; x(n); m] = \sum_{i_1=-\infty}^{+\infty} \dots \sum_{i_{p-2m}=-\infty}^{+\infty} \sum_{s_1=-\infty}^{+\infty} \dots \sum_{s_m=-\infty}^{+\infty} k_p(i_1, \dots, i_{p-2m}, s_1, \dots, s_m) \cdot x(n-i_1) \dots x(n-i_{p-2m}) \quad (2.29)$$

D'aquest resultat cal destacar dues qüestions. En primer lloc, s'observa que el G-funcional d'ordre  $p$  només depèn de la funció  $k_{p(p)}$  que es coneix pel nom de *nucli de Wiener d'ordre  $p$*  i es defineix de nou com  $k_p(i_1, i_2, \dots, i_p)$ . Per altra banda, els coeficients que acompanyen a cada terme del sumatori (Eq.2.28) es poden identificar com els coeficients dels polinomis normalitzats d'Hermite [Sch89]. Aquest darrer resultat està íntimament relacionat amb el fet de que el que s'està ortogonalitzant amb el mètode de Gram-Schmidt són funcions polinòmiques de variables gaussianes [Sch81].

Abans de finalitzar aquest apartat caldria deixar clar que la definició dels G-funcionals com un polinomi d'Hermite normalitzat (Eq.2.28) però aplicat sobre la funció  $f(\cdot)$  (Eq.2.29), força a que G-funcionals de diferent ordre però calculats a partir de la mateixa entrada  $x(n)$  siguin ortonormals (Eq.2.26). Cal destacar però, que aquesta propietat d'ortonormalitat es mantindrà sempre i quan l'entrada sigui gaussiana, i és en aquesta peculiaritat que ja es va

perfilant la idea de que el model de Wiener està condicionat per l'estadística del senyal d'entrada.

A part d'això, i abans d'entrar en detall en el model de Wiener, és interessant comentar que hi ha diferents mètodes que permeten identificar els nuclis de Wiener. D'entre ells, el més conegut és el proposat per Lee-Schetzen [Lee65, Sch89] que es basa en la correlació creuada entre el G-funcional d'ordre  $p$  i el monomi multivariable de també ordre  $p$  del senyal d'entrada. De tota manera, aquest mètode presenta problemes per identificar el nucli  $k_p(i_1, i_2, \dots, i_p)$  sempre que dos o més índexs qualsevols coincideixin.

### 2.3.2 Descomposició en funcions ortonormals del nucli de Wiener

Un cop s'ha arribat a aquest punt, el que cal és determinar el nucli de Wiener del que depenen els G-funcionals. Wiener va proposar una descomposició de  $k_p(i_1, i_2, \dots, i_p)$  en termes d'un conjunt de funcions ortonormals que denotarem per  $\{\phi_i(n) \forall i=1, \dots, \infty\}$ . De fet, ja es comprovarà més endavant que el que es pretén aconseguir és la incorrelació entre parells de sortides obtingudes de convolucionar l'entrada  $x(n)$  amb les funcions  $\{\phi_i(n) \forall i=1, \dots, \infty\}$ ,

$$E\left\{[x(n) * \phi_i(n)] \cdot [x(n) * \phi_j(n)]\right\} = \int_{-1/2}^{1/2} \Phi_i(e^{j2\pi f}) \Phi_j^*(e^{-j2\pi f}) S_{xx}(e^{j2\pi f}) df \equiv A \delta_{ij} \quad (2.30)$$

essent  $\Phi_i(e^{j2\pi f})$  la transformada de Fourier de la funció  $\phi_i(n)$ ,  $S_{xx}(e^{j2\pi f})$  la densitat espectral de potència del procés d'entrada  $x(n)$  i  $A$  una constant qualsevol. Si aquest procés és blanc, tal i com suposava Wiener, la densitat espectral de potència és constant i la incorrelació entre senyals de l'expressió (Eq.2.30) es tradueix en ortogonalitat entre funcions (Eq.2.31).

$$\sum_{n=0}^{+\infty} \phi_i(n) \cdot \phi_j(n) = \int_{-1/2}^{1/2} \Phi_i(e^{j2\pi f}) \cdot \Phi_j^*(e^{-j2\pi f}) df = \delta_{ij} \equiv \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \quad (2.31)$$

Wiener va desenvolupar el seu model escollint aquestes funcions ortogonals com les funcions de Laguerre, indicades per  $\{l_i(n) \forall i=0, \dots, \infty\}$  (Eq.2.32). Això és com a conseqüència de que es poden implementar fàcilment amb circuits elèctrics en cascada.

*Funció de Laguerre*

$$\phi_i(n) \equiv l_i(n) \xrightarrow{\text{Transformada } Z} L_i(z) = L_0(z) \cdot [B(z)]^i \quad 0 \leq a < 1 \quad (2.32.a)$$

$$L_0(z) = \frac{\sqrt{1-a^2}}{1-az^{-1}} \quad ; \quad B(z) = \frac{z^{-1}-a}{1-az^{-1}} \quad (2.32.b)$$

De totes maneres, qualsevol altre conjunt de funcions ortogonals seria igualment vàlid pel desenvolupament que presentarem [Rugh81]. Una possibilitat força interessant, tot i que no deixa de ser un cas particular de les funcions de Laguerre ( $a=0$ ), és el conjunt ortogonal format per les funcions originades a partir d'una línia de retards (Eq.2.33). De fet, en aquest conjunt està basat el mètode d'identificació dels nuclis de Wiener citat anteriorment [Lee65,Sch89].

$$\phi_i(n) = \delta(n-i) \xrightarrow{\text{Transformada } Z} \Phi_i(z) = z^{-i} \quad (2.33)$$

Un cop escollit el conjunt de funcions ortogonals i després d'èsser normalitzades, Wiener va descomposar el nucli de Wiener d'ordre  $p$  com a combinació lineal de totes les possibles formes de multiplicar  $p$  funcions de Laguerre.

$$k_p(i_1, i_2, \dots, i_p) = \sum_{m_1=0}^{+\infty} \sum_{m_2=0}^{+\infty} \dots \sum_{m_p=0}^{+\infty} c_{m_1, m_2, \dots, m_p} \cdot l_{m_1}(i_1) l_{m_2}(i_2) \dots l_{m_p}(i_p) \quad (2.34)$$

Com a conseqüència de l'ortonormalitat de les funcions, cada coeficient es pot expressar tal i com segueix.

$$c_{m_1, m_2, \dots, m_p} = \sum_{i_1=-\infty}^{+\infty} \sum_{i_2=-\infty}^{+\infty} \dots \sum_{i_p=-\infty}^{+\infty} k_p(i_1, i_2, \dots, i_p) \cdot l_{m_1}(i_1) l_{m_2}(i_2) \dots l_{m_p}(i_p) \quad (2.35)$$

### 2.3.3 El treball originari de Wiener

Amb aquesta descomposició del nucli de Wiener d'ordre  $p$  en funcions ortonormals i a partir de la sèrie de Wiener truncada fins a un ordre finit  $P$ , es pot interpretar el model de Wiener com un estimador de la relació entrada/sortida d'un SNL que és lineal amb els paràmetres.

$$y(n) \equiv \hat{y}(n) = \sum_{p=0}^P G_p[k_p; x(n)] \quad (2.36)$$

En efecte, si es substitueix l'equació (Eq.2.34) en el model (Eq.2.36) es pot comprovar que arribem a una expressió que combina linealment els coeficients  $c_{m_1, m_2, \dots, m_p}$  amb els G-funcionals d'ordre  $p$  aplicats sobre el producte de funcions de Laguerre associat a cada coeficient.

$$\hat{y}(n) = \hat{g}[x(n)] = \sum_{p=0}^P \left\{ \sum_{m_1=0}^{+\infty} \sum_{m_2=0}^{+\infty} \dots \sum_{m_p=0}^{+\infty} c_{m_1, m_2, \dots, m_p} \cdot G_p[l_{m_1} \cdot l_{m_2} \dots l_{m_p}; x(n)] \right\} \quad (2.37)$$

Ara per ara, el model de Wiener està expressat com un estimador que depèn linealment dels coeficients  $c_{m_1, m_2, \dots, m_p}$  però, a part d'això, ens interessaria que el conjunt de les funcions  $G_p[l_{m_1} l_{m_2} \dots l_{m_p}; x(n)]$  que es combinen amb aquests coeficients fos ortogonal. A l'apartat anterior ja s'ha raonat que els G-funcionals de diferent ordre són ortogonals però, i entre G-funcionals del mateix ordre? La demostració de que dos G-funcionals del mateix ordre, aplicats sobre una combinació diferent de funcions de Laguerre, són ortogonals es discuteix al següent punt però no és immediata (Eq.2.38). Els detalls de la demostració els obviarem, referint-nos a [Sch89] com a complement.

$$E \left\{ G_p[l_{m_1} l_{m_2} \dots l_{m_p}; x(n)] \cdot G_p[l_{s_1} l_{s_2} \dots l_{s_p}; x(n)] \right\} = 0 \quad \begin{array}{l} \text{si la combinació } (m_1, m_2, \dots, m_p) \\ \text{és diferent de la } (s_1, s_2, \dots, s_p) \end{array} \quad (2.38)$$

En resum, les propietats d'ortogonalitat tant entre G-funcionals de diferent ordre (Eq.2.26) com entre G-funcionals del mateix ordre però aplicats sobre un conjunt de funcions de Laguerre diferents (Eq.2.38), fa que els paràmetres del model de Wiener en un problema d'identificació es puguin mesurar per separat un dels altres tal i com segueix,

$$c_{m_1, \dots, m_p} = \frac{E \left\{ y(n) \cdot G_p[l_{m_1} \dots l_{m_p}; x(n)] \right\}}{p! \eta^p} \quad ; \forall m_i = 0, \dots, +\infty; \forall i = 1, \dots, p; \forall p = 0, \dots, P \quad (2.39)$$

on  $y(n)$  és la sortida del SNL real que es pretèn identificar i  $\eta$  és la densitat espectral de potència del procés blanc de l'entrada.

Aquesta és, en definitiva, la forma en que Wiener proposa determinar els paràmetres del seu model, solventant així el problema de disseny dels nuclis de Volterra. Per tal de donar una visió de conjunt, seria interessant interpretar la proposta de Wiener (Eq.2.39) com un cas particular de la solució de MMSE (Eq.2.21) en la que la inversió de la matriu d'autocorrelació de les funcions de l'estimador s'ha evitat gràcies a que aquestes funcions són ortogonals entre elles (Eq.2.23).

### 2.3.4 Arquitectura del model de Wiener

En aquest apartat es presenta un esquema de la implementació pràctica del model de Wiener d'ordre  $P$  (Eq.2.40). Donat que es tracta de la suma de  $(P+1)$  G-funcionals, es començarà deduint l'arquitectura per un G-funcional d'ordre genèric  $p$ .



$$\hat{y}(n) = \sum_{p=0}^P G_p[k_p; x(n)] \quad (2.40)$$

A partir de la descomposició del nucli de Wiener,  $k_p(i_1, i_2, \dots, i_p)$ , en termes de les funcions de Laguerre, ja s'ha vist que el G-funcional es pot escriure com la següent combinació lineal.

$$\begin{aligned} G_p[k_p; x(n)] &= G_p \left[ \sum_{m_1=0}^{+\infty} \dots \sum_{m_p=0}^{+\infty} c_{m_1, \dots, m_p} \cdot l_{m_1} \dots l_{m_p}; x(n) \right] = \\ &= \sum_{m_1=0}^{+\infty} \dots \sum_{m_p=0}^{+\infty} c_{m_1, \dots, m_p} \cdot G_p[l_{m_1} \dots l_{m_p}; x(n)] \end{aligned} \quad (2.41)$$

Cada sumand inclou un G-funcional d'ordre  $p$  que s'aplica al producte de  $p$  funcions de Laguerre. Cadascun d'aquests G-funcionals es pot expressar tal i com segueix [Sch89].

$$G_p[l_{m_1} \dots l_{m_p}; x(n)] = G_{\sum_{j=1}^{NJ} s_j} \left[ \prod_{j=1}^{NJ} (l_{q_j})^{s_j}; x(n) \right] \stackrel{\text{Propietat}}{\equiv} \prod_{j=1}^{NJ} G_{s_j} \left[ (l_{q_j})^{s_j}; x(n) \right] \quad (2.42)$$

L'enter  $NJ$  representa el nombre de funcions de Laguerre diferents que estan implicades en el producte  $(l_{m_1} l_{m_2} \dots l_{m_p})$ . Així,  $s_j$  és el nombre de cops que la funció de Laguerre  $l_{q_j}$  apareix en l'argument del G-funcional, i per tant la suma de tots els  $s_j$  des de  $j=1$  fins a  $j=NJ$  equival a l'ordre del G-funcional  $p$ . Notar també que els enters  $q_j \forall j=1, \dots, NJ$  són diferents per valors diferents de  $j$  ja que s'han agrupat les funcions de Laguerre del mateix ordre. Proposem un exemple que il·lustri la igualtat anterior.

**Exemple 2.1.** Exemple de la propietat de l'expressió (Eq.2.42)

El G-funcional d'ordre 6 aplicat sobre el producte de funcions de Laguerre  $l_0.l_1.l_3.l_0.l_2.l_2$  és equivalent a

$$\begin{aligned} G_6[l_0 l_1 l_3 l_0 l_2 l_2; x(n)] &= G_{2+1+2+1} \left[ (l_0)^2 l_1 (l_2)^2 l_3; x(n) \right] = \\ &= G_2 \left[ (l_0)^2; x(n) \right] G_1[l_1; x(n)] G_2 \left[ (l_2)^2; x(n) \right] G_1 \left[ (l_3)^2; x(n) \right] \end{aligned}$$

♣

En definitiva, aquesta propietat ens permet expressar cada G-funcional d'ordre  $p$  aplicat sobre el producte de funcions ortogonals  $[(l_{q_1})^{s_1} \dots (l_{q_{NJ}})^{s_{NJ}}]$ , com el producte de  $NJ$  G-funcionals, cadascun d'ordre  $s_j$ , aplicats sobre cadascuna de les funcions  $(l_{q_j})^{s_j}$  per separat. Per

tant, tenint en compte aquesta igualtat dels G-funcionals (Eq.2.42) i substituint-la a (Eq.2.41), s'obté la següent expressió pel G-funcional d'ordre  $p$ .

$$G_p[k_p; x(n)] = \sum_{m_1=0}^{+\infty} \dots \sum_{m_p=0}^{+\infty} c_{m_1, \dots, m_p} \cdot \prod_{j=1}^{NJ} G_{s_j}[(l_{q_j})^{s_j}; x(n)] \quad (2.43)$$

És important adonar-se de que cada sumant de l'expressió (Eq.2.43) inclou un producte de G-funcionals d'ordre menor o igual a  $p$  (de manera que entre tots sumen  $p$ ), i que cadascun d'ells està aplicat sobre una funció de Laguerre diferent (elevada al mateix ordre del G-funcional en el que es troba). Arribats a aquest punt, només ens cal saber com implementar cadascun d'aquests G-funcionals, representats genèricament per  $G_s[(l_i)^s; x(n)]$ .

Recordem que el G-funcional d'ordre  $s$  es pot expressar en termes dels coeficients del polinomi d'Hermite normalitzat d'ordre  $s$  (Eq.2.28), combinats linealment amb el conjunt de funcions  $\{f((l_i)^s; x; q), \forall q=1, \dots, [s/2]\}$ .

$$G_s[(l_i)^s; x(n)] = \sum_{q=1}^{\lfloor \frac{s}{2} \rfloor} \frac{(-1)^q s! \eta^q}{(s-2q)! q! 2^q} \cdot f((l_i)^s; x; q) \quad (2.44)$$

Tenint en compte la propietat d'ortogonalitat de les funcions de Laguerre (Eq.2.31) i la definició de la funció  $f((l_i)^s; x; q)$  (Eq.2.29), es pot comprovar que aquesta funció depèn de la convolució de l'entrada  $x(n)$  amb la funció de Laguerre  $l_i(n)$ .

$$f((l_i)^s; x; q) = \left( \sum_{m=0}^{+\infty} l_i(m) \cdot x(n-m) \right)^{s-2q} \equiv [u_i(n)]^{s-2q} \quad (2.45)$$

Es pot concloure que el G-funcional d'ordre  $s$ ,  $G_s[(l_i)^s; x(n)]$ , es pot expressar com un polinomi normalitzat d'Hermite d'ordre  $s$ , denotat per  $PH_s[.]$ , però aplicat sobre potències del senyal  $u_i(n)$  (Eq.2.46).

$$G_s[(l_i)^s; x(n)] = \sum_{q=1}^{\lfloor \frac{s}{2} \rfloor} \frac{(-1)^q s! \eta^q}{(s-2q)! q! 2^q} \cdot [u_i(n)]^{s-2q} \equiv PH_s[u_i(n)] \quad (2.46)$$

Amb aquesta darrera expressió en termes dels polinomis d'Hermite i substituint-la a (Eq.2.43), s'arriba a que el G-funcional d'ordre  $p$  es pot implementar amb una arquitectura com la que segueix.

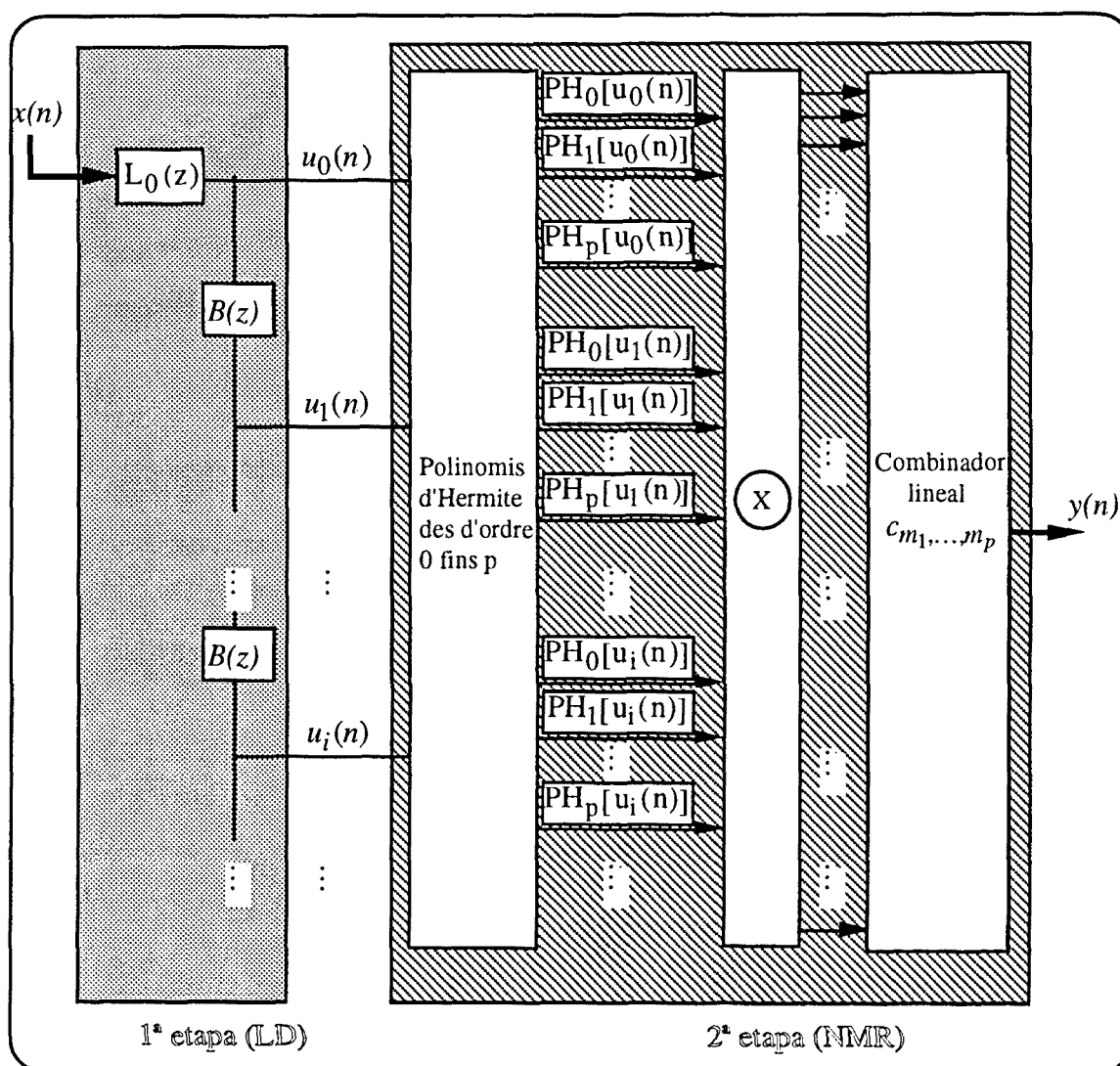


Figura 2.3 Implementació del G-funcional d'ordre  $p$  del model de Wiener

La primera etapa implementa la convolució de l'entrada amb cadascuna de les funcions de Laguerre, generant el vector de pesos  $u(n)$  que conté els senyals  $u_i(n)$ . Tot i que a la figura (Fig.2.3) s'ha considerat un nombre infinit d'aquestes funcions per congruència amb la notació utilitzada fins ara, el nombre de funcions de Laguerre que s'implementarien en un cas real hauria de ser finit,  $\{l_i(n) \ i=0, \dots, N-1\}$ , essent llavors el vector  $u(n)$  també de dimensió finita. A part d'això, és interessant adonar-se de que aquesta primera part representa una transformació lineal del senyal i que, a més a més, conté tota la dinàmica del model. És per aquesta raó que es pot interpretar com la part LD del model més general que s'havia presentat anteriorment (Fig.2.1).

La segona etapa es subdivideix alhora en tres parts. A partir de cada senyal  $u_i(n)$ , la primera part genera els polinomis d'Hermite des d'ordre 0 fins el d'ordre  $p$  (Eq.2.46). A

continuació, un mòdul multiplicador s'encarrega d'obtenir totes les possibles combinacions de polinomis d'Hermite de diferents  $u_i(n)$ , de manera que la suma dels ordres sigui  $p$  (Eq.2.42). En darrer lloc, es situa un bloc que combina linealment cadascun d'aquests productes de polinomis d'Hermite juntament amb els pesos corresponents  $c_{m1, \dots, mp}$  (Eq.2.43). En conjunt, aquesta segona etapa es pot identificar amb el mòdul NMR (Fig.2.1) ja que inclou transformacions no lineals i estàtiques, i una posterior combinació lineal amb un conjunt de coeficients.

Un cop ja s'ha implementat el G-funcional d'ordre  $p$ , és possible presentar el model de Wiener seguint les mateixes pautes que a la figura anterior (Fig.2.3).

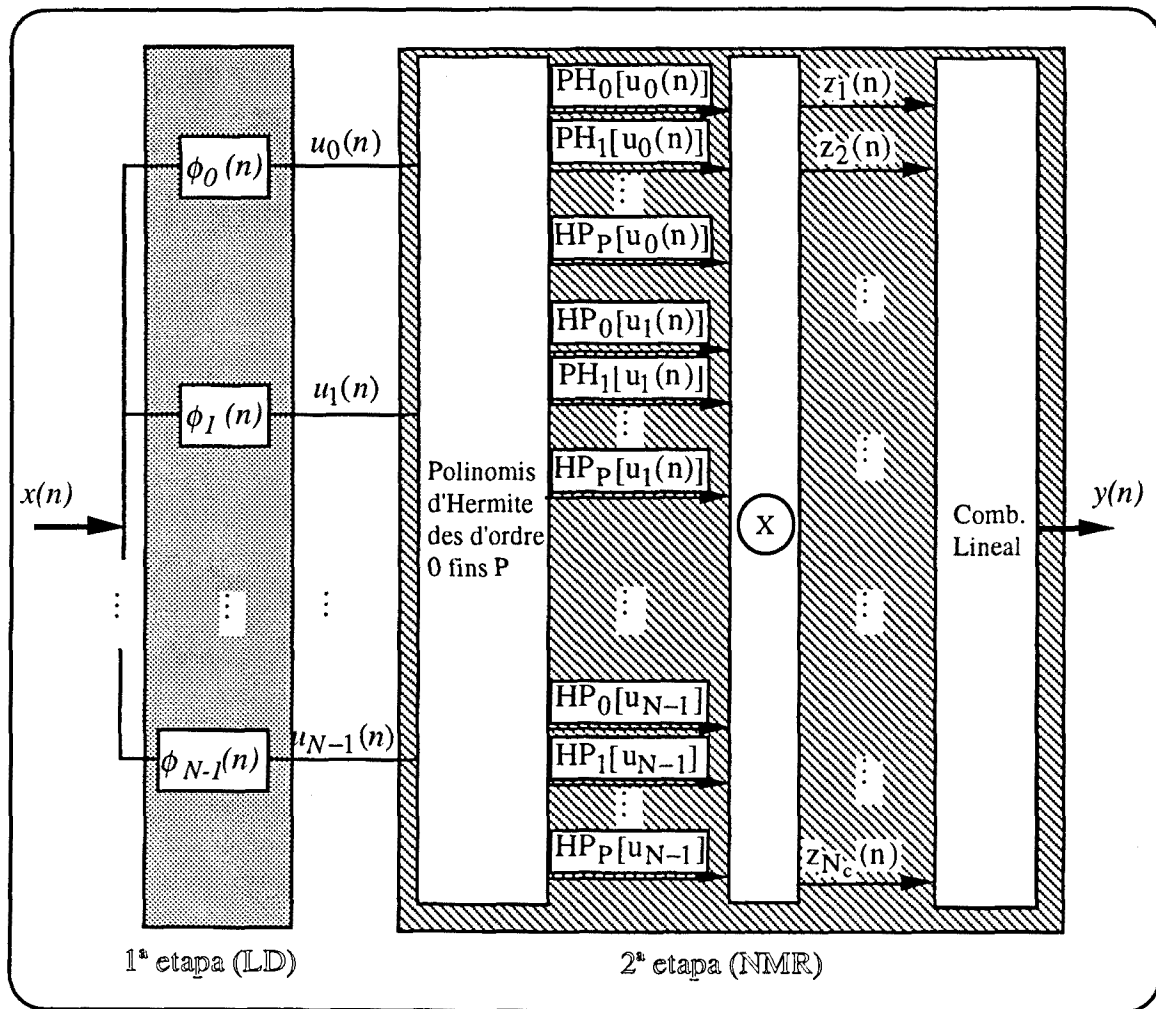


Figura 2.4 Model general de Wiener( $P,N$ ) de SNLs invariants amb memòria.

Per tal de poder-se realitzar, el model de Wiener s'ha de truncar en dos sentits. En primer lloc, es suposa que l'ordre dels G-funcionals arriba fins a  $P$  (Eq.2.40) i, per altra banda, també s'admet que només  $N$  funcions de Laguerre intervenen en el desenvolupament dels nuclis

de Wiener. De tota manera, ja s'ha comentat abans que el desenvolupament presentat en termes de les funcions de Laguerre es podia haver fet utilitzant qualsevol grup complet de funcions ortogonals (Eq.2.31). És en aquest sentit que es presenta el model més general de Wiener en termes de  $N$  funcions ortogonals qualsevols  $\{\phi_i(n) \forall i=0, \dots, N-1\}$ .

Consisteix també en una arquitectura de dues etapes, força semblant a la de la figura (Fig.2.3). Cal destacar però, que la primera etapa (LD) ha perdut l'estructura en cascada, pròpia de les funcions de Laguerre, en utilitzar un conjunt de funcions ortogonals diferent.

En segon lloc es troba l'etapa (NMR). Donat que el model de Wiener inclou els G-funcionals des d'ordre 0 fins a  $P$ , els polinomis d'Hermite que s'han de calcular a partir de cada  $u_i(n)$ , van des del d'ordre 0 fins al d'ordre  $P$ . Per altra banda, el bloc multiplicador de la figura (Fig.2.4) ha de proporcionar tots els productes de polinomis d'Hermite de diferents  $u_i(n)$ , de manera que la suma dels ordres vagi des de 0 fins a  $P$ . El resultat són un grup de  $N_c$  funcions, indicades per  $\{z_j(n) \forall j=1, \dots, N_c\}$ , que es combinen linealment amb els coeficients  $\{c_{m1, \dots, mp} \forall m_j=0, \dots, N-1, \forall p=0, \dots, P\}$ , que seguint la notació de l'apartat (Ap.2.2) s'agruparien en el vector  $w$ .

És interessant remarcar que, tot i que el model de Wiener soluciona el problema d'identificació dels nuclis que patia el model de Volterra, no ha disminuït gens la càrrega computacional. En efecte, el nombre de coeficients que inclou el model de Wiener( $P, N$ ) continua essent el mateix que el del model de Volterra( $P, N$ ), és a dir, de l'ordre de  $O(N^P)$  quan  $P < N$ . De totes maneres, hi ha treballs [Fej94b] que semblen recolzar la idea de que el model de Wiener( $P, N$ ) utilitzant funcions de Laguerre i amb un disseny adequat de la variable  $a$  (Eq.2.32), pot presentar una reducció de la memòria  $N$ , assolint les mateixes prestacions que un model de Volterra( $P, N$ ). Això es tradueix doncs en una reducció considerable del nombre de coeficients. El problema està en que el valor òptim de  $a$  serà específic a cada problema i tampoc es coneix un mètode sistemàtic que el proporioni.

Wiener va ser el primer en treballar sobre un model de SNLs que es pogués descomposar en una part que fos lineal i que contingués tota la dinàmica del sistema ( etapa LD), i una segona part no lineal i estàtica (generador de polinomis d'Hermite i multiplicador), que combina linealment les funcions  $\{z_j(k) \forall j=1, \dots, N_c\}$  amb els coeficients o paràmetres del model (part NMR). És per això que els models que segueixen les pautes de la figura LD/NMR (Fig.2.1) es califiquen de model de Wiener, tot i que el treball originari de Wiener correspon al desenvolupament presentat al llarg de tota aquesta secció (Fig.2.4, Ap.2.3).

### 2.3.5 Validació del model de Wiener

Wiener va utilitzar la sèrie de Volterra per crear un model de SNLs que superés els problemes de convergència i determinació del model de Volterra. La qüestió de la convergència ja s'ha comentat a l'inici de l'apartat que queda justificada amb el teorema de Stone-Weierstrass. Per altra banda, la determinació dels nuclis es supera entenent el model de Wiener com un estimador lineal en termes dels seus paràmetres  $\{c_{m_1, \dots, m_p} \forall m_i=0, \dots, N-1 \forall i=1, \dots, p \forall p=0, \dots, P\}$  que es combinen linealment amb un conjunt de funcions ortogonals  $\{z_j(n) \forall j=1, \dots, N_c\}$ . Això permet que, sota un criteri de disseny tal com la minimització de l'error quadràtic mig, el càlcul de cada coeficient es pugui efectuar independentment dels altres (Eq.2.39).

De tota manera, el model de Wiener presenta uns problemes que fan que no sigui un model de SNLs massa extès. El primer de tot és la complexitat del model, no tan sols en nombre de coeficients (que continua essent del mateix ordre de magnitud que el model de Volterra, és a dir, creixement exponencial), sino que la pròpia arquitectura de generació del model (Fig.2.4) és considerablement feixuga. El problema més important però que presenta el model de Wiener és que la seva principal característica (l'ortogonalitat entre les funcions  $z_j(n)$ ) es veu alterada quan l'entrada deixa d'ésser un procés gaussià i blanc.

La necessitat de que el procés d'entrada sigui blanc ve donada per la necessitat de garantir que les funcions  $\{u_i(n) \forall i=0, \dots, N-1\}$ , que s'obtenen de convolucionar l'entrada amb la corresponent funció  $\phi_i(n)$ , siguin incorrelades (Eq.2.30) sempre i quan el conjunt de funcions  $\{\phi_i(n) \forall i=0, \dots, N-1\}$  siguin ortogonals (Eq.2.31). A més a més, per ser  $x(n)$  un procés gaussià i les transformacions de la primera etapa lineals, els senyals  $\{u_i(n) \forall i=0, \dots, N-1\}$  són també gaussians i la incorrelació entre ells es tradueix en independència estadística. Aquesta independència estadística i la característica gaussiana del conjunt  $\{u_m(n) \forall m=0, \dots, N-1\}$  fa que les funcions  $\{z_j(n) \forall j=1, \dots, N_c\}$  siguin ortogonals entre elles.

En efecte, ja s'ha vist anteriorment que cada funció  $z_j(n)$  està formada per un conjunt de polinomis d'Hermite, generats a partir de diferents  $u_i(n)$ , que es multipliquen entre ells. La correlació creuada de dues funcions  $z_j(n)$  serà nul·la per dues raons. En primer lloc, la correlació creuada entre dos polinomis d'Hermite de qualsevol ordre generats a partir de diferents entrades és nul·la perquè són independents (ja que  $u_i(n)$  són també independents) i de mitja zero. En segon lloc, la correlació creuada de dos polinomis d'Hermite de diferent ordre però generats a partir de la mateixa entrada  $u_i(n)$  gaussiana és nul·la per definició (de fet, els polinomis d'Hermite es calculen en voler generar un conjunt de polinomis que siguin ortogonals quan l'entrada és gaussiana).

En conseqüència, en el moment en que es perd o bé la propietat de gaussianitat de l'entrada o bé el fet de que sigui un procés blanc, les funcions del model de Wiener deixen de ser ortogonals. Per ambdós problemes hi ha solucions teòriques. Així, en cas de que el procés d'entrada no sigui blanc s'haurà de sotmetre el senyal d'entrada prèviament a un procés de blanqueig (de l'anglès “whitening”) [Sch89]. En el cas de que la part LD consisteixi en una cadena de retards  $\{\Phi_i(z)=z^{-i} \forall i=0,\dots,N-1\}$ , aquest procés de blanqueig es pot incorporar al mateix model mitjançant una estructura *lattice* [Hay91] que implementi aquesta primera etapa. Cal destacar en aquest sentit, que també existeix l'estructura *lattice* generalitzada a funcions de Laguerre, però que només és vàlida per senyals estacionaris [Fej94a]. Per altra banda, si el senyal d'entrada no és gaussià haurà de passar abans per un filtre no lineal que el transformi en gaussià [Sch89]. Aquesta solució no sembla gaire viable ja que, en primer lloc, dita transformació no sempre existeix i, fins i tot en el cas de que existeixi, en general no es coneix.

Es pot concloure que Wiener va idear un nou model que superava els problemes del model de Volterra però que alhora és molt rígid i poc robust en front de l'estadística del senyal d'entrada. De tota manera, Wiener va posar de manifest l'atractiu del model de Volterra pel fet de ser lineal amb els coeficients i poder-se dissenyar amb un criteri de MMSE. El desenvolupament dels mètodes adaptatius per aquest tipus d'estimadors va cridar l'atenció sobre el model de Volterra i, és en aquest camp on bàsicament es troben les seves aplicacions. D'alguna manera, els problemes d'identificació dels nuclis de Volterra es traslladen i s'intenten superar mitjançant algorismes adaptatius adients que condueixin el model de Volterra cap a la solució òptima de Wiener de MMSE. És per això que s'ha considerat oportú dedicar un darrer apartat d'aquest capítol al disseny adaptatiu, bàsicament del model de Volterra, presentant una visió potser superficial dels diferents algorismes però suficient per fer palesa la problemàtica d'aquests mètodes en ésser aplicats a models no lineals tal i com els presentats.

## 2.4 Disseny adaptatiu

Un cop fixat l'ordre del model  $P$  i la diversitat temporal del mateix  $N$ , tant el model de Volterra com el de Wiener són lineals amb els paràmetres a dissenyar. Aquest fet permet que la solució de mínim error quadràtic mig (MMSE)

$$w_{opt} = E[z(x,n) \cdot z^H(x,n)]^{-1} \cdot E[z(x,n) \cdot d^*(n)] \quad (2.47)$$

es pugui obtenir adaptativament d'una forma totalment paral·lela a la dels models lineals. En aquest apartat es discutiran els dos grans mètodes adaptatius: els mètodes de gradient (centrant-nos en l'algorisme “Least-Mean Square” (LMS) per ser el més àmpliament utilitzat) i els

mètodes recursius (bàsicament, l'algorisme "Recursive-Least Squares" (RLS)). El propòsit no és doncs una anàlisi exhaustiva dels mateixos, que per altra banda es pot trobar per exemple a [Hay91], sino que tan sols es pretén esmentar les particularitats i problemes que presenten aquests algorismes alhora d'ésser aplicats a l'entrenament de sistemes no lineals com els presentats [Mat91, Mat93].

Com a conseqüència de que el model de Volterra és el més estès i és també el que s'utilitzarà més endavant per comparar-lo amb el model que es proposa al capítol 3, farem més èmfasi en el disseny adaptatiu del model de Volterra que en el de Wiener. Respecte d'aquest darrer, caldria destacar el treball realitzat per Fejzo en la seva tesi doctoral [Fej94b] on agrupa sota una mateixa arquitectura LD/NMR (Fig.2.1) el model de Wiener i el de Volterra. Aquesta perspectiva li permet enfrontar-se als problemes del model de Wiener, que només és vàlid per senyals gaussianes i blanques, d'una forma semblant als problemes de Volterra (introducció de blanqueig als filtres de Laguerre, utilització d'algorismes eficients i adequats a l'estadística del senyal d'entrada).

Per altra banda, és important destacar que en l'àmplia bibliografia de disseny adaptatiu de sistemes de Volterra, es treballa generalment amb un filtre de segon ordre, és a dir,  $P=2$  (eq.2.12). La raó radica en el fet de que es dona per suposat que la millora que s'obténdria en l'estimació en utilitzar sistemes de Volterra d'ordre superior a 2, no compensaria la complexitat i cost computacional que s'hauria de pagar; en definitiva, el nombre de coeficients d'un model de Volterra( $P,N$ ) de l'ordre de  $O(N^P)$  fa inviable una implementació pràctica en aplicacions en temps real per valors de  $P$  elevats. Tot i que s'intentarà presentar una anàlisi del disseny adaptatiu d'un filtre de Volterra( $P,N$ ) general, es faran referències al cas de segon ordre Volterra( $2,N$ ) per tal de comparar entre la bibliografia existent.

### 2.4.1 Mètodes de gradient: l'algorisme LMS

Els mètodes de gradient es basen en que la funció de cost d'error quadràtic mig es pot expressar com una forma quadràtica en termes de l'error de coeficients (Eq.2.19, [Hay91]). Assumint que la matriu d'autocorrelació de l'espai de senyal,  $\mathbf{R}_{zz}$ , és definida positiva, la funció de cost té un comportament convex que assegura la unicitat de la solució en buscar un mínim d'aquesta funció.

$$\xi_{\mathbf{w}} \Big|_{MINIM} \approx \nabla_{\mathbf{w}} \mathbf{H} \xi_{\mathbf{w}} \equiv 0 \Rightarrow \mathbf{w}_{opt} = \mathbf{R}_{zz}^{-1} \cdot \mathbf{r}_{zd} \quad (2.48)$$

Sense entrar en detall, es pot generalitzar que els mètodes de gradient modifiquen els pesos o coeficients de l'estimador a cada iteració en un sentit oposat al que marca el gradient de la funció



de cost (Eq.2.49). D'aquesta manera es pretén que els nous pesos estiguin cada cop més a prop de la solució òptima.

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \frac{1}{2} \mu \cdot \nabla_{\mathbf{w}(n)^H} \xi_{\mathbf{w}}(n) \quad (2.49)$$

L'algorisme LMS (Tau.2.1) consisteix en substituir el gradient, que és un valor esperat, pel valor instantani o, el que és el mateix, per una estimació del mateix amb una sola mostra. L'estudi estadístic d'aquest algorisme està fora de les expectatives d'aquesta tesi, referint-nos a [Hay91] per detalls.

$\mu = \text{constant de pas} \Rightarrow 0 < \mu < \frac{2}{\text{Potència total d'entrada}}$ <p>Condicions inicials: <math>\mathbf{w}(0) = \mathbf{0}</math></p> <p>Algorisme: Per <math>n = 0, 1, 2, \dots</math></p> $e(n) = d(n) - \mathbf{w}^H(n) \cdot \mathbf{z}(x, n)$ $\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \cdot \mathbf{z}(x, n) \cdot e^*(n)$
---

**Taula 2.1** L'algorisme adaptatiu LMS

L'èxit del LMS radica en la seva explicació senzilla i intuïtiva, així com el baix cost computacional que representa envers les prestacions que se n'obtenen. En efecte, es tracta d'un algorisme que realitza de l'ordre de  $O(N_c)$  operacions a cada iteració, entenent per operacions les multiplicacions i divisions reals, i essent  $N_c$  el nombre de coeficients de l'estimador. Així, en el cas d'un sistema de Volterra(P,N) la complexitat de l'algorisme es tradueixi en  $O(N^P)$  operacions cada iteració.

Una característica important del LMS, i que esdevé encara més decisiva en els sistemes no lineals, és la dependència que pateix respecte de l'estadística del senyal d'entrada. És ben conegut que la velocitat de convergència del LMS depen de la dispersió d'autovalors de la matriu d'autocorrelació  $\mathbf{R}_{zz}$ , de manera que en la mesura en que aquesta sigui més gran, l'algorisme convergirà més lentament. Aquest problema es presenta, per exemple, en el disseny d'un model lineal FIR quan el senyal d'entrada no és un procés blanc. El problema està en que la correlació entre components del vector de dades fa que la matriu d'autocorrelació no sigui la identitat, augmentant en conseqüència la dispersió d'autovalors. En cas de dissenyar un model de Volterra, el problema s'agreuja ja que, en general, la correlació entre components del vector de dades  $\mathbf{z}(x, n)$  és elevada i origina una dispersió d'autovalors de la matriu  $\mathbf{R}_{zz}$  força gran. En aquest sentit, cal destacar el treball de [Koh85] en el que es planteja la identificació d'un sistema

amb un model de Volterra de segon ordre, essent el senyal d'entrada gaussià. Es pot comprovar que el mode més lent del filtre de Volterra té una constant de temps proporcional al quadrat de la dispersió d'autovalors de la matriu d'autocorrelació del procés d'entrada  $R_{xx}$ .

En resum, es pot dir que el LMS és un algorisme amb un cost computacional baix, proporcional al nombre de coeficients del filtre en qüestió, però que en contrapartida la velocitat de convergència també serà baixa com a conseqüència de que la dispersió d'autovalors de la matriu d'autocorrelació  $R_{zz}$  és elevada.

#### 2.4.1.1 L'algorisme LMS Normalitzat (NLMS)

Tal i com es pot comprovar a la taula (Tau.2.1), la constant de pas  $\mu$  està limitada al rang  $0 < \mu < 2 / (\text{Potència de } z)$  per tal d'assegurar la convergència en mitja dels coeficients de l'estimador. Una modificació del LMS consisteix en redefinir aquesta constant de pas en funció d'una nova  $\bar{\mu}$  i de la potència instantànea de  $z$  (Eq.2.44) donant lloc a l'anomenat LMS Normalitzat (NLMS).

$$\mu_{NLMS} = \frac{\bar{\mu}}{z^H(x,n) \cdot z(x,n)} \quad (2.50)$$

Darrerament han sorgit treballs tals com [Slo93] on s'analitza el comportament estadístic del NLMS mitjançant un model de senyal no gaire realista, però que sembla estar recolzat pel resultat de les simulacions. L'autor utilitza aquest model de senyal per justificar el comportament superior que mostra el NLMS en front del LMS en quant a velocitat i desajust final.

#### 2.4.1.2 L'algorisme NLMS amb $\bar{\mu}$ variable

És ben conegut el compromís en el LMS entre velocitat de convergència i desajust final de l'error quadràtic mig, i com aquest compromís és controlat per la constant de pas  $\mu$ . Així, una  $\mu$  elevada farà que l'algorisme convergeixi ràpidament però que, per altra banda, assoleixi una variança de l'estimació també elevada, i viceversa. Una solució a aquest problema seria utilitzar una  $\mu$  variable amb el temps de manera que a l'inici de l'algorisme fos gran (assegurant així una velocitat de convergència elevada), però que anés disminuint per tal d'assolir un desajust final petit.

Algunes implementacions pràctiques opten per aquesta solució escullint dos valors de  $\mu$ , un valor alt adequat per l'estat transitori i un altre de baix per l'estat estacionari. El principal problema és el criteri per decidir quan s'efectua la transició. De tota manera, la solució més

lògica sembla que hagi de ser la de dissenyar una constant de pas variable amb el temps i que en tot moment minimitzi l'error quadràtic mig. D'entre certs algorismes de gradient que opten per modificar la  $\mu$  mitjançant una regla més o menys heurística [Sic87], caldria destacar el treball de Mathews i Xie [Mat90] on la  $\mu$  s'actualitza mitjançant un algorisme de gradient LMS per disminuir l'error quadràtic mig, i, especialment, el treball de Slock [Slo93]. En aquest darrer es proposa un valor variable de la  $\bar{\mu}$  del NLMS tal i com segueix,

$$\bar{\mu}(n) = \bar{\mu}(n-1) \frac{1 - \frac{\bar{\mu}(n-1)}{N_c}}{1 - \frac{(\bar{\mu}(n-1))^2}{N_c}} \quad (2.51)$$

essent el valor inicial de la constant de pas del NLMS delimitada per  $0 < \bar{\mu}(0) < 1$ . Tot i que aquest criteri d'actualització de la constant de pas només minimitza l'error quadràtic mig quan el procés d'entrada  $(z(x,n))$  és blanc, les simulacions realitzades pel mateix Slock semblen recolzar la validesa de la seva proposta (Eq.2.51) per accelerar la convergència del LMS sense penalitzar el desajust final, fins i tot en situacions on aquesta premissa no es compleix.

#### 2.4.2 Mètodes recursius: l'algorisme RLS

Els mètodes recursius també estan desenvolupats a l'entorn dels estimadors lineals amb els paràmetres (Eq.2.17) però la funció de cost, a diferència dels mètodes de gradient, és ara una funció determinista

$$\xi(n) = \sum_{i=1}^n \lambda^{n-i} \cdot |e(i)|^2 \quad ; 0 < \lambda < 1 \quad (2.52)$$

que pondera, generalment d'una forma exponencial, l'error quadràtic comès per l'estimador. El conjunt de pesos òptim que minimitza  $\xi(n)$  correspon al següent sistema d'equacions

$$\hat{R}_{zz}(n) \cdot w_{opt}(n) = \hat{r}_{zd}(n) \quad (2.53)$$

on  $\hat{R}_{zz}(n)$  (Eq.2.54.a) i  $\hat{r}_{zd}(n)$  (Eq.2.54.b) es poden interpretar com estimacions de la matriu d'autocorrelació i del vector de correlació creuada, respectivament.

$$\hat{R}_{zz}(n) = \sum_{i=1}^n \lambda^{n-i} z(x,i) \cdot z^H(x,i) = \lambda \hat{R}_{zz}(n-1) + z(x,n) \cdot z^H(x,n) \quad (2.54.a)$$

$$\hat{r}_{zd}(n) = \sum_{i=1}^n \lambda^{n-i} z(x,i) \cdot d^*(i) = \lambda \hat{r}_{zd}(n-1) + z(x,n) \cdot d^*(n) \quad (2.54.b)$$

Hi ha diverses formes d'arribar a l'algorisme recursiu de mínims quadrats (Recursive Least Squares, RLS). El camí tradicional, però potser alhora més mancat d'una explicació intuïtiva, és el d'interpretar el RLS com un algorisme que troba la solució de mínims quadrats

$$w_{opt}(n) = \hat{R}_{zz}^{-1}(n) \cdot \hat{r}_{zd}(n) \quad (2.55)$$

utilitzant el lema de la matriu inversa [Hay91] per calcular de manera adaptativa la inversa de la matriu d'autocorrelació  $\hat{R}_{zz}(n)$ . L'algorisme resultant es recull a la taula (Tau.2.2), on  $k(n)$  és l'anomenat *vector de guany* i  $\alpha(n)$  representa l'*error a priori* de l'estimador. L'anàlisi estadístic d'aquest algorisme està fora del propòsit d'aquest treball tot i que es pot trobar a [Hay91].

$0 < \lambda < 1 \quad \text{generalment, proper a } 1$ $0 < \delta \quad \text{generalment, proper a } 0$ <p>Condicions inicials: <math>w(0) = \theta</math></p> $P(0) = \delta^{-1} I \quad ; P \equiv \hat{R}_{zz}^{-1}$ <p>Algorisme: Per <math>n = 1, 2, \dots</math></p> $a'(n) = z^H(x, n) P(n-1)$ $b(n) = \lambda + a'(n) z(x, n)$ $k(n) = \frac{a^*(n)}{b(n)}$ $\alpha(n) = d(n) - w^H(n-1) z(x, n)$ $w(n) = w(n-1) + k(n) \alpha^*(n)$ $T(n-1) = k(n) a'(n)$ $P(n) = \frac{1}{\lambda} (P(n-1) - T(n-1))$
---

**Taula 2.2** L'algorisme adaptatiu RLS

És interessant destacar que es pot deduir el RLS com un cas particular del filtre de Kalman [Hay91] i que és aquesta relació la que dota a l'algorisme recursiu d'un comportament notablement millor que els mètodes de gradient. En efecte, el RLS és més robust envers l'estadística del senyal que es combina linealment amb els pesos, mostrant una velocitat de convergència considerablement major que el LMS, sense empitjorar el desajust final de l'error. Aquesta propietat del RLS esdevé molt important en sistemes no lineals ja que, la correlació elevada entre components del vector de senyal  $z(x, n)$  fa que el LMS presenti una velocitat de

convergència prohibitivament baixa, essent el RLS una opció a tenir en compte perquè no és tan sensible a la dispersió d'autovalors de la matriu d'autocorrelació  $R_{zz}$ . De tota manera, es fàcil comprovar (Tau.2.2) com la complexitat del RLS és de l'ordre de  $O(N_c^2)$  operacions per iteració, en front de  $O(N_c)$  que presenta el LMS. A part d'aquesta limitació també és un algorisme numèricament inestable, el que fa que presenti problemes en ésser implementat amb precisió finita.

Donat aquestes característiques del RLS, els esforços s'encaminen a trobar uns algorismes alternatius de la solució de mínims quadrats que siguin menys costosos computacionalment parlant i que alhora presentin bones propietats de robustesa numèrica. En aquest sentit, cal destacar primer de tot treballs tals com [Dav87] on es presenta un algorisme recursiu RLS de complexitat  $O(N^2)$  per sistemes de Volterra de segon ordre, però que només és vàlid per senyals d'entrada gaussianes i que, a més a més, té problemes numèrics.

Per altra banda, tot i que tampoc té robustesa numèrica, s'han desenvolupat algorismes RLS ràpids ("Fast-RLS" de l'ordre de  $O(N^3)$  operacions per iteració per sistemes de Volterra(2,N)) sense cap restricció sobre l'estadística de l'entrada [Lee93]. Aquest treball es podria generalitzar per un sistema de Volterra de qualsevol ordre, essent la complexitat de l'algorisme resultant de  $O(N^{2P-1})$ .

Per últim, caldria assenyalar que s'han desenvolupat algorismes recursius per sistemes de Volterra basats en la descomposició QR per a la implementació del RLS [Sye93]. Aquests mètodes assoleixen un cost computacional relativament baix, de l'ordre de  $O(N^3)$  per sistemes de Volterra de segon ordre, i alhora mostren una robustesa numèrica major que els altres algorismes RLS ràpids.

### 2.4.3 Estructures en gelosia o "lattice"

Tot i que no està dins les expectatives d'aquesta tesi entrar en un estudi profund dels diferents mètodes adaptatius de disseny de sistemes de Volterra, sí que es pretén donar una visió global de la problemàtica que comporta i cap a on s'està avançant. És per això que s'ha considerat interessant dedicar un darrer punt a les estructures *lattice* per filtres no lineals.

Originàriament, el filtre *lattice* sorgeix com una estructura no directa que implementa un filtre transversal de retards [Hay91]. La filosofia d'aquesta estructura consisteix en, a partir d'un vector de dades que genera un subespai de senyal determinat (en un filtre FIR seria el vector  $x(n)$  (Eq.2.14)), es tracta d'obtenir un altre vector que sigui capaç de generar el mateix subespai de senyal però que les seves components siguin ortogonals o incorrelades entre elles. Tal i com s'havia introduït anteriorment (Apt.2.3.5), es pot utilitzar com a filtre blanquejador de senyals gaussianes. Les principals prestacions dels filtres *lattice* són unes propietats numèriques

millors que les versions directes, així com gaudir d'una capacitat modular que les fa adequades per implementacions VLSI (de l'anglès "Very-Large Scale Integration") pel fet d'ésser recursives en l'ordre.

Les estructures lattice desenvolupades per sistemes no lineals de Volterra persegueixen bàsicament la mateixa fita: a partir del vector de senyal  $z(x,n)$ , construir un vector nou que generi el mateix subespai de senyal i que les seves components siguin incorrelades. Aquesta propietat d'ortogonalitat entre components del nou vector de senyal permet utilitzar mètodes adaptatius amb un cost computacional baix tals com els algorismes LMS o NLMS, i que, a més a més, presenten bones prestacions ja que la dispersió d'autovalors de la nova matriu d'autocorrelació haurà disminuït considerablement.

Un exemple d'aquesta versió lattice per un sistema de Volterra de segon ordre amb una entrada gaussiana es pot trobar a [Koh83]. De tota manera, la tendència dels filtres lattice no lineals és la d'interpretar el sistema de Volterra com un filtre lineal multicanal amb la particularitat de que cada canal conté un nombre diferent de retards. Sobre aquestes architectures es pot trobar un tutorial a [Mat91], que va elaborar partint del treball de [Lin84] i [Zar85].

El problema que presenten els filtres lattice no lineals és que, en general, estan sobredimensionats, és a dir, necessiten d'un major nombre de coeficients que les versions directes. Per exemple, un sistema de Volterra(2,N) de segon ordre conté de l'ordre de  $O(N^2)$  coeficients i, en canvi, la implementació lattice esmentada genera un ordre de magnitud més,  $O(N^3)$ . Aixó significa que si s'utilitza un algorisme d'adaptació per actualitzar aquests coeficients tipus el LMS (algorisme "Gradient Adaptive Lattice" (GAL) desenvolupat per Griffiths [Hay91]), el cost computacional de la implementació lattice d'un model de Volterra(2,N) és de l'ordre de  $O(N^3)$ , és a dir, del mateix ordre de magnitud que les versions ràpides del RLS comentades anteriorment. Per tant, és important adonar-se que les prestacions d'una estructura lattice en sistemes no lineals de Volterra no suposa una millora tant important com succeix en filtratge lineal.

Per últim, caldria destacar que hi ha d'altres implementacions modulars dels sistemes de Volterra però que es basen en una actualització dels pesos no amb mètodes de gradient sino amb mètodes recursius [Zar85, Sye90].

## 2.5 Conclusions

En aquest segon capítol de la present tesi s'ha pretès donar una visió general documentada, i alhora també molt personal, del que entenem que és l'estat de l'art actual pel que fa a modelatge de SNLs amb sistemes de Volterra.

El model de Volterra presenta propietats molt interessants pel fet bàsicament d'ésser un model de SNLs que és lineal en termes dels seus coeficients. Aquesta característica és primordial per a explicar l'èxit que ha experimentat aquesta família de models ja que permet aplicar mètodes ja desenvolupats en l'entorn de processament lineal del senyal, directament sobre ells. El principal problema que sorgeix és la dificultat en el disseny dels coeficients com a conseqüència de la forta correlació existent entre les funcions que generen l'espai de senyal utilitzat pel model de Volterra.

Aquesta dificultat va ser la que va motivar a Wiener a idear un model de SNLs basat en el model de Volterra però que implementa un procés d'ortogonalització de Gram-Schmidt sota la hipòtesi de que el senyal d'entrada és Gaussià. En el nostre parer, el model de Wiener, tot i que ha tingut una forta repercussió a nivell d'arquitectura, no ha experimentat una continuïtat ni molta difusió bàsicament com a conseqüència de dos aspectes: en primer lloc, l'escassa flexibilitat del model per a adequar-se a senyals que no compleixin la restricció de Gaussianitat i, en segon lloc, la complexitat del model resultant. Una altra raó fonamental per a explicar la proliferació de mètodes relacionats amb el model de Volterra en detriment del de Wiener es pot trobar en el fet de que el model de Volterra ha situat les seves principals aplicacions en el camp de modelatge de SNLs però amb un disseny adaptatiu. Darrerament, les accions de recerca més rellevants s'han encaminat a desenvolupar algorismes adaptatius adequats al disseny de sistemes de Volterra. Per una banda, aquests algorismes han d'intentar reduir al màxim el cost computacional perquè, tal i com s'ha vist, representa un problema important en models de Volterra ja que el nombre de coeficients creix exponencialment. Per altra banda, els algorismes han de proporcionar unes prestacions acceptables de velocitat de convergència i desajust final, fet que redueix les opcions a l'ús de mètodes recursius atesa l'alta correlació entre els nuclis de Volterra, i alhora assegurar-ne l'estabilitat.

A continuació, el capítol III es centra més en el cos de la tesi, presentant un model de SNLs alternatiu al model de Volterra i que, tal i com s'anirà deixant palès, té unes propietats que el fan superior al model de Volterra, precisament en dissenys adaptatius.

## 2.6 Referències

- [And79] Anderson B.D.O., Moore J.B., *Optimal Filtering*, Prentice Hall, Englewood Cliffs, New Jersey, 1979.
- [Dav87] Davila C.E., Welch A.J., Rylandes H.G., "A Second-Order Adaptive Volterra Filter with Rapid Convergence", *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. 35, N° 9, pp. 1259-1263, Setembre 1987.

- [Fej94a] Fejzo Z, Lev-Ari H., "Adaptive Laguerre Filters with Lattice Orthogonalization", *IEEE Proc. Workshop on Statistical Signal and Array Processing SSAP'94*, Quebec City, Quebec, Canada, pp. 285-288, Juny 1994.
- [Fej94b] Fejzo Z, *Adaptive Non-Linear Laguerre-Based Filtering*, Tesi Doctoral, Northeastern University, Boston, 1994.
- [Hay91] Haykin S., *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, New Jersey, 1991.
- [Koh83] Koh T., Powers E.J., "An Adaptive Nonlinear Digital Filter with Lattice Orthogonalization", *IEEE Proc. International Conference on Acoustics, Speech and Signal Processing ICASSP'83*, Boston, Massachusetts, pp. 37-40, Abril 1983.
- [Koh85] Koh T., Powers E.J., "Second-Order Volterra Filtering and its Application to Nonlinear System Identification", *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. 33, N° 6, pp. 1445-1455, Desembre 1985.
- [Lee65] Lee Y.W., Schetzen M., "Measurement of the Wiener Kernel in a Nonlinear System by Crosscorrelation", *International Journal of Control*, Vol. 2(3), pp. 237-254, Setembre 1965.
- [Lee93] Lee J., Mathews V.J., "A Fast Recursive Least Squares Adaptive Second-Order Volterra Filter and its Performance Analysis", *IEEE Trans. on Signal Procesing*, Vol. 41, N° 3, pp. 1087-1101, Març 1993.
- [Lin84] Ling F., Proakis J.G., "A Generalized Multichannel Least Squares Lattice Algorithm based on Sequential Processing Stages", *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. 32, N° 2, pp. 381-390, Abril 1984.
- [Mat88] Mathews V.J., Lee J., "A Fast Recursive Least-Squares Second Order Volterra Filter", *IEEE Proc. International Conference on Acoustics, Speech and Signal Processing ICASSP'88*, New York, pp. 1383-1386, Abril 1988.
- [Mat90] Mathews V.J., Xie Z., "Stochastic Gradient Adaptive Filters with Gradient Adaptive Step-Sizes", *IEEE Proc. International Conference on Acoustics, Speech and Signal Processing ICASSP'90*, Albuquerque, New Mexico, pp. 1385-1388, Abril 1990.
- [Mat91] Mathews V.J., "Adaptive Polynomial Filters", *IEEE Signal Procesing Magazine*, Vol. 8, N° 3, pp. 10-26, Juliol 1991.
- [Mat93] Mathews V.J., Sicuranza G.L., "Volterra and General Polynomial Related Filtering", *Proc. IEEE Winter Workshop on Nonlinear Digital Signal Processing*, pp. T2.1-T2.8, Tampere, Finland, Gener 1993.



- [Pag95] Pagès-Zamora A., Lagunas M.A., Nájjar M., Pérez-Neira A., “The K-filter: a new architecture to model and design non-linear systems from the Kolmogorov's theorem”, *EURASIP Signal Processing*, Vol. 44, N° 3, pp. 249-267, Juliol 1995.
- [Rug81] Rugh W.J., *Nonlinear System Theory: the Volterra/Wiener Approach*, The Johns Hopkins University Press, Baltimore, 1981.
- [Sch81] Schetzen M., “Nonlinear System Modeling Based on the Wiener Theory”, *Proc. of the IEEE*, Vol. 69, n° 12, Desembre 1981.
- [Sch89] Schetzen M., *The Volterra/Wiener Theories of Nonlinear Systems*, Robert E. Krieger Publishing Company, Florida, 1989. (1<sup>a</sup> edició 1980).
- [Sic87] Sicuranza G.L., Ramponi G., “A variable-Step Adaptation Algorithm for Memory-Oriented Volterra Filters”, *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. 35, N° 10, pp. 1492-1494, Octubre 1987.
- [Sic92] Sicuranza G.L., “Quadratic Filters for Signal Processing”, *Proc. of the IEEE*, Vol. 80, N° 8, pp. 1262-1285, Desembre 1992.
- [Slo93] Slock D.T.M., “On the Convergence Behavior of the LMS and Normalized LMS Algorithms”, *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. 41, N° 9, pp. 2811-2825, Setembre 1993.
- [Sye90] Syed M.A., Mathews V.J., “Lattice and QR Decomposition-based Algorithms for Recursive Least Squares Adaptive Nonlinear Filters”, *IEEE Proc. International Symposium Circuits and Systems*, New Orleans, Louisiana, pp. 262-265, Maig 1990.
- [Sye93] Syed M.A., Mathews V.J., “QR-Decomposition Based Algorithms for Adaptive Volterra Filtering”, *IEEE Trans. on Circuits and Systems-I: Fundamental Theory and Applications*, Vol. 40, N° 6, pp. 372-381, Juny 1993.
- [Vid78] Vidyasagar M., *Nonlinear Systems Analysis*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1978.
- [Vol13] Volterra V., *Leçons sur les fonctions de lignes*, Gauthier-Villars, Paris, 1913.
- [Wie58] Wiener N., *Nonlinear problems in random theory*, The Technology Press, M.I.T. and John Wiley & Sons, Inc., New York, 1958.
- [Zar85] Zarzycki J., *Nonlinear Prediction Ladder Filters for Higher Order Stochastic Sequences*, Springer-Verlag, Berlin, 1985.

## Capítol 3

### Model de Fourier de sistemes no lineals basat en la sèrie trigonomètrica

En el camp de teoria del senyal, altres sèries de funcions s'han anat imposant per sobre de la sèrie polinòmica o de Taylor a l'hora d'aproximar senyals, tant si es tractava de senyals en el domini temporal com en el domini espacial. Un exemple clar és la ben coneguda sèrie trigonomètrica, que forma part del que s'anomenen *sèries de Fourier* [Chu69].

Basant-nos en aquest fet, el principal objectiu d'aquest capítol consisteix en estudiar les possibilitats de la sèrie trigonomètrica com a eina matemàtica per modelar SNLs [Gra95], tot analitzant els avantatges i limitacions que presenta en front del model de Volterra. La comparació del model de Volterra i el model basat en la sèrie trigonomètrica (que anomenarem *model de Fourier*) s'efectuarà des de tres perspectives diferents: a nivell de sèrie, d'arquitectura i de disseny. Així, és necessari primer de tot examinar la capacitat de la sèrie trigonomètrica per a aproximar funcions, cosa que en determinades ocasions pot justificar l'ús d'aquesta sèrie en lloc de la polinòmica. Per altra banda, per tal d'obtenir una visió global de les prestacions del model de Fourier es farà imprescindible abordar aspectes de caire més pràctic com són la implementació i disseny del model.

L'organització del capítol segueix les mateixes pautes que l'anterior, és a dir, en primer lloc es presenta el model de Fourier per SNLs sense memòria i, a continuació, l'apartat (3.2) es

dedica al model de Fourier multidimensional per SNLs amb memòria. En tots dos casos es discuteixen aspectes relacionats amb la formulació, implementació i disseny del model, tant en bloc com adaptatiu. Per últim, i abans d'acabar amb aquest capítol íntegrament dedicat a la presentació del model de Fourier, a la tercera secció (Ap.3.3) es desenvolupa una implementació en gelosia (o "lattice" en anglès) del model de Fourier sense memòria que dota al sistema resultant d'unes prestacions especialment adequades per a un disseny adaptatiu. El desenvolupament que es presenta fa palesa la possibilitat d'implementar versions no directes del model de Fourier (com, per exemple, l'estructura "lattice" pel cas del model de Fourier sense memòria) d'una manera totalment paral·lela a la realitzada pels sistemes de Volterra [Zar85].

### 3.1 Model de Fourier per a sistemes no lineals sense memòria

El model que proposem en aquest capítol per modelar sistemes no lineals sense memòria es basa en una aproximació de la relació entrada/sortida del SNL a partir d'una combinació lineal de les següents funcions,

$$\{1, \cos p\hat{x}, \sin p\hat{x}\} \quad (p = 1, 2, \dots) \quad (3.1)$$

que formen la base de l'anomenada sèrie trigonomètrica, essent  $\hat{x}$  en el cas que ens ocupa el senyal d'entrada al SNL normalitzat dins el rang  $[-\pi, +\pi]$ .

Malgrat que la sèrie trigonomètrica és, estrictament parlant, un cas particular de les anomenades *sèries de Fourier generalitzades* [Chu69], s'ha optat per denotar el model resultant que es presenta en aquest capítol amb el nom genèric de *model de Fourier*. Això ha estat fonamentalment motivat per l'àmplia acceptació que en el món de tractament del senyal hi ha sobre el reconeixement de la següent sèrie (Eq.3.2) com la sèrie de Fourier (homòloga a la trigonomètrica però formulada en termes de funcions exponencials complexes).

$$g(\hat{x}) = \sum_{p=-\infty}^{+\infty} c_p \cdot e^{j p \hat{x}} \quad (3.2)$$

En aquest sentit és imprescindible remarcar que, a diferència d'altres autors [Rug81], el nom de "model de Fourier" no implica l'ortogonalitat entre les funcions involucrades en el model de SNLs, sino que només denota el fet de que aquest es construeix a partir d'una sèrie de Fourier. De totes maneres, abans d'entrar en detalls del model, el primer pas consisteix en presentar la sèrie en la que es basa, analitzant les propietats més rellevants que condicionaran directament al model de Fourier.

### 3.1.1 Sèries de Fourier. Principals propietats

Les sèries de Fourier es desenvolupen entorn d'un espai de funcions  $f(x)$ , denotat per  $\Theta$ , finites i delimitades dins d'un interval  $a < x < b$ , definint el *producte escalar* de dues funcions qualsevols d'aquest espai, per exemple  $f(x)$  i  $g(x)$ , tal i com segueix.

$$\langle f(x), g(x) \rangle = \int_a^b f(x) \cdot g^*(x) dx \quad (3.3)$$

En conseqüència, la *norma* de tota funció que pertanyi a aquest espai en resulta un nombre real i no negatiu.

$$\|f(x)\| \equiv \text{NORMA} \equiv \sqrt{\langle f(x), f(x) \rangle} = \sqrt{\int_a^b |f(x)|^2 dx} \quad (3.4)$$

Donat un conjunt ordenat i possiblement infinit de funcions  $\{\phi_p(x) \in \Theta\}$  que siguin ortonormals entre elles segons la norma establerta,

$$\langle \phi_p, \phi_q \rangle = \begin{cases} 0 & q \neq p \\ 1 & q = p \end{cases} \quad \forall p, q = 1, 2, \dots \quad (3.5)$$

es defineix la *sèrie de Fourier* d'una funció  $f(x)$  respecte d'aquest conjunt de funcions ortonormals com la següent combinació lineal

$$\sum_{p=1}^{\infty} c_p \cdot \phi_p(x) \quad (3.6)$$

que aproxima la funció  $f(x)$ , de manera que els coeficients  $c_p$  compleixen la següent igualtat.

$$c_p = \langle f(x), \phi_p^*(x) \rangle = \int_a^b f(x) \cdot \phi_p^*(x) dx \quad \forall p = 1, 2, \dots \quad (3.7)$$

Aquests coeficients  $\{c_p \forall p=1, 2, \dots\}$ , anomenats normalment *constants o coeficients de Fourier*, són a més a més uns coeficients que minimitzen la distància entre la funció  $f(x)$  i la combinació lineal que formen amb les funcions ortonormals. És a dir, no existeix cap altre combinació lineal diferent de les funcions  $\{\phi_p(x) \forall p=1, 2, \dots\}$ ,

$$\sum_{p=1}^{\infty} \gamma_p \cdot \phi_p(x) \quad \gamma_p \neq c_p \quad (3.8)$$

tal que la distància a la funció  $f(x)$ , segons la norma definida per l'espai de funcions, sigui menor que l'assolida per la combinació lineal amb les constants de Fourier  $\{c_p \forall p=1,2,\dots\}$ .

Es diu que el conjunt de funcions  $\{\phi_p(x) \forall p=1,2,\dots\}$  és *complet* si no existeix cap funció que pertanyi a l'espai, que sigui ortogonal a totes les funcions del conjunt. Per altra banda, si ens definim  $S_p(x)$  com la suma parcial de la sèrie de Fourier generalitzada següent,

$$S_p(x) = \sum_{q=1}^p c_q \cdot \phi_q(x) \quad (3.9)$$

llavors es diu que  $\{\phi_p(x)\}$  és un conjunt *tancat en el sentit de convergència en la mitja* si el següent límit tendeix a zero:

$$\lim_{p \rightarrow \infty} \int_a^b |S_p(x) - f(x)|^2 dx = 0 \quad (3.10)$$

Desenvolupant el quadrat d'aquest límit i tenint present la definició dels coeficients de Fourier (Eq.3.7), fàcilment s'arriba a la següent relació anomenada *equació de Parseval*.

$$PARSEVAL \Rightarrow \sum_{p=1}^{\infty} |c_p|^2 = \int_a^b |f(x)|^2 dx \quad (3.11)$$

És important notar que, en general, la propietat (3.10) de la sèrie de Fourier no implica una convergència uniforme i punt a punt de la suma parcial  $S_p(x)$  cap a la funció  $f(x)$ . En altres paraules, no és equivalent al següent límit

$$\lim_{p \rightarrow \infty} S_p(x) = f(x) \quad (3.12)$$

que, per altra banda, quan es compleix es diu que  $\{\phi_p(x)\}$  és un conjunt *tancat en el sentit de convergència puntual*.

Un cop establerts els principis de les sèries de Fourier generalitzades, ens centrarem en el cas particular de la sèrie trigonomètrica o, ja directament, la sèrie exponencial complexa de Fourier. Abans però, caldrà definir el que s'entén per una funció *contínua per seccions*.

**Definició 3.1:** *Funció contínua per seccions.* Suposem que  $f(x)$  és una funció contínua en tots els punts d'un interval finit  $a \leq x \leq b$  excepte possiblement en un conjunt finit  $\{a, x_1, \dots, x_n, b\}$  tals que  $a \leq x_1 \leq \dots \leq x_n \leq b$ . Llavors  $f(x)$  és contínua a cadascun dels intervals oberts  $\{a < x_1, x_1 < x_2, \dots, x_n < b\}$ , no essent necessari que estigui definida als extrems. De

totes maneres, si a cada subintèrval,  $f(x)$  tendeix a un valor finit a mesura que  $x$  s'apropa a l'extrem per l'interior, llavors es diu que  $f(x)$  és contínua per seccions dins l'interval  $(a,b)$ .

♣

Un cop definit el que s'entén per funcions contínues per seccions, es pot formular el següent teorema [Chu69] que determina quines són les funcions que admeten un desenvolupament en sèrie de Fourier.

**Teorema 3.1.** Per a tota funció  $f(x)$  tal que sigui contínua per seccions dins l'interval  $(-X_0, X_0)$ , i periòdica amb període  $2X_0$ , la següent sèrie de Fourier

$$\sum_{p=-\infty}^{\infty} c_p \cdot e^{\frac{jp\pi x}{X_0}} \quad \text{on} \quad c_p = \frac{1}{2X_0} \int_{-X_0}^{X_0} f(x) \cdot e^{-\frac{jp\pi x}{X_0}} dx \quad (3.13)$$

convergeix al valor

$$\frac{1}{2}[f(x+0) + f(x-0)] \quad -\infty < x < +\infty \quad (3.14)$$

per tot valor  $x$  sempre i quan existeixi la derivada lateral per la dreta i per l'esquerra de la funció  $f(x)$ . (La notació utilitzada, és a dir,  $f(x+0)$  i  $f(x-0)$  indica límit lateral per la dreta i per l'esquerra de la funció  $f(\cdot)$  en  $x$ , respectivament).

♣

Aquest teorema és fonamental per a establir quines funcions són susceptibles d'ésser desenvolupades en sèrie de Fourier i d'aquesta manera ser capaços d'establir comparacions entre el model de Fourier i el de Volterra. En aquest sentit, en el teorema s'especifica que la funció que admet un desenvolupament en sèrie de Fourier ha de ser periòdica. En general, es pot afirmar que la sèrie de Fourier proposada (Eq.3.13) convergeix també per a qualsevol funció limitada a l'interval  $(-X_0, X_0)$ , cumplint-se llavors l'expressió (Eq.3.14) per valors de  $x$  inclosos dins d'aquest rang (i no a tot el rang dels reals). En aquest cas, la sèrie de Fourier reproduïx fidelment la funció  $f(x)$  dins el rang que interessa, és a dir,  $(-X_0, X_0)$ , repetint-la a fora periòdicament ja que l'expressió de la sèrie de Fourier (Eq.3.13) és inherentment periòdica amb període  $2X_0$ . Per altra banda, cal destacar que el fet de que el teorema (3.1) estigui enunciat suposant un interval de l'entrada simètric respecte de l'origen no suposa cap restricció per funcions definides dins d'un interval que no ho sigui. En efecte, la condició de convergència de l'expressió (Eq.3.14) es compleix també per qualsevol subinterval  $(X_{min}, X_{max})$  no simètric que estigui inclòs dins el rang  $(-X_0, X_0)$ .

Tot i que la sèrie de Taylor (Eq.2.2) no imposa cap condició respecte del rang d'entrada  $i$ , en canvi, la de Fourier sí, ja es va comentar a l'apartat (Ap.2.1.2) que el teorema de Weierstrass només assegura l'existència d'un model polinòmic d'ordre finit per a aproximar funcions amb un error absolut arbitràriament baix quan aquestes són: reals, contínues i definides dins d'un interval finit. Per tant, la necessitat d'haver de delimitar el rang de la variable d'entrada per assegurar la convergència de la sèrie de Fourier, no suposa cap limitació del model de Fourier envers el de Volterra, ja que a la pràctica el model de Volterra també requereix d'aquesta condició.

L'aspecte més rellevant que posa de manifest el teorema és l'existència de la sèrie de Fourier per funcions fins i tot discontinües, sempre i quan siguin limitades (és a dir, discontinuïtats de primera espècie per tal que la funció pertanyi a l'espai de funcions normat (Eq.3.4)). Així, la sèrie de Fourier convergeix a  $f(x)$  quan aquesta funció és contínua, i quan no ho és, convergeix al valor promig dels límits laterals als que tendeix la funció  $f$  quan s'acosta a la discontinuïtat  $x$ . Per exemple, un cas particular és el comportament de la sèrie als extrems del rang que tendeix al valor promig dels límits laterals interiors als quals tendeix  $f(x)$ .

$$\frac{1}{2}[f(-X_0 + 0) + f(X_0 - 0)] \quad (3.15)$$

Aixó fa que, a priori, el model que presentarem basat en la sèrie de Fourier abracci unes funcions amb determinades discontinuïtats que el model de Volterra en principi no és capaç de modelar. De totes maneres, és evident que per qüestions de caire pràctic el model de Fourier es construirà a partir de la suma parcial  $S_p(x)$  (Eq.3.9) i no de la sèrie original amb un nombre infinit de termes. Llavors, en el cas de la sèrie trigonomètrica la convergència de la suma parcial cap a  $f(x)$  per tot  $x$ , només està assegurada allà on la funció que s'està desenvolupant és contínua [Chu69]. En cas contrari, la suma parcial  $S_p(x)$  té un comportament al voltant de les discontinuïtats que es coneix pel nom de *fenòmen de Gibbs*. Aquest fenòmen consisteix en que, mentres que al punt de discontinuïtat  $x$ , la sèrie de Fourier pren el valor mig de la discontinuïtat, a cada costat la sèrie de Fourier té un comportament oscil·latori amb un període inversament proporcional a l'ordre  $p$  de l'aproximació però amb un pic inicial que només depèn del desnivell de la discontinuïtat [Car88] i que, per tant, per més que s'elevi l'ordre no desapareixerà. En definitiva, només en cas de que la funció  $f(x)$  que s'està aproximant sigui contínua, es pot assegurar que la suma parcial  $S_p(x)$  tendeix uniformement cap a  $f(x)$  per a tot  $x$  a mesura que  $p$  augmenta. Per altra banda, és interessant tenir en consideració el fenòmen de Gibbs que apareixerà sempre que hi hagi una discontinuïtat o també, quan el valor de  $f(x)$  als límits del rang d'entrada sigui diferent (Eq.3.15).

### 3.1.2 Model de Fourier sense memòria

Procedint d'una forma similar al capítol anterior pel cas de Volterra, el model de Fourier per a un determinat SNL es defineix a partir del desenvolupament en sèrie de Fourier de la relació entrada/sortida que caracteritza a aquest sistema, que es denota per la funció  $g(\cdot)$ . Prenent en consideració que el SNL no té memòria, això significa que la sortida  $y(n)$  només depèn del valor actual de l'entrada  $x(n)$  i que per tant la funció  $g(x)$  és unidimensional.

$$y(n) = g[x(n)] \quad (3.16)$$

Sota la hipòtesi de que el senyal d'entrada pertany a un rang limitat, per exemple  $(X_{min}, X_{max})$ , i que la funció  $g(x)$  admet un desenvolupament de Fourier tal i com el que segueix,

$$g(x) = \sum_{p=-\infty}^{+\infty} c_p \cdot e^{-j p \frac{\pi}{X_0} x} \quad 2X_0 \geq (X_{max} - X_{min}) \quad (3.17)$$

es proposa que el *model de Fourier d'ordre P per SNLs sense memòria* (Eq.3.18) sigui una versió truncada de la sèrie de Fourier per tal de que sigui realitzable.

$$\hat{g}(x) = \sum_{p=-P}^{+P} c_p \cdot e^{-j p \frac{\pi}{X_0} x} \quad (3.18)$$

En cas de que  $2X_0 > (X_{max} - X_{min})$ , l'aproximació implementada pel model de Fourier (Eq.3.18) continuarà essent vàlida sempre i quan el desenvolupament en sèrie de Fourier (Eq.3.17) sigui igual a la relació entre l'entrada i sortida del SNL real per a tot  $x$  que pertanyi al rang d'entrada  $(X_{min}, X_{max})$ . És en aquest sentit que el valor que prengui la sèrie de Fourier fora d'aquest rang no perjudica l'aproximació del model final però que, per altra banda, si està ben dissenyada pot ajudar a disminuir el fenòmen de Gibbs en els extrems així com disminuir el cost computacional del model resultant forçant simetries [Pag94].

Ja s'ha vist anteriorment que, per tal de minimitzar la norma de l'error entre la funció que implementa el model,  $\hat{g}(x)$ , i el sistema real,  $g(x)$ , els coeficients  $c_p$  han d'estar definits segons (Eq.3.7). Una manera alternativa d'expressar els coeficients de Fourier és mitjançant la funció  $G_e(\omega)$ ,

$$c_p = \frac{1}{2X_0} \cdot G_e(\omega) \Big|_{\omega=p\omega_0} \quad ; \omega_0 = \frac{\pi}{X_0} \quad (3.19)$$

que representa la transformada de Fourier de la funció  $g_e(x)$  respecte de  $x$ , essent  $g_e(x)$  una versió en finestrada  $g(x)$  nul·la fora de l'interval  $(-X_0, X_0)$  i que ha de coincidir amb  $g(x)$  pels valors de  $x$  que pertanyin al rang d'entrada.



$$\begin{aligned}
G_e(\omega) &\equiv \int_{-\infty}^{+\infty} g_e(x) \cdot e^{-j\omega x} dx = \\
&= \int_{X_0}^{X_{\min}} g_e(x) e^{-j\omega x} dx + \int_{X_{\min}}^{X_{\max}} g(x) e^{-j\omega x} dx + \int_{X_{\max}}^{X_0} g_e(x) e^{-j\omega x} dx
\end{aligned} \tag{3.20}$$

En el cas en que  $g_e(x)$  també s'esculli nul·la fora del rang d'entrada, es pot interpretar que els coeficients  $c_p$  s'obtenen de mostrejar la transformada de Fourier de  $g(x)$  a múltiples enters de  $\omega_0 = \pi/X_0$ , paràmetre que d'ara endavant s'anomenarà *frequència principal*. Aquest paràmetre del model de Fourier esdevé crucial ja que controla el període de repetició de  $g(x)$  quan es desenvolupa en sèrie de Fourier. Així, si  $G_e(\omega)$  es mostra a una cadència massa elevada, és a dir,  $\omega_0 > 2\pi/(X_{\max} - X_{\min})$ , la funció que reproduïx el model de Fourier proposat és  $g(x)$  repetida amb un període menor que el rang d'entrada, de manera que es superposa cometent el que en anglès s'anomena "aliasing" i proporcionant una aproximació distorsionada de la veritable funció  $g(x)$ . Al contrari, si  $G_e(\omega)$  es mostra amb un període molt més petit que el mínim, és a dir  $\omega_0 \ll 2\pi/(X_{\max} - X_{\min})$ , es necessitarà incrementar l'ordre  $P$  del model, i per tant la seva complexitat, per assolir el mateix llindar d'error.

Aquest grau d'aproximació que és capaç d'assolir el model de Fourier d'ordre  $P$  depèn exclusivament de la relació entrada/sortida del SNL particular que es vol modelar. El que sí que es pot afirmar és que el model de Fourier proposat (Eq.3.18) serà millor en la mesura que la funció  $g(x)$  tingui un comportament "suau" en  $x$ , és a dir, que  $G_e(\omega)$  tingui un contingut més important al voltant de l'origen (per  $\omega$  baixes) que no a freqüències més elevades. El fet de que certs dispositius no lineals típics (per exemple, els amplificadors de RF o bé els díodes) tinguin un clar comportament suau, així com el resultat de les simulacions que es presenten més endavant han encoratjat l'aprofundiment d'aquest model. En cas de que la funció  $G_e(\omega)$  tingui un contingut freqüencial important en un altre marge de freqüències, seria convenient modificar els valors que pren l'índex  $p$  del sumatori (Eq.3.18) per adequar-se a les característiques del SNL particular amb el que ens enfrontem (malgrat que això és poc realista i difícil de determinar).

La teoria de Fourier ens proporciona un criteri de disseny dels coeficients  $c_p$  de manera que minimitzin la norma de la funció error  $e(x)$ , definida com la diferència entre la funció real  $g(x)$  i l'aproximació  $\hat{g}(x)$ .

$$\left. \int_{-X_0}^{X_0} |g(x) - \hat{g}(x)|^2 dx \right|_{\text{MINIM}} \Rightarrow c_p = \int_{-X_0}^{X_0} g(x) \cdot e^{-j p \omega_0 x} dx \quad ; p = -P, \dots, P \tag{3.21}$$

En general, els SNLs que es tracten en aquesta tesi segueixen una relació entrada/sortida real i per tant és desitjable que l'aproximació proposada pel model de Fourier,  $\hat{g}(x)$ , també ho

sigui. Malgrat que els coeficients de Fourier són complexos, aquesta propietat s'assoleix gràcies a l'hermiticitat de les funcions exponencials complexos, forçant alhora una simetria hermítica en els coeficients  $c_p$ .

$$c_p = c_{-p}^* \quad p = 1, \dots, P \quad (3.22)$$

Tot i que queda fora de l'abast d'aquesta tesi, el model de Fourier també seria adequat per modelar SNLs amb una relació entrada/sortida que fos una funció complexa sempre que depengués d'una variable real.

Un cas particular que sovint es presenta en el modelat de SNLs és la propietat de que la funció  $g(x)$  segueix una simetria determinada, per exemple una simetria parell o senar respecte de l'origen. Així, en cas d'existir una simetria parell, la sèrie de Fourier, i per tant també el model de Fourier, es redueix a una combinació lineal de funcions cosinus més una component de contínua,

$$g(x) = g(-x) \Rightarrow \hat{g}(x) = \frac{1}{2}a_0 + \sum_{p=1}^P a_p \cdot \cos\left(p \frac{\pi}{X_0} x\right) \quad (3.23)$$

o bé, una combinació lineal de sinus en cas de simetria senar.

$$g(x) = -g(-x) \Rightarrow \hat{g}(x) = \sum_{p=1}^P b_p \cdot \sin\left(p \frac{\pi}{X_0} x\right) \quad (3.24)$$

Els coeficients d'ambdós models, denotats per  $\{a_p \forall p=1,2,\dots\}$  i  $\{b_p \forall p=0,1,\dots\}$ , esdevenen ara reals, guardant una estreta relació amb els coeficients de Fourier de la sèrie basada en exponencials complexos.

$$a_p = 2 \cdot \Re[c_p] \quad p = 0, \dots, P \quad (3.25.a)$$

$$b_p = 2 \cdot \Im[c_p] \quad p = 1, \dots, P \quad (3.25.b)$$

Tal i com es veurà a les simulacions, no és del tot inusual enfrontar-se a un problema pràctic en el que el SNL a modelar compleixi alguna d'aquestes simetries. Per altra banda, aquests models també resulten útils en cas de que el rang del senyal d'entrada pertanyi als reals positius  $[0, X_{max}]$  (o bé, d'una forma totalment paral·lela als negatius). Així, en una situació general en la que  $g(0) \neq g(X_{max})$ , el SNL es podria modelar mitjançant un model parell (Eq.3.23) amb una freqüència de  $\omega_0 = \pi/X_{max}$  de manera que el model reproduïx la funció del SNL pel rang de valors de l'entrada  $[0, X_{max}]$ . Quan, a més a més, es compleix que  $g(0) = g(X_{max})$  també seria vàlid un model senar basat en funcions senoidals (Eq.3.24). En

aquests casos, i en general sempre que es pugui fins i tot recorrent a solucions artificials com les presentades, serà interessant tenir en compte aquestes versions simplificades del model de Fourier (Eq.3.23,3.24) ja que comporten una reducció considerable tant del cost computacional com de la memòria necessària. En efecte, són models que utilitzen funcions i paràmetres reals en lloc de complexos com en el cas del model de Fourier amb exponencials complexos (Eq.3.18), sense penalitzar el grau d'aproximació que és capaç d'assolir.

De totes maneres, tant per a la implementació del model de Fourier així com pel seu disseny, és més interessant desenvolupar el model més general en funció de les exponencials complexos en lloc de les versions simplificades. Això està motivat en part pel caràcter general del resultat i, a més a més, perquè d'aquesta manera sorgeixen aspectes rellevants del model de Fourier en relació a possibles aplicacions que motiven el capítol quart de la present tesi.

### 3.1.3 Arquitectura per implementar el model de Fourier

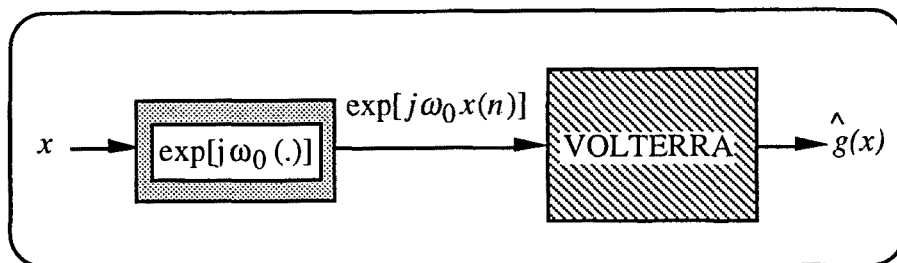
Tal i com es veurà més endavant en el cas de SNLs amb memòria, el model de Fourier també admet una arquitectura en dues etapes (LD/NMR) que segueix les mateixes pautes descrites a l'apartat (Ap.2.1.4) pel model de Volterra. Malgrat això, en el cas sense dinàmica, la primera etapa LD és eliminada i, per tant, no resulta adient recórrer a aquesta arquitectura per a establir comparacions entre ambdós models.

Sense una anàlisi gaire exhaustiva, pot semblar que el model de Fourier comporta una complexitat força més elevada que un model de Volterra del mateix ordre ja que en lloc de potències conté exponencials complexos aplicades sobre el senyal d'entrada.

$$\hat{g}^{FOURIER}[x(n)] = \sum_{p=-P}^P c_p \cdot e^{jp\omega_0 x} \equiv \sum_{p=-P}^P c_p \cdot z_p^{FOURIER}(x, n) \quad (3.26.a)$$

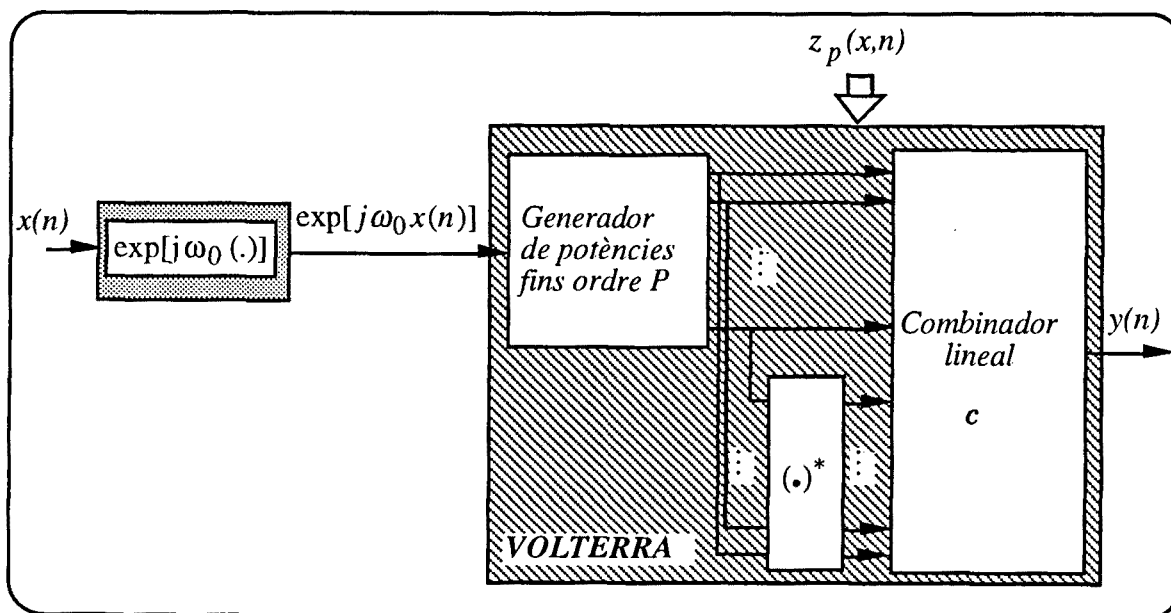
$$\hat{g}^{VOLTERRA}[x(n)] = \sum_{p=0}^P c_p \cdot x^p \equiv \sum_{p=0}^P c_p \cdot z_p^{VOLTERRA}(x, n) \quad (3.26.b)$$

De totes maneres, el model de Fourier es pot interpretar, fonamentalment, com un model de Volterra precedit d'un modulador de freqüència (Fig.3.1) que fonamentalment delimita l'estadística del senyal d'entrada al rang [-1,1].



**Figura 3.1** Interpretació del model de Fourier com una transformació exponencial complexa seguida d'un sistema de Volterra.

Aquesta representació posa de manifest el fet de que tant el cost computacional del model de Fourier com la complexitat de la seva implementació són molt semblants als d'un model de Volterra del mateix ordre (sense considerar el cost d'obtenir l'exponencial complexa inicial). Per tal d'èsser més rigurosos, l'esquema anterior es pot detallar tal i com segueix, on els coeficients  $\{c_p \forall p=-P, \dots, P\}$  s'han ajuntat en el vector  $c$ .



**Figura 3.2** Implementació del model de Fourier útil per comparar complexitat. L'asterisc representa una transformació de complex conjugat.

La comparació de la complexitat del model de Fourier envers la del model de Volterra s'estableix amb vista a una implementació discreta, en termes del nombre de multiplicacions necessàries per a generar un valor del senyal de sortida. Per tal de comptabilitzar aquesta càrrega computacional cal considerar, fonamentalment, dues operacions (Fig.3.2): per una banda, la generació de l'espai de senyal format per les funcions  $z(x,n)$  corresponents a cada

model, i per altra banda, la combinació lineal d'aquestes funcions amb els respectius coeficients,  $c$ .

Respecte de la generació de l'espai de funcions  $z(x,n)$  del model de Volterra sense memòria d'ordre  $P$ , es necessiten un total de  $(P-1)$  multiplicacions per a generar els  $P$  monomis que inclou el model (Eq.3.26.b). Per altra banda, en el cas del model de Fourier només cal generar les  $P$  potències positives de l'exponencial d'ordre 1 (és a dir,  $\forall p=1,2,\dots,P$ ) ja que els termes corresponents a un índex  $p$  negatiu es poden obtenir aplicant el complex conjugat (Fig.3.2). Per tal d'obtenir la càrrega computacional d'aquests  $P$  termes positius s'han de calcular potències de nombres complexos i per tant cada multiplicació suposa 4 multiplicacions reals. En definitiva, es pot trobar fàcilment que són necessàries en total  $(4P-4)$  multiplicacions reals per obtenir totes les exponencials complexes necessàries per a implementar un model de Fourier sense memòria d'ordre  $P$  (Eq.3.26.a). Així, és evident que el model de Fourier tot i que comporta un nombre més elevat d'operacions que el de Volterra (4 cops més, sense considerar el càlcul de la primera exponencial complexa), es manté dins el mateix ordre de magnitud, és a dir,  $O(P)$ .

En aquest sentit però, és important destacar la diferència entre el paper que juga l'ordre  $P$  en un model de Volterra o bé en un model de Fourier. En efecte, mitjançant el desenvolupament en sèrie de Taylor (Eq.2.2) de la funció exponencial complexa d'ordre genèric  $p$ ,

$$e^{jp\omega_0 x} = 1 + jp\omega_0 x + \frac{(jp\omega_0 x)^2}{2} + \dots \quad (3.27)$$

es pot comprovar que el model de Fourier d'ordre  $P$  (Eq.3.18) en realitat inclou potències del senyal d'entrada fins a un ordre infinit, tot i que té un nombre limitat de graus de llibertat.

En cas de que el SNL que s'ha de modelar segueixi una simetria parell o senar, tant el model de Volterra com el de Fourier admeten una simplificació que repercuteix en els respectius costos computacionals. Així, el model de Volterra necessita aproximadament la meitat de multiplicacions per a generar el seu nou espai de dades format ara per potències parells del senyal d'entrada si  $g(x)$  té simetria parell, o bé per potències senars si  $g(x)$  té simetria senar. El model de Fourier, en canvi, no experimenta cap reducció del nombre d'operacions necessàries per obtenir el nou espai de dades ja que, en principi, sembla igual de costós calcular les funcions cosinus del model parell (Eq.3.23), o bé les funcions sinus del model senar (Eq.3.24), com la part real i imaginària, respectivament, de les exponencials complexes del model més general (Eq.3.26.a, Fig.3.2).

Pel que fa al cost computacional de la combinació lineal dels coeficients  $c$  amb les funcions  $z(x,n)$ , el model de Volterra d'ordre  $P$  (Eq.3.26.b) requereix  $(P+1)$  multiplicacions,

mentres que el de Fourier (Eq.3.26.a) en necessita  $(8P+1)$  (els seus coeficients són complexos i cada multiplicació de complexos equival a 4 de reals). De totes maneres, si s'aprofita la simetria hermítica dels coeficients i funcions del model de Fourier, és ben conegut que la mateixa sortida  $y(n)$  es pot obtenir només a partir dels sumands corresponents a índexs positius,

$$y(n) = \Re e \left\{ c_0 + 2 \sum_{p=1}^P c_p \cdot e^{j p \omega_0 x(n)} \right\} \quad (3.28)$$

de forma que estrictament només serien necessàries  $(4P+1)$  multiplicacions reals. Tot i que formalment les expressions anteriors (Eq.3.26.a,3.27) coincideixen, experimentalment s'ha pogut observar que l'esquema corresponent a la primera formulació resulta ésser més robust a errors numèrics.

Per altra banda, respecte dels models que contemplen una certa simetria de la funció del SNL,  $g(x)$ , el model de Volterra necessita de l'ordre de  $O(P/2)$  multiplicacions ja que el nombre de potències i coeficients s'ha vist reduït a la meitat, mentre que el de Fourier en requereix  $P$  o  $(P+1)$  segons el model sigui senar o parell, respectivament, pel fet de que en aquest supòsit els coeficients són reals.

En resum, es pot concloure que el model de Fourier necessita en general 4 cops més multiplicacions que un model de Volterra del mateix ordre  $P$  i que, per tant, la complexitat es manté dins el mateix ordre de magnitud (sense considerar el cost de calcular l'exponencial complexa inicial corresponent a  $p=1$ ). Per últim, i un cop establerta la comparació entre ambdós models respecte de llur complexitat, és interessant comparar-los en l'etapa de disseny dels paràmetres on, tal i com es veurà a continuació, el model de Fourier presenta unes característiques superiors al model de Volterra.

### 3.1.4 Disseny del model de Fourier

El model de Fourier, a l'igual del model de Volterra, és fonamentalment un model paràmetric per a SNLs i per tal de completar el disseny de l'arquitectura que l'implementa (Fig.3.2), cal determinar el valor dels paràmetres que el formen. Tal i com es pot comprovar a l'expressió del model de Fourier següent (Eq.3.29), es poden distingir tres grups diferents de paràmetres: l'ordre del model  $P$ , la freqüència principal  $\omega_0$ , i els coeficients  $w_p$  (que per ser, en general, diferents als coeficients de Fourier  $c_p$ , es denoten a partir d'ara per  $w_p$  i s'ajunten en el vector  $w$ ).

$$\hat{g}(x) = \sum_{p=-P}^{+P} w_p^* \cdot e^{-j p \frac{\pi}{X_0} x} = w^H \cdot z(x, n) \quad (3.29)$$

En primer lloc es troba l'ordre  $P$  que condiona enormement el grau d'aproximació que el model serà capaç d'assolir. En principi, ja s'ha vist que si es tracta d'un SNL amb una funció  $g(x)$  contínua, un increment de l'ordre del model de Fourier ha de repercutir en una disminució de la norma de l'error (Eq.3.10, 3.12). De totes maneres, fins i tot en cas de que  $g(x)$  sigui contínua, aquesta convergència uniforme de la norma de l'error cap a zero a mesura que  $P$  augmenta es deixa de complir si els coeficients del model que formen el vector  $w$  no coincideixen amb els de la sèrie de Fourier  $c_p$  (Eq.3.21). En general, en una situació pràctica no es disposa de la funció  $g(x)$  necessària per a calcular els coeficients de Fourier, de manera que interessa un ordre el més reduït possible per tal de minimitzar el cost computacional que requereix el model, però alhora suficientment elevat com per a assolir un grau d'aproximació satisfactori per a l'aplicació concreta. De la mateixa manera que en el model de Volterra, l'ordre  $P$  en el model de Fourier es decideix prèviament prenent en consideració limitacions tecnològiques referents al nombre d'operacions per mostra o bé, en base a unes simulacions efectuades anteriorment que determinen un ordre adequat a partir del qual, la millora que s'obté pel fet d'incrementar-lo no compensa l'augment d'operacions necessàries que comporta. En aquest sentit, també existeix el test de Fisher [She76] que proporciona un criteri per a la selecció del nombre d'harmònics que inclou un model d'exponencials complexes.

Per altra banda, es troba l'anomenada freqüència principal que està limitada per un valor inversament proporcional al rang del senyal d'entrada per tal d'evitar una superposició de  $g(x)$  en la funció  $\hat{g}(x)$  implementada pel model de Fourier.

$$\omega_0 \leq \frac{2\pi}{2X_0} \quad (3.30)$$

Aquest paràmetre també es determina inicialment a partir del valor del marge dinàmic del senyal d'entrada del SNL que es vol modelar. En determinats problemes però, no es disposa d'aquest rang i s'haurà de fixar un valor  $X_0$  elevat i conservador, que en principi repercutirà en una aproximació pobre en resolució. Tal i com es veurà a les simulacions, el més adequat generalment és escollir una freqüència principal propera al valor màxim permès (Eq.3.28), i en cas de que els valors de la funció  $g(x)$  als extrems del rang d'entrada siguin diferents sovint és interessant una  $\omega_0$  menor per evitar el fenòmen de Gibbs que allà s'hi produeix.

Per últim, però potser la part més interessant ja que és on el model de Fourier presenta unes característiques rellevants, és el disseny dels coeficients  $w_p$  que formen el vector de coeficients  $w$  i que, pel cas de la sèrie de Fourier, resulten iguals als coeficients de Fourier conjugats ja que produeixen la combinació lineal de les exponencials complexes que minimitza la norma de l'error (Eq.3.21)

$$w_p \equiv c_p^* \quad p = -P, \dots, P \Rightarrow \int_{-X_0}^{X_0} |e(x)|^2 dx \Big|_{MINIM} \quad (3.31)$$

En un cas general, el desconeixement de  $g(x)$  no permet utilitzar l'expressió (3.21) per a calcular aquests coeficients i, en conseqüència, serà necessari establir un nou criteri de disseny per tal de determinar-los. En definitiva, es pot dir que la teoria de les sèries de Fourier (a l'igual del teorema de Stone-Weierstrass envers el model de Volterra), el que proporciona és la certesa de que existeix un conjunt de coeficients  $w_p = (c_p)^*$  que minimitzen la norma de la funció error sempre i quan la sèrie de Fourier de  $g(x)$  existeixi.

En aquest contexte del criteri de disseny dels coeficients, cal destacar el treball fet per Mulgrew [Mul94] en el modelat de SNLs a partir també d'una sèrie de Fourier però dins d'un espai de funcions on la norma es defineix com,

$$\langle f(x), f(x) \rangle = \int_{-\infty}^{+\infty} f(x) \cdot f^*(x) \cdot p_X(x) dx \quad (3.32)$$

essent  $p_X(x)$  la funció densitat de probabilitat del senyal d'entrada. Les funcions  $\{\phi_p(x)\}$  que formen la base del model que proposa no són doncs exactament les exponencials complexes sino que estan dividides per l'arrel quadrada de  $p_X(x)$  per tal de que siguin ortonormals en la norma definida. Per tant, Mulgrew proposa un model de SNLs sense memòria en el que les funcions que utilitza són ortonormals i que, per tant, el disseny dels coeficients es simplifica molt sota un criteri de MMSE de Wiener (Ap.2.2). El problema del model radica en el fet de que és necessari conèixer l'estadística del senyal o, en el seu defecte, estimar-la per a calcular les funcions del model. Per altra banda, l'extensió al modelat de SNLs amb memòria encara es preveu més complicada ja que en aquest cas es necessitaria saber o estimar la probabilitat conjunta de la diversitat temporal del senyal d'entrada.

En el cas que ens ocupa, els coeficients del model de Fourier  $\{w_p \quad p = -P, \dots, P\}$  es dissenyaran de manera que minimitzin la mateixa norma de la funció error que la utilitzada a [Mul94].

$$\int_{-X_0}^{X_0} |g(x) - \hat{g}(x)|^2 \cdot p_X(x) dx \Big|_{MINIM} \Rightarrow w_p = w_{p,optim} \quad (3.33)$$

A diferència del treball de Mulgrew on es proposa un conjunt de funcions ortonormals segons la norma definida (Eq.3.32), el model de Fourier per SNLs continua essent una combinació lineal de les exponencials complexes amb els nous coeficients,



$$\hat{g}(x) = \sum_{p=-P}^P w_p^* \cdot e^{jp\omega_0 x} \quad (3.34)$$

permetent la següent formulació vectorial,

$$\hat{g}(x) = \mathbf{w}^H \cdot \mathbf{z} \quad (3.35)$$

que inclou el producte del vector que agrupa els coeficients  $\mathbf{w}^H$  i el vector  $\mathbf{z}$  que conté les exponencials d'índexs succesius.

$$\mathbf{w} = [w_{-P} \quad w_{-P+1} \quad \dots \quad w_P] \quad (3.36.a)$$

$$\mathbf{z} = [e^{-jP\omega_0 x} \quad e^{-j(P-1)\omega_0 x} \quad \dots \quad e^{jP\omega_0 x}] \quad (3.36.b)$$

En definitiva, es pot veure com el model de Fourier es tracta d'un model de SNLs però lineal amb els paràmetres  $w_p$ , aconseguint d'aquesta manera que la funció de cost (Eq.3.37, 3.33) que s'ha de minimitzar sigui una forma quadràtica en termes de  $\mathbf{w}$ .

$$\xi_{\mathbf{w}} = E \left[ |g(x) - \mathbf{w}^H \cdot \mathbf{z}|^2 \right] = E \left[ |g(x)|^2 \right] - \mathbf{w}^H E \left[ \mathbf{z} \cdot g^*(x) \right] - E \left[ g(x) \cdot \mathbf{z}^H \right] \mathbf{w} + \mathbf{w}^H E \left[ \mathbf{z} \cdot \mathbf{z}^H \right] \mathbf{w} \quad (3.37)$$

Tal i com ja es va comentar a la secció (2.2) del capítol anterior, la funció de cost  $\xi_{\mathbf{w}}$  té una forma convexa en funció dels coeficients  $\mathbf{w}$  sempre i quan la matriu d'autocorrelació del nou espai de dades generat pel vector  $\mathbf{z}$ , és a dir  $\mathbf{R}_{\mathbf{z}\mathbf{z}} = E[\mathbf{z}^H \cdot \mathbf{z}]$ , sigui definida positiva. Sota aquesta hipòtesi, que més tard es discutirà, el conjunt de pesos que minimitzen aquesta funció de cost s'obté igualant a zero el gradient de  $\xi_{\mathbf{w}}$  respecte del vector de coeficients,

$$\xi_{\mathbf{w}} \Big|_{MINIM} \Rightarrow \nabla_{\mathbf{w}^H} \xi_{\mathbf{w}} = \mathbf{R}_{\mathbf{z}\mathbf{z}} \cdot \mathbf{w} - E \left[ g^*(x) \cdot \mathbf{z} \right] \equiv 0 \quad (3.38)$$

arribant a la següent expressió pels coeficients que es denoten amb el subíndex *opt* ja que són òptims en el sentit dictat per Wiener.

$$\mathbf{w}_{opt} = \mathbf{R}_{\mathbf{z}\mathbf{z}}^{-1} \cdot E \left[ g^*(x) \cdot \mathbf{z} \right] \quad (3.39)$$

Tot i que aquesta expressió ja s'havia trobat pel cas general d'un model lineal amb els paràmetres (Ap.2.2), s'ha considerat oportú desenvolupar-la de nou pel cas concret del model de Fourier ja que es tracta de l'eix principal d'aquesta tesi. A la solució de pesos òptima trobada es posa de relleu novament la necessitat de la no singularitat de la matriu d'autocorrelació  $\mathbf{R}_{\mathbf{z}\mathbf{z}}$ . Donat que en el cas particular del model de Fourier, el vector  $\mathbf{z}$  està format per les exponencials

complexes (Eq.3.36.b), cada component de la matriu d'autocorrelació resulta ésser el producte escalar de dues d'aquestes funcions complexes,

$$\mathbf{R}_{zz}(m,l) = E\left[e^{j(m-l)\omega_0 x}\right] \equiv \psi_X(j(m-l)\omega_0) \quad (3.40)$$

coincidint amb l'anomenada *funció característica* d'una variable aleatòria  $X$  denotada per  $\psi_X(jv)$  [Pap65], però avaluada a un múltiple enter de  $\omega_0$ .

$$\psi_X(jv) \equiv \left\{ \begin{array}{l} \text{Funció} \\ \text{Característica} \end{array} \right\} = E\left[e^{jvx}\right] = \int_{-\infty}^{+\infty} e^{jvx} \cdot p_X(x) dx \quad (3.41)$$

Aquesta relació entre les components de la matriu d'autocorrelació del model de Fourier (que també anomenarem *matriu característica*) i la funció característica de  $x$  és mol interessant ja que dota a  $\mathbf{R}_{zz}$  de determinades propietats que la diferencien d'una manera rellevant de la matriu d'autocorrelació generada pel model de Volterra. Aquesta relació, així com tot el que se'n pot deduir d'ella, és el que ha motivat el quart capítol d'aquesta tesi dedicat fonamentalment a l'estimació de funcions densitat de probabilitat amb mètodes d'estimació espectral. De totes maneres, ara per ara el que interessa és estudiar la matriu d'autocorrelació que s'ha d'invertir per trobar els pesos òptims del model de Fourier. No és gaire complicat comprovar que la matriu d'autocorrelació definida per (Eq.3.40) consisteix en una matriu hermítica gràcies a l'hermiticitat de la funció característica i a més a més és Toeplitz.

$$\begin{aligned} \text{Hermiticitat} &\Rightarrow \mathbf{R}_{zz}(m,l) = \mathbf{R}_{zz}^*(l,m) \\ \text{Toeplitz} &\Rightarrow \mathbf{R}_{zz}(m,l) = \mathbf{R}_{zz}(m+1,l+1) \end{aligned} \quad (3.42)$$

Per altra banda, la funció característica es pot interpretar com la transformada inversa de Fourier de la funció densitat de probabilitat de  $x$  (Eq.3.41) de manera que pel fet d'ésser aquesta darrera real i positiva, la funció característica actua com una funció d'autocorrelació. D'entre elles, de moment, destaquem la particularitat d'ésser una funció no creixent i hermítica,

$$\psi_X(0) = 1 \geq |\psi_X(jv)| \quad \forall v \in \mathfrak{R} \quad (3.43.a)$$

$$\psi_X(-jv) = \psi_X^*(jv) \quad \forall v \in \mathfrak{R} \quad (3.43.b)$$

propietats que assegurin que la matriu  $\mathbf{R}_{zz}$  és una matriu definida no negativa. Aquesta propietat de la funció característica juntament amb la particularitat d'ésser una matriu Toeplitz fan que en general, almenys aixó és el que s'ha experimentat a les simulacions, la matriu d'autocorrelació  $\mathbf{R}_{zz}$  del model de Fourier estigui molt millor condicionada que la generada pel model de Volterra que conté moments d'ordre superior de la variable d'entrada  $x$ . Malgrat que

cada component de la matriu d'autocorrelació del model de Fourier es pot expressar en funció dels moments de la variable d'entrada  $x$ ,

$$\psi_X(j\nu) = E[e^{j\nu x}] = E\left[\sum_{i=0}^{+\infty} \frac{1}{i!} (j\nu)^i x^i\right] = \sum_{i=0}^{+\infty} \frac{1}{i!} (j\nu)^i E[x^i] \quad (3.44)$$

aquesta darrera relació sembla que no aporta res sobre aquest suposat millor condicionament de la matriu d'autocorrelació del model de Fourier envers la del model de Volterra. Cal esmentar però, que aquest comportament del model de Fourier observat en les simulacions implica un aprenentatge més efectiu en dissenys adaptatius. De totes maneres, abans de discutir aspectes relacionats amb el disseny adaptatiu seria interessant parar atenció al disseny del model de Fourier en bloc, tant des d'un punt de vista computacional com prenent en consideració l'estimació de les funcions necessàries.

#### 3.1.4.1. Disseny de mínims quadrats del model de Fourier sense memòria

Suposem un cas simple d'identificació d'un SNL sense memòria per tal d'il·lustrar el disseny de mínims quadrats del model de Fourier d'ordre  $P$  (Fig.3.3)

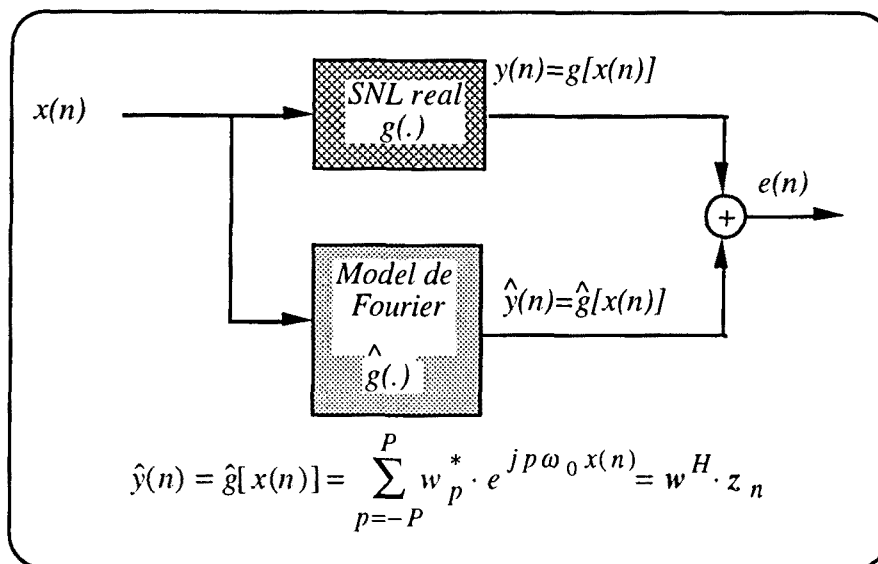


Figura 3.3 Identificació d'un SNL sense memòria mitjançant un model de Fourier d'ordre  $P$ .

En aquest cas, la variació temporal de l'entrada i sortida del SNL real (suposarem que es disposa de  $M$  mostres) proporciona una seqüència d'error denotada per  $e(n)$ ,

$$e(n) = y(n) - \sum_{p=-P}^{+P} w_p^* \cdot e^{j p \omega_0 x(n)} = y(n) - \mathbf{w}^H \cdot \mathbf{z}_n \quad (3.45)$$

en funció del vector  $\mathbf{z}_n$  format per les exponencials complexes de l'expressió (3.36.b) però avaluades a  $x=x(n)$ . El criteri de disseny de mínims quadrats (o, en anglès "Least Squares") del conjunt de coeficients  $\{w_p \ p=-P, \dots, P\}$  del model de Fourier es basa en la minimització de la funció  $\xi_M^{LS}$ , que representa la suma dels quadrats de la seqüència d'error generada pel model.

$$\xi_M^{LS} = \frac{1}{M} \sum_{n=0}^{M-1} |e(n)|^2 \quad (3.46)$$

El vector de coeficients  $\mathbf{w}$  que minimitza la funció de cost  $\xi_M^{LS}$  ha de complir el següent sistema lineal d'equacions que sorgeix d'igualar a 0 el gradient de  $\xi_M^{LS}$  respecte del vector de coeficients hermític  $\mathbf{w}^H$ ,

$$-\frac{1}{M} \sum_{n=0}^{M-1} \left[ \left( y^*(n) - \mathbf{z}_n^H \cdot \mathbf{w} \right) \cdot \mathbf{z}_n \right] = 0 \quad (3.47)$$

i que agrupant els termes convenientment es pot expressar tal i com segueix.

$$\left( \frac{1}{M} \sum_{n=0}^{M-1} \mathbf{z}_n \cdot \mathbf{z}_n^H \right) \cdot \mathbf{w} - \left( \frac{1}{M} \sum_{n=0}^{M-1} y^*(n) \cdot \mathbf{z}_n \right) = 0 \quad (3.48)$$

El primer terme, que multiplica al vector de pesos  $\mathbf{w}$ , correspon a una matriu que denotarem per  $\hat{\mathbf{R}}_{zz}$ ,

$$\hat{\mathbf{R}}_{zz} = \frac{1}{M} \sum_{n=0}^{M-1} \mathbf{z}_n \cdot \mathbf{z}_n^H \quad (3.49)$$

ja que en efecte es pot interpretar com una aproximació de la matriu d'autocorrelació  $\mathbf{R}_{zz}$  del model de Fourier (Eq.3.40) sempre que l'estadística del senyal d'entrada no canviï amb el temps i els succesius valors de  $x(n)$  siguin independents. Així, cada component de la matriu  $\hat{\mathbf{R}}_{zz}$  coincideix amb l'estimador de la funció característica de  $x$  conegut amb el nom de *Funció Característica Empírica* (FCE) [Stu94]. (les propietats estadístiques del qual es detallen a l'apèndix (Apèn.4.A) del següent capítol),

$$\hat{\mathbf{R}}_{zz}(l, m) = \hat{\psi}[(l-m)\omega_0] = \left\{ \begin{array}{l} \text{Estimador} \\ \text{FCE} \end{array} \right\} \equiv \frac{1}{M} \sum_{n=0}^{M-1} e^{j(l-m)\omega_0 x(n)} \quad \forall l, m = 1, \dots, N_c \quad (3.50)$$

que utilitza  $M$  valors de la variable aleatòria  $x$ , denotats per  $x(n)$  per a realitzar l'estimació. És important destacar que aquest estimador (Eq.3.50) no està esbiaixat i, a més a més, manté les propietats d'hermiticitat i de simetria Toeplitz de la matriu d'autocorrelació, importants des d'un punt de vista de cost computacional i estabilitat numèrica. En efecte, el fet de que la matriu  $\hat{\mathbf{R}}_{zz}$  proporcionada per l'estimador FCE sigui hermítica i Toeplitz permet que el càlcul de la seva inversa requereixi només de l'ordre de  $O(N_c^2)$  operacions (amb mètodes clàssics el creixement seria cúbic), essent  $N_c$  el nombre de coeficients del model, és a dir  $2P+1$ . Cal esmentar l'existència d'altres mètodes d'estimació de la funció característica [Arn95] però que queden fora de l'abast d'aquesta tesi.

Per altra banda, resta el segon terme de la solució de mínims quadrats (Eq.3.48) que es representa pel vector  $\hat{\gamma}_{gz}$ ,

$$\hat{\gamma}_{gz} = \frac{1}{M} \sum_{n=0}^{M-1} y^*(n) \cdot z_n \quad (3.51)$$

ja que quan l'estadística del senyal d'entrada i de sortida no canvien amb el temps aquest terme es pot interpretar com un estimador no esbiaixat de la correlació creuada teòrica,  $\gamma_{gz}$ , entre la funció del SNL real i el vector d'exponencials complexes (Eq.3.39). En definitiva, la solució de mínims quadrats

$$w_{opt}^{LS} = \hat{\mathbf{R}}_{zz}^{-1} \cdot \hat{\gamma}_{gz} \quad (3.52)$$

pot interpretar-se, en un entorn estacionari, com un estimador de la solució de Wiener (Eq.3.39) i que en qualsevol cas assoleix un valor mínim en la funció de cost igual a la següent expressió.

$$\xi_M^{LS} \Big|_{MINIM} = \frac{1}{M} \sum_{n=0}^{M-1} |y(n)|^2 - \hat{\gamma}_{gz}^H \cdot \hat{\mathbf{R}}_{zz}^{-1} \cdot \hat{\gamma}_{gz} \quad (3.53)$$

És important remarcar que, en presència d'un soroll additiu a la sortida del SNL real, la solució de mínims quadrats (Eq.3.52) manté les propietats estadístiques que se li atribueixen en el disseny de sistemes lineals (veure per exemple [Hay91]). A l'apèndix (Apèn.3.A) es reproduïxen les que fan referència a l'esbiaix i covariància. Destaca el fet de que l'estimació del vector de coeficients és no esbiaixada si el soroll és de mitja nul·la, així com la següent expressió per a la covariància,

$$\text{cov}[\hat{w}^{LS}] = \sigma_\epsilon^2 \cdot \mathbf{R}_{zz}^{-1} \quad (3.54)$$

vàlida sempre que el soroll sigui, a més a més, independent de l'entrada i blanc, amb una variància denotada per  $\sigma_\varepsilon^2$  [Hay??] (el vector  $\hat{w}^{LS}$  representa l'estimació dels coeficients proporcionada per l'estimador de mínims quadrats (Eq.3.52) en presència de soroll additiu). En aquest sentit, sorgeix un altre cop el fet de que un mal condicionament de la matriu  $R_{zz}$  repercuteix en una variància elevada en l'estimació dels coeficients.

En vistes a les simulacions, l'apèndix (Apèn.3.B) inclou l'error quadràtic mínim que assolix l'estimador de mínims quadrats, denotat per  $\hat{\xi}_M^{LS}$ , en presència de soroll de mitjana nul·la, blanc i independent del senyal d'entrada. La mitjana d'aquest valor equival a la següent expressió (vàlida per a  $M \geq (2P+1)$ ),

$$E[\hat{\xi}_M^{LS}] = \sigma_\varepsilon^2 \cdot \left(1 - \frac{2P+1}{M}\right) \quad (3.55)$$

en la que es pot observar com a mesura que el nombre de dades disponibles  $M$  augmenta, el valor esperat de l'error quadràtic tendeix cap a la potència de soroll. Aquest resultat serà d'utilitat en les simulacions (Ap.5.1) on es treballa sota la premisa d'un soroll additiu a la sortida del SNL real.

Deixant de banda la influència del soroll, el disseny de mínims quadrats del model de Fourier (Eq.3.52) representa un disseny en bloc ja que es necessita d'un recull de  $M$  mostres del senyal d'entrada i sortida del SNL que s'identifica (Fig.3.3) per a obtenir el vector de coeficients que determinen el model de Fourier emprat. Aquesta mateixa solució també es pot obtenir de forma recursiva a mesura que noves dades de  $x(n)$  i  $y(n)$  són disponibles, tot i que aquests mètodes es classificarien dins el camp de dissenys adaptatius que es presenta a continuació.

### 3.1.4.2 Disseny adaptatiu del model de Fourier

El treball de Mathews [Mat91] va posar en evidència la possibilitat d'utilitzar mètodes adaptatius clàssics, ja desenvolupats en problemes de processament lineal del senyal, però aplicats al disseny adaptatiu de models polinòmics o de Volterra. Així, al capítol anterior (Ap.2.4) s'ha mostrat que el model de Volterra, directe o indirecte mitjançant estructures lattice, pot dissenyar-se amb algorismes adaptatius pel sol fet d'ésser lineal amb els coeficients. De la mateixa manera, el model de Fourier presentat (Eq.3.34) també és lineal amb els coeficients, fet que permet l'ús de mètodes adaptatius tant recursius com de gradient per conduir el disseny dels seus coeficients cap al valor òptim proporcionat per la solució de Wiener (Eq.3.39) o la de mínims quadrats (Eq.3.52).

Tal i com indica el seu nom, els algorismes adaptatius recursius calculen recursivament la solució de mínims quadrats i, en general, es caracteritzen per una convergència més independent de l'estadística del senyal, en comparació amb els mètodes de gradient. Això fa que en el cas dels sistemes de Volterra, i com a conseqüència del mal condicionament de la seva matriu d'autocorrelació  $\mathbf{R}_{zz}$ , és freqüent l'ús de mètodes adaptatius recursius ja que els de gradient sovint proporcionen un comportament inadequat. En contrapartida, el cost computacional és força elevat, de l'ordre de  $O(N_c^2)$  essent  $N_c$  el nombre de coeficients del model, i a més a més, en general els mètodes recursius tenen problemes d'inestabilitat numèrica.

En el disseny adaptatiu del model de Fourier només es parlarà atenció als mètodes de gradient, concretament l'algorisme NLMS, tot i que per descomptat també es podrien utilitzar de recursius. La raó està en el fet que en aquest cas es poden aprofitar certes propietats del model de Fourier per sobrepassar les prestacions del model de Volterra pel que fa a velocitat i desajust final sense un increment important del cost computacional. Així, els pesos en el model de Fourier són modificats a cada iteració en el sentit invers del gradient de la funció de cost estadística,

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu \cdot \nabla_{\mathbf{w}} \xi \quad (3.56)$$

que en el cas de l'algorisme LMS adopta el seu valor instantani en funció del senyal error  $e(n)$  (Eq.3.45).

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu \cdot e^*(n) \cdot \mathbf{z}_n \quad (3.57)$$

Per tal de fer menys dependent l'evolució dels pesos de la variància del senyal d'entrada, el LMS Normalitzat (NLMS) inclou una normalització de la constant de pas  $\mu$  respecte d'una estimació de la potència del senyal d'entrada, denotada per  $p(n)$ .

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \frac{\mu}{p(n)} \cdot e^*(n) \cdot \mathbf{z}_n \quad (3.58)$$

Aquesta estimació de potència generalment es duu a terme mitjançant un filtre passa-baix dels successius valors de potència instantània del senyal d'entrada, i en funció del que comunament s'anomena factor de memòria  $\beta$ .

$$p(n) = \beta \cdot p(n-1) + (1-\beta) \cdot \mathbf{z}_n^H \mathbf{z}_n \quad (3.59)$$

S'han de destacar dos aspectes rellevants del model de Fourier en relació a un disseny adaptatiu amb el NLMS. Per començar, ja s'ha comentat anteriorment que la matriu d'autocorrelació associada al model de Fourier està, en general i pel que es podrà observar a les

simulacions, molt millor condicionada que la del model de Volterra. Això fa que la dispersió d'autovalors no sigui tan elevada, proporcionant així un millor comportament respecte de la velocitat de convergència i desajust final assolit. Per altra banda, està el fet de que les funcions exponencials complexes que inclou el model de Fourier són totes de potència unitat, propietat que facilita l'estimació de la potència  $p(n)$  que esdevé igual al nombre de coeficients del model. Aquesta característica és decisiva en un disseny adaptatiu ja que, és ben conegut, com la dispersió dels valors de potència de les diferents funcions (tal i com succeeix en el model de Volterra) dificulta enormement la selecció de paràmetres en el disseny adaptatiu (concretament les variables  $\mu$  i  $\beta$ ), així com el propi funcionament de l'etapa d'aprenentatge.

### 3.1.5 Perspectiva general del model de Fourier per SNLs sense memòria

El model de Fourier per SNLs sense memòria està basat en l'aproximació de la relació entrada/sortida del SNL a modelar mitjançant la sèrie trigonomètrica. Malgrat que els nuclis d'aquest model (és a dir, les exponencials complexes) no són ortogonals en la norma definida ja que el producte escalar depèn de l'estadística del senyal d'entrada, el model de Fourier presenta certes propietats que el configuren com un model alternatiu i capaç de competir amb el model polinòmic o de Volterra en determinades situacions. Bàsicament dos aspectes són els que determinen aquesta conclusió. Per una banda està el fet de que el cost computacional del model de Fourier, tot i ésser major que el de Volterra, es manté dins el mateix ordre de magnitud. L'altre aspecte és la menor correlació que, en general, s'observa experimentalment entre els diferents nuclis del model de Fourier que entre els del model de Volterra, i la major uniformitat de llur normes. Tal i com es mostra a l'apartat (5.1.1) del cinquè capítol, íntegrament dedicat a simulacions, tot això fa que el model de Fourier mostri un millor comportament que el model de Volterra, especialment en dissenys adaptatius amb mètodes de gradient i sense un increment substancial del cost computacional.

A part d'això, també és molt important destacar que a l'igual del model de Volterra, el model de Fourier permet un ampli ventall de transformacions per millorar les seves prestacions o bé adequar-se a cada problema concret. En aquest darrer apartat (Ap.3.2) s'ha vist com, per exemple, certes simetries en la funció  $g(x)$  repercuteixen en una major simplicitat del model. Així, i en un intent d'equiparar les possibilitats del model de Fourier al de Volterra, a continuació es presenta un desenvolupament equivalent al presentat però generant un model de Fourier per a modelar SNLs amb memòria.



## 3.2 Model de Fourier per sistemes no lineals amb memòria

El fet de que el SNL que s'ha de modelar contingui memòria implica que el valor de la sortida no només depèn del valor actual de l'entrada  $x(n)$  sino que també depèn de certs valors anteriors de la mateixa entrada.

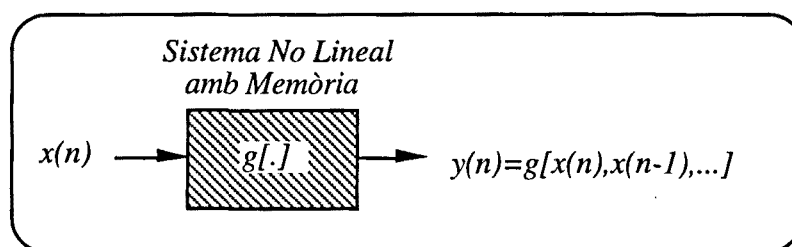


Figura 3.4 Dependència temporal en la relació entrada/sortida d'un Sistema No Lineal amb Memòria

Concretament, quan es diu que la memòria d'un SNL és d'ordre  $N$  significa que el senyal de sortida d'aquest SNL depèn tant de  $x(n)$  com dels  $(N-1)$  valors anteriors, de manera que aquesta diversitat temporal de l'entrada s'agrupa en el que d'ara endavant s'anomena vector d'entrada  $\mathbf{x}$ .

$$\mathbf{x}_n = [x(n) \quad x(n-1) \quad \dots \quad x(n-N+1)]^T \quad (3.60)$$

Tal i com mostra la figura (Fig.3.4), el sistema amb el qual ens enfrontem consta d'una entrada múltiple, denotada per  $\mathbf{x}$ , i només una sortida,  $y(n)$ . És interessant destacar, tot i que no s'entrarà en més detalls, que el model de Fourier que tot seguit es presenta és igualment vàlid per a SNLs de múltiple entrada i una única sortida, encara que l'entrada no representi les diferents components de la diversitat temporal d'un mateix senyal. És més, fins i tot en el cas d'un SNL amb més d'una sortida (el que en anglès es coneix pel nom de sistema MIMO, "Multiple Input / Multiple Output"), el model de Fourier també serviria si es considerés cada sortida com una funció de les múltiples entrades i cadascuna es modelés amb el model de Fourier basat en la sèrie de Fourier multidimensional que a continuació es presenta.

### 3.2.1 Sèrie de Fourier multidimensional

Donada la funció denotada per  $g[\cdot]$  que s'aplica sobre un vector  $\mathbf{x}$  de dimensió genèrica  $N$  que conté les variables  $\{x_m \forall m=1, \dots, N\}$ , i sota les mateixes hipòtesis de funció contínua per seccions que en el teorema (3.1) pel cas unidimensional, aquesta funció permet un desenvolupament en sèrie de Fourier tal i com el que segueix.

$$g[\mathbf{x}] = \sum_{p_1=-\infty}^{+\infty} \sum_{p_2=-\infty}^{+\infty} \dots \sum_{p_N=-\infty}^{+\infty} \left\{ c_{p_1, p_2, \dots, p_N} \cdot e^{j(p_1 \omega_{0(1)} x_1 + p_2 \omega_{0(2)} x_2 + \dots + p_N \omega_{0(N)} x_N)} \right\} \quad (3.61)$$

Cada freqüència principal  $\{\omega_{0(m)} \forall m=1, \dots, N\}$  està condicionada pel rang de la variable d'entrada corresponent i per tant, a l'igual del cas unidimensional, les variables d'entrada  $x_m$  han de pertànyer a un rang limitat  $[-X_m, X_m]$ , que sense pèrdua de generalitat es suposa simètric.

$$\omega_{0(m)} = \frac{\pi}{X_{0(m)}} \leq \frac{\pi}{X_m} \quad \forall m = 1, \dots, N \quad (3.62)$$

Tal i com s'observa a l'expressió (3.61), la sèrie de Fourier està formada a partir d'un conjunt infinit de funcions exponencials complexes que són ortogonals entre elles en la següent norma,

$$\|f(\mathbf{x})\|^2 = \langle f(\mathbf{x}), f(\mathbf{x}) \rangle \equiv \int_{-X_{0(1)}}^{X_{0(1)}} \int_{-X_{0(2)}}^{X_{0(2)}} \dots \int_{-X_{0(N)}}^{X_{0(N)}} f(\mathbf{x}) \cdot f^*(\mathbf{x}) dx_1 dx_2 \dots dx_N \equiv \int_{X_0^N} |f(\mathbf{x})|^2 dx \quad (3.63)$$

on  $X_0^N$  representa la regió d'integració delimitada inferiorment pels extrems del rang dinàmic de cada variable d'entrada  $x_m$ . A partir de la propietat d'ortogonalitat, els anomenats coeficients o constants de Fourier pel cas multidimensional es poden obtenir com el producte escalar entre la funció  $g[\mathbf{x}]$  i l'exponencial associada a aquell coeficient.

$$c_p = \left\langle g(\mathbf{x}), e^{j \sum_{m=1}^N p_m \omega_{0(m)} x_m} \right\rangle \equiv \int_{X_0^N} g(\mathbf{x}) \cdot e^{-j(p_1 \omega_{0(1)} x_1 + p_2 \omega_{0(2)} x_2 + \dots + p_N \omega_{0(N)} x_N)} dx \quad (3.64)$$

Des d'un altra perspectiva, els coeficients  $c_p = c(p_1, p_2, \dots, p_N)$  també poden interpretar-se com els valors obtinguts de mostrejar la funció  $G_e(\omega)$  a múltiples de les freqüències principals, on  $G_e(\omega)$  representa la transformada de Fourier multidimensional de la funció  $g_e[\mathbf{x}]$ .

$$G_e(\omega) = G_e(\omega_1, \omega_2, \dots, \omega_N) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} g_e[\mathbf{x}] \cdot e^{-j\omega' \cdot \mathbf{x}} dx \quad (3.65)$$

La funció  $g_e[\mathbf{x}]$  representa a l'igual del cas unidimensional una prolongació de la relació entrada/sortida del SNL real,  $g[\mathbf{x}]$ , essent zero per tot valor de  $\mathbf{x}$  que es surti del recinte delimitat per  $X_0^N$ .

Es pot observar com la sèrie de Fourier presentada és una extrapol·lació fidel de la sèrie de Fourier d'una sola dimensió al cas multidimensional, mantenint alhora totes les característiques pròpies de la sèrie d'una dimensió [Chu69]. Així, per exemple, la definició dels

coeficients de Fourier (Eq.3.64) està minimitzant la norma de la funció error que aconseguix la sèrie de Fourier. Per altra banda, està el fet de que en cas de que la funció  $g[\mathbf{x}]$  sigui contínua, la sèrie de Fourier parcial (generada igual que a l'expressió (Eq.3.59) però truncant els sumatoris infinits a un nombre finit de termes) convergeix uniformement per a qualsevol vector  $\mathbf{x}$  cap al valor de la funció  $g[\mathbf{x}]$  a mesura que els sumatoris inclouen més i més termes.

De la mateixa manera que pel cas de SNLs sense memòria, aquest desenvolupament en sèrie de Fourier d'una funció multidimensional és la base a partir de la qual es desenvolupa l'anomenat *model de Fourier amb memòria*, que tal i com el seu nom indica és capaç de modelar SNLs amb memòria.

### 3.2.2 Model de Fourier amb memòria

Per tal de generar un model de SNLs amb memòria que sigui realitzable a partir de la sèrie de Fourier multidimensional (Eq.3.61) serà necessari truncar els sumatoris infinits fins a un ordre finit que denotarem per  $P$ . En un cas més general, aquest ordre  $P$  podria ser diferent per cadascun dels índexs  $p_m$ , però per comoditat, i donat que no es perd generalitat en els resultats del desenvolupament, els considerarem tots iguals. Per altra banda, està el fet de que les components del vector d'entrada  $\mathbf{x}$  coincideixen amb valors succesius temporals del senyal d'entrada (Eq.3.60). Amb aquestes premisses i tenint en compte que, a l'igual del cas sense memòria, els pesos del model de Fourier són en general diferents de les constants de Fourier (Eq.3.64), el model de Fourier per modelar un SNL d'ordre  $P$  i memòria  $N$  es defineix de la següent manera.

$$\begin{aligned} \hat{y}(n) &= \hat{g}[\mathbf{x}_n] = \\ &= \sum_{p_1=-P}^{+P} \sum_{p_2=-P}^{+P} \dots \sum_{p_N=-P}^{+P} \left\{ w_{p_1, p_2, \dots, p_N}^* \cdot e^{j\omega_0 [p_1 x(n) + p_2 x(n-1) + \dots + p_N x(n-N+1)]} \right\} \end{aligned} \quad (3.66)$$

És interessant notar que en aquest cas totes les freqüències  $\omega_0(m)$  de la sèrie de Fourier multidimensional són iguals ja que les variables que componen el vector  $\mathbf{x}$  tenen totes el mateix rang.

$$\omega_0 = \frac{2\pi}{2X_0} \quad \text{essent } 2X_0 \geq (X_{\max} - X_{\min}) \quad (3.67)$$

El nombre de coeficients que inclou aquest model creix de forma exponencial a l'igual que succeïa amb el model Volterra per SNLs amb memòria. En efecte, el nombre de coeficients del model és de  $P^N$  (resultat de mostrejar un hipercub  $N$ -dimensional amb  $P$  punts per aresta). Una forma de disminuir el cost computacional que aquest model suposa, és tenir en compte

possibles simetries de la funció  $g[\mathbf{x}]$ . De la mateixa manera que en SNLs sense memòria, si la relació entrada/sortida del SNL és parella,

$$g[\mathbf{x}] = g[-\mathbf{x}] \quad (3.68)$$

el model de Fourier es redueix a una combinació lineal de funcions cosinus més un terme de contínua.

$$\hat{g}[\mathbf{x}_n] = \sum_{p_1=0}^{+P} \sum_{p_2=-P}^{+P} \dots \sum_{p_N=-P}^{+P} \left\{ a_{p_1, p_2, \dots, p_N} \cdot \cos\{\omega_0[p_1 x(n) + p_2 x(n-1) + \dots + p_N x(n-N+1)]\} \right\} \quad (3.69)$$

En canvi, si  $g[\mathbf{x}]$  és senar,

$$g[\mathbf{x}] = -g[-\mathbf{x}] \quad (3.70)$$

només inclou els sinus (per tant, a l'expressió (3.71)  $b_{0,0,\dots,0}=0$ ).

$$\hat{g}[\mathbf{x}_n] = \sum_{p_1=0}^{+P} \sum_{p_2=-P}^{+P} \dots \sum_{p_N=-P}^{+P} \left\{ b_{p_1, p_2, \dots, p_N} \cdot \sin\{\omega_0[p_1 x(n) + p_2 x(n-1) + \dots + p_N x(n-N+1)]\} \right\} \quad (3.71)$$

D'aquesta manera, sempre que es compleixi una simetria parell o senar en el SNL, serà molt interessant aprofitar-la per a utilitzar els models simplificats de funcions cosinus (Eq.3.69) o bé de sinus (Eq.3.71) ja que el nombre de coeficients es redueix a la meitat i, a més a més, també són reals a l'igual que les funcions sinus i cosinus que conformen les funcions que s'utilitzen en el model de Fourier amb simetria.

### 3.2.3 Arquitectura per implementar el model de Fourier amb memòria

En un treball previ, concretament [Pag95], el model de Fourier va ser utilitzat per a desenvolupar una arquitectura per modelar SNLs que seguís en la mesura del possible les pautes marcades pel *teorema de Kolmogorov* [Kol63]. Aquest teorema demostra l'existència d'una expressió basada només en sumes i composicions de funcions univariades vàlida per a modelar qualsevol funció contínua multivariable. Aquest teorema amaga una arquitectura en dues etapes de manera que el nombre de variables intermitjes que connecten la primera amb la segona etapa és de  $(2N+1)$ , essent  $N$  el nombre de variables de les que depèn la funció que es vol reproduir.

Kolmogorov va publicar aquest teorema l'any 1957 [Kol57, Kol63], rebatint el tretzè d'un total de vint postulats que Hilbert va augurar l'any 1901 a París [Hil00] que serien les

grans qüestions a resoldre pels matemàtics del segle XX. L'expectació que, en un principi, el teorema va desencadenar (ja que posava de manifest la possibilitat d'aproximar funcions diguem-ne *complicades*, és a dir multivariables, en termes de funcions més *senzilles*) contrasta amb l'escassetat d'aplicacions que han sorgit ja que Kolmogorov no va proposar quines havien de ser les funcions que conformaven el model. Es tracta doncs d'un teorema d'existència que diversos investigadors, especialment [Spr65], han fet que evolucioni en el sentit d'intentar determinar les funcions que inclou. En l'actualitat i en l'àmbit de processament del senyal, el camp més ampli d'aplicació del teorema resideix en els models basats en xarxes neuronals [Kur92].

En el citat article [Pag95], els esforços van encaminats a desenvolupar un model no lineal anomenat *K-filter* a partir del model de Fourier amb memòria i que segueixi les pautes d'estructura i nombre de variables a les que fa referència el teorema. L'arquitectura resultant, tot i no ser totalment fidel al teorema de Kolmogorov, va resultar satisfactòria gràcies, en gran part, a les característiques del model de Fourier.

Per altra banda, i amb un esquema diferent per tal de comparar el model de Volterra amb el de Fourier, aquest darrer també es pot implementar en base a la configuració LD/NMR en dues etapes presentada al capítol anterior (Fig.2.1), però que reproduïm tot seguit per comoditat.

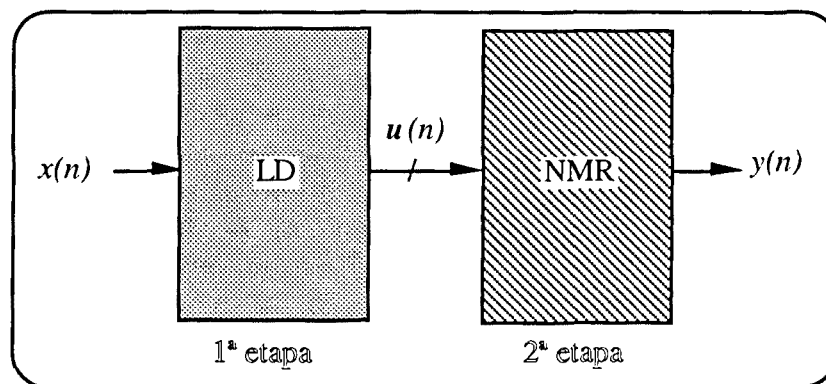


Figura 3.5 Configuració de dues etapes per models no lineals paramètrics ("Linear-Dynamic/Nonlinear-Memoryless-Readout").

Tot i que també es tracta d'una arquitectura en dues etapes com la proposada a [Pag95] pel *K-filter*, la filosofia d'ambdues representacions és força diferent. En el cas que ens ocupa, la primera part ha de contenir la dinàmica o memòria del sistema mitjançant transformacions lineals efectuades sobre l'entrada, mentre que la segona etapa agrupa tot el contingut no lineal del model. En el cas de tractar-se del model de Fourier amb memòria, i seguint les mateixes

directrius que pel cas sense memòria, l'arquitectura resultant sota aquesta representació LD/NMR es pot implementar tal i com segueix.

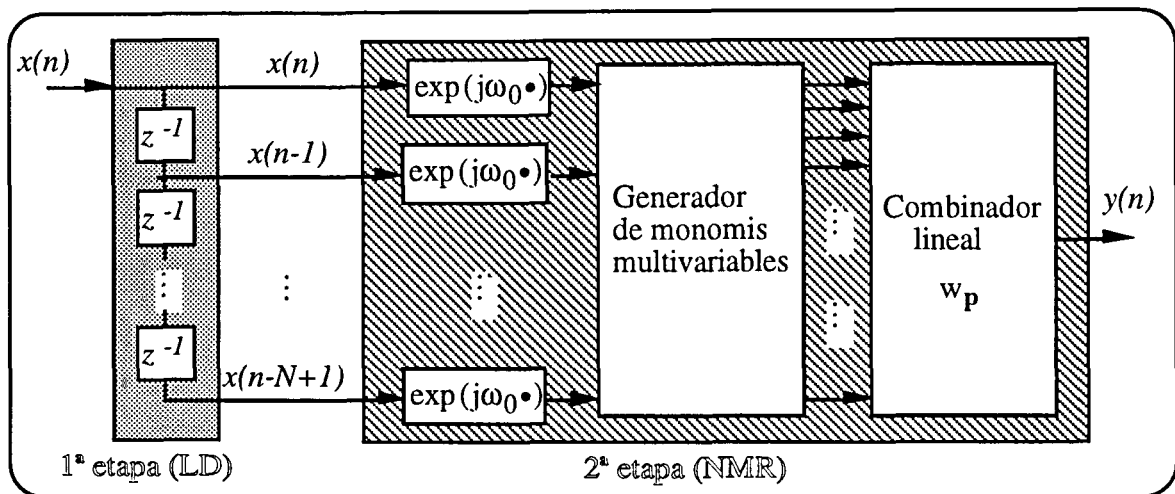


Figura 3.6 Arquitectura LD/NMR per a implementar el model de Fourier amb memòria

La part Lineal i Dinàmica (LD) està dedicada a generar la diversitat temporal del senyal d'entrada, mentre que a la segona etapa es poden diferenciar tres parts. En primer lloc, es modula cadascuna de les components de la diversitat temporal de l'entrada amb un modulador FM (o transformació exponencial complexa en implementacions discretes) de freqüència  $\omega_0$ . Tot seguit, s'introdueixen al generador de monomis multivariables que obté totes les possibles combinacions d'exponencials complexes que conté l'expressió del model (Eq.3.66). En darrer lloc està situat el mòdul que combina linealment aquests termes amb els coeficients del model, denotats per  $\{w_p\}$ , on  $p$  és un vector de dimensió  $N$ ,  $p=[p_1 p_2 \dots p_N]'$ , que pren totes les possibles combinacions dels enters  $\{p_i=-P, \dots, P \forall i=1, \dots, N\}$ .

La semblança entre aquest darrer esquema i el que implementa el model de Volterra ( $P, N$ ) genèric (Fig.2.2) fa preveure que el cost computacional del model de Fourier per a obtenir les exponencials complexes a la sortida del generador de monomis, tot i ser major, no s'incrementa en l'ordre de magnitud de les operacions que necessita un model de Volterra del mateix ordre per generar els monomis (sense considerar el cost de les exponencials d'ordre 1).

Suposem, per exemple, un model de Volterra (3,2) (memòria  $N=2$  i ordre màxim  $P=3$ ). La figura (3.7.a) mostra els monomis generats per aquest model en un gràfic bidimensional, en el que cada eix representa l'ordre al que s'eleva la component  $x(n)$  i  $x(n-1)$ , respectivament. El resultat són 10 termes diferents (és a dir, 10 coeficients reals) d'acord amb l'expressió general (Eq.2.13). Al seu costat, (Fig.3.7.b), es mostra un possible esquema de generació de monomis, semblant al utilitzat en implementacions multicanal per a sistemes de Volterra. En aquest cas particular es comprova com el nombre de multiplicacions necessàries per a generar

l'espai de senyal d'un model de Volterra tendeix, aproximadament i quan  $N$  i  $P$  creixen, al nombre de coeficients del model.

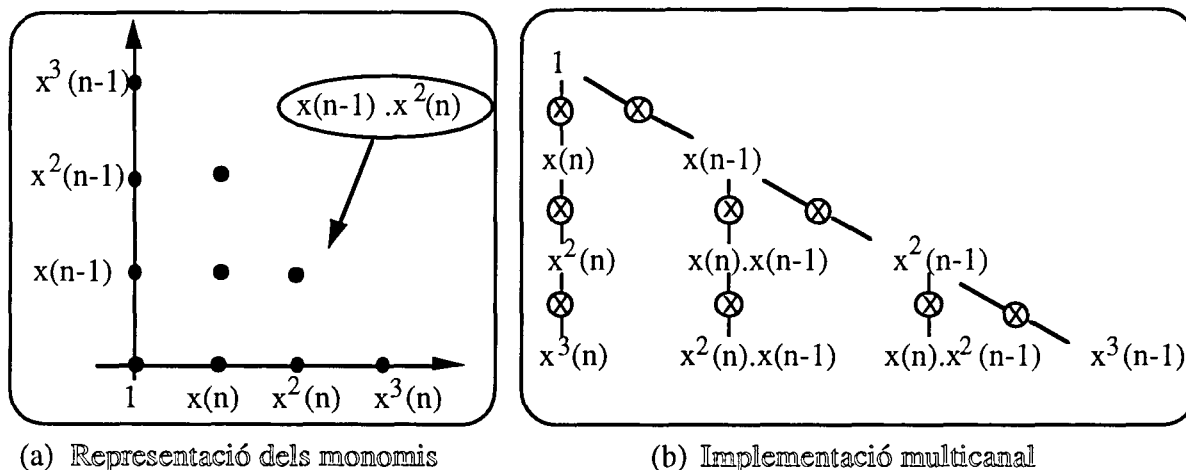


Figura 3.7 (a) Monomis generats en el model de Volterra (3,2) i (b) implementació amb estructura multicanal.

Considerem ara un model de Fourier ( $P=3, N=2$ ) tal i com el de l'expressió (3.66). Per coherència amb Volterra, no es consideren les exponencials corresponents a coeficients  $w(p_1, p_2)$  on  $|p_1+p_2| > P$ . Així, la figura (Fig.3.8) inclou els productes d'exponencials que es tenen en compte en el model, un total de 25 termes complexos, tot i que són hermítics de dos en dos.

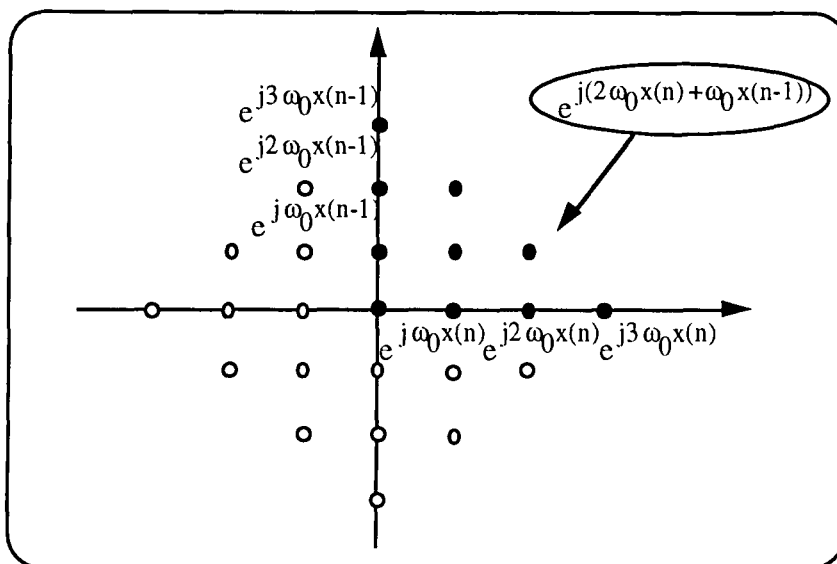


Figura 3.8 Representació bidimensional de les exponencials complexes generades en un model de Fourier(3,2).

La generació de les exponencials del primer quadrant (senyalades a la figura (Fig.3.8) amb negreta) es pot obtenir mitjançant un esquema semblant al cas de Volterra (Fig.3.9 amb

Fig.3.7.a) en el que ara cada multiplicació, en ésser els operands complexos, implica 4 multiplicacions reals.

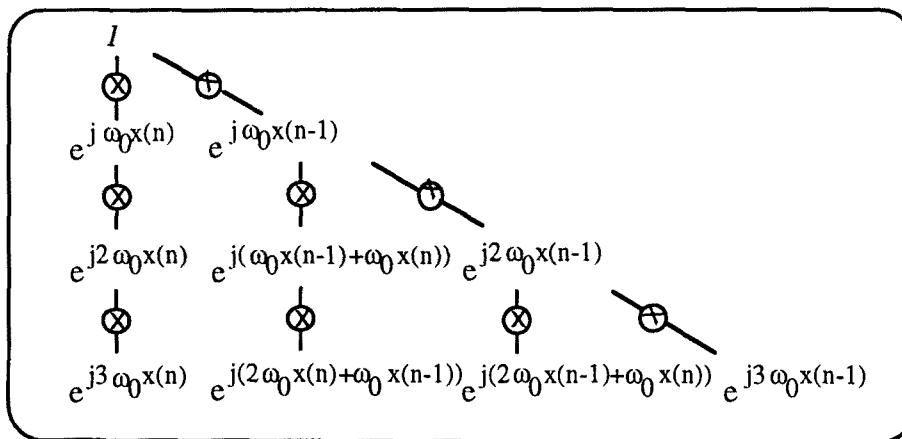


Figura 3.9 Generació de l'espai de senyal del model de Fourier mitjançant una estructura multicanal.

És interessant constatar que les altres exponencials complexos (en aquest cas particular de dues dimensions les corresponents als quadrants II, III i IV) no necessiten cap multiplicació addicional ja que es poden obtenir com a combinacions lineals diferents de les que s'ha hagut d'utilitzar pel primer quadrant. Veiem el següent exemple que il.lustra com amb les exponencials del primer quadrant es poden obtenir les del quadrant II.

$$\begin{aligned}
 e^{j[2\omega_0 x(n)+\omega_0 x(n-1)]} &= \cos(2\omega_0 x_n)\cos(\omega_0 x_{n-1}) - \sin(2\omega_0 x_n)\sin(\omega_0 x_{n-1}) + \\
 &\quad j[\sin(2\omega_0 x_n)\cos(\omega_0 x_{n-1}) + \cos(2\omega_0 x_n)\sin(\omega_0 x_{n-1})] \\
 e^{j[2\omega_0 x(n)-\omega_0 x(n-1)]} &= \cos(2\omega_0 x_n)\cos(\omega_0 x_{n-1}) + \sin(2\omega_0 x_n)\sin(\omega_0 x_{n-1}) + \\
 &\quad j[\sin(2\omega_0 x_n)\cos(\omega_0 x_{n-1}) - \cos(2\omega_0 x_n)\sin(\omega_0 x_{n-1})] \quad (3.72)
 \end{aligned}$$

Respecte de les exponencials complexos del 3<sup>er</sup> i 4<sup>rt</sup> quadrant es poden obtenir a partir del complex conjugat de les dels 1<sup>er</sup> i 2<sup>on</sup>, respectivament.

En general, es pot assegurar que el nombre de multiplicacions reals necessàries per generar l'espai de senyal d'un model de Fourier (P, N) (és a dir, les exponencials complexos) és 4 cops superior al que necessita un model de Volterra del mateix ordre (P, N). Per contra, i com a conseqüència de que els pesos del model de Fourier són complexos i a part experimenten un augment de 2<sup>N</sup> respecte del nombre de coeficients que necessita el model de Volterra, els requeriments en quant a memòria són també més elevats. Així, un model de Volterra (P, N) consta de N<sub>c</sub> = (P+N)! / (P!N!) coeficients i ha d'efectuar de l'ordre de O(N<sub>c</sub>) multiplicacions cada iteració per calcular els monomis. Per altra banda, un model de Fourier del mateix ordre



$(P, N)$  comporta  $2^{N+1}N_c$  coeficients reals diferents (assumint una simetria hermítica, és a dir,  $g[x]$  real), i requereix de l'ordre de  $O(4N_c)$  multiplicacions reals per generar les exponencials complexes.

### 3.2.4 Disseny del model de Fourier amb memòria

La memòria del model controlada per  $N$  és un element nou en el disseny del model de Fourier amb memòria respecte del disseny del model de Fourier per SNLs sense memòria. De la mateixa manera que l'ordre  $P$  i la freqüència principal  $\omega_0$ , el paràmetre  $N$  es determina prèviament en funció de les característiques del SNL real que es pretén modelar i del problema amb el que ens enfrontem. Delimitant superiorment els valors de  $N$  i  $P$  està la complexitat del model que incrementa en quant aquests ho fan (el nombre de coeficients i per tant, el de multiplicacions necessàries per iteració en un disseny discret creix ara al ritme de  $O(P^N)$ ).

Un cop fixats els ordres  $(P, N)$  així com la freqüència principal, que depèn del rang del senyal d'entrada (Ap.3.1.4), el model de Fourier resulta ésser un model paramètric i lineal en termes dels seus coeficients denotats per  $\{w_p\}$ . Per tal de determinar aquests coeficients és interessant per comoditat tornar a l'expressió del model de Fourier amb memòria que inclou tots els coeficients  $w_p$  obtinguts de mostrejar un hipercub de dimensió  $N$  i amb  $(2P+1)$  mostres per aresta.

$$\begin{aligned} \hat{g}[x] &= \sum_{p_1=-P}^{+P} \sum_{p_2=-P}^{+P} \dots \sum_{p_N=-P}^{+P} \left\{ w_{p_1, p_2, \dots, p_N}^* \cdot e^{j\omega_0 [p_1 x(n) + p_2 x(n-1) + \dots + p_N x(n-N+1)]} \right\} \\ &\equiv \sum_{p_1=-P}^{+P} \sum_{p_2=-P}^{+P} \dots \sum_{p_N=-P}^{+P} \left\{ w_p^* \cdot e^{j\omega_0 p' \cdot x} \right\} \end{aligned} \quad (3.73)$$

El fet d'ésser lineal amb els coeficients permet una representació vectorial del model,

$$\hat{g}[x] = \mathbf{w}^H \cdot \mathbf{z} \quad (3.74)$$

en funció dels vectors  $\mathbf{w}$  i  $\mathbf{z}$  cadascun de dimensió  $N_c = (2P+1)^N$  que agrupen adequadament els coeficients i les exponencials complexes del model, respectivament.

$$\mathbf{w} = \begin{bmatrix} \underbrace{w_{-P, -P, \dots, -P}}_{P_1} & \underbrace{w_{-P+1, -P, \dots, -P}}_{P_2} & \dots & \underbrace{w_{+P, -P, \dots, -P}}_{P_{2P+1}} \\ \underbrace{w_{-P+1, -P, \dots, -P}}_{P_{2P+2}} & \underbrace{w_{-P+1, -P+1, \dots, -P}}_{P_{2P+3}} & \dots & \underbrace{w_{+P, -P+1, \dots, -P}}_{P_{2(2P+1)}} \\ \dots & \dots & \dots & \dots \\ \underbrace{w_{-P, +P, \dots, +P}}_{P_{N_c - (2P+1)}} & \underbrace{w_{-P+1, +P, \dots, +P}}_{P_{N_c - 2P}} & \dots & \underbrace{w_{+P, +P, \dots, +P}}_{P_{N_c}} \end{bmatrix} \quad (3.75.a)$$

$$\mathbf{z} = \begin{bmatrix} e^{j\omega_0 P_1' \cdot x} & e^{j\omega_0 P_2' \cdot x} & \dots & e^{j\omega_0 P_{2P+1}' \cdot x} \\ e^{j\omega_0 P_{2P+2}' \cdot x} & e^{j\omega_0 P_{2P+3}' \cdot x} & \dots & e^{j\omega_0 P_{2(2P+1)'} \cdot x} \\ \dots & \dots & \dots & \dots \\ e^{j\omega_0 P_{N_c - (2P+1)'} \cdot x} & e^{j\omega_0 P_{N_c - 2P}' \cdot x} & \dots & e^{j\omega_0 P_{N_c}' \cdot x} \end{bmatrix} \quad (3.75.b)$$

A més a més, la linealitat del model respecte dels coeficients fa que es pugui dissenyar sota un criteri de Wiener que minimitza la funció de cost  $\xi_w$ , que representa la norma de la funció error,  $e_x$ ,

$$\xi_w = E[|e[x]|^2] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |g[x] - \hat{g}[x]|^2 \cdot p_X(x) \, dx(n) dx(n-1) \dots dx(n-N+1) \quad (3.76)$$

i que ara depèn de la funció  $p_X(x)$ , que denota la probabilitat conjunta de les variables que conformen el vector  $x$ .

De la mateixa manera que pel model de Fourier de SNLs sense memòria, el vector de coeficients “òptim” si existeix és únic, i s’obté igualant a zero el gradient de la funció de cost respecte del vector de coeficients. Així, s’arriba a una expressió totalment paral·lela a la obtinguda en el cas sense memòria,

$$\mathbf{w}_{opt} = \mathbf{R}_{zz}^{-1} \cdot E[g^*(x) \cdot \mathbf{z}] \quad (3.77)$$

que es diferencia però, en el fet de que ara la matriu d’autocorrelació del vector  $\mathbf{z}$ , denotada per  $\mathbf{R}_{zz}$  i de dimensions  $N_c \times N_c$ , deixa d’èsser Toeplitz però continua essent hermítica. A part d’aquest aspecte, important però alhora de calcular la inversió de la matriu, cada component de la matriu  $\mathbf{R}_{zz}$  està ara relacionada amb la funció característica conjunta de les components del vector  $x$  mostrejada a múltiples de la freqüència principal.

$$\mathbf{R}_{zz}(m, l) = E \left[ e^{j\omega_0(p_m - p_l) \cdot \mathbf{x}} \right] = \psi_X(j\omega_0(p_m - p_l)) \quad (3.78)$$

La funció característica conjunta és una funció  $N$ -dimensional que equival a la transformada inversa de Fourier (excepte en un factor d'escala) de la funció densitat de probabilitat conjunta.

$$\begin{aligned} \psi_X(j\mathbf{v}) &= E \left[ e^{j\mathbf{v}' \cdot \mathbf{x}} \right] = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{j[v_1 x(n) + v_2 x(n-1) + \dots + v_N x(n-N+1)]} \cdot p_X(\mathbf{x}) \, d\mathbf{x} \end{aligned} \quad (3.79)$$

És fàcil doncs comprovar que en el cas particular de que el senyal d'entrada tingui una distribució uniforme dins el rang  $[-X_0, X_0]$  i que cada component del vector  $\mathbf{x}$  sigui independent,

$$p_X(\mathbf{x}) = \prod_{i=1}^N p_{X_i}(x(n-i)) \quad \text{on} \quad p_{X_i}(x(n-i)) = \frac{1}{2X_0} \quad ; \forall i = 1, \dots, N \quad (3.80)$$

la matriu d'autocorrelació del model de Fourier amb memòria (Eq.3.78) esdevé diagonal ja que les exponencials complexes resulten ortonormals entre elles. És només en aquest cas quan el vector de pesos òptim (Eq.3.77) coincideix amb les constants o coeficients de Fourier (era d'esperar ja que sota la premisa (Eq.3.80) els productes escalars (Eq.3.76) i (Eq.3.63) esdevenen idèntics).

En un cas general amb qualsevol  $p_X(\mathbf{x})$ , els coeficients òptims que proporciona el criteri de Wiener són diferents dels de Fourier. És important, però, adonar-se de que la solució de disseny del model de Fourier amb memòria (Eq.3.77) està íntimament relacionada amb la solució de disseny de Wiener d'un filtre lineal MA multidimensional (de l'anglès "Moving Average") en el que, en lloc d'utilitzar la funció d'autocorrelació de les diferents components del vector d'entrada, surgeix la funció característica conjunta. El millor condicionament o no de la matriu d'autocorrelació de les exponencials complexes del model de Fourier,  $\mathbf{R}_{zz}$ , en comparació amb la matriu d'autocorrelació que genera el disseny MMSE d'un model de Volterra depèn, a l'igual que pel cas sense memòria, de l'estadística del senyal d'entrada. De totes maneres, el fet de que la funció d'autocorrelació conjunta sigui màxima a l'origen afavoreix un bon condicionament de la matriu  $\mathbf{R}_{zz}$  del model de Fourier. Aixó és com a conseqüència de que, en aquest cas, els valors de la diagonal d'aquesta matriu són tots iguals a  $\psi_X(\mathbf{0})=1$  i majors que les altres components, condició necessària (tot i que no suficient) per tal de que una matriu no sigui singular. Aquesta és una propietat important que condiciona tant el

disseny en bloc o de mínims quadrats del model de Fourier així com, molt especialment, el disseny amb mètodes adaptatius.

### 3.2.4.1 Disseny de mínims quadrats del model de Fourier amb memòria

En el disseny en bloc del model de Fourier amb memòria, el mateix desenvolupament fet pel cas sense memòria (Ap.3.1.4.1) és vàlid. En aquest cas però, el model respon a l'expressió (3.72), on ara el vector  $z$  depen del temps, denotant-se per  $z_n$ . La minimització de l'error quadràtic promitjat temporalment,

$$\xi_M^{LS} = \frac{1}{M} \sum_{n=0}^{M-1} |e(n)|^2 = \frac{1}{M} \sum_{n=0}^{M-1} |g[x(n)] - w^H \cdot z_n|^2 \quad (3.81)$$

conduïx a un conjunt de pesos òptim,

$$\hat{w}_{opt} = \hat{R}_{zz}^{-1} \cdot \hat{\gamma}_{gz} \quad (3.82)$$

expressat en termes de la matriu  $\hat{R}_{zz}$  i el vector  $\hat{\gamma}_{gz}$  que responen a les següents expressions.

$$\hat{R}_{zz} = \frac{1}{M} \sum_{n=0}^{M-1} z_n \cdot z_n^H \quad (3.83.a)$$

$$\hat{\gamma}_{gz} = \frac{1}{M} \sum_{n=0}^{M-1} g^*[x_n] \cdot z_n \quad (3.83.b)$$

En cas de que el conjunt de dades del senyal d'entrada,  $x(n)$ , siguin independents, la matriu  $\hat{R}_{zz}$  i el vector  $\hat{\gamma}_{gz}$  es poden interpretar, respectivament, com a estimadors de la matriu d'autocorrelació (Eq.3.78) (funció característica empírica) i del vector de correlació creuada, ambdós presents a la solució de Wiener (Eq.3.77).

De totes maneres, el vector de pesos que minimitza el valor mig de l'error quadràtic comporta la inversió d'una matriu de  $N_c \times N_c$  essent  $N_c = (2P+1)^N$  en el cas del model de Fourier amb memòria. Els mètodes tradicionals d'inversió de matrius necessiten de l'ordre de  $O(N_c^3)$  tot i que si s'organitzen adequadament les components del vector  $z$ , la matriu  $\hat{R}_{zz}$  resulta ésser Toeplitz per blocs i el cost computacional de l'inversió es pot reduir a  $O(N_c^2)$ .

Pel que fa a l'estimació que proporciona aquest estimador en presència de soroll additiu de mesura a la sortida del SNL real, es mantenen les mateixes propietats comentades anteriorment pel cas sense memòria (Ap.3.1.4.1).

### 3.2.4.1 Disseny adaptatiu del model de Fourier amb memòria

Com a conseqüència de l'elevat nombre de coeficients, de l'ordre de  $O(P^N)$ , un disseny adaptatiu del model de Fourier amb memòria requerirà més que mai un algorisme adaptatiu que sigui poc costós. El NLMS és un d'ells, amb un cost del mateix ordre de magnitud que el nombre de coeficients, i ja s'ha vist que modifica els pesos en el sentit invers del que marca el gradient de la funció de cost.

$$w(n+1) = w(n) - \frac{\mu}{p(n)} \cdot e^*(n) \cdot z_n \quad (3.84)$$

En el cas del model de Fourier, és interessant notar que la potència de cadascuna de les components del vector de dades  $z$  és per a totes igual a 1, de manera que això propicia un millor comportament del model de Fourier envers el model de Volterra.

En aplicacions on el NLMS no proporcioni la qualitat desitjada i la capacitat de càlcul disponible ho permeti, altres mètodes adaptatius poden ésser utilitzats per augmentar la velocitat de convergència. Així, l'apartat anterior mostra la possibilitat d'utilitzar mètodes adaptatius recursius (com, per exemple, el RLS) que facin tendir el model de Fourier cap a la solució de mínims quadrats.

### 3.2.5 Perspectiva general del model de Fourier per SNLs sense memòria

Aquesta secció s'ha dedicat fonamentalment a la generalització del model de Fourier sense memòria al cas amb memòria, i també a posar en relleu quines són les propietats i característiques que es continuen mantenint. En aquest sentit, i amb la corroboració de les simulacions realitzades a l'apartat (Ap.5.1.2) en el capítol 5, es pot comprovar com el model de Fourier amb memòria continua presentant un comportament (especialment en dissenys adaptatius) més adequat que el model de Volterra amb memòria.

## 3.3 Disseny en gelosia o "lattice"

En les dues seccions prèvies d'aquest capítol, el model de SNLs basat en la sèrie trigonomètrica o de Fourier s'ha presentat com un model alternatiu al clàssic basat en sistemes de Volterra. S'han discutit les seves principals característiques en aspectes relacionats amb el disseny i cost computacional, i és en aquests termes en els que s'ha establert la comparació en vers el model de Volterra. Un dels punts més atractius del disseny del model de Fourier és el fet

de que la correlació entre les funcions exponencials complexes que conformen el subespai de dades del model és, en general, menor que la correlació entre els monomis multivariables que determinen el subespai de dades en el model de Volterra. Aquesta propietat, tal i com es comprovarà en el capítol dedicat a simulacions, dota al model de Fourier amb un millor comportament especialment en dissenys adaptatius.

Malgrat això, les exponencials complexes del model de Fourier són lluny d'ésser ortogonals, depenent el grau de correlació d'aquestes de l'estadística del senyal d'entrada. De fet, a tall d'introducció, tot seguit es mostrarà com només en cas de tractar-se d'un senyal uniformement distribuït, el conjunt de funcions exponencials representen una base ortonormal de l'espai de funcions contínues, i els coeficients òptims obtinguts sota un criteri de disseny de Wiener coincideixen amb les anomenades constants de Fourier.

En el cas del model de Volterra, on també es dona una correlació elevada entre les potències que conformen l'espai de dades [Sch81], sovint es recorre a una implementació lattice que, tal i com es va introduir al capítol anterior (Ap.2.4.3), ortogonalitza el conjunt de funcions que defineixen el model. De la mateixa manera, per tal de solucionar el problema de la correlació entre funcions exponencials complexes del model de Fourier, aquest també permet una implementació no directa mitjançant una estructura lattice tal i com es mostra a continuació.

### 3.3.1 Implementació lattice del model de Fourier sense memòria

Amb vista a resoldre un problema de modelatge o identificació de SNLs, una solució eficient i que facilita enèrmement el disseny és la d'utilitzar un model de SNLs tal que els seus coeficients poguessin determinar-se directament com la correlació creuada entre el senyal de referència i la corresponent funció del model (Ap.2.2).

$$w_{optim} = E[z(x,n) \cdot d^*(n)] \quad (3.85)$$

Aixó implicaria que en un disseny en bloc s'ha evitat la costosa inversió de la matriu d'autocorrelació de les funcions del model o bé, en un disseny adaptatiu, que algorismes adaptatius amb una càrrega computacional baixa (p.e., el NLMS) poden ésser utilitzats sense perill de que hi hagi problemes en la convergència com a conseqüència de les diferents velocitats a les que convergeixen els diferents modes de la matriu d'autocorrelació quan la dispersió d'autovalors és elevada (és a dir, quan  $R_{zz}$  està mal condicionada).

En el model de Fourier sense memòria aquesta situació (Eq.3.85) es dona sempre que la funció densitat de probabilitat del senyal d'entrada respon a una distribució uniforme dins el rang de senyal,

$$p_X(x) = \begin{cases} \frac{1}{2X_0} & |x| \leq X_0 \\ 0 & |x| > X_0 \end{cases} \quad (3.86)$$

ja que llavors les funcions exponencials complexes que utilitza el model esdevenen ortonormals entre elles,

$$E[e^{jl\omega_0 x(n)} \cdot e^{-jm\omega_0 x(n)}] = \begin{cases} 1 & l = m \\ 0 & l \neq m \end{cases} \quad (3.87)$$

i el vector de coeficients òptim proporcionat pel criteri de mínim error quadràtic mig coincideix amb les anomenades constants de Fourier (Eq.3.21).

$$w_p \equiv \left\{ \begin{array}{l} \text{Constants} \\ \text{de Fourier} \end{array} \right\} = \frac{1}{2X_0} \int_{-X_0}^{X_0} g(x) \cdot e^{-jp\omega_0 x} dx \quad ; p = -P, \dots, P \quad (3.88)$$

És interessant remarcar que aquesta solució (Eq.3.85) s'assoleix sempre que la matriu d'autocorrelació (Eq.3.40) és diagonal, cosa que succeeix quan la funció característica de l'entrada mostrejada amb una velocitat de  $\omega_0$  esdevé una funció delta de Kronecker. De forma equivalent, això implicaria que  $p_X(x)$  és una funció tal que repetida al ritme de  $2\pi/\omega_0$  resulta plana. Un exemple clar, és el cas en que el senyal d'entrada té una distribució uniforme dins el rang determinat. En definitiva, i des d'un punt de vista més general, per tal de que la matriu d'autocorrelació sigui diagonal, caldria que  $p_X(x)$  fos un pols de Nyquist [Car88]. A excepció del cas uniforme, aquesta propietat no és gaire comuna en les distribucions que habitualment s'utilitzen i, per altra banda, no ens interessa en el modelatge de SNLs ja que llavors s'estaria cometent aliasing i l'aproximació de  $g[x]$  proporcionada pel model de Fourier resultaria distorsionada.

Aquest comportament del model de Fourier ens condueix a dues possibles alternatives per tal d'aconseguir un model amb funcions ortonormals. La primera d'elles consistiria en transformar la funció  $p_X(x)$  de l'entrada en una funció uniforme dins d'un rang determinat i després aplicar el model de Fourier directament. Aquesta opció es discutirà al proper capítol en el que el problema d'estimació de funcions densitat de probabilita és adreçat.

La segona opció que se'ns presenta es basa en la construcció d'un conjunt nou de funcions a partir del conjunt d'exponencials complexes, denotades per  $\{\phi_p(x, P)\}$ , de manera que les funcions que formen aquesta nova base són ortonormals entre elles en la norma definida

$$\langle \phi_l(x, P), \phi_l(x, P) \rangle = \int_{-\infty}^{+\infty} \phi_l(x, P) \cdot \phi_m^*(x, P) \cdot p_X(x) dx = \begin{cases} 1 & l = m \\ 0 & l \neq m \end{cases} \quad (3.89)$$

i també són capaces de generar el mateix subespai de senyal que les funcions exponencials complexes. Aquesta darrera premisa ens assegura que aquest nou conjunt de funcions pot obtenir exactament la mateixa aproximació de la funció del SNL que aconseguien les funcions exponencials del model de Fourier.

$$\hat{g}[x(n)] = \sum_{p=-P}^P w_p^* \cdot e^{jP\omega_0 x(n)} \equiv \sum_{p=0}^{2P} v_p^* \cdot \phi_p(x, n) \quad (3.90)$$

Per tal d'obtenir el conjunt de funcions  $\{\phi_p(x, P) \ p=0, \dots, 2P\}$  ortonormals en la norma definida (Eq.3.89), s'aplicarà el conegut mètode de Gram-Schmidt que condueix en aquest cas a una arquitectura lattice per implementar el model de Fourier basat en les funcions ortonormals (Eq.3.90). Suposem així que es disposa del conjunt d'exponencials complexes que el model de Fourier utilitza per a generar la seva aproximació.

$$\{e^{-jP\omega_0 x}, e^{-j(P-1)\omega_0 x}, \dots, 1, \dots, e^{jP\omega_0 x}\} \quad (3.91)$$

La dependència temporal de les exponencials, en realitat de la variable  $x=x(n)$ , s'ha suprimit per tal de no enterbolir la formulació, quedant aquesta de totes maneres implícita ja que es reprenrà en seccions posteriors. La primera funció del conjunt ortogonal, denotada per  $b_0(x, P)$  abans d'èsser normalitzades, s'escull igual a l'exponencial de major ordre

$$b_0(x, P) \equiv e^{jP\omega_0 x} \quad (3.92)$$

que, per altra banda, ja té norma unitat i per tant coincideix amb la primera funció del conjunt ortonormal.

$$E[|b_0(x, P)|^2] = 1 \Rightarrow \phi_0(x, P) = b_0(x, P) \quad (3.93)$$

La següent funció,  $b_1(x, P)$ , esdevé una combinació de les dues exponencials de major grau,

$$b_1(x, P) \equiv e^{j(P-1)\omega_0 x} + a_1^* \cdot e^{jP\omega_0 x} \quad (3.94)$$

on el coeficient  $a_1^*$  es determina de manera que la funció  $b_1(x, P)$  sigui ortogonal a l'anterior funció  $b_0(x, P)$ . La representació vectorial de la figura (3.10) il·lustra el procediment, obtenint-se que el coeficient depèn de la projecció de la funció  $\exp[j(P-1)\omega_0 x]$  sobre la direcció de la funció  $\exp[jP\omega_0 x]$ .



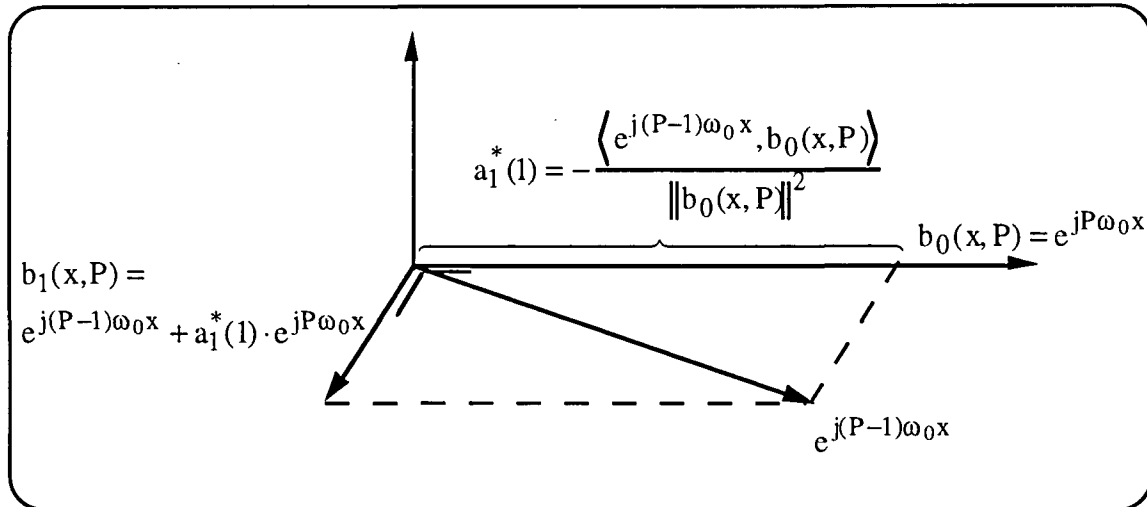


Figura 3.10 Procés d'ortogonalització de Gram-Schmidt de les funcions exponencials complexes del model de Fourier sense memòria

Un cop s'ha trobat el coeficient cal dividir la funció  $b_l(x, P)$  respecte de la seva norma, definida per  $\eta_l^b$ , per tal de que la funció tingui norma 1.

$$\phi_1(x, P) = \frac{b_1(x, P)}{\|b_1(x, P)\|} \equiv \frac{b_1(x, P)}{\sqrt{\eta_1^b}} \quad (3.95)$$

Aquest procediment s'itera de forma que la funció d'un ordre genèric  $b_p(x, P)$  resulta ésser una combinació lineal de totes les exponencials complexes des d'ordre  $P$  fins  $(P-p)$ .

$$b_p(x, P) = e^{j(P-p)\omega_0 x} + \sum_{i=1}^p a_p^*(i) \cdot e^{j(P-p+i)\omega_0 x} \quad ; \forall p = 1, \dots, 2P \quad (3.96)$$

A l'igual que anteriorment, ara els coeficients  $\{a_p^*(i) \forall i=1, \dots, p\}$  es determinen forçant que la funció  $b_p(x, P)$  sigui ortogonal a totes i cadascuna de les exponencials d'ordre inferior a  $(P-p)$  que conté. La combinació lineal resultant esdevé, en conseqüència, ortogonal a totes les anteriors funcions del conjunt ortogonal.

$$\langle b_p(x, P), e^{j(P-p+l)\omega_0 x} \rangle = 0 ; \forall l = 1, \dots, p \Rightarrow \langle b_p(x, P), b_q(x, P) \rangle = 0 ; \forall q < p \quad (3.97)$$

La imposició d'ortogonalitat amb les exponencials d'ordre inferior (Eq.3.97) és equivalent a  $p$  equacions lineals en les que apareix novament implicada la funció característica del procés d'entrada  $x$ .

$$\psi_X(-j l \omega_0) + \sum_{i=1}^p a_p^*(i) \cdot \psi_X(-j(l-i)\omega_0) = 0 \quad \forall l = 1, \dots, p \quad (3.98)$$

Aquestes equacions ja són suficients per a determinar els coeficients, que un cop calculats serviran també per a obtenir la norma de la funció  $b_p(x, P)$ , denotada per  $\eta_p^b$ ,

$$\eta_p^b \equiv E\left[|b_p(x, P)|^2\right] = \langle b_p(x, P), e^{j(P-p)\omega_0 x} \rangle = \psi_X(0) + \sum_{i=1}^p a_p^*(i) \cdot \psi_X(ji\omega_0) \quad (3.99)$$

necessària per a normalitzar.

$$\phi_p(x, P) \equiv \frac{b_p(x, P)}{\sqrt{\eta_p^b}} \quad \forall p = 0, \dots, 2P \quad (3.100)$$

La forma més estesa d'expressar les condicions que han de complir els coeficients, és ajuntar en forma matricial les  $p$  equacions lineals d'ortogonalitat (Eq.3.98) i la de la norma (Eq.3.99).

$$\begin{bmatrix} \psi_X(0) & \psi_X(1) & \dots & \psi_X(p) \\ \psi_X(-1) & \psi_X(0) & \dots & \psi_X(p-1) \\ \dots & & \ddots & \dots \\ \psi_X(-p) & \dots & & \psi_X(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_p^*(1) \\ \dots \\ a_p^*(p) \end{bmatrix} = \begin{bmatrix} \eta_p^b \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (3.101)$$

Aquest sistema d'equacions lineals és molt semblant a les equacions de Yule-Walker [Kay88], però en lloc d'estar expressada en termes de la funció d'autocorrelació, apareix la funció característica. Donat que la funció característica compleix unes propietats molt semblants a les d'una funció d'autocorrelació (ja es detallaran al proper capítol), el sistema anterior pot ésser resolt pel mètode de Levinson-Durbin [Kay88]. Aquest algorisme proporciona de manera recursiva els coeficients  $\{a_p^*(i) \forall i=1, \dots, p\}$  que fan que la funció  $b_p(x, P)$  sigui ortogonal a les d'ordre inferior,

$$a_p^*(i) = \begin{cases} a_{p-1}^*(i) + k_p^* \cdot a_{p-1}(p-i) & i = 1, \dots, p-1 \\ k_p^* & i = p \end{cases} \quad (3.102)$$

així com també la norma de la funció per tal de que es pugui normalitzar.

$$\eta_p^b = \left(1 - |k_p|^2\right) \cdot \eta_{p-1}^b \quad (3.103)$$

El coeficient de menor ordre  $a_p(0)$  és igual a la unitat per definició, mentre que el de major ordre,  $a_p(p)$  es denota pel paràmetre  $k_p$  que es coneix pel nom de *coeficient de reflexió* o

“PARCOR”, i que es pot expressar en funció de certs valors de la funció característica del senyal d’entrada.

$$a_p^*(p) \equiv k_p^* = - \frac{\psi_X(-p) + \sum_{l=1}^{p-1} a_{p-1}^*(l) \cdot \psi_X(l-p)}{\psi_X(0) + \sum_{l=1}^{p-1} a_{p-1}^*(l) \cdot \psi_X(l)} \quad (3.104)$$

L’algorisme de Levinson-Durbin és, en definitiva, un mètode eficient d’invertir la matriu de l’expressió (Eq.3.101), proporcionant d’una forma recursiva les funcions  $b_p(x, P)$ , i per tant les funcions  $\phi_p(x, P)$ , ja que el conjunt de pesos  $\{a_p(i) \forall i=1, \dots, p\}$  associats a aquesta funció es troben a partir dels coeficients de la funció d’ordre inferior  $b_{p-1}(x, P)$ . Aquesta recursió és la que ens condueix cap a una arquitectura en cascada capaç de generar les funcions ortonormals  $\{\phi_p(x, P) \forall i=1, \dots, p\}$ .

En efecte, si es substitueixen els coeficients en la funció d’ordre  $p$  (Eq.3.96) per l’expressió que proporciona l’algorisme de Levinson-Durbin (Eq.3.102) s’arriba a la següent relació.

$$\begin{aligned} b_p(x, P) &= e^{j(P-p)\omega_0 x} + \sum_{i=1}^p a_p^*(i) \cdot e^{j(P-p+i)\omega_0 x} = \\ &= e^{j(P-p)\omega_0 x} + \sum_{i=1}^{p-1} \left( a_{p-1}^*(i) + k_p^* \cdot a_{p-1}(p-i) \right) \cdot e^{j(P-p+i)\omega_0 x} + k_p^* \cdot e^{j(P)\omega_0 x} \\ &= e^{j(P-p)\omega_0 x} + \sum_{i=1}^{p-1} a_{p-1}^*(i) \cdot e^{j(P-p+i)\omega_0 x} + k_p^* \cdot \sum_{i=1}^p a_{p-1}(p-i) \cdot e^{j(P-p+i)\omega_0 x} \\ &= \sum_{i=0}^{p-1} a_{p-1}^*(i) \cdot e^{j(P-p+i)\omega_0 x} + k_p^* \cdot \sum_{i=0}^{p-1} a_{p-1}(i) \cdot e^{j(P-i)\omega_0 x} \end{aligned} \quad (3.105)$$

El primer terme coincideix amb la funció d’ordre  $(p-1)$  però que inclou fins l’exponencial d’ordre  $(P-1)$ ,

$$b_{p-1}(x, P-1) = \sum_{i=0}^{p-1} a_{p-1}^*(i) \cdot e^{j(P-p+i)\omega_0 x} \quad (3.106)$$

mentres que el sumatori del segon terme es defineix com la funció  $f_{p-1}(x, P)$ .

$$f_{p-1}(x, P) = \sum_{i=0}^{p-1} a_{p-1}(i) \cdot e^{j(P-i)\omega_0 x} \quad (3.107)$$

La notació d'aquestes funcions, els noms de  $b$  i  $f$ , corresponen a la notació d'aquestes expressions quan l'estructura lattice s'utilitza per a predicció lineal. En efecte, tal i com mostra la figura (3.11) les funcions  $b_p(x, P)$  i  $f_p(x, P)$  es poden relacionar amb els errors backward i forward d'ordre  $p$ , respectivament, igual que si es tractès d'una predicció lineal de la funció  $z(p) = \exp[jp\omega_0 x]$ .

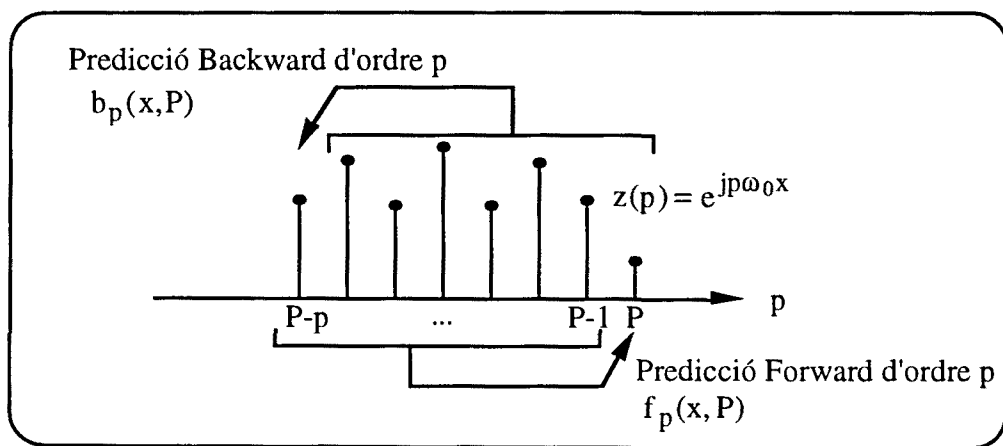


Figura 3.11 Interpretació de les funcions  $b_p(x, P)$  i  $f_p(x, P)$  com a errors de predicció d'ordre  $p$  backward i forward de la funció  $z(p) = \exp[jp\omega_0 x]$

Aprofitant les definicions anteriors (Eq.3.106, 107), la funció ortogonal d'ordre  $p$  es pot expressar en una forma més compacte tal i com segueix.

$$b_p(x, P) = b_{p-1}(x, P-1) + k_p^* \cdot f_{p-1}(x, P) \quad (3.108)$$

Procedint de la mateixa manera per a la funció forward d'ordre  $p$ ,

$$f_p(x, P) = \sum_{i=0}^p a_p(i) \cdot e^{j(P-i)\omega_0 x} \quad (3.109)$$

i substituint els coeficients per la solució proporcionada per l'algorisme de Levinson-Durbin, s'arriba a una relació semblant.

$$f_p(x, P) = f_{p-1}(x, P) + k_p \cdot b_{p-1}(x, P-1) \quad (3.110)$$

Aquesta equació juntament amb la trobada per la funció backward,  $b_p(x, P)$  (Eq.3.108), serveixen per construir la cel.la bàsica de l'estructura lattice. Abans però, és necessari

determinar la relació entre la funció  $b_p(x,P)$  i la  $b_p(x,P-1)$ . Tenint en consideració la definició de la funció genèrica  $b_p(x,P)$  (Eq.3.96) i el fet de que la solució  $\{a_p(i) \forall i=1,\dots,p\}$  proporcionada per l'equació de Yule-Walker (Eq.3.101) no depèn de l'ordre  $P$ , s'arriba a la següent relació entre ambdues funcions.

$$b_p(x, P-1) = e^{-j\omega_0 x} \cdot b_p(x, P) \quad (3.111)$$

Un cop es té aquesta relació, ja es pot definir la cel.la elemental del filtre lattice (Fig.3.12) que obté les funcions  $b_p(x,P)$  i la  $f_p(x,P)$  a partir de les d'ordre inferior  $b_{p-1}(x,P)$  i la  $f_{p-1}(x,P)$ .

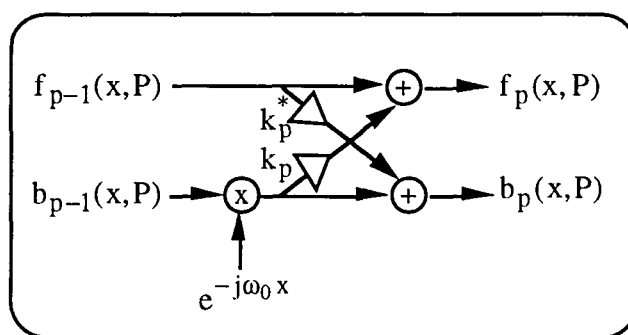


Figura 3.12 Cel.la elemental del filtre lattice.

Abans d'encadenar cel.les com l'anterior (Fig.3.12), és necessari determinar el valor de les funcions d'ordre 0, és a dir  $b_0(x,P)$  i  $f_0(x,P)$ , així com llur norma.

$$b_0(x,P) = f_0(x,P) = e^{jP\omega_0 x} \Rightarrow \eta_0^b = \eta_0^f = 1 \quad (3.112)$$

Un cop ja s'han determinat aquestes funcions només cal encadenar successivament cel.les elementals del tipus representat a la figura (3.12) que van obtenint les funcions ortogonals  $b_p(x,P)$  d'ordres superiors. El resultat consisteix en una estructura en cadena que en total consta de  $2P$  cel.les i que per tant proporciona les funcions ortogonals  $\{b_p, p=0,\dots,2P\}$  del procés d'ortogonalització de Gram-Schmidt.

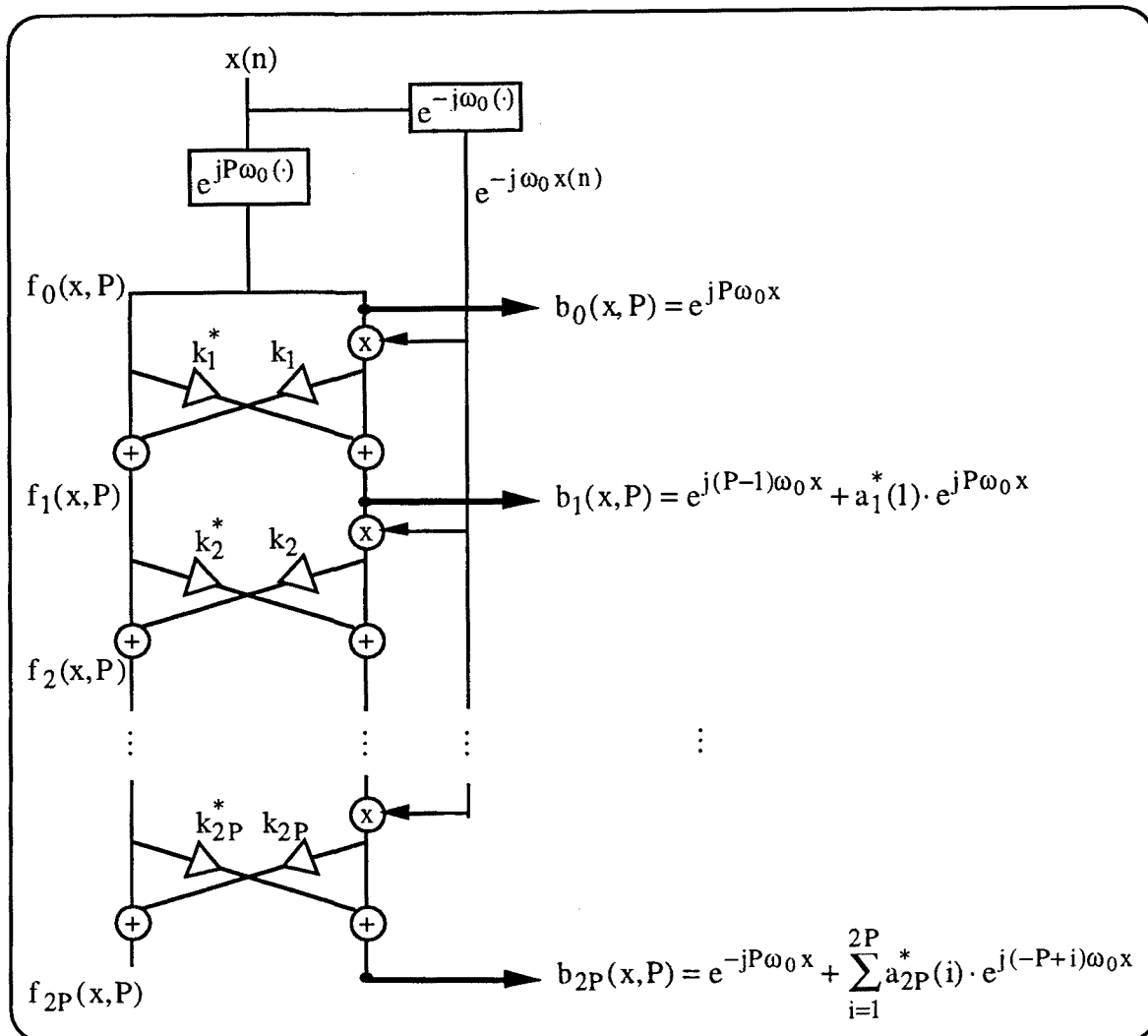
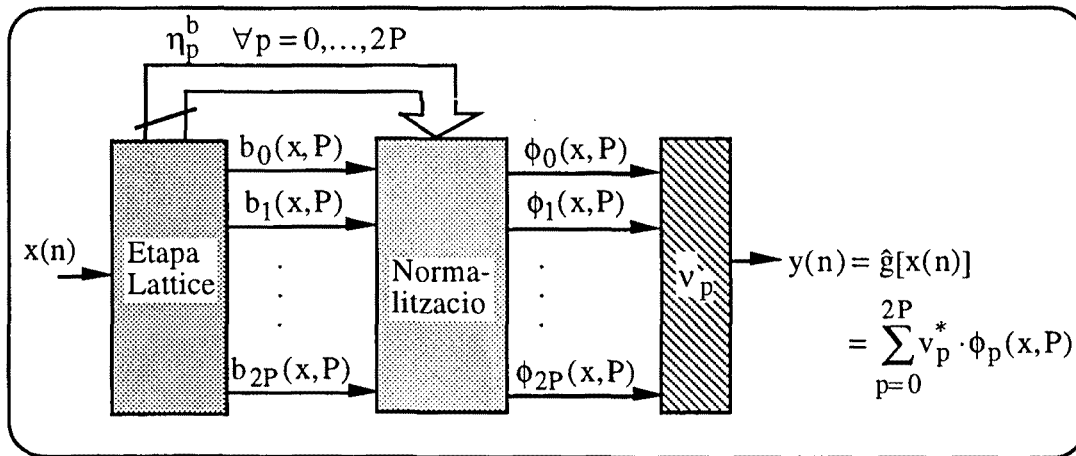


Figura 3.13 Estructura lattice d'ordre  $P$  que genera les funcions ortogonals  $\{b_p(x, P) \ p=0, \dots, 2P\}$

Aquestes funcions han d'èsser posteriorment normalitzades per l'arrel quadrada de la seva norma d'acord amb l'expressió (3.100). Donat que el conjunt resultant de funcions ortonormals, format per  $\{\phi_p(x, P) \ p=0, \dots, 2P\}$ , no és res més que combinacions lineals de les exponencials complexes (Eq.3.91) del model de Fourier originari (3.90), tots dos conjunts inclouen la mateixa informació de manera que són capaços de generar la mateixa aproximació de la funció  $g[x]$ . La implementació del model de Fourier mitjançant una estructura lattice-ladder esdevé en conseqüència el representat a la següent figura.



**Figura 3.14** Implementació no directa del model de Fourier sense memòria d'ordre  $P$  mitjançant una arquitectura tipus lattice-ladder.

L'avantatge d'utilitzar les funcions ortonormals radica, fonamentalment, en que el disseny dels coeficients  $\{v_p \ p=0, \dots, 2P\}$  no necessita invertir la matriu d'autocorrelació a diferència dels pesos relacionats amb la implementació directa model de Fourier  $w_p$ .

$$v_p = E[g^*[x] \cdot \phi_p(x, P)] \quad \forall p = 0, \dots, 2P \quad (3.113)$$

A més a més, tal i com s'ha comentat abans, el disseny adaptatiu dels pesos  $\{v_p \ p=0, \dots, 2P\}$  pot ésser realitzat per un algorisme adaptatiu senzill com el NLMS ja que no presentarà problemes de convergència com a conseqüència de la dispersió d'autovalors de la matriu d'autocorrelació.

De totes maneres, és important parar atenció al fet de que d'alguna manera el cost computacional de la inversió de la matriu que ens estalviem, s'ha traslladat a la primera part de la figura (3.14) d'on s'obtenen les funcions ortogonals. És fàcil comprovar com el cost computacional necessari per calcular tots els paràmetres és de l'ordre de  $O(4P^2)$  (mètode de Levinson-Durbin per un sistema de  $2P+1$  coeficients). A més a més, les estructures lattice són un model sobredimensionat ja que el nombre de coeficients que utilitza incrementa respecte de la implementació directa, i de fet en aquest cas particular es duplica ja que a part dels coeficients  $v_p$  també estan els paràmetres,  $k_p$ . Per tant, es pot dir que tot i que, en general i basant-nos en resultats de processament lineal del senyal, les implementacions lattice d'un determinat filtre doten al sistema resultant d'un comportament més adequat en dissenys adaptatius això s'aconsegueix a costa d'un increment important en el cost computacional. Abans però, d'entrar en detalls respecte del cost computacional del model de Fourier-Lattice, és interessant ressaltar les propietats de l'estructura lattice desenvolupada pel cas particular de modelatge de SNLs sense memòria.

Per una banda, és interessant adonar-se de que l'estructura en cascada (Fig.3.13) està basada a partir de l'anomenada cel.la elemental i que, per tant, un increment en l'ordre del model de Fourier (per exemple, de  $P$  a  $(P+1)$ ) només suposa encadenar dues cel.les més al final de l'estructura i incrementar en 1 l'ordre de la primera exponencial, és a dir, passar de  $\exp[jP\omega_0x]$  a  $\exp[j(P+1)\omega_0x]$ . Tot i que hi haurien dos coeficients de reflexió més, els que ja existeixen no canvien de valor ja que tal i com mostra l'expressió (3.104) el coeficient de reflexió  $p$ -èssim només depèn de les  $p$  primeres mostres (tant positives com negatives) de la funció característica mostrejada amb un període de  $\omega_0$ . Una altra propietat dels PARCORs és que el seu mòdul està limitat a ésser estrictament menor que la unitat [Kay88], cosa que els fa adequats per implementacions discretes del model de Fourier mitjançant microprocessadors o arquitectures d'aritmètica finita.

En definitiva, es pot assegurar que s'ha desenvolupat una arquitectura lattice-ladder que implementa d'una forma no directa el model de Fourier sense memòria. Cal però, proposar un disseny dels coeficients associats a aquesta arquitectura, tant dels coeficients de reflexió com dels coeficients que en darrer lloc es combinen linealment amb les funcions ortonormals (Fig.3.14). Aquest disseny, a l'igual que en implementacions anteriors, pot ésser en bloc o també adaptatiu.

### 3.3.1.1 Disseny en bloc del model de Fourier-Lattice sense memòria

Tenint en consideració que el coeficient de reflexió,  $k_p$ , està definit com el coeficient  $a_p(p)$  de la funció backward, una forma equivalent d'arribar a la solució de Yule-Walker consisteix en minimitzar la funció  $b_p(x,P)$  (interpretada com un "error" de predicció backward) respecte del conjunt de coeficients  $\{a_p(i) \ i=1, \dots, p\}$ .

$$E\left[|b_p(x,P)|^2\right]_{MINIM} \Rightarrow \text{Equació de Yule - Walker} \Rightarrow \text{Obtenció dels parcors, } k_p \quad (3.114)$$

Així, és fàcil comprovar com el criteri de minimització de  $\eta_p^b$  equival a l'ortogonalitat entre la funció error backward,  $b_p(x,P)$ , i les funcions exponencials construïdes a partir de les dades (Eq.3.988). Donat que la norma de la funció backward i la funció forward són iguals,

$$E\left[|f_p(x,P)|^2\right] = E\left[|b_p(x,P)|^2\right] \equiv \eta_p \quad (3.115)$$

un criteri habitual de disseny dels coeficients de reflexió consisteix en minimitzar la següent funció de cost, proposada inicialment per Burg (per exemple a [Hay91]),



$$E\left[|f_p(x,P)|^2 + |b_p(x,P)|^2\right]_{MINIM} \Rightarrow k_p \quad (3.116)$$

conduint-nos a la següent relació dels PARCORs que és equivalent a la trobada anteriorment (Eq.3.104).

$$k_p = -\frac{2E[f_{p-1}(x,P) \cdot b_{p-1}^*(x,P-1)]}{E\left[|f_{p-1}(x,P)|^2 + |b_{p-1}(x,P-1)|^2\right]} \quad \forall p = 1, \dots, 2P \quad (3.117)$$

A partir d'aquesta expressió pels coeficients de reflexió (Eq.3.117) i en un entorn de processament lineal del senyal, és usual substituir les mitjes estadístiques per promitjos temporals amparant-se en l'estacionarietat del senyal d'entrada. És en aquest punt on convé recuperar la dependència temporal del procés d'entrada,  $x=x(n)$ , així com de totes les funcions que es formen a partir d'ell (Fig.3.13). Suposant que es disposa de  $M$  valors de l'entrada, aquesta diversitat temporal proporciona un estimador pels succesius valors dels coeficients de reflexió (denotat amb un circumflex),

$$\hat{k}_p = -\frac{2 \sum_{n=1}^M f_{p-1}(x(n), P) \cdot b_{p-1}^*(x(n), P-1)}{\sum_{n=1}^M \left[ |f_{p-1}(x(n), P)|^2 + |b_{p-1}(x(n), P-1)|^2 \right]} \quad \forall p = 1, \dots, 2P \quad (3.118)$$

així com de les succesives potències de les funcions ortogonals.

$$\hat{\eta}_p = \left(1 - |\hat{k}_p|^2\right) \cdot \hat{\eta}_{p-1} \quad \forall p = 1, \dots, 2P \quad (3.119)$$

Aquests estimadors juntament amb les expressions de la cel.la elemental de l'arquitectura lattice (Fig.3.12) i la següent definició dels valors inicials,

$$b_0(x(n), P) = f_0(x(n), P) \equiv e^{jP\omega_0 x(n)} \quad (3.120.a)$$

$$\eta_0 = \psi_X(0) \Rightarrow \hat{\eta}_0 = 1 \quad (3.120.b)$$

proporcionen una de les opcions més utilitzades en el disseny en bloc de l'estructura lattice. És interessant adonar-se de que aquesta solució equival a una solució de mínims quadrats, que s'obtidria minimitzant la següent funció de cost,

$$\xi^{LS}(k_p) = \frac{1}{M} \sum_{n=1}^M \left( |b_p(x(n), P)|^2 + |f_p(x(n), P)|^2 \right) \quad (3.121)$$

respecte del coeficient de reflexió d'ordre  $p$ , tenint en compte les relacions de la cel.la elemental (Fig.3.12).

Tot i que sota un criteri de disseny de mínims quadrats no és necessari fer cap suposició sobre l'estadística del senyal d'entrada, en cas de que les mostres  $x(n)$  del senyal d'entrada siguin independents entre elles però que, a més a més, totes elles tinguin la mateixa funció característica (o funció densitat de probabilitat), el disseny de mínims quadrats dels PARCORs es pot interpretar com una aproximació de l'expressió estadística (3.104) mitjançant l'estimador FCE de la funció característica. Per contra, en entorns no estacionaris serà necessari l'ús d'algorismes adaptatius que condueixin els coeficients cap a la solució òptima de mínims quadrats, que en aquests casos inclou un enfilament temporal.

Abans però, d'entrar en el disseny adaptatiu de la implementació lattice del model de Fourier sense memòria, cal parar atenció als coeficients de la darrera etapa del model (Fig.3.14),  $\{v_p \ p=0, \dots, 2P\}$ , que es combinen linealment amb les funcions ortonormals proporcionades per l'etapa lattice. Ja s'ha vist abans, que sota un criteri de minimització del valor mig de l'error quadràtic, la solució òptima d'aquests pesos esdevé la correlació creuada entre la funció del SNL  $g[x]$  i la funció ortonormal corresponent.

$$v_p = E[g^*[x] \cdot \phi_p(x, P)] \quad \forall p = 0, \dots, 2P \quad (3.122)$$

De la mateixa manera que per a la versió directa del model de Fourier sense memòria i suposant un problema d'identificació, aquests valors de correlació es poden estimar a partir de la diversitat temporal de l'entrada. Així, anteriorment ja s'ha vist com l'etapa lattice proporciona els successius valors de les funcions ortonormals a cada instant de temps a partir dels parcos estimats, de manera que es poden utilitzar en l'estimació dels coeficients de l'etapa final,

$$\hat{v}_p = \frac{1}{M} \sum_{n=1}^M g^*[x(n)] \cdot \phi_p(x(n), P) \quad \forall p = 0, \dots, 2P \quad (3.123)$$

juntament amb els valors de la sortida del SNL real, denotats per  $y(n)=g[x(n)]$ .

### 3.3.1.2 Disseny adaptatiu del model de Fourier-Lattice sense memòria

En el disseny adaptatiu de la implementació lattice-ladder del model de Fourier sense memòria (Fig.3.14), s'han d'encadenar dos algorismes adaptatius diferents, un pel disseny dels parcos  $\{k_p \ p=1, \dots, 2P\}$ , i un altre pels coeficients  $\{v_p \ p=0, \dots, 2P\}$  que es combinen amb les funcions obtingudes de l'etapa lattice.

Respecte del disseny dels coeficients de reflexió, tant l'ús de mètodes de gradient (per exemple, Gradient Adaptive Lattice algorithm desenvolupat per Griffiths [Gri78]), com

mètodes recursius ([Pro92, Fej94], són adequats pel disseny adaptatiu dels parcors del model de Fourier-lattice. En el nostre cas (veure les simulacions al capítol 5 (Ap.5.1.1)), s'ha optat per un disseny recursiu (Tau.3.1, Apèn.3.C) que, tot i que requereix més operacions per iteració, ha donat bons resultats en les simulacions.

Per altra banda, tal i com es pot comprovar a la taula anterior, el mètode utilitzat per obtenir els pesos  $\{v_p \ p=0, \dots, 2P\}$  ha estat el NLMS. És interessant notar que l'ús del NLMS en lloc del LMS no normalitzat és necessari ja que, fins que els parcors de l'etapa lattice no hagin convergit a un valor adequat, les funcions  $\{\phi_p(x, P) \ p=0, \dots, 2P\}$  són lluny de tenir norma 1.

Val a dir de totes maneres que, en principi, qualsevol mètode adaptatiu, tant de gradient com recursiu, ja desenvolupat per a un filtre lattice en l'entorn de filtres lineals continua essent vàlid per dissenyar el model de Fourier lattice presentat. No s'ha considerat oportú prestar una major dedicació a aquests mètodes ja que tampoc aportariem res al que ja es coneix sobre ells i, a més a més, aquests aspectes més particulars queden fora de l'abast d'aquesta tesi.

### 3.4 Valoracions del model de Fourier

En aquest capítol s'ha presentat un model de SNLs alternatiu al model de Volterra o polinòmic, i desenvolupat a partir de la sèrie trigonomètrica. El resultat ha estat un model anomenat model de Fourier i que presenta certes peculiaritats que el fan especialment adequat en l'etapa de disseny.

En primer lloc, es tracta d'un model de SNLs però que guarda una dependència lineal amb els seus paràmetres cosa que proporciona una expressió tancada pel disseny dels paràmetres sota un criteri d'error quadràtic mig mínim (MMSE). Aquesta linealitat del model, a més a més, permet l'ús de mètodes ja desenvolupats en sistemes lineals d'una forma totalment paral·lela en l'entorn del model de SNLs (criteri de mínims quadrats, mètodes adaptatius tant recursius com de gradient). De totes maneres, aquesta propietat és també compartida amb el model de Volterra i per tant s'han destacat altres propietats del model de Fourier que aporten avantatges addicionals.

Una de les principals propietats del model de Fourier radica en el paral·lelisme de la solució de MMSE amb la d'un sistema lineal, ja que el conjunt de pesos trobat s'expressa en termes de l'anomenada matriu característica que té unes propietats idèntiques a les de la matriu d'autocorrelació que apareix en el disseny de MMSE en sistemes lineals. Aquest paral·lelisme permet aplicar mètodes eficients de resolució per a la solució de MMSE del model de Fourier (tant en el cas sense memòria com amb memòria), que ja han estat utilitzats en sistemes lineals. A més a més, donades les característiques de la matriu característica, en general, està molt millor condicionada que la matriu que surgeix en el disseny de MMSE del model de Volterra i

aquesta propietat dota al model de Fourier d'un millor comportament en dissenys adaptatius. En aquest sentit, també hi contribueix el fet de que les funcions exponencials complexes que genera el model de Fourier a partir del senyal d'entrada tenen norma unitat, a diferència del model de Volterra on la dispersió de la potència de les funcions polinòmiques agreuja encara més el seu comportament en dissenys adaptatius.

És necessari remarcar que el cost que s'ha de pagar per aquest millor comportament (veure capítol 5, secció (Ap.5.1)) és un increment en el nombre d'operacions necessàries, tant motivat per la major complexitat de les funcions exponencials així com el caràcter complex dels coeficients i funcions. De totes maneres, ja s'ha justificat que aquest cost computacional és del mateix ordre de magnitud que el que necessita un model de Volterra, sense prendre en consideració el càlcul de l'exponencial complexa de primer ordre.

Finalment, també s'ha emfatitzat la versatilitat del model en el sentit de que és capaç d'adequar-se a particularitats del SNL que es pretén modelar. Per una banda, s'han proposat unes versions del model que contemplen una simetria parell o senar en la funció del SNL real i que disminueixen la càrrega computacional sense perjudicar el grau d'aproximació assolit. Per una altra banda, s'ha mostrat com el model sense memòria admet una implementació no directa amb una estructura lattice, d'una forma totalment paral·lela a l'arquitectura lattice d'un filtratge FIR. Com a conseqüència de la manca d'una arquitectura lattice per un filtratge multidimensional lineal, no s'ha pogut desenvolupar una versió pel model de Fourier amb memòria. Malgrat això, creiem que és possible arribar a una implementació no directa del model de Fourier amb memòria que orthogonalitzi les funcions exponencials multidimensionals considerant com a patró arquitectures ja desenvolupades pel model de Volterra amb memòria i basades amb sistemes lineals multicanal [Zar85].

En conclusió, aquest capítol s'ha centrat en el sistema de Fourier com a model de SNLs, destacant aspectes relacionats fonamentalment amb el disseny i la flexibilitat del mateix. Per contra, el següent capítol es centra més en explotar el paral·lisme que mostra el model de Fourier amb determinats sistemes lineals i que condueixen cap a resultats força interessants.

### 3.5 Referències

- [Arn95] Arnold M.J., Zoubir A.M., Boashash B., "Estimation of the Characteristic Function", *Proc. of the IEEE-ATHOS Workshop on Higher-Order Statistics*, Begur, Girona, Spain, pp. 231-235, 1996.
- [Car88] Carlson A.B., *Communication Systems. An introduction to Signals and Noise in Electrical Communication*, McGraw-Hill, Inc., USA, 1988 (3<sup>a</sup> edició).

- [Chu69] Churchill R.V., *Fourier Series and Boundary Value Problems*, McGraw-Hill, Inc., USA, 1969 (2<sup>a</sup> edició).
- [Fej94] Fejzo Z, *Adaptive Non-Linear Laguerre-Based Filtering*, Tesi Doctoral, Northeastern University, Boston, 1994.
- [Gra95] Gray R.M., Goodman J.W., *Fourier Transforms. An Introduction for Engineers*, Kluwer Academic Publishers, 1995.
- [Gri78] Griffiths L.J., "An adaptive lattice structure for noise-cancelling applications", *Proc. of the IEEE-ICASP*, Tulsa, Okla., pp.87-90, 1978.
- [Hay91] Haykin S., *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, New Jersey, 1991.
- [Hil00] Hilbert, D., "Mathematische Probleme", *Nachr. Akad. Wiss. Göttingen*, pp. 290-329, 1900
- [Kay88] Kay S.M., *Modern Spectral Estimation. Theory and Applications*, Prentice-Hall, Eng. Cliff, NJ, 1988.
- [Kol57] Kolmogorov A.N., "On the Representation of Continuous Functions of Many Variables by Superposition of Continuous Functions of One Variable and Addition", *Doklady Akademii Nauk. SSSR*, 114(5), pp. 953-956, 1957.
- [Kol63] Kolmogorov A.N., "On the Representation of Continuous Functions of Many Variables by Superposition of Continuous Functions of One Variable and Addition", *American Mathematical Society Translation*, 28, pp 55-59, 1963.
- [Kur92] K<sup>o</sup>urková V., "Kolmogorov's Theorem and Multilayer Neural Networks", *Neural Networks*, Vol. 5, pp. 501-506, 1992.
- [Mul94] Mulgrew B., "Orthogonal Functions for Nonlinear Signal Processing and Adaptive Filtering", *Proc. of IEEE-ICASSP*, Adelaide, Australia, Vol.3, pp.509-512, 1994.
- [Pag94] Pagès-Zamora A., Lagunas M.A., "The K-filter: Design alternatives to model nonlinear systems", *Proc. of the EURASIP-EUSIPCO*, Edimburgh, Scotland, U.K., Vol.III, pp.1563-1566, Setembre 1994.
- [Pag95] Pagès-Zamora A., Lagunas M.A., Nájjar M., Pérez-Neira A., "The K-filter: a new architecture to model and design non-linear systems from the Kolmogorov's theorem", *EURASIP Signal Processing*, Vol. 44, N<sup>o</sup> 3, pp. 249-267, July 1995.
- [Pap65] Papoulis A., *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, Inc., USA, 1965.
- [Pro92] Proakis J.G., Rader C.M., Ling F., Nikias C.L., *Advanced Digital Signal Processing*, Pacmillan Publishing Company, NY, 1992.
- [Rug81] Rugh W.J., *Nonlinear System Theory: the Volterra/Wiener Approach*, The Johns Hopkins University Press, Baltimore, 1981.

- [Sch81] Schetzen M., "Nonlinear System Modeling Based on the Wiener Theory", *Proc. of the IEEE*, Vol. 69, nº 12, Desembre 1981.
- [She76] Sheldon M.R., *Introduction to Probability and Statistics for Engineers and Scientist*, Wiley, New York, 1976.
- [Spr65] Sprecher D.A., "On the Structure of Continuous Functions of Several Variables", *Transactions American Mathematical Society*, vol. 115, pp. 340-355, 1965.
- [Stu94] Stuart A., Ord K., *Kendall's Advanced Theory of Statistics. Vol.I: Distribution Theory*, Edward Arnold, London, 6th Edition, 1994.
- [Zar85] Zarzycki J., *Nonlinear Prediction Ladder Filters for Higher Order Stochastic Sequences*, Springer-Verlag, Berlin, 1985.

### Apèndix 3A Estadística de l'estimador de mínims quadrats en presència de soroll additiu

Suposant que el senyal real de sortida del SNL a identificar respon a la següent expressió,

$$y(n) = g[x(n)] = \mathbf{w}_{opt}^H \cdot \mathbf{z}(n) = \mathbf{z}(n)^H \cdot \mathbf{w}_{opt}$$

i que en la mesura del mateix s'hi afegeix un soroll,  $\varepsilon(n)$ , el senyal de referència denotat per  $d(n)$  es pot expressar com:

$$d(n) = \mathbf{z}(n)^H \cdot \mathbf{w}_{opt} + \varepsilon(n)$$

Si es disposa de  $M$  mostres del senyal d'entrada i de la referència per a identificar el sistema, es pot optar per una notació vectorial que agrupa les  $M$  mostres de cada senyal en els respectius vectors de referència ( $d$ ) i soroll ( $\varepsilon$ ).

$$d = A \cdot \mathbf{w}_{opt} + \varepsilon \quad ; \text{on la matriu } A = [z_1 \quad z_2 \quad \dots \quad z_M]^H$$

L'estimació dels coeficients proporcionada pel criteri de mínims quadrats equival a:

$$\hat{\mathbf{w}}^{LS} = \hat{\mathbf{R}}_{zz}^{-1} \cdot \hat{\gamma}_{dz} \equiv (A^H A)^{-1} A^H d$$

Sota la hipòtesi de que el soroll és de mitja nul·la, s'arriba a que l'estimador de mínims quadrats és no esbiaixat.

$$\begin{aligned} E[\hat{\mathbf{w}}^{LS}] &= E\left[(A^H A)^{-1} A^H d\right] = \{d = A \cdot \mathbf{w}_{opt} + \varepsilon\} = E\left[(A^H A)^{-1} A^H A \cdot \mathbf{w}_{opt}\right] + \\ &+ E\left[(A^H A)^{-1} A^H \varepsilon\right] = E[\mathbf{w}_{opt}] + (A^H A)^{-1} A^H E[\varepsilon] = \mathbf{w}_{opt} \end{aligned}$$

Respecte de la covariància de l'estimador, cal suposar que el soroll és blanc i amb una variància  $\sigma_\varepsilon^2$  per tal d'arribar a la següent expressió tancada.

$$\begin{aligned} \text{cov}(\hat{w}^{LS}) &= E\left[\left(\hat{w}^{LS} - w_{opt}\right) \cdot \left(\hat{w}^{LS} - w_{opt}\right)^H\right] = \left\{ \begin{array}{l} \hat{w}^{LS} = (A^H A)^{-1} A^H d \\ d = A \cdot w_{opt} + \varepsilon \end{array} \right\} \\ &= (A^H A)^{-1} A^H E[\varepsilon \cdot \varepsilon^H] A (A^H A)^{-1} = \left\{ E[\varepsilon \cdot \varepsilon^H] = \sigma_\varepsilon^2 I \right\} = \sigma_\varepsilon^2 \cdot \hat{R}_{zz}^{-1} \end{aligned}$$

### Apèndix 3B Error quadràtic assolit per l'estimador de mínims quadrats en presència de soroll

Sota les mateixes hipòtesis considerades a l'apèndix (Apèn.3.A), l'error quadràtic assolit per l'estimador de mínims quadrats es pot expressar tal i com segueix.

$$\hat{\xi}_{MIN}^{LS} = \frac{1}{M} \left( d^H d - d^H A (A^H A)^{-1} A^H d \right)$$

Expressant el senyal de referència en funció dels coeficients,

$$d = A \cdot w_{opt} + \varepsilon$$

s'arriba al següent resultat per a l'error quadràtic assolit per l'estimador.

$$\hat{\xi}_{MIN}^{LS} = \frac{1}{M} \left( \varepsilon^H \varepsilon - \varepsilon^H A (A^H A)^{-1} A^H \varepsilon \right)$$

Tal i com es mostra a continuació, la mitjana d'aquest valor equival a:

$$\begin{aligned} E\left[\hat{\xi}_{MIN}^{LS}\right] &= E\left[\frac{1}{M} \sum_{n=1}^M |\varepsilon(n)|^2 - \frac{1}{M} \sum_{l=1}^M \sum_{n=1}^M \varepsilon(n) \varepsilon^*(l) \cdot z_l^H (A^H A)^{-1} z_n\right] = \\ &= \sigma_\varepsilon^2 - \frac{\sigma_\varepsilon^2}{M} \sum_{n=1}^M z_n^H (A^H A)^{-1} z_n = \sigma_\varepsilon^2 - \frac{\sigma_\varepsilon^2}{M} \text{traça}\left[A (A^H A)^{-1} A^H\right] = \\ &= \left\{ \begin{array}{l} \text{Commutativitat} \\ \text{en la traça} \end{array} \right\} = \sigma_\varepsilon^2 - \frac{\sigma_\varepsilon^2}{M} \text{traça}\left[(A^H A)^{-1} A^H A\right] = \sigma_\varepsilon^2 - \frac{\sigma_\varepsilon^2}{M} \text{traça}[I] \\ &= \sigma_\varepsilon^2 \left(1 - \frac{2P+1}{M}\right) \end{aligned}$$

### Apèndix 3C Algorisme adaptatiu en el disseny lattice-ladder del model de Fourier sense memòria

Valors inicials etapa Lattice

Per a  $p = 1, \dots, 2P$ :  $f_p(x(0), P) = b_p(x(0), P) = 0$ ;  $\Delta_p(0) = 0$ ;  $Q_i(0) = \varepsilon$   $0 < \varepsilon \ll 1$ ;  $\lambda \ll 1$

Valors inicials etapa Ladder

$Pref(0) = 1$ ;  $\beta \ll 1$ ; Per a  $p = 0, \dots, 2P$ :  $v_p(0) = 0$ ;

**Bucle principal**

Per a  $n = 1, \dots, M$ :

$$\{ f_0(x(n), P) = b_0(x(n), P) = \phi_0(x(n), P) = e^{jP\omega_0 x(n)}$$

Per a  $p = 1, \dots, 2P$ :

$$\{ b_{p-1}(x(n), P-1) = b_{p-1}(x(n), P) \cdot e^{-j\omega_0 x(n)}$$

$$\Delta_p(n) = \lambda \cdot \Delta_p(n-1) + f_{p-1}(x(n), P) \cdot b_{p-1}^*(x(n), P-1)$$

$$Q_{p-1}(n) = \lambda \cdot Q_{p-1}(n-1) + \frac{1}{2} \left( |f_{p-1}(x(n), P)|^2 + |b_{p-1}(x(n), P-1)|^2 \right)$$

$$k_p(n) = \frac{\Delta_p(n)}{Q_{p-1}(n)} \quad ; \quad \phi_{p-1}(x(n), P) = \frac{b_{p-1}(x(n), P)}{\sqrt{(1-\lambda) \cdot Q_{p-1}(n)}}$$

$$f_p(x(n), P) = f_{p-1}(x(n), P) + k_p(n) \cdot b_{p-1}(x(n), P-1)$$

$$b_p(x(n), P) = b_{p-1}(x(n), P-1) + k_p^*(n) \cdot f_{p-1}(x(n), P)$$

}

$$b_{2P}(x(n), P-1) = b_{2P}(x(n), P) \cdot e^{-j\omega_0 x(n)}$$

$$Q_{2P}(n) = \lambda \cdot Q_{2P}(n-1) + \frac{1}{2} \left( |f_{2N}(x(n), P)|^2 + |b_{2N}(x(n), P-1)|^2 \right)$$

$$\phi_{2P}(x(n), P) = \frac{b_{2P}(x(n), P)}{\sqrt{(1-\lambda) \cdot Q_{2P}(n)}}$$

Disseny NLMS ( $d(n) \equiv$  senyal de referència)

$$\hat{y}(n) = \sum_{p=0}^{2P} v_p^*(n-1) \cdot \phi_p(x(n), P); \quad e(n) = d(n) - \hat{y}(n)$$

$$Pref(n) = \beta \cdot Pref(n-1) + (1-\beta) \cdot \sum_{p=0}^{2P} \|\phi_p(x(n), P)\|^2$$

$$Per a p = 0, \dots, 2P: v_p^H(n) = v_p^H(n-1) + \frac{\mu}{Pref(n)} \cdot e^*(n) \cdot \phi_p(x(n), P)$$

}



