



**UNIVERSIDAD POLITECNICA DE MADRID
E.T.S.I. DE TELECOMUNICACION**

TESIS DOCTORAL

**APLICACIONES
DE NORMAS NO CUADRATICAS
EN
TRATAMIENTO DIGITAL DE LA SEÑAL**

**Autor: José Manuel Paez Borrallo
Director: Anibal R. Figueiras Vidal**

Madrid 1987

87/15

TESIS DOCTORAL:

APLICACIONES DE NORMAS NO CUADRATICAS EN TRATAMIENTO DIGITAL DE LA SEÑAL

AUTOR: José Manuel Páez Borrallo

DIRECTOR: D. Aníbal R. Figueiras Vidal

El Tribunal nombrado para juzgar la Tesis Doctoral arriba citada, compuesto por los siguientes señores:

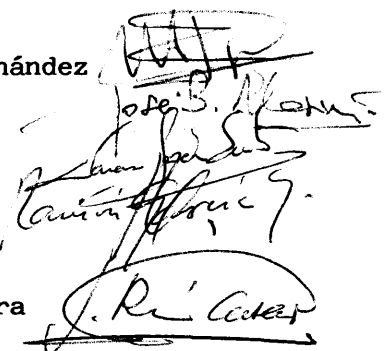
PRESIDENTE: Miguel A. Lagunas Hernández

VOCALES: José B. Mariño Acebal

Narciso García Santos

Ramón García Gómez

SECRETARIO: José R. Casar Corredera



Acuerda otorgarle la calificación de...AP.T.O..C.U.M..LAUDE

R. 12.903

Madrid, 22 DE MAYO DE 1987

PROLOGO

Hace ya algunos años decidí iniciar mi actividad profesional en el marco de la empresa privada, pero poco tiempo después empezaba a echar en falta un entorno de estudio e investigación donde saciar mis verdaderas inquietudes profesionales. Afortunadamente encontré el lugar idóneo donde se me permitía llevar a cabo aquello que deseaba en el Grupo de Teoría y Tratamiento de la Señal de la ETSIT de Madrid. Aquí, más que un lugar de trabajo, encontré a amigos que facilitaron que muchas de mis lagunas de conocimiento se fueran desvaneciendo. Por tanto, quiero agradecer individualmente a todos ellos, el estímulo que han supuesto para mí en estos años, así como la paciencia de que han hecho gala en las numerosas discusiones que han hecho posible la elaboración de este trabajo.

Finalmente agradecer la confianza que el Profesor Aníbal R. Figueiras Vidal, director de este trabajo, puso en mí desde mi llegada al grupo.

Febrero, 1987

A Ana

RESUMEN

La mayoría de los problemas de estimación están basados en la minimización de una función de coste cuadrático del error de estimación. Este criterio de calidad posee la propiedad de proporcionar el mejor y más sencillo estimador posible en un gran número de aplicaciones. Sin embargo, existen situaciones concretas donde la selección de funciones de costo más generales ofrecen similares o incluso mejores prestaciones que la minimización cuadrática. El presente trabajo introduce criterios de optimización alternativos, los cuales están basados en la minimización de la norma de orden k del error de estimación.

Con objeto de investigar estas nuevas funciones de costo, se han elegido dos campos de aplicación diferentes: Estimación Espectral Paramétrica y Filtrado Adaptativo. Dentro del primero, se han analizado dos métodos clásicos y eficientes, que necesitan la resolución de sistemas lineales sobredimensionados, por medio de procedimientos de procesado en bloque. Sus formulaciones y soluciones respectivas provienen de las minimizaciones norma L_1 (de error absoluto) y norma L_2 (de error cuadrático). En este análisis comparativo, la solución L_1 se comporta mejor en muchos casos debido al carácter picudo del vector de error a minimizar.

En relación con el Filtrado Adaptativo, se estudia una familia general de algoritmos adaptativos de gradiente estocástico. Está basada en la minimización del gradiente instantáneo de la k -ésima potencia absoluta del error total. El análisis generalizado se apoya en un esquema de Identificación de Planta, tanto en el dominio del tiempo como en el de la frecuencia. Dicho análisis proporciona las ecuaciones teóricas recurrentes para la varianza del error bajo supuestos diferentes de ruido de planta

según el orden del algoritmo adaptativo. También se evalúan las condiciones de convergencia del correspondiente paso de adaptación y el desajuste final del error. Los resultados teóricos del análisis son verificados por medio de simulaciones en el ordenador, que demuestran que las prestaciones de convergencia del algoritmo L_k dependen principalmente de la distribución estadística del ruido aditivo de planta y/o del tamaño del error. Se han simulado, además, utilizando algoritmos adaptativos L_k , algunos sistemas reales que hacen uso de filtros adaptativos. Los resultados obtenidos permiten considerar algunos de los algoritmos L_k como una clara y ventajosa alternativa al clásico algoritmo LMS ($k=2$) bajo ciertas condiciones de aplicación.

Finalmente, se presenta un resumen con las principales conclusiones generales y nuevas líneas de investigación suscitadas por este nuevo criterio de optimización

SUMMARY

Most of the estimation problems are based on the minimization of a quadratic cost function of the estimation error. This quality measure has the property of providing the "best" and easiest estimator (linear) in many applications. Nevertheless, there are other specific situations where the selection of more general cost functions offer similar or even better performance than the quadratic minimization. This work introduces alternative optimization criteria based on the minimization of a k th-norm of the estimation error.

In order to investigate these new cost functions, two different fields of application have been selected: Parametric Spectral Estimation and Adaptive Filtering. Within the first one, two efficient and classical methods, that need to solve overdetermined linear systems, are analyzed using two batch processing procedures. Their formulations and solutions come from the minimization of the L_1 -norm (absolute error) and L_2 -norm (quadratic error), respectively. In this comparative analysis, the L_1 -solution performs better in many cases because of the spiky character of the error vector to be minimized.

In relation with the Adaptive Filtering field, a general family of stochastic gradient adaptive algorithms is studied. It is based on the minimization of the instantaneous gradient of the k th-absolute-power of the overall error. The generalized analysis is provided for a Plant Identification scheme in both the time and frequency domains; it gives the theoretical recursive equations for the error variance under different plant-noise assumptions and according to the order of the adaptive algorithm. The convergence conditions for the corresponding adaptation step

and the steady state misadjustments are also evaluated. The theoretical results are verified by means of some computer simulations. They show that the convergence performance of the Lk-algorithm mainly depends on the statistical distribution of the additive plant-noise and/or the error magnitude. Furthermore, some real systems, which make use of adaptive filters, have been simulated by using different Lk-adaptive algorithms. The obtained results allow to consider some of these Lk-algorithms as a clearly advantageous alternative to the classical LMS algorithm ($k=2$) in some cases.

Finally, a summary of the main general conclusions and new lines of research open from these new optimization criteria are presented.

INDICE

I.- <u>CAPITULO INTRODUCTORIO</u>	1
I.1.- Introducción.....	2
I.2.- Revisión de los métodos de Prony-Hildebrand y Cadzow.....	8
I.2.1.- Desarrollo original de ambos métodos.....	9
I.2.2.- Desarrollo en el dominio de la frecuencia.....	13
I.3.- Revisión de los métodos adaptativos de gradiente.....	16
I.4.- Contenido de la Tesis.....	22
II.- <u>ESTIMADORES ESPECTRALES L2 vs. L1</u>	25
II.1.- Introducción: Formulación Simplex de la minimización norma L1.....	26
II.2.- Prony-Hildebrand en el tiempo.....	28
II.3.- Segundo método de Cadzow en el dominio de la autocorrelación	33
II.4.- Prony-Hildebrand en la frecuencia.....	38
II.5.- Segundo método de Cadzow en la frecuencia.....	43
II.6.- Conclusiones.....	47
III.- <u>IDENTIFICACION ADAPTATIVA DE PLANTA EN EL DOMINIO DEL TIEMPO</u>	49
III.1.- Introducción del algoritmo.....	50
III.2.- Análisis del error residual cuadrático medio.....	53
III.2.1.- Expresión recurrente generalizada para la varianza del error residual.....	54
III.2.2.- Caso 1: Planta sin ruido	55
III.2.3.- Caso 2: Planta ruidosa	56
III.3.- Condiciones necesarias para la convergencia del error cuadrático medio.....	64
III.3.1.- Condicionamiento del polinomio $P(u,n)$	64

III.3.2.- Ausencia de ruido de planta.	66
III.3.3.- Presencia de ruido de planta.	70
III.4.- Evaluación del desajuste final y relación con las constantes de tiempo de adaptación.....	76
III.4.1.- Análisis de la convergencia en media.	77
III.4.2.- Desajustes medios finales	79
III.5.- Optimización del paso de adaptación.	84
III.5.1.- Ausencia de ruido de planta.	85
III.5.2.- Presencia de ruido de planta.	87
III.6.- Simulación de los algoritmos y conclusiones finales.	93
III.6.1.- Convergencia con ausencia de ruido de planta.	94
III.6.2.- Convergencia con ruido de planta.	97
III.6.3.- Convergencia óptima con ruido de planta.	107
III.6.4.- Idéntica velocidad y diferente desajuste final.	112
III.6.5.- Conclusiones.	114
IV.- <u>IDENTIFICACION ADAPTATIVA DE PLANTA EN EL DOMINIO DE LA FRECUENCIA</u>	117
IV.1.- Introducción del algoritmo.	118
IV.2.- Análisis del error residual cuadrático medio.	124
IV.2.1.- Expresión recurrente generalizada para la varianza del error residual.	125
IV.2.2.- Caso 1: Planta sin ruido.	126
IV.2.3.- Caso 2: Planta ruidosa.	127
IV.3.- Condiciones necesarias para la convergencia del error cuadrático medio.	131
IV.3.1.- Condicionamiento del polinomio $P(m,1)$	131
IV.3.2.- Ausencia de ruido de planta.	133

IV.3.3.- Presencia de ruido de planta.	134
IV.4.- Simulación de los algoritmos y conclusiones finales.	138
IV.4.1.- Convergencia con ruido de planta.	140
IV.4.2.- Conclusiones.	149
V.- <u>APLICACIONES DE LA NUEVA FAMILIA DE ALGORITMOS ADAPTATIVOS</u>	152
V.1.- Introducción.	153
V.2.- Cancelación de ecos con transmisión de voz.	155
V.3.- Cancelación de ecos con transmisión de datos.	162
VI.- <u>CONCLUSIONES FINALES Y LINEAS DE TRABAJO</u>	166
VI.1.- Conclusiones.	167
VI.2.- Sugerencias de futuras líneas de trabajo.	174
VII.- <u>REFERENCIAS</u>	176
VIII.- <u>APENDICES</u>	186
A1.- Cálculo del mínimo de la función objetivo	187
A2.- Estudio de la convexidad de la función objetivo.	189
A3.- Cálculo de $S_1(n)$ para ruido de planta Gaussiano.	191
A4.- Cálculo de $S_1(n)$ para ruido de planta Laplaciano.	195
A5.- Cálculo de $S_1(n)$ para ruido de planta Uniforme.	199
A6.- Cálculo de $S_1(n)$ para ruido de planta Binario.	202
A7.- Cálculo de las integrales I_A, I_B e I_C utilizadas en el Apéndice A.3.	205
A8.- Cálculo de las integrales I_D, I_E e I_F utilizadas en el Apéndice A.4.	208
A9.- Cálculo de las integrales I_G, I_H, I_I e I_J utilizadas en los Apéndices A.5 y A.6.	210
A10.- Cálculo de $S_2(n)$ con ruido de planta.	213
A11.- Caracterización de la v.a. compleja $X(m,l)$ y cálculo de su momento absoluto de orden r	216

A12.- Evaluación de la esperanza $E[E+U ^{2k-2}]$	219
A13.- Evaluación de la esperanza $E[E+U ^{k-2} (E ^2 + \text{Re}(EU^*))]$	221
A14.- Evaluación de Q_1 para los algoritmos L2 y L4.....	227
A15.- Evaluación de Q_1 para los algoritmos L1 y L3.....	229

CAPITULO I

CAPITULO INTRODUCTORIO

I.1.- INTRODUCCION

Uno de los principales retos del Procesado de Señal es la extracción de información de un registro disponible de señal. Para ello se utiliza un sistema o procesador adecuado que entrega a su salida una señal deseada (modelo señal-sistema-señal). Este modelo permite estudiar de una forma directa la transformación o proceso que sufre el registro conocido de señal con el objeto de obtener una determinada información. En un gran número de situaciones prácticas esta información, bien del procesador o de una de las señales, puede ser requerida en forma paramétrica ya que previamente se posee conocimiento sobre su estructura básica. Al proceso de extracción de los parámetros definitorios del sistema y/o señal a partir de observaciones sobre el registro de datos disponibles se le conoce como estimación paramétrica.

Concretando más, el problema de estimación paramétrica se fundamenta en la minimización de una función (estadístico) de una señal (proceso) considerada como de error, que depende del conjunto de parámetros buscados y está sujeta a una serie de restricciones. Generalmente, se elige un función par del error, minimizando así por igual los errores positivos y negativos. No obstante, la elección de tal función (llamada de coste) es la que condiciona el resultado del proceso de estimación, puesto que al enfatizar más unos valores de error que otros se fuerza a la solución que minimiza dicha función de coste a ser, en principio, distinta a la que minimizaría otra enfatización diferente del error (KLEI, 1964). Como ejemplo, si se toma una potencia del valor absoluto del error, los valores de error superiores a la unidad estarán más enfatizados (costarán más)

cuanto mayor sea la potencia y viceversa con los menores a la unidad. Sin embargo, en la mayoría de los problemas de estimación, se elige una función cuadrática ya que bajo la habitual hipótesis Gaussiana para los datos, el estimador de máxima verosimilitud correspondiente es lineal en los datos, lo cual hace recomendable su elección en la mayoría de los problemas de estimación paramétrica. Otro aspecto favorable a la elección de una función cuadrática (o coste cuadrático) del error, frente a otras, es el de obtener el mejor estimador bayesiano posible bajo condiciones no muy restrictivas para la f.d.p. de los datos (p. ej. simetría), según el Teorema de Invarianza [VAN TREES,1968], [SRINATH,1979]].

No obstante, la robustez de la solución cuadrática en el problema de estimación no es siempre la adecuada, ya que las condiciones del problema pueden diferir de las que hacen recomendable su elección (asumidas en el anterior Teorema). Como ejemplo, en [LEVY,1981] se indicó la ventaja de la minimización norma L_1 (o valor absoluto) a la de norma L_2 (o valor cuadrático) cuando la solución esperada (en ese caso un espectro de potencia de sinusoides en ruido) presentaba picos muy marcados. Esto es debido a que la solución L_2 posee un carácter más distribuido debido al alto coste (su valor al cuadrado) que deben pagar dichos picos en la solución final. Para ilustrar tal afirmación, en el siguiente ejemplo numérico se demuestra dicha ventaja al comparar las correspondientes soluciones L_1 y L_2 .

Considerese el siguiente sistema subdimensionado (con más de una solución posible) de dos ecuaciones y cuatro incógnitas:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + x_4 &= 3 \\ 3x_1 + x_2 - x_3 + x_4 &= 2 \end{aligned} \quad (\text{i.l.l})$$

y se desea minimizar :

$$\sum_{i=1}^4 x_i^2 \quad (\text{solución L2})$$

$$\sum_{i=1}^4 |x_i| \quad (\text{solución L1})$$

sujeto a (i.l.l). Sus correspondientes soluciones son:

$$\text{Para L2: } x_1=0.544, x_2=0.474, x_3=0.404, x_4=0.298$$

$$\text{Para L1: } x_1=0.9, x_2=0, x_3=0.7, x_4=0.$$

Aunque la carga computacional requerida por la solución L1 (en este caso el método Simplex) es superior a la correspondiente L2, se verifica que en el caso L2 la solución es prácticamente plana y para L1 aparecen dos picos bien marcados (al considerar los nulos x_2 y x_4). Trabajos posteriores al de Levy han utilizado el mismo razonamiento básico para la elección de una solución L1 en problemas de estimación espectral paramétrica ([KEYBL,1983], [FIGUEIRAS-VIDAL,1985a]) cuando se esperaban espectros con picos muy pronunciados. Tales alternativas han demostrado ofrecer mejores resultados que los correspondientes a minimizaciones cuadráticas. Este hecho justifica el interés de introducir nuevos criterios de optimización para estimadores o algoritmos en aquellos problemas con condicionamientos particulares distintos a los que hacen recomendable la utilización de una minimización cuadrática.

Si se adopta un estimador lineal, dada su sencillez de implementación, la señal de error es, en general, alguna combinación lineal de señales relacionadas con la entrada y salida del procesador. El problema de estimación consiste en encontrar los parámetros definitorios que minimicen la función elegida del error actuando sobre los registros de señales de entrada. Esta actuación puede producirse de una sola vez y forma única sobre dicho registro (procedimiento bloque) o de forma iterativa reduciendo progresivamente la función del error (procedimiento recurrente). Ambos procedimientos son aplicables, dependiendo fundamentalmente de la disponibilidad del registro de entrada completo y/o del volumen de cómputo necesitado según el criterio de optimización (función del error elegida).

Las técnicas recurrentes, inicialmente procedentes en su mayoría de la ingeniería de control, son cada vez más utilizadas debido, en gran medida, a la aparición de procesadores digitales apropiados a los algoritmos existentes y en desarrollo, los cuales facilitan un adecuado ajuste y/o seguimiento de la evolución de un conjunto de parámetros [CLAASEN,1983]. Sin embargo, al igual que en los procedimientos bloque de estimación paramétrica, debido a su robustez, la función objetivo a minimizar más explotada ha sido la cuadrática.

El objetivo genérico del presente trabajo es la exploración de posibles ventajas en el uso de funciones de coste alternativas a la cuadrática, tanto en procedimientos de estimación bloque como en estimación adaptativa. Para ello, y debido a su probado interés, se han seleccionado campos tan significativos de TDS (Tratamiento Digital de la Señal) como son los de estimación espectral paramétrica e identificación adaptativa de planta.

Como problema de estimación bloque, se analizan un par de procedimientos ya conocidos en el ámbito de la estimación espectral paramétrica, planteando formulaciones alternativas a las usuales tanto en el dominio del tiempo como en el de la frecuencia. Tales formulaciones son originadas por la elección de una solución mínimo valor absoluto del registro de error. Para su resolución se utilizan técnicas de programación lineal (método Simplex). En concreto, se analizan los algoritmos debidos a Cadzow [CADZOW,1982a], [BRY,1982] y Prony-Hildebrand [HILDEBRAND,1956], [HOLTZ,1973] donde la minimización del módulo del error (norma L1) ofrece un menor sesgo en la localización de los picos espectrales, así como una mayor resolución espectral que la correspondiente a la minimización del error al cuadrado (norma L2). En el Apartado 1.2 se revisan ambos procedimientos.

A continuación, se completa y analiza una nueva familia de algoritmos adaptativos, ya introducida por Walach y Widrow [WALACH,1984], basada en métodos de gradiente estocástico, en el que se proponen algoritmos derivados de la minimización del gradiente instantáneo de una potencia par del error residual de adaptación (el ya conocido LMS introducido por Widrow [WIDROW,1960] es un caso particular de los anteriores). La generalización consiste en minimizar el gradiente instantáneo de cualquier potencia genérica del módulo de una señal de error (se incluyen las potencias impares), considerando formulaciones de los algoritmos en los dominios del tiempo y la frecuencia. Como soporte del planteamiento y análisis del problema se elige un esquema de identificación de planta transversal, dada la versatilidad de dicho sistema para modelar diferentes situaciones típicas en comunicaciones o ingeniería de control [CLAASEN,1983]. Fundamentalmente se analizan las condiciones necesarias para la

convergencia y la comparación de prestaciones de la familia de algoritmos dentro de un amplio marco de situaciones, tales como diferentes ruidos de planta, elección del paso de adaptación, longitud de la planta, etc. En el Apartado 1.3 se revisan someramente los métodos de gradiente.

Finalmente, en el Apartado 1.4 de este Capítulo introductorio, se plantean con detalle los objetivos específicos de la Tesis, presentando además la distribución por Capítulos del trabajo.

1.2.- REVISION DE LOS METODOS DE PRONY-HILDEBRAND Y CADZOW

El análisis espectral se fundamenta en encontrar un estimador de la densidad espectral de potencia de un proceso estocástico del que se conoce un registro limitado de puntos [HAYKIN,1979], [CHILDERS,1978], [MAKHOUJ,1975], [KAY,1981], [LACOSS,1971]. Frente a los métodos tradicionales de estimación espectral, como puede ser el periodograma (equivalente a un ajuste cuadrático de los datos a los obtenidos mediante un modelo armónico de generación de señales [BLOOMFIELD,1976]), que necesitan registros largos de puntos para resolver en frecuencia dos rayas próximas, existen los métodos paramétricos, los cuales, en función de algún conocimiento previo, tratan de ajustar un modelo de generación de señal al registro de datos, infiriendo así en el comportamiento de los datos fuera del registro. Este hecho permite que, incluso con registros cortos de datos, la resolución en frecuencia de estos métodos sea superior a la de los métodos tradicionales. Los algoritmos propuestos por Prony-Hildebrand y Cadzow, están incluidos dentro de los paramétricos, y son de conocido interés para la estimación de la localización espectral de sinusoides contaminadas con ruido aditivo. Ambos métodos necesitan la resolución de un sistema sobredimensionado de ecuaciones lineales, con la diferencia que mientras el método de Prony está definido en el dominio del tiempo (trabaja directamente con los datos), el de Cadzow lo está en el del desplazamiento de la autocorrelación.

I.2.1.- Desarrollo original de ambos métodos

Ambos métodos se basan en la propiedad siguiente:

Sea $s(n)$ un registro de N puntos compuesto por la suma de p sinusoides

$$s(n) = \sum_{k=1}^p A_k \cos(\omega_k n + \theta_k) \quad (i.2.1.1)$$

donde θ_k es una fase aleatoria distribuida uniformemente en $[0, 2\pi]$. Es bien conocido [KAY, 1981]

$$\sum_{i=0}^{2p} a_i s(n-i) = 0 \quad (i.2.1.2)$$

y en [CADZOW, 1982] :

$$\sum_{i=0}^{2p} a_i r(n-i) = 0 \quad (i.2.1.3)$$

con

$$r(n) = \sum_{k=1}^p (A_k^2 / 2) \cos(\omega_k n) \quad (i.2.1.4)$$

donde a_0 ($a_i=1$, a_{2p-i}) son los coeficientes de un polinomio cuyas raíces son $\exp(\pm j\omega_k)$, siendo ω_k las pulsaciones de las sinusoides.

Añadiendo cualquier ruido $w(n)$ al registro $s(n)$

$$x(n) = s(n) + w(n) \quad (i.2.1.5)$$

se verifica (para Prony):

$$\sum_{i=0}^{2p} a_i x(n-i) = \sum_{i=0}^{2p} a_i w(n-i) = e(n) \quad (i.2.1.6)$$

donde $x(n)$ es la señal más ruido y $e(n)$ una versión filtrada del ruido; y para Cadzow (asumiendo $w(n)$ blanco de media nula y varianza σ^2):

$$\sum_{i=0}^{2p} a_i r(n-i) = \sigma^2 \sum_{i=0}^{2p} a_i \delta(n-i) \quad (i.2.1.7)$$

Esta es la expresión formal de las ecuaciones de Yule-Walker de un modelo ARMA (2p, 2p) donde los coeficientes de las partes AR y MA son idénticos. Este es el motivo por el que también se conoce a este procedimiento como modelo paso-todo de sinusoides en ruido blanco.

Habitualmente, no se suele disponer del verdadero registro de autocorrelación $r(n)$, sino de los propios datos $x(n)$. Para paliar este hecho, se usa una estimación $r'(n)$ de la secuencia de autocorrelación $r(n)$ que incorpora obviamente un error (ruido) de estimación. Por tanto (i.2.1.7) se puede reescribir como:

$$\sum_{i=0}^{2p} a_i r'(n-i) = e(n) \quad (i.2.1.8a)$$

donde el registro $e(n)$ está compuesto por la suma del error de estimación de $r'(n)$ (presente en todas las ecuaciones) y por una cantidad $\sigma^2 a_i$ (sólo en las ecuaciones centrales) debido a la varianza del ruido presente en los datos $x(n)$. Explícitamente $e(n)$ se puede expresar en forma vectorial como:

$$\underline{e} = \underline{e}' + \underline{e}_w \quad (i.2.1.8b)$$

siendo \underline{e}' el ruido de estimación y \underline{e}_w debido a la presencia de $w(n)$, donde:

$$\underline{e}_w = \sigma^2 [0 \dots 0 \ 1 \ a_1 \dots a_{2p} \ 0 \dots 0]^t \quad (i.2.1.8c)$$

En este punto se observa que (i.2.1.6) e (i.2.1.8a) son formalmente idénticas salvo el dominio de actuación de cada una de ellas y la composición de su término independiente.

El objetivo final de estos dos métodos es encontrar el conjunto de valores a_i 's que minimice una función del error total $e(n)$ para posteriormente resolver la ecuación $A(z) = a_0 + a_1 z + \dots + a_{2p} z^{2p} = 0$, ya que los argumentos de sus raíces han de estar próximos a los de las sinusoides $s(n)$.

a) Método de Prony-Hildebrand

El método de Prony-Hildebrand propone la resolución de (i.2.1.6) utilizando una minimización L2. Debido a que las soluciones de $A(z)=0$ deseadas han de encontrarse en el círculo unidad se impone, además, simetría par en los coeficientes a_i 's. Este hecho, al forzar pares de raíces reciprocas, ayuda a la aparición de raíces conjugadas de módulo unidad (de hecho, sólo excepcionalmente se observan otras [KAY,1981], [SULLIVAN,1978]).

Las ecuaciones formadas por (i.2.1.6) usando un registro de N puntos ($N > 2p$) pueden escribirse:

$$\underline{X} \underline{a} = \underline{e} \quad (i.2.1.9)$$

$$\underline{X} = \begin{bmatrix} x(p+1) & \dots & x(1) \\ \dots & \dots & \dots \\ x(N-p) & \dots & x(N-2p) \end{bmatrix} + \begin{bmatrix} x(p+1) & \dots & x(2p+1) \\ \dots & \dots & \dots \\ x(N-p) & \dots & x(n) \end{bmatrix} \quad (i.2.1.10a)$$

$$\underline{a} = [a_{p/2} \quad a_{p-1} \quad \dots \quad 1]^t \quad (i.2.1.10b)$$

$$\underline{e} = [e(2p+1) \quad e(2p+2) \quad \dots \quad e(N)]^t \quad (i.2.1.10c)$$

b) Método de Cadzow.

Cadzow propone dos alternativas para la solución del sistema sobredimensionado de ecuaciones: el primero excluye las ecuaciones que incorporan el efecto de la varianza del ruido aditivo y blanco $w(n)$ disponiendo así tan solo de las ecuaciones con error de estimación y el segundo las acepta todas, considerando el efecto de la varianza del ruido como parte del error total a minimizar.

El primer procedimiento es preferible cuando la longitud del registro de datos $x(n)$ es suficientemente grande, ya que se puede disponer de un número alto de ecuaciones con buenas estimas de la secuencia de autocorrelación $r(n)$ y por consiguiente una buena estimación de las posiciones de las rayas espectrales al eliminar el efecto perturbador de las ecuaciones de ruido. Sin embargo, cuando el registro $x(n)$ es corto no se deben despreciar esas ecuaciones, ya que la estimación de la secuencia $r(n)$ estaría seriamente limitada y el número total de ecuaciones sería reducido. Este método, por otro lado, permite su utilización bajo hipótesis de ruido coloreado, ya que se tolera el error adicional introducido al no poder determinar exactamente las ecuaciones donde el efecto perturbador del ruido coloreado actúa.

En este trabajo se ha elegido la segunda versión propuesta por Cadzow debido a un hecho fundamental: el efecto de ruido producido por la varianza en el registro de error $e(n)$, semejante a ruido impulsivo, el cual justifica, como se verá más adelante, la minimización del error absoluto total.

Se adopta el modelo original de Cadzow, sin simetría par en los coeficientes con $N+1$ desplazamientos ($r'(-N/2), \dots, r'(0), \dots, r'(N/2)$), ya que el forzar simetría en este procedimiento facilita la aparición de máximos espectrales más picudos pero más sesgados [FIGUEIRAS-VIDAL, 1985b].

Las ecuaciones formadas por (i.2.1.8), usando un registro estimado de autocorrelaciones de $N+1$ puntos ($N > 2p$), pueden escribirse:

$$R' a = e \quad (i.2.1.11)$$

donde

$$\underline{R}' = \begin{bmatrix} r'(-N/2+2p) & \dots & r'(-N/2) \\ \dots & \dots & \dots \\ r'(N/2) & \dots & r'(N/2-2p) \end{bmatrix} \quad (i.2.1.12a)$$

$$\underline{a} = [1 \ a_1 \ \dots \ a_{2p}]^t \quad (i.2.1.12b)$$

$$\underline{e} = \underline{e}' + \underline{e}_w \quad (i.2.1.12c)$$

La solución de los sistemas (i.2.1.9) e (i.2.1.11) bajo la restricción de error total cuadrático $\sum_m e^2(m)$ mínimo es bien conocida [KAY,1981]

$$\underline{a} = [\underline{Y}^H \ \underline{Y}]^{-1} \alpha \underline{u} \quad (i.2.1.13a)$$

con

$$\underline{u} = [1 \ 0 \ \dots \ 0]^t \quad (i.2.1.13b)$$

donde

$$\underline{Y} = \begin{cases} \underline{X} & \text{(Prony)} \\ \underline{R}' & \text{(Cadzow)} \end{cases} \quad (i.2.1.13c)$$

(y α es seleccionado para que $a_0=1$).

I.2.2.- Desarrollo en el dominio de la frecuencia

Ambos métodos admiten un desarrollo paralelo en el dominio de la frecuencia conduciendo a sistemas de ecuaciones formalmente similares [PAEZ-BORRALLLO,1985b].

Para el método de Prony-Hildebrand, tomando transformadas de Fourier en ambos lados de (i.2.1.6) se obtiene:

$$X(\omega) \sum_{i=0}^{2p} a_i e^{-j\omega i} = E(\omega) \quad (i.2.2.1)$$

discretizando ω uniformemente a lo largo de la circunferencia como

$\omega = 2\pi k/N$ y descomponiendo (i.2.2.1) en partes real e imaginaria:

$$E_r(k) = \sum_{i=0}^{2p} a_i \phi_r(k,i) \quad (i.2.2.2a)$$

$$(0 \leq i \leq 2p), (0 \leq k \leq N-1)$$

$$E_i(k) = \sum_{i=0}^{2p} a_i \phi_i(k,i) \quad (i.2.2.2b)$$

donde

$$\phi_r(k,i) = X_r(k) \cos(2\pi ki/N) + X_i(k) \sin(2\pi ki/N) \quad (i.2.2.3a)$$

$$\phi_i(k,i) = X_i(k) \cos(2\pi ki/N) - X_r(k) \sin(2\pi ki/N) \quad (i.2.2.3b)$$

Considerando que el registro $x(n)$ es real, existe simetría par según k en $\phi_r(k,i)$ y $\phi_i(k,i)$ y utilizando $a_0=1$ y $a_i=a_{2p-i}$ se puede reescribir (i.2.2.2a,b) en forma vectorial como:

$$\begin{pmatrix} \underline{\phi}_r \\ \underline{\phi}_i \end{pmatrix} \underline{a} = \begin{pmatrix} \underline{E}_r \\ \underline{E}_i \end{pmatrix} \quad (i.2.2.4)$$

donde denominando indistintamente (ϕ, E) a (ϕ_r, E_r) y (ϕ_i, E_i) :

$$\underline{\phi}_{k,i} = \phi(k,i) + \phi(k, 2p-i) \quad (i.2.2.5a)$$

$$\underline{E} = \begin{pmatrix} \dots \\ E(k) \\ \dots \end{pmatrix} \quad (i.2.2.5b)$$

para $(0 \leq k \leq N/2)$ y $(0 \leq i \leq p)$.

De forma equivalente para el método de Cadzow, tomando transformadas en (i.2.1.8) se obtiene:

$$S(\omega) = \sum_{i=0}^{2p} a_i e^{-j\omega i} = E(\omega) \quad (i.2.2.6)$$

donde la transformada $S(\omega)$ de $r'(n)$ es una función real y par, por lo que al discretizar igualmente $\omega = 2\pi k/N$ solo se necesitan $N/2+1$ valores de $S(k)$. Descomponiendo (i.2.2.6) en partes real e imaginaria:

$$E_r(k) = \sum_{i=0}^{2p} a_i \phi_r(k,i) \quad (i.2.2.7a)$$

(0 ≤ i ≤ 2p), (0 ≤ k ≤ N/2)

$$E_i(k) = \sum_{i=0}^{2p} a_i \phi_i(k,i) \quad (i.2.2.7b)$$

donde

$$\phi_r(k,i) = S(k) \cos(2\pi ki/N) \quad (i.2.2.8a)$$

$$\phi_i(k,i) = -S(k) \sin(2\pi ki/N) \quad (i.2.2.8b)$$

que se puede escribir en forma vectorial como:

$$\begin{pmatrix} \phi_r \\ \phi_i \end{pmatrix} \underline{a} = \begin{pmatrix} E_r \\ E_i \end{pmatrix} \quad (i.2.2.9)$$

donde

$$\phi_{k,i} = \phi(k,i) \quad (i.2.2.10a)$$

$$\underline{E} = \begin{pmatrix} \dots \\ E(k) \\ \dots \end{pmatrix} \quad (i.2.2.10b)$$

para (0 ≤ k ≤ N/2) y (0 ≤ i ≤ 2p).

Ambos procedimientos presentan en el dominio de la frecuencia sistemas de ecuaciones formalmente equivalentes a los obtenidos anteriormente. Sin embargo, al igual que ocurre en el dominio del tiempo, la solución L2 (i.2.1.13), planteada por sus autores, no siempre conduce a la mejor solución posible, ya que debido a que el espectro de la señal (sinusoides más ruido) es de tendencia picuda, la transformada del error presenta igualmente picos de error espectral relativamente grandes en las proximidades de las sinusoides, aconsejando una solución norma L1 [LEVY,1981], [PAEZ-BORRALLA,1985a], [MARTINELLI,1986], [FIGUEIRAS-VIDAL,1985b].

I.3.- REVISION DE LOS METODOS ADAPTATIVOS DE GRADIENTE

Una de las técnicas adaptativas de estimación (o seguimiento) de un conjunto de parámetros más usada (por su sencillez de implementación como contrapartida a las técnicas RLS "Recursive Least Squares" [PROAKIS,1983], [HASTINGS,1977], [FALCONER,1978], [LJUNG,1983]) es la basada en la búsqueda iterativa del mínimo de una función de dichos parámetros según la dirección opuesta a la marcada por el gradiente de dicha función. Este método exige, generalmente, funciones objetivos con un único mínimo absoluto (superficies convexas de forma de taza en los parámetros) para evitar enganches (gradientes nulos) en posibles mínimos locales (ver textos generales [EYKOFF,1974], [HAYKIN,1984], [HONIG,1984], [WIDROW,1985], [ALEXANDER,1986] para una perspectiva global de los métodos adaptativos de gradiente).

Tal método es también conocido como "descenso de máxima pendiente" (en inglés "steepest descent") y se caracteriza porque los sucesivos desplazamientos necesarios para el correcto ajuste de los parámetros están totalmente determinados, ya que se supone conocido el gradiente de la función objetivo en todo punto. Evidentemente, tal conocimiento no es habitual y en la práctica se suelen utilizar métodos adaptativos de gradiente estocástico, los cuales utilizan una estima del gradiente calculada a partir de los observables.

En la actualidad y dentro del campo de aplicaciones de los algoritmos de gradiente estocástico (para función de coste cuadrático) sobresalen fundamentalmente dos bien conocidas: algoritmos adaptativos para sistemas con estructura transversal [WIDROW,1960], [WIDROW,1976] y para sistemas con

estructuras en celosía [GRIFFITHS,1977], [GIBSON,1980], [HONIG,1981]. En el primer caso, se adaptan los coeficientes del filtro y en el segundo los coeficientes de reflexión de cada celosía. Como diferencias fundamentales entre ellos cabe destacar que, debido a la implementación en celosía, la velocidad de convergencia de estos algoritmos es independiente de la dispersión de autovalores de la matriz de autocorrelación de la entrada [MEDAUGH,1981] y superior a la de los algoritmos implementados con estructura transversal a costa de un moderado aumento en la carga computacional.

En el presente trabajo se plantea el estudio con sistemas de estructura transversal. El sistema mostrado en la Fig. F1.1 ayuda a una formulación transversal del algoritmo. Se supone el instante temporal n .

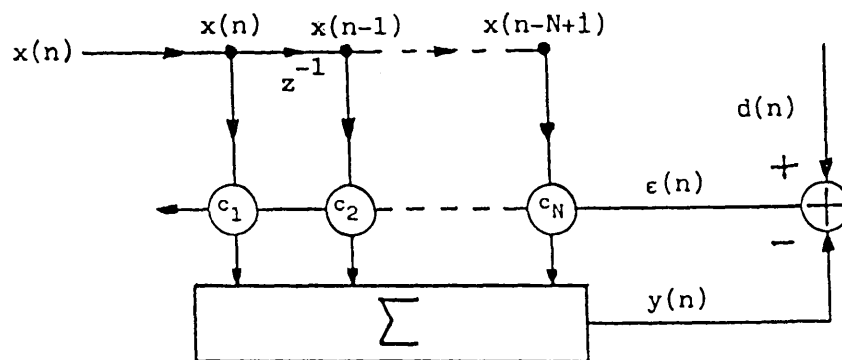


Figura F1.1

con:

$$\epsilon(n) = d(n) - y(n) \quad (i.3.1)$$

$$y(n) = \sum_{i=1}^N c_i(n) x(n-i+1) \quad (i.3.2)$$

donde: x es el proceso de entrada, c el vector de coeficientes adaptativos, d la referencia deseada y ϵ el error cometido.

Si $C(\epsilon)$ es la función (estadístico) objetivo a minimizar, en general función de x , d , c o de estadísticos suyos, el método de máxima pendiente o gradiente determinista se define formalmente:

Descenso de máxima pendiente:

$$\underline{c}(n+1) = \underline{c}(n) - \mu \underline{\nabla} [C(\epsilon)] \quad (i.3.3)$$

donde

$$\underline{\nabla} (.) = \frac{d(.)}{dc} = \left[\frac{d(.)}{dc_1} \quad \frac{d(.)}{dc_2} \quad \dots \quad \frac{d(.)}{dc_N} \right]^t \quad (i.3.5)$$

es la definición de gradiente y μ es el coeficiente de adaptación o paso que controla la estabilidad y velocidad de convergencia del proceso.

Particularizando la función objetivo como error cuadrático medio $C(\epsilon) = E[\epsilon^2]$ y calculando su gradiente, el algoritmo resulta ser [WIDROW,1985] :

$$\underline{c}(n+1) = \underline{c}(n) - 2\mu \underline{Rc}(n) + 2\mu \underline{P} \quad (i.3.6)$$

con el mínimo del gradiente verificando:

$$\underline{\nabla} (.) = -2\underline{P} + 2\underline{Rc} = 0 \quad (i.3.7)$$

de donde se deriva la solución óptima (expresión matricial de la ecuación de Wiener-Hopf que resuelve el problema de estimación lineal cuadrática media mínima) [WIENER,1949] cuya solución es:

$$\underline{c}_{opt} = \underline{R}^{-1} \underline{P} \quad (i.3.8)$$

donde \underline{R} y \underline{P} son los estadísticos: $\underline{R} = E[\underline{X}_n \underline{X}_n^t]$ y $\underline{P} = E[d(n) \underline{X}_n]$ con $\underline{X}_n = [x(n) \dots x(n-N+1)]^t$.

Sin embargo, habitualmente, no se dispone del vector gradiente en cada iteración y éste debe de ser estimado directamente de los datos X_n y $d(n), d(n-1), \dots$ o a partir de estimadores de los estadísticos \underline{R} y \underline{P} , lo que lleva a los métodos de gradiente estocástico. Estos algoritmos se expresan formalmente:

Gradiente estocástico:

$$\underline{c}(n+1) = \underline{c}(n) - \mu \hat{\underline{v}} [C(\epsilon)] \quad (i.3.9)$$

donde $\hat{\underline{v}}(\cdot)$ denota un estimador del gradiente.

Particularizando para el caso de superficie cuadrática, Widrow presentó en [WIDROW,1960] el conocido algoritmo LMS. Este procedimiento utiliza como estimador del gradiente de $E[\epsilon^2]$ el gradiente del error cuadrático instantáneo ϵ^2 resultando, según prueba la experiencia, un algoritmo robusto y de simple manejo, esto es:

$$\underline{c}(n+1) = \underline{c}(n) + 2\mu\epsilon(n)\underline{X}_n \quad (i.3.10)$$

Posteriormente Griffiths [GRIFFITHS,1969], presentó una nueva alternativa, ya que en su problema no conocía la señal de referencia $d(n)$ pero si el vector de correlación cruzada \underline{P} . Con ello, introdujo el siguiente algoritmo:

$$\underline{c}(n+1) = \underline{c}(n) - \mu \underline{X}_n \underline{X}_n^t \underline{c}(n) + 2\mu \underline{P} \quad (i.3.11)$$

Este algoritmo puede ser considerado una versión intermedia entre el de máxima pendiente y el LMS y ofrece mejores prestaciones que el LMS cuando

la correlación cruzada entre el proceso de entrada y la referencia es baja [FEUER,1985]. No obstante, al ser el algoritmo de Widrow una aproximación instantánea resulta más eficiente computacionalmente, hecho que ha conducido a la utilización casi generalizada de algoritmos LMS en aplicaciones muy diversas.

Posteriormente, muchos y completos trabajos han sido presentados analizando las características dinámicas y comportamiento estadístico de dicho algoritmo con estructura transversal. Algunos de ellos son: [WIDROW,1976] y [WIDROW,1984] que estudian su comportamiento con entradas no estacionarias; [FARDEN,1981], [DANIELL,1970] y [KIM,1975] con entradas correladas; [FEUER,1985] con entradas incorreladas y gaussianas; [GILLIN,1973] optimización del paso de adaptación, y recientemente un trabajo muy completo sobre su estabilidad y condiciones generales de convergencia en [GARDNER,1984].

Igualmente, dicho algoritmo ha sido objeto de algunas modificaciones interesantes que, aunque dificultan la identificación de la función objetivo de origen, aportan algunas ventajas. Entre las diferentes modificaciones realizadas cabe destacar las dos siguientes: NLMS (Normalized LMS), el cual reemplaza el paso de adaptación μ por $\mu / \|\underline{x}_n\|^2$ (norma cuadrática de x) con el fin de independizar la elección de μ de la potencia del proceso entrante [BITMEAD,1980], [WIDROW,1976], [HSIA,1983], [BERSHAD,1986]; y el "clipped" LMS, el cual sustituye el error o los datos en (i.3.10) por su signo aritmético con el fin de simplificar su implementación, evitando una cuantificación más fina [SARI,1982], [MONTAGNA,1982], [CLAASEN,1981].

Aún más, en [WIDROW,1975a] se extendió la aplicación del algoritmo LMS al campo complejo, dando este hecho origen a posteriores propuestas de utilización de dicho algoritmo sobre las rayas espectrales obtenidas con transformaciones FFT de bloques de datos [DENTINO,1978], [BERSHAD,1979], [BERSHAD,1982]. Tal implementación aprovecha además las ventajas de la ortogonalización de la señal que introduce el procesado DFT, la menor carga computacional que la formulación en bloques del algoritmo adaptativo conlleva [FERRARA,1980] y su separabilidad en frecuencia.

Finalmente, hay que mencionar un par de familias de esquemas adaptativos de gradiente estocástico, los cuales están basados en superficies de error distintas de la cuadrática. Una de ellas [WALACH,1984] ha sido ya nombrada en el apartado 1.1 de este Capítulo y es la que ha incitado parte de esta Tesis. La otra, conocida como CMA ("Constant Modulus Algorithm") es estudiada en [LARIMORE,1983] y minimiza el estadístico $E[||y(k)||^p - d^p |^q]$ (con $d > 0$). Sin embargo, aunque no se ha demostrado aun la unicidad de un mínimo global de la anterior expresión para cualquier par (p,q) , si se ha comprobado la eficiencia de alguno de ellos en problemas de igualación y supresión de interferencias particularmente con observables de envolvente constante.

I.4.- CONTENIDO DE LA TESIS

Tal como han sido presentados en el Apartado I.1, los objetivos fundamentales de esta Tesis son los de examinar los resultados ofrecidos por la aplicación de criterios de optimización diferentes al cuadrático en problemas de estimación, centrándose en casos representativos para inducir de ellos reglas prácticas.

En el Capítulo II se analizan los dos problemas propuestos de estimación espectral. Mientras que en el método de Prony-Hildebrand el ruido aditivo $w(n)$ no está sujeto a ninguna restricción previa, con lo que la presencia de ruido impulsivo puede contaminar de 'outliers' el registro de error $e(n)$, en el método propuesto por Cadzow, la presencia de los términos $\sigma^2 a$ en el registro de error simulan el mismo efecto cuando el nivel del ruido aditivo $w(n)$ es suficiente relevante. Este método admite además la posibilidad de utilizar diferentes estimadores para el registro de autocorrelación. Igualmente, en la formulación en frecuencia de dichos métodos, el registro de error tiene una estructura picuda, ya que el espectro original de los datos utilizados (sinusoides más ruido) presenta claramente picos espectrales. En este Capítulo se contrastan los resultados de las soluciones L1 y L2 de los métodos propuestos en ambos dominios, discutiendo en cada caso la mejora espectral obtenida debido a la estimación de los coeficientes del polinomio $A(z)$.

En el Capítulo III se introduce la nueva familia de algoritmos adaptativos, en el dominio del tiempo, basados en la búsqueda del mínimo de una potencia entera del valor absoluto del error residual de adaptación

(algoritmos Lk's). Como sistema soporte para el análisis, se utiliza una planta transversal ruidosa de N coeficientes centrado el problema en la identificación adaptativa de la planta. Se analiza la convergencia en varianza del error residual, estableciendo para el algoritmo genérico Lk las condiciones necesarias para que aquella exista. Se evalúan la velocidad de convergencia y varianza del error residual en régimen permanente, atendiendo a condicionantes externos tales como potencia de los procesos, distribución del ruido de planta y longitud de la planta, observando la mejor adecuación de cada uno de ellos a diferentes situaciones. Se plantea una optimización determinista local del paso de adaptación, obteniendo con ello algoritmos con mejores prestaciones. Y por último, se establecen unas guías útiles para la elección de algoritmos y/o pasos de adaptación adecuados en función de un error residual o velocidad de convergencia deseados. Todos los resultados y conclusiones teóricas derivados del análisis son contrastados convenientemente con simulaciones realizadas en ordenador.

En el Capítulo IV se analiza el comportamiento de la familia de algoritmos anterior en el dominio de la frecuencia. Partiendo de las transformadas discretas de Fourier de los datos, del error residual y del ruido de planta, se establece un procedimiento adaptativo de gradiente estocástico para cada componente espectral. El algoritmo generalizado Lk se define en el campo complejo y por ello presenta ligeras modificaciones al definido en el campo real. Los criterios de evaluación de estos algoritmos son similares a los ya expuestos en el punto anterior e igualmente los resultados del análisis son contrastados con simulaciones en ordenador.

En el Capítulo V se presentan simulaciones de sistemas adaptativos reales en el dominio del tiempo (modelables como identificadores de planta), tales como los de cancelación de ecos en línea telefónica (voz y datos). Con estas aplicaciones se pretende comprobar algunas de las ventajas de utilizar algoritmos adaptativos alternativos al clásico LMS.

Finalmente, en el Capítulo VI, se resumen las principales conclusiones prácticas derivadas de los métodos y algoritmos desarrollados en esta Tesis, al mismo tiempo que se señalan algunas de las vías de investigación paralelas a o abiertas por este trabajo.

Existe un conjunto de Apéndices donde se han recogido los principales cálculos y desarrollos matemáticos necesarios. Su objeto es el de facilitar la lectura y comprensión del trabajo evitando cálculos intermedios tediosos.

CAPITULO II

ESTIMADORES ESPECTRALES L2 vs. L1

II.1.- INTRODUCCION: FORMULACION SIMPLEX DE LA MINIMIZACION NORMA L_1

Inicialmente, para dar generalidad al problema de minimización norma L_1 sujeto a un conjunto de restricciones, se utiliza la siguiente expresión como sistema de ecuaciones representativo de ambos procedimientos:

$$\underline{Y} \underline{a} = \underline{e} \quad (\text{ii.1.1})$$

donde

\underline{Y} ($nf \times nc$) es la matriz de elementos conocidos

\underline{a} ($nc \times 1$) el vector de nc incognitas

\underline{e} ($nf \times 1$) el vector error (desconocido igualmente)

$\sum_m |e_m|$ ($1 \leq m \leq nf$) la función a minimizar.

Para poder utilizar el método de Simplex [LUENBERGER,1973], [KUNZI,1971],[GOLDSTEIN,1977] en la resolución del problema anterior, es necesario reformular convenientemente (ii.1.1) y la función objetivo. Debido a que la formulación típica del Simplex no admite incognitas negativas, se pueden separar estas como la diferencia de dos nuevas variables positivas, es decir:

$$a_i = a_i^+ - a_i^- \quad (1 \leq i \leq nc) \quad (\text{ii.1.2a})$$

$$e_m = e_m^+ - e_m^- \quad (1 \leq m \leq nf) \quad (\text{ii.1.2b})$$

con $a_i^+, a_i^-, e_m^+, e_m^- \geq 0$.

En este caso, la función objetivo queda $\sum_m |e_m^+ - e_m^-|$, pero debido a la forma de actuar del Simplex en que el mínimo se obtiene cuando al menos una de las dos variables es nula, se puede reescribir:

$$|e_m| = |e_m^+ - e_m^-| = e_m^+ + e_m^- \quad (\text{ii.1.3})$$

y por tanto la nueva formulación de (ii.1.1) queda :

función objetivo

$$\min \sum_m e_m^+ + e_m^- \quad (\text{ii.1.4a})$$

sujeto a

$$\underline{Y} (\underline{a}^+ - \underline{a}^-) = \underline{e}^+ - \underline{e}^- \quad (\text{ii.1.4b})$$

Originando un nuevo sistema de ecuaciones con doble número de incógnitas $(2nc + 2nf)$ resoluble mediante el método de Simplex.

II.2.- PRONY-HILDEBRAND EN EL TIEMPO

Este método presenta la ventaja de trabajar directamente con el registro de datos ([WEISS,1963],[VAN BLARICUM,1978], [MARPLE,1979]) sin imponer ninguna restricción al color del ruido aditivo $w(n)$. Este hecho es el que permite adoptar la solución L1 en vez de la L2 como más adecuada en casos donde el ruido sea de tipo impulsivo.

La formulación y solución L2 de dicho procedimiento para un registro $x(n)$ de N puntos y p sinusoides ha sido ya presentada en (i.2.1.9), (i.2.1.10) e (i.2.1.13) y no se repetirá en este Apartado. Sin embargo, atendiendo a la formulación L1 generalizada del Apartado anterior y particularizándola para este caso, se obtiene:

funcion objetivo

$$\min \sum_{n=p+1}^{N-p} e^+(n) + e^-(n) \quad (\text{ii.2.2.1a})$$

sujeto a

$$x(n) + \sum_{i=0}^{p-1} (a_i^+ - a_i^-) [x(n-i) + x(n+i)] = e^+(n) - e^-(n) \quad (\text{ii.2.2.1b})$$

con $(p+1 \leq n \leq N-p)$. En este caso se ha normalizado el vector de coeficientes por $2a_p=1$, ya que este hecho no tiene efecto alguno en las raíces del polinomio formado por los a_i 's.

Debido al carácter simétrico de los coeficientes, Keybl y Eichmann propusieron en [KEYBL,1983] una acotación adicional de los coeficientes a_i 's, derivada de las relaciones de Cardano (para $p=2$). Sin embargo, la complejidad del cálculo de estas relaciones es considerable cuando el orden p es alto, verificándose además que estas ligaduras son inactivas en la mayoría de los casos analizados [FIGUEIRAS-VIDAL,1985a].

En [FIGUEIRAS-VIDAL,1985b] se verifica que los resultados ofrecidos por la solución L1 son ligeramente peores que los obtenidos por la correspondiente L2 para detectar sinusoides en ruido Gaussiano blanco y que aumentar la longitud del registro de puntos no mejora sustancialmente la solución L1, ya que (a diferencia de lo que ocurriría en el dominio de la autocorrelación con registros largos) los efectos del ruido serían similares en cualquier región del registro y solo la aparición de una nueva ecuación mejor condicionada ayudaría a la solución final. Sin embargo, esto no ocurre cuando aparecen 'outliers' en el registro, ya que la minimización norma L1 soporta mejor tales apariciones sin distorsionar la solución final. La Tabla T2.1 confirma lo anterior para el siguiente ejemplo:

EJEMPLO E2.1:

- 10 registros de 32 puntos formados por
- * Dos sinusoides de amplitud 1., frecuencias normalizadas 0.25 y 0.30 y fases 0 y 90 grados.
- * Ruido Gaussiano blanco 30 dB por debajo de la potencia de señal.
- * Outliers aditivos de aparición aleatoria (1.5% del tiempo) 8 dB por encima de la potencia de señal.
- * Se toma un modelo de 2 sinusoides.

En la Tabla T2.1, las columnas centrales indican la posición de las rayas espectrales una vez resuelto el polinomio $A(z)=0$. La última columna indica la posición donde ha aparecido el pico dentro del registro de 32 puntos. En las Figuras E2.1a y E2.1b se muestran las curvas

correspondientes a las dos primeras realizaciones de la Tabla. Representan $1/|A(\omega)|^2$, en escala logarítmica, donde $A(\omega)$ es el denominador de la función de transferencia del filtro formado por los coeficientes a_i 's y en la Figura F2.1c se presentan los periodogramas correspondientes a dichas realizaciones.

Ens.	(L2)		(L1)		Pico
	f_1	f_2	f_1	f_2	
1	.250	.300	.248	.302	-
2	.266	.266	.246	.295	2
3	.250	.299	.248	.297	-
4	.250	.299	.250	.299	-
5	.250	.298	.248	.297	-
6	.263	.263	.245	.295	11
7	.250	.301	.251	.300	-
8	.265	.265	.235	.291	8
9	.251	.299	.248	.298	-
10	.250	.300	.247	.302	-

TABLA T2.1

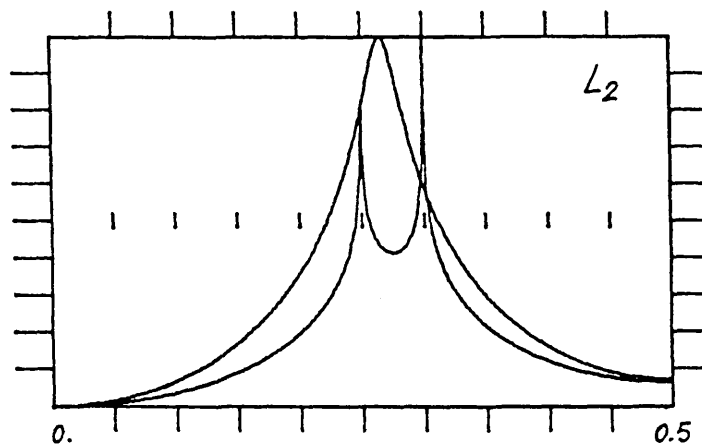


Figura F2.1a

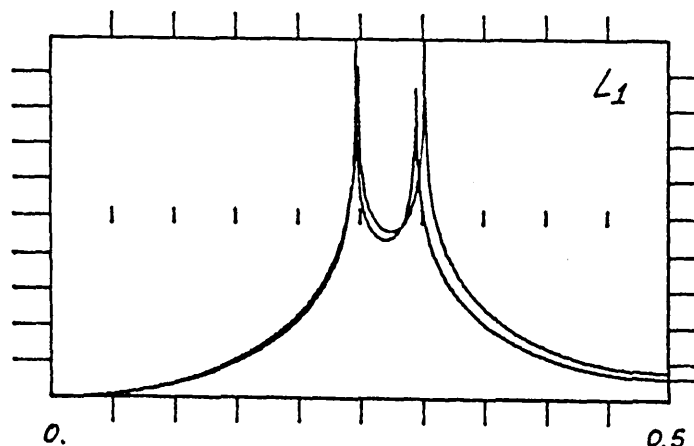


Figura F2.1b

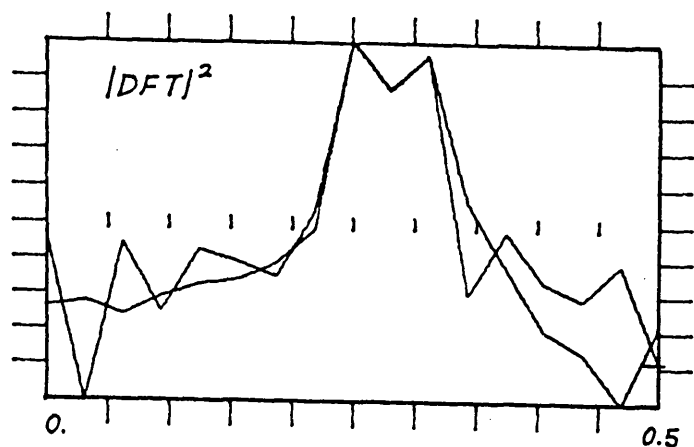


Figura F2.2c

Los resultados anteriores suscitan los siguientes comentarios: a) El periodograma es incapaz de resolver espectralmente ambas realizaciones debido a la escasez de datos, aunque presenta una cierta robustez frente a la presencia de outliers.

b) En aquellas realizaciones donde no aparecen picos de ruido, la minimización L2 ofrece resultados menos desviados espectralmente que la L1, aunque con igual capacidad de resolución en frecuencia para esta S/N y separación de los picos.

c) La pérdida de capacidad resolutive espectral de la solución L2 ante la aparición de un único pico en el registro de la señal.

Según los resultados anteriores, como conclusión final de este Apartado se recomienda el uso de la solución L1 frente al de la L2 cuando la longitud del registro de datos sea corta, la S/N alta, el orden del modelo sea igual o próximo al número de sinusoides presentes y se presuma la presencia de un registro de datos contaminado con outliers.

II.3.- SEGUNDO METODO DE CADZOW EN EL DOMINIO DE LA AUTOCORRELACION

Este procedimiento de análisis espectral, tal como se indicó en el capítulo introductorio, está basado en la resolución de un sistema sobredimensionado de ecuaciones (ecuaciones de Yule-Walker correspondientes a un modelo ARMA paso-todo para sinusoides en ruido blanco) definidas en el dominio del desplazamiento de la autocorrelación, en el que no se rechazan las ecuaciones centrales donde aparece el efecto de la varianza del ruido en la forma $\sigma^2 a_i$'s [CADZOW,1979],[CADZOW,1982a],[CADZOW,1982b]. Esto es equivalente a considerar que la secuencia de autocorrelación está compuesta de una serie de sinusoides más un ruido aditivo (error de estimación y términos $\sigma^2 a_i$'s) que hacen el problema similar al de Prony-Hildebrand estudiado anteriormente. Este hecho justifica, en algunos casos, una solución L1, ya que el efecto debido a σ^2 puede ser considerado como de tipo impulsivo en el registro de error. Por tanto, la utilización del Simplex en la minimización del error absoluto total puede ofrecer una alternativa de mantener dichos valores centrales prácticamente inalterados, distorsionando así menos los coeficientes a_i 's y por consiguiente la localización espectral de las rayas.

Para favorecer este comportamiento del Simplex es conveniente condicionar el problema de forma tal que se enfatice el efecto de las ecuaciones centrales. Esto ocurrirá cuando:

- La relación señal a ruido no sea excesivamente alta.
- El orden del modelo no sea mucho mayor que el número total de sinusoides presente, ya que en caso contrario aparecerían valores de los a_i 's relativamente pequeños enmascarables por el error.

- Obtener buenas estimas de la autocorrelación $r'(n)$, ya que esto fuerza un error nulo en las ecuaciones no centrales resaltando el efecto picudo de las centrales.

Evidentemente todo este condicionamiento del problema reduce la robustez de la solución L1, pero no por ello deja de ser preferible su aplicación en esos casos concretos.

La formulación L1 de este método, utilizando el algoritmo Simplex, es según el Apartado II.1 de este Capítulo y la expresión (i.2.1.11) :

funcion objetivo

$$\min \sum_{n=-N/2+p}^{N/2} e^+(n)+e^-(n) \quad (\text{ii.3.1a})$$

sujeto a

$$r'(n) + \sum_{i=1}^{2p} (a_i^+ - a_i^-) r'(n-i) = e^+(n) - e^-(n) \quad (\text{ii.3.1b})$$

$$(-N/2+2p \leq n \leq N/2)$$

donde se ha supuesto ($a_0^+ = 1, a_0^- = 0$), $N/2$ desplazamientos para la autocorrelación y un modelo de p sinusoides.

Con objeto de resaltar la dependencia del resultado final de la bondad del estimador elegido para la secuencia de autocorrelación, se utilizarán el clasico estimador insesgado y la secuencia real de autocorrelación $r(n)$ en el calculo de los coeficientes. Con ello se apreciará la sensibilidad del procedimiento en relación con dicha elección. El estimador insesgado es habitualmente definido como:

$$r'(n) = (1/L-n) \sum_{m=1}^{L-n} x(m)x(m+n) \quad (\text{ii.3.2})$$

el cual es un estimador consistente de $r(n)$.

El ejemplo siguiente E2.2 permite confirmar la dependencia del resultado de la elección del estimador de $r(n)$.

EJEMPLO E2.2

- * Dos sinusoides de amplitud 1., frecuencias normalizadas 0.235 y 0.255 y fase 0.
- * Ruido Gaussiano blanco 10 dB por debajo de la potencia de señal.
- * 10 desplazamientos (17 ecuaciones) para la autocorrelación en un modelo de 2 sinusoides.
- * Secuencia de autocorrelación estimada a partir de un registro de 128 puntos y secuencia de autocorrelación teórica (con objeto de comparación).

Los resultados del ejemplo anterior se presentan evaluando la expresión $1/|A(\omega)|^2$ en el margen de frecuencias normalizado $[0, .5]$. Las figuras siguientes los recogen:

- Figuras F2.2a y F2.2b: Soluciones L1 y L2 respectivamente con autocorrelación estimada.

- Figuras F2.3a y F2.3b: Soluciones L1 y L2 respectivamente con autocorrelación teórica.

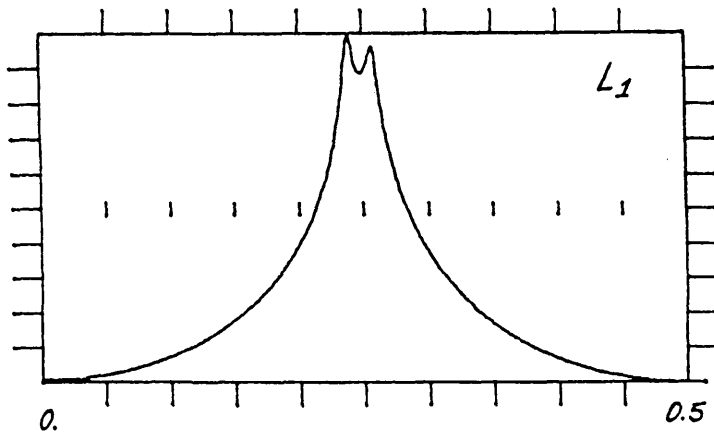


Figura F2.2a

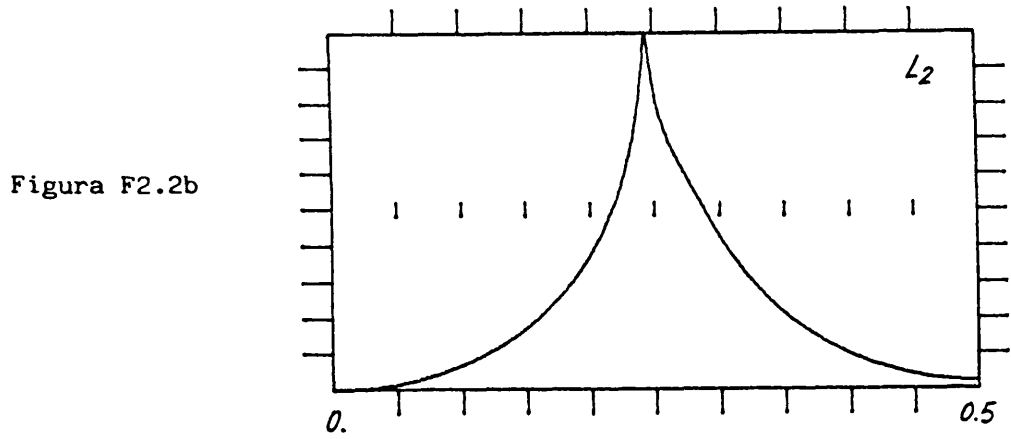


Figura F2.2b

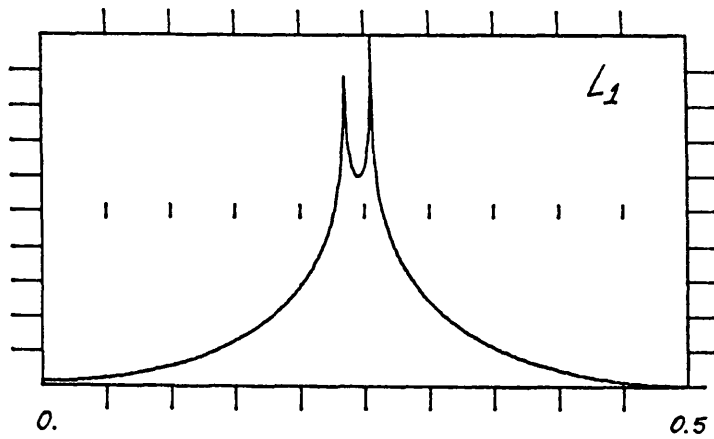


Figura F2.3a

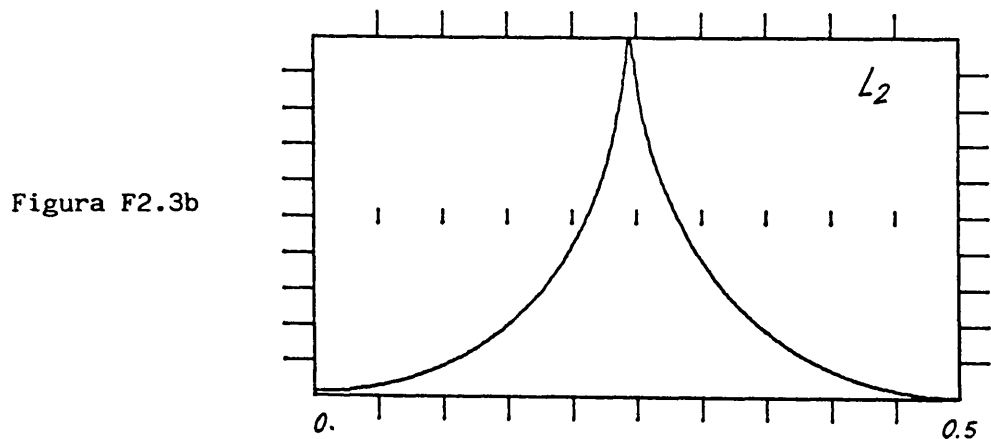


Figura F2.3b

Se observa que la solución L2 no resuelve espectralmente las dos sinusoides existentes, mientras que la L1 lo hace en ambos casos, presentando obviamente una mejor resolución en el caso de la autocorrelación teórica. En este caso presenta además un menor sesgo en la localización espectral de los máximos.

Esta mejora de la solución L1 sobre la L2 ha aparecido en más de un ensayo (además del presentado en el ejemplo) observándose un sesgo en la posición de las líneas espectrales a medida que estas se van cerrando. Esto ocurre especialmente en el caso de la solución L2 ya que al tener menor resolución espectral tiende a confundir las dos líneas en una.

Se ha de mencionar nuevamente la dependencia del resultado de la calidad del estimador, ya que cuanto mejor sea la estima de la secuencia $r(n)$, mejor se marca el carácter picudo de los errores de las ecuaciones centrales potenciando así la solución L1. Haciendo referencia a este punto, se comenta que además del estimador insesgado, se ha utilizado un estimador de $r(n)$ calculado como la transformada inversa de una versión limitada en banda de un preestimador espectral (periodograma) del registro. En los resultados obtenidos no se aprecia una mejora sustancial de la solución L1 sobre la L2. Esto es debido a que la limitación en banda elimina ruido y hace desaparecer el efecto $\sigma^2 a_i$ en el error de las ecuaciones centrales.

II.4.- PRONY-HILDEBRAND EN LA FRECUENCIA

Tal como se indicó en el capítulo introductorio, el método original de Prony-Hildebrand admite un desarrollo paralelo en el dominio de la DFT de los datos. Según (i.2.2.5) el sistema sobredimensionado de ecuaciones en el dominio de la frecuencia es formalmente idéntico al original, con la salvedad que se separan las partes real e imaginaria. Este queda como:

$$E_r(k) = X_r(k) + \sum_{i=1}^P a_i (\phi_r(k,i) + \phi_r(k,2p-i)) \quad (ii.4.1a)$$

$$E_i(k) = X_i(k) + \sum_{i=1}^P a_i (\phi_i(k,i) + \phi_i(k,2p-i)) \quad (ii.4.1b)$$

$$(0 \leq k \leq N/2)$$

donde los subíndices r e i denotan partes real e imaginaria respectivamente, y los coeficientes genéricos son:

$$\phi_r(k,i) = X_r(k) \cos(2\pi ki/N) + X_i(k) \sin(2\pi ki/N) \quad (ii.4.2a)$$

$$\phi_i(k,i) = X_i(k) \cos(2\pi ki/N) - X_r(k) \sin(2\pi ki/N) \quad (ii.4.2b)$$

Este procedimiento supone una carga computacional mayor con respecto a la versión original L1 en el dominio del tiempo, pero:

a) Ofrece una solución que resalta los picos del registro de error $E(k)$, por lo que será interesante cuando el espectro de $x(n)$ presente picos, ya que, en estos casos, la DFT del error $e(n)$ tiende a aumentar de una manera abrupta en el entorno de la localización de los picos.

b) Permite incluir directamente algún conocimiento a priori sobre el espectro de la señal. Se puede restringir el campo de validez de k desde N_1 a N_2 siendo $[2\pi N_1/(N/2), 2\pi N_2/(N/2)]$ la banda espectral a considerar. Esto reduce además la carga computacional, ya que se reduce el número total de ecuaciones, eliminando aquellas menos significativas.

Llamando por comodidad de notacion:

$$b_r(k,i) = \phi_r(k,i) + \phi_r(k,2p-i) \quad (ii.4.3a)$$

$$b_i(k,i) = \phi_i(k,i) + \phi_i(k,2p-i) \quad (ii.4.3b)$$

la formulacion Simplex correspondiente queda:

funcion objetivo

$$\min \sum_{k=N1}^{N2} E_r^+(k) + E_r^-(k) + E_i^+(k) + E_i^-(k) \quad (ii.4.4a)$$

sujeto a

$$X_r(k) + \sum_{i=1}^p (a_i^+ - a_i^-) b_r(k,i) = E_r^+(k) - E_r^-(k) \quad (ii.4.4b)$$

$$X_i(k) + \sum_{i=1}^p (a_i^+ - a_i^-) b_i(k,i) = E_i^+(k) - E_i^-(k) \quad (ii.4.4c)$$

$$(N1 \leq k \leq N2)$$

donde se ha supuesto $a_0^+ = 1$ y $a_0^- = 0$.

La formulación equivalente L2 es también posible, pero es interesante resaltar que los resultados obtenidos serían muy similares a los del dominio del tiempo (debido a la relación de Parseval), ya que la única diferencia introducida es la periodización de los datos originales forzada por la DFT, esto es, el procedimiento es equivalente a minimizar en el tiempo la expresión $\sum_{n=0}^{N-1} |e(n)|^2$ con $e(n) = \text{DFT}^{-1}(E(K))$. Evidentemente se están incluyendo puntos de error ($e(0), \dots, e(2p-1)$) no existentes en la versión original, pero esto no tiene importancia relevante si se tienen registros largos de puntos. En la minimización L1 no existe una relación directa entre la minimización en el tiempo y en la frecuencia.

En el siguiente ejemplo, se observa ventaja de la solución L1 sobre la L2. Como base para un conocimiento previo se utiliza la DFT del registro de datos $x(n)$, la cual indica claramente donde se concentra la mayor parte

de energia.

EJEMPLO E2.3:

- * Dos sinusoides de amplitud 1., frecuencias normalizadas 0.235 y 0.255 y fase 0.
- * Ruido blanco Gaussiano de 0 dB.
- * Un registro de $N=64$ puntos.
- * Una limitación de banda desde $N1=12$ a $N2=19$ que se corresponde con la banda normalizada $[.188,.297]$.
- * Un modelo de 2 sinusoides.

En las Figuras siguientes se recogen los resultados:

- Figura F2.4: DFT de 64 puntos de los datos, donde se marca la banda elegida $[N1, N2]$.
- Figura F2.5: Solución L2 en el dominio del tiempo.
- Figura F2.6a: Solución L1 con limitación de banda en el dominio de la DFT.
- Figura F2.6b: Solución L2 con limitación de banda en el dominio de la DFT.

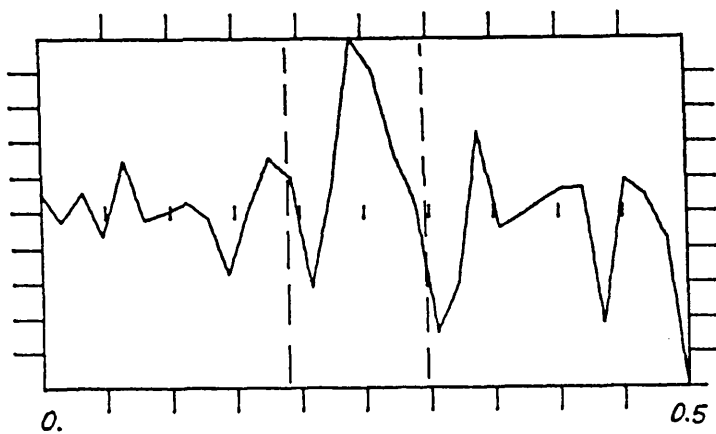


Figura F2.4

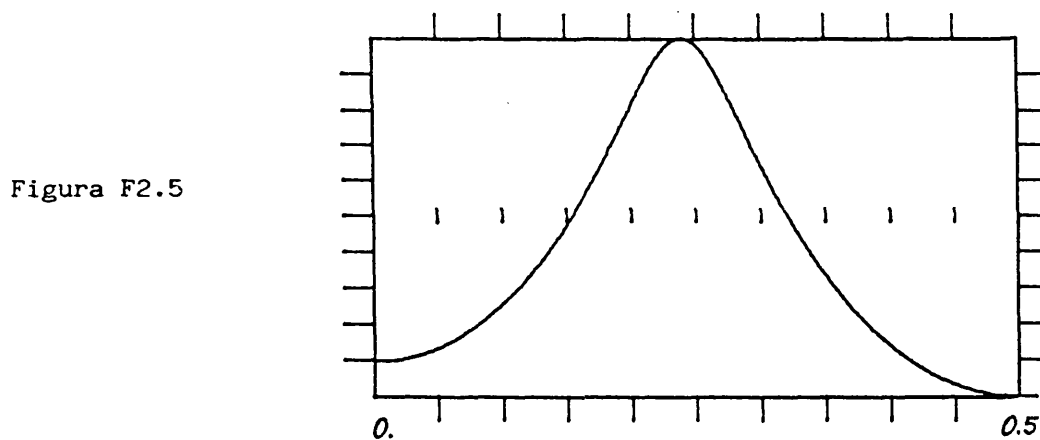


Figura F2.5

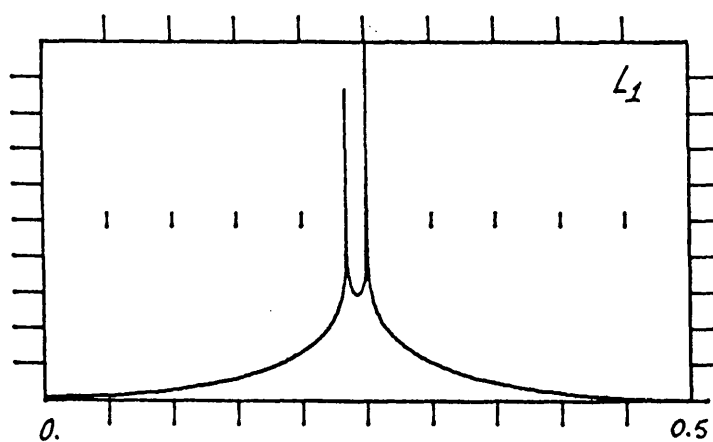


Figura F2.6a

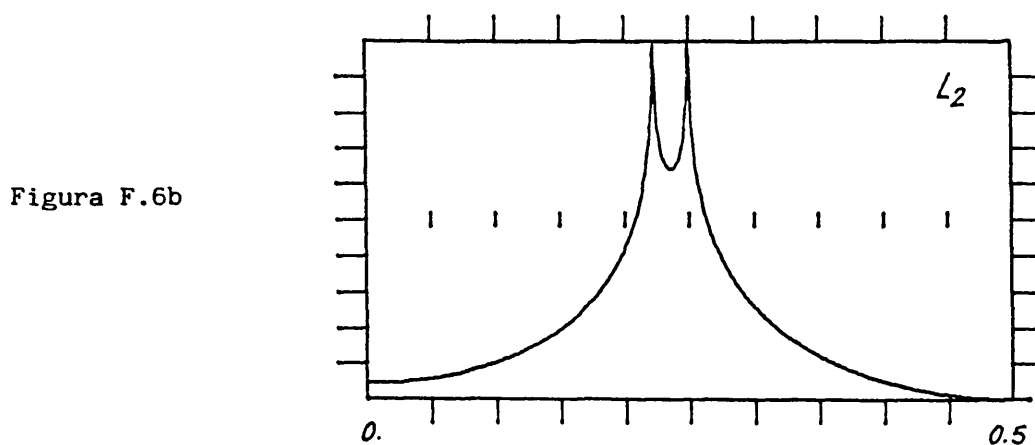


Figura F.6b

Se observa, la falta de resolución de la DFT así como de la solución temporal L2. Esta última al no disponer de la restricción en banda incorpora más ruido a las ecuaciones. Con respecto a los métodos en el dominio de la DFT hay que resaltar, que aunque la solución L2 resuelve dos rayas el sesgo introducido es bastante grande, hecho que no sucede con la solución L1.

Como resumen final se resalta que el hecho de restringir la banda de análisis a $[N1, N2]$, evita el incorporar ecuaciones, debidas fundamentalmente a ruido fuera de banda no relevantes, en la solución buscada. Este efecto fuerza implícitamente a que se eleve la relación señal/ruido real de los datos utilizados.

II.5.-SEGUNDO METODO DE CADZOW EN LA FRECUENCIA

El procedimiento original de Cadzow admite igualmente al de Prony-Hildebrand un desarrollo similar en el dominio de la frecuencia, tal como se indicó en el capítulo introductorio. Sin embargo, en este caso la transformada de la secuencia de autocorrelación $r(n)$ (o cualquier otro estimador del espectro de potencia usado) es una función real y par y por tanto las expresiones finales se reducen notablemente. Según (i.2.2.7) el sistema queda como:

$$E_r(k) = \sum_{i=0}^{2p} a_i S(k) \cos(2\pi ki/N) \quad (ii.5.1a)$$

$$E_i(k) = \sum_{i=0}^{2p} -a_i S(k) \sin(2\pi ki/N) \quad (ii.5.1b)$$

$$(0 \leq k \leq N/2)$$

donde los subíndices r e i denotan partes real e imaginaria respectivamente.

Este procedimiento admite igualmente la aplicación directa del conocimiento espectral a priori, reduciendo así el número de ecuaciones utilizadas. El resto de comentarios es similar al realizado en el Apartado anterior.

La formulación Simplex de este método queda:

funcion objetivo

$$\min \sum_{k=N1}^{N2} E_r^+(k) + E_r^-(k) + E_i^+(k) + E_i^-(k) \quad (ii.5.2a)$$

sujeto a

$$S(k) + \sum_{i=1}^{2p} (a_i^+ - a_i^-) S(k) \cos(2\pi ki/N) = E_r^+(k) - E_r^-(k) \quad (ii.5.2b)$$

$$\sum_{i=1}^{2p} (a_i^+ - a_i^-) S(k) \sin(2\pi ki/N) = E_i^+(k) - E_i^-(k) \quad (ii.5.2c)$$

$$(N1 \leq k \leq N2)$$

donde nuevamente se ha supuesto $a_0^+ = 1$ y $a_0^- = 0$.

De nuevo se utiliza un ejemplo simulado para apreciar la ventaja de la solución norma L1 sobre la solución norma L2. Como estimador espectral de potencia $S(k)$ se utiliza el periodograma. El ejemplo es el siguiente:

EJEMPLO E2.4:

- * Dos sinusoides de amplitud 1., frecuencias normalizadas .235 y .255 y fase 0.
- * Ruido blanco Gaussiano de 0 dB.
- * Un registro de $N=64$ puntos.
- * Una limitación de banda desde $N_1=12$ a $N_2=19$ que se corresponde con la banda normalizada [.188,.297].
- * Un modelo de 2 sinusoides.

En las figuras siguientes se recogen los resultados del ejemplo anterior.

- Figura F2.7: Periodograma de 64 puntos realizado a partir de los datos $x(n)$. Se marca la banda elegida.

- Figura F2.8: Solución L2 al método de Cadzow en el dominio de la autocorrelación.

- Figura F2.9a: Solución L1 con limitación en banda en el dominio de la frecuencia.

- Figura F2.9b: Solución L2 con limitación en banda en el dominio de la frecuencia.

Al igual que en el método de Prony-Hildebrand en frecuencia, se observa la falta de resolución en el periodograma y en la solución L2 en el dominio de la autocorrelación. En los métodos donde se ha aplicado una restricción de banda los resultados aparecen con dos líneas espectrales bien definidas, pero hay que destacar que la solución L2 es menos abrupta y más sesgada que la correspondiente L1. Estos resultados son similares a los obtenidos anteriormente y resaltan la ventaja intrínseca de una minimización norma L1 sobre la correspondiente L2 en casos de soluciones con picos espectrales marcados.

II.6.- CONCLUSIONES

En este Capítulo se ha presentado la solución de mínimo error absoluto de un sistema sobredimensionado de ecuaciones como una alternativa a la ampliamente usada de error cuadrático mínimo. Dicha alternativa se plantea cuando se presupone o intuye que la solución buscada posee una estructura picuda (o efecto equivalente al de outliers), esto es, algunos valores están claramente desplazados del resto. Esta elección es propiciada debido a que la solución L2 penaliza más los valores altos que los pequeños, dado el carácter cuadrático de la función de coste, mientras que la solución L1 tolera más dichos valores admitiéndolos más fácilmente como posible solución.

Como confirmación concreta de las aseveraciones anteriores se han analizado dos procedimientos conocidos de estimación espectral paramétrica (para sinusoides en ruido) basados en la resolución de un sistema sobredimensionado de ecuaciones lineales. Estos procedimientos fueron elegidos por su interés intrínseco (alta resolución y robustez) y por que sus planteamientos son susceptibles de ser desarrollados con métodos bloque para solución L1, aunque en su origen fueron definidos y analizados con la solución L2. En ambos procedimientos, el registro de error puede poseer una estructura picuda y poco distribuida, la cual favorece una minimización norma L1 sobre la norma L2. En particular, esto ocurre con el método de Prony-Hildebrand cuando los datos están contaminados con ruido impulsivo y con el segundo método de Cadzow con buenas estimas de la autocorrelacion y un ruido de fondo significativo. En ambos casos el orden de los modelos no debe exceder en mucho el número de sinusoides presentes. El condicionar de esta forma a ambos métodos reduce la robustez de la solución L1, pero evidentemente sus prestaciones aumentan, tal y como se reveló en los

ejemplos de los Apartados anteriores.

Adicionalmente se han desarrollado las formulaciones de los métodos estudiados anteriormente en el dominio de la frecuencia con el propósito de potenciar un registro de error espectral picudo, ya que el espectro original de la señal (sinusoides en ruido) es de tendencia afilada debido a la presencia de las sinusoides. Este nuevo planteamiento permite además la inclusión directa de conocimiento previo sobre la localización espectral de las rayas, disminuyendo así el número de ecuaciones del sistema al eliminar aquellas debidas fundamentalmente a ruido, con lo que se favorece una mejor solución y se reduce la carga computacional.

En general, se puede concluir que la solución norma L_1 es más apropiada que la de norma L_2 cuando el registro de error a minimizar posee un carácter poco distribuido conteniendo además valores claramente desplazados del resto. Esto es debido a que la función de coste de valor absoluto enfatiza menos los errores grandes que la cuadrática, permitiendo desplazamientos sustancialmente mayores sin elevar fuertemente el error final. En esta Tesis se han presentado un par de aplicaciones donde, bajo determinadas condiciones, la solución norma L_1 mínima ofrece mejores prestaciones que la correspondiente L_2 . Queda abierta, no obstante, la búsqueda de aplicaciones y situaciones concretas basadas en unas ligaduras lineales, donde la estructura del vector a minimizar presente un carácter poco distribuido y picudo que favorezca su resolución norma L_1 por el procedimiento Simplex o algún otro alternativo de Programación Lineal.

CAPITULO III

IDENTIFICACION ADAPTATIVA DE PLANTA

EN EL DOMINIO DEL TIEMPO

III.1.- INTRODUCCION DEL ALGORITMO

Tal como se indicó en el capítulo introductorio, la familia de algoritmos bajo estudio es una extensión a la introducida por Walach y Widrow en [WALACH,1984], ya que no se restringe al análisis de algoritmos derivados de la búsqueda del mínimo de una potencia par del error residual, sino que la amplía a cualquier potencia entera positiva del módulo de dicho error. En [PAEZ-BORRALLA,1985b] se presentó una breve introducción de la familia completa de algoritmos.

El soporte elegido para el estudio es el de un identificador transversal de planta, ya que es un modelo generalizado que puede configurar adecuadamente sistemas adaptativos típicos, como por ejemplo en comunicaciones (cancelador de ecos) y en ingeniería de control (identificación de sistemas). En la Figura F3.0 se esquematiza el sistema.

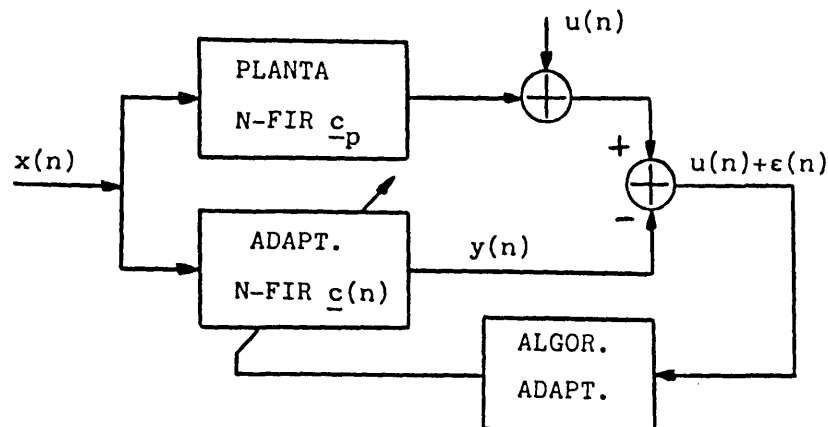


Figura F3.0

Se adopta igual longitud (FIR de N coeficientes) para la planta y el filtro adaptativo y se denominan: $x(n)$, al proceso de entrada a ambos sistemas; $u(n)$, al ruido a la salida de la planta; $\epsilon(n)$, al error residual de identificación; y \underline{c}_p y $\underline{c}(n)$ a los conjuntos de coeficientes de los filtros.

El análisis teórico de la convergencia y de las condiciones necesarias para que ésta exista sigue dos planteamientos. El primero estudia el comportamiento de la convergencia del error cuadrático medio en todo su margen dinámico de variación, así como las condiciones necesarias para que exista sin aproximaciones de irrelevancia de unas variables frente a otras. Y el segundo, reproduce las ecuaciones de Walach y Widrow [WALACH,1984], las cuales permiten la evaluación de las constantes de tiempo y los desajustes finales de los procesos adaptativos en la familia completa de algoritmos (excepto para $k=1$) cuando el error residual de adaptación es pequeño comparado con el ruido de planta.

El algoritmo generalizado Lk bajo estudio trata de minimizar, en cada instante n , el estadístico:

$$f(\underline{c}(n)) = E[|\epsilon(n)+u(n)|^k] \quad (\text{iii.1.1})$$

usando el gradiente del estimador instantáneo de la función objetivo, esto es:

$$\hat{\nabla} E[|\epsilon(n)+u(n)|^k] = \nabla[|\epsilon(n)+u(n)|^k] \quad (\text{iii.1.2})$$

Antes de proseguir con la expresión final del algoritmo adaptativo, es necesario encontrar el mínimo de la función objetivo (iii.1.1) y garantizar su carácter convexo, demostrando así la unicidad de su mínimo. Dicho mínimo $f_{\min}(\cdot)$ se calcula resolviendo el sistema :

$$\nabla f(\underline{c}) = \underline{0} \quad (\text{iii.1.3})$$

En el Apéndice A.1 se desarrolla (iii.1.3) obteniendo:

$$f_{\min}(\underline{c}) = f(\underline{c}_p) = E[|u|^k] \quad (\text{iii.1.4})$$

donde se observa que el mínimo ocurre cuando $\underline{c}_p = \underline{c}$ y que depende exclusivamente del ruido de planta y del orden del algoritmo elegido.

El carácter convexo de la superficie de error (iii.1.1) se analiza en el Apéndice A.2, obteniendo que el Hessiano de la superficie es definido positivo, lo cual supone convexidad estricta en el dominio de los coeficientes garantizando que el mínimo calculado es único.

Una vez, encontrado el único punto posible de convergencia del algoritmo adaptativo, su aproximación estocástica se define:

$$\underline{c}(n+1) = \underline{c}(n) - \mu \nabla (|\epsilon(n)+u(n)|^k) \quad (\text{iii.1.5})$$

donde μ es el paso de adaptación.

Como el error residual $\epsilon(n)$ se puede expresar:

$$\epsilon(n) = \underline{v}(n)^t \underline{x}(n) \quad (\text{iii.1.6})$$

con

$$\underline{v}(n) = \underline{c}_p - \underline{c}(n) \quad (\text{iii.1.7})$$

vector diferencia del vector óptimo con el de coeficientes, el gradiente instantáneo resulta:

$$\nabla (|\underline{v}^t(n)\underline{x}(n)+u(n)|^k) = -k(\epsilon(n)+u(n))|\epsilon(n)+u(n)|^{k-2} \underline{x}(n) \quad (\text{iii.1.8})$$

con lo que, finalmente, el algoritmo Lk se expresa:

$$\underline{c}(n+1) = \underline{c}(n) + \mu k(\epsilon(n)+u(n))|\epsilon(n)+u(n)|^{k-2} \underline{x}(n) \quad (\text{iii.1.9})$$

donde se observa que para $k=2$, se obtiene el clásico LMS.

III.2.- ANALISIS DEL ERROR RESIDUAL CUADRATICO MEDIO

Para analizar el comportamiento del algoritmo definido en (iii.1.9) se utilizan las siguientes hipótesis simplificadoras habituales [WIDROW,1976],[CLAASEN,1981] :

- H1) $x(n)$ es un proceso Gaussiano estacionario, blanco, de media nula y varianza σ_x^2 . Se asume una entrada blanca con objeto de excitar todos los modos de la planta y poder así determinar completamente el sistema en su banda de trabajo.

- H2) $u(n)$ es igualmente un proceso estacionario, blanco, de media nula, varianza σ_u^2 y de distribución par en u .

- H3) $x(n)$ y $u(n)$ son independientes entre si.

- H4) $\underline{c}(n)$ y $\underline{x}(n)$ son vectores aproximadamente independientes.

- H5) El grado de convergencia alcanzado es suficiente para considerar que $x(n)$ y $\epsilon(n)$ están aproximadamente incorrelados. Se acepta esta incorrelación como independencia para poder desarrollar la media de un producto como el producto de las medias correspondientes, simplificando así el análisis final.

Evidentemente las hipótesis H4 y H5 son falsas, pero H4 no difiere demasiado de la realidad, ya que el paso de adaptación suele ser suficiente pequeño y H5, aunque es débil en casos de error residual grande, se comporta bastante bien al comparar los resultados simulados con los teóricos.

III.2.1.- Expresión recurrente generalizada para la varianza del error residual

Apoyándose en las hipótesis H1 y H4, el análisis del error cuadrático medio $E[\epsilon^2(n)]$ coincide con el de la varianza $\sigma_\epsilon^2(n)$, ya que su valor medio es nulo al poderse expresar la esperanza de (iii.1.6), por H4, como producto de valores medios y ser $x(n)$ de media nula.

Si se sustrae \underline{c}_p en las dos partes de (iii.1.9) y su resultado se multiplica término a término por el traspuesto de él mismo, se obtiene:

$$\begin{aligned} \underline{v}^t(n+1)\underline{v}(n+1) &= \underline{v}^t(n)\underline{v}(n) - 2\mu k(\epsilon(n)+u(n))|\epsilon(n)+u(n)|^{k-2} \underline{v}^t(n)\underline{x}(n) + \\ &+ \mu^2 k^2 |\epsilon(n)+u(n)|^{2k-2} \underline{x}^t(n)\underline{x}(n) \end{aligned} \quad (\text{iii.2.1.1})$$

Considerando nuevamente H1 y H4 y utilizando (iii.1.6) la varianza del error residual se puede expresar como:

$$\begin{aligned} \sigma_\epsilon^2(n) &= E[|\underline{v}^t(n)\underline{x}(n)|^2] = E[(\sum_{i=1}^N v_i(n)x(n-i+1))^2] = \\ &= \sigma_x^2 E[\underline{v}^t(n)\underline{v}(n)] = \sum \sigma_x^2 E[v_i^2(n)] \end{aligned} \quad (\text{iii.2.1.2})$$

como según H1, la esperanza:

$$E[\underline{x}^t(n)\underline{x}(n)] = E[\sum_{i=1}^N x^2(n-i+1)] = N\sigma_x^2 \quad (\text{iii.2.1.3})$$

Si se toman valores medios en (iii.2.1.1), se multiplican ambas partes por σ_x^2 y se considera la hipótesis H5, según (iii.2.1.2) se llega a la conclusión:

$$\sigma_\epsilon^2(n+1) = \sigma_\epsilon^2(n) (1 - \mu S_1(n) + \mu^2 S_2(n)) \quad (\text{iii.2.1.4})$$

expresión recurrente para la varianza del error residual dependiente solamente del instante anterior n , donde:

$$S_1(n) = 2k\sigma_x^2 E[\epsilon(n)(\epsilon(n)+u(n))|\epsilon(n)+u(n)|^{k-2}] / \sigma_\epsilon^2(n) \quad (\text{iii.2.1.5})$$

$$S_2(n) = k^2 N \sigma_x^4 E[|\epsilon(n)+u(n)|^{2k-2}] / \sigma_\epsilon^2(n) \quad (\text{iii.2.1.6})$$

Por comodidad en la notación se introduce el polinomio cuadrático en μ definido como:

$$P(\mu, n) = 1 - \mu S_1(n) + \mu^2 S_2(n) \quad (\text{iii.2.1.7})$$

y sustituyéndolo en (iii.2.1.4) queda:

$$\sigma_\epsilon^2(n+1) = \sigma_\epsilon^2(n) P(\mu, n) \quad (\text{iii.2.1.8})$$

siendo fácil estudiar la convergencia en varianza del algoritmo sobre el polinomio $P(\cdot)$.

Con el propósito de facilitar el análisis, la evaluación de $S_1(n)$ y $S_2(n)$ se realizará bajo dos supuestos: planta sin ruido ($u(n)=0$) y planta ruidosa. En este segundo caso, se necesita conocer la distribución estadística del ruido de planta para la evaluación de $S_1(n)$ y $S_2(n)$.

III.2.2.- Caso 1 : Planta sin ruido

Debido al carácter Gaussiano de $x(n)$, el error residual $\epsilon(n)$ puede ser considerado Gaussiano sin incurrir en grave error, ya que, si la longitud de planta N es suficientemente larga, la aproximación Gaussiana para $\epsilon(n)$, por el Teorema Central del Límite, no es disparatada, justificando así el siguiente análisis.

En ausencia de ruido de planta $S_1(n)$ y $S_2(n)$ son:

$$S_1(n) = 2k \sigma_x^2 E[|\epsilon(n)|^{k-2}] / \sigma_\epsilon^2(n) \quad (\text{iii.2.2.1})$$

$$S_2(n) = k^2 N \sigma_x^4 E[|\epsilon(n)|^{2k-2}] / \sigma_\epsilon^2(n) \quad (\text{iii.2.2.2})$$

Pero, según [PAPOULIS,1984], los valores medios de las potencias enteras del valor absoluto de una v.a. Gaussiana z son:

$$E[z^{2m}] = (2m-1)!! \sigma_z^{2m} \quad (\text{iii.2.2.3})$$

$$E[|z|^{2m-1}] = \sqrt{2/\pi} 2^{m-1} (m-1)! \sigma_z^{2m-1} \quad (\text{iii.2.2.4})$$

con lo que sustituyendo en (iii.2.2.1) y (iii.2.2.2)

$$S_1(n) = \begin{cases} 2k (k-1)!! \sigma_x^2 \sigma_\epsilon^{k-2} (n) & ; k \text{ par} & (\text{iii.2.2.5a}) \\ 2\sqrt{2/\pi} k (k-1)!! \sigma_x^2 \sigma_\epsilon^{k-2} (n) & ; k \text{ impar} & (\text{iii.2.2.5b}) \end{cases}$$

$$S_2(n) = \begin{cases} N \sigma_x^4 / \sigma_\epsilon^2 (n) & ; k=1 & (\text{iii.2.2.6a}) \\ N k^2 (2k-3)!! \sigma_x^4 \sigma_\epsilon^{2k-4} (n) & ; k>1 & (\text{iii.2.2.6b}) \end{cases}$$

Estas expresiones son igualmente una buena aproximación para aquellos casos en que el ruido de planta sea despreciable frente al error residual, como suele ocurrir al inicio del proceso de adaptación si el vector $\underline{c}(n)$ esta bastante desadaptado de \underline{c}_p .

III.2.3.- Caso 2 : Planta ruidosa

En este caso las expresiones para $S_1(n)$ y $S_2(n)$ se pueden reformular como:

$$S_1(n) = \begin{cases} (2k \sigma_x^2 / \sigma_\epsilon^2 (n)) \sum_{j=0}^{k-1} \binom{k-1}{j} E[\epsilon^{j+1}(n)] E[u^{k-1-j}] & ; k \text{ par} & (\text{iii.2.2.7a}) \\ (2k \sigma_x^2 / \sigma_\epsilon^2 (n)) \sum_{j=0}^{k-1} \binom{k-1}{j} \phi_j(n) & ; k \text{ impar} & (\text{iii.2.2.7b}) \end{cases}$$

donde

$$\phi_j(n) = E[\epsilon^{j+1}(n) u^{k-1-j} \text{sgn}(\epsilon(n)+u)] \quad (\text{iii.2.2.8})$$

y

$$S_2(n) = (k^2 N \sigma_x^4 / \sigma_\epsilon^2(n)) \sum_{j=0}^{2k-2} \binom{2k-2}{j} E[\epsilon^j(n)] E[u^{2k-2-j}] \quad (\text{iii.2.2.9})$$

Se ha utilizado el carácter estacionario de $u(n)$ (hipótesis H2) para eliminar la dependencia con n de su valores medios.

El cálculo de (iii.2.2.7a) y (iii.2.2.9) es relativamente fácil, ya que sólo se necesitan los momentos centrados de $\epsilon(n)$ y u .

Para simplificar la resolución de (iii.2.2.7b) a través de (iii.2.2.8) se reescribe este valor medio como:

$$\phi_j(n) = \int_{-\infty}^{\infty} \epsilon^{j+1}(n) g_j(\epsilon(n)) f(\epsilon(n)) d\epsilon(n) \quad (\text{iii.2.2.10})$$

donde

$$\begin{aligned} g_j(\epsilon(n)) &= \int_{-\infty}^{\infty} u^{k-1-j} \text{sgn}(\epsilon(n)+u) f(u) du = \\ &= \int_{-\infty}^{-\epsilon(n)} u^{k-1-j} f(u) du + \int_{-\epsilon(n)}^{\infty} u^{k-1-j} f(u) du \end{aligned} \quad (\text{iii.2.2.11})$$

y

$$f(\epsilon(n)) = (1/\sqrt{2\pi} \sigma_\epsilon(n)) \exp(-\epsilon^2(n)/2\sigma_\epsilon^2(n)) \quad (\text{iii.2.2.12})$$

la f.d.p. Gaussiana de media nula y varianza $\sigma_\epsilon^2(n)$.

Dado el carácter de simetría par de las funciones de densidad de probabilidad de u (hipótesis H2), la expresión anterior queda:

$$g_j(n) = \begin{cases} 2 \text{sgn}(\epsilon(n)) \int_0^{|\epsilon(n)|} u^{k-1-j} f(u) du ; k-1-j \text{ par} & (\text{iii.2.2.13a}) \\ 2 \int_{|\epsilon(n)|}^{\infty} u^{k-1-j} f(u) du , k-1-j \text{ impar} & (\text{iii.2.2.13b}) \end{cases}$$

donde se llamará por comodidad en la notación:

$$I_1(m,t) = \int_0^t u^m f(u) du \quad (\text{iii.2.2.14a})$$

$$I_2(m,t) = \int_t^\infty u^m f(u) du \quad (\text{iii.2.2.14b})$$

Para el cálculo de las expresiones anteriores se requiere la distribución estadística del ruido de planta $u(n)$. En esta tesis se ha elegido un conjunto de ellas que cubren una amplia gama tanto en el campo de las comunicaciones como en el de ingeniería de control. Se hace además especial énfasis en su grado de concentración en torno a su valor central: distribuciones de colas cortas y largas, debido a que esta característica es representativa en el momento de elegir el orden k del algoritmo según las prestaciones de convergencia requeridas. Esta última afirmación será verificada en el Apartado III.6 de este Capítulo.

Las distribuciones bajo estudio son:

- Gaussiana, desarrollada en el Apéndice A.3

$$f(u) = (1/\sqrt{2\pi}\sigma_u) \exp(-u^2/2\sigma_u^2) \quad (\text{iii.2.2.15})$$

- Laplaciana, desarrollada en el Apéndice A.4

$$f(u) = (1/\sqrt{2}\sigma_u) \exp(-2|u|/\sigma_u) \quad (\text{iii.2.2.16})$$

- Uniforme, desarrollada en el Apéndice A.5

$$f(u) = \begin{cases} 1/2\sqrt{3}\sigma_u & ; |u| \leq \sqrt{3}\sigma_u \\ 0 & ; |u| > \sqrt{3}\sigma_u \end{cases} \quad (\text{iii.2.2.17})$$

- Binaria, desarrollada en el Apéndice A.6

$$f(u) = (\delta(u+\sigma_u) + \delta(u-\sigma_u))/2 \quad (\text{iii.2.2.18})$$

todas ellas con media nula y potencia σ_u^2 .

En los Apéndices referenciados anteriormente se analiza inicialmente la expresión $g_j(n)$ para cualquier j ($0 \leq j \leq k-1$) de las distribuciones mencionadas y posteriormente se procede con el cálculo de $S_1(n)$, pero debido a la tediosidad que supone el análisis de esta función para cualquier k impar arbitrario, se ha particularizado el análisis solamente para los algoritmos L1 y L3, los cuales son claramente representativos y permiten extrapolar conclusiones con respecto a miembros de la subfamilia impar.

Los valores particulares de $S_2(n)$ para los algoritmos L1 a L4 son desarrollados en el Apéndice A.10. En la Tablas 13.1a a 13.1d se resumen todos los valores de las expresiones $S_1(n)$ y $S_2(n)$ para los 4 primeros algoritmos de la familia, en los cuales se ha normalizado la potencia de error residual por la potencia de ruido de planta, esto es se ha utilizado:

$$s^2 = \sigma_e^2(n)/\sigma_u^2 \quad (\text{iii.2.2.20})$$

Las expresiones de dichas Tablas sirven para la evaluación del comportamiento dinámico de la varianza del error residual como función de parámetros intrínsecos a la planta (su longitud N y su ruido) y de la potencia del proceso de entrada.

k

DISTRIBUCION GAUSSIANA

$S_1(n)$	1	$2\sqrt{2/\pi} \sigma_x^2 / \sigma_u \sqrt{1+s^2}$
	2	$4\sigma_x^2$
	3	$6\sqrt{2/\pi} \sigma_x^2 \sigma_u (1+2s^2+s^2/(1+s^2)) / \sqrt{1+s^2}$
	4	$24\sigma_x^2 \sigma_u^2 (1+s^2)$
$S_2(n)$	1	$N\sigma_x^4 / \sigma_u^2 s^2$
	2	$4N\sigma_x^4 (1+s^2) / s^2$
	3	$27N\sigma_x^4 \sigma_u^2 (1+s^2)^2 / s^2$
	4	$240N\sigma_x^4 \sigma_u^2 (1+s^2)^3 / s^2$

Tabla T3.1a

k

DISTRIBUCION LAPLACIANA

S (n)	1	$(2\sqrt{2}\sigma_x^2/\sigma_u)(1-\text{erf}(s))\exp(s^2)$
	2	$4\sigma_x^2$
	3	$6\sqrt{2}\sigma_x^2\sigma_u(2s/\sqrt{\pi}+(1-\text{erf}(s))\exp(s^2))$
	4	$24\sigma_x^2\sigma_u^2(1+s^2)$
S (n)	1	$N\sigma_x^4/\sigma_u^2s^2$
	2	$4N\sigma_x^4(1+s^2)/s^2$
	3	$27N\sigma_x^4\sigma_u^2(2+2s^2+s^4)/s^2$
	4	$240N\sigma_x^4\sigma_u^2(6+6s^2+3s^4+s^6)/s^2$

Tabla I3.1b

k

DISTRIBUCION UNIFORME

$S_1(n)$	1	$2\sigma_x^2 \operatorname{erf}(\sqrt{3/2}s^2)/\sqrt{3}\sigma_u$
	2	$4\sigma_x^2$
	3	$6\sigma_x^2\sigma_u((\sqrt{3}+s^2/\sqrt{3})\operatorname{erf}(\sqrt{3/2}s^2)+\sqrt{2/\pi}\exp(-3s^2/2))$
	4	$24\sigma_x^2\sigma_u^2(1+s^2)$
$S_2(n)$	1	$N\sigma_x^4/\sigma_u^2s^2$
	2	$4N\sigma_x^4(1+s^2)/s^2$
	3	$27N\sigma_x^4\sigma_u^2(3/5+2s^2+s^4)/s^2$
	4	$48N\sigma_x^4\sigma_u^4(9/5+15s^2+15s^4+5s^6)/s^2$

Tabla T3.1c

k

DISTRIBUCION BINARIA

$S_1(n)$	1	$(2\sqrt{2/\pi} \sigma_x^2 / \sigma_u) \exp(-1/2s^2) / s$
	2	$4\sigma_x^2$
	3	$12\sigma_x^2 \sigma_u (\operatorname{erf}(1/2s) + \sqrt{2/\pi} s \exp(-1/2s^2))$
	4	$24 \sigma_x^2 \sigma_u^2 (1+s^2)$
$S_2(n)$	1	$N \sigma_x^4 / \sigma_u^2 s^2$
	2	$4N \sigma_x^4 (1+s^2) / s^2$
	3	$9N \sigma_x^4 \sigma_u^2 (1+6s^2+3s^4) / s^2$
	4	$16N \sigma_x^4 \sigma_u (1+15s^2+45s^4+15s^6) / s^2$

Tabla T3.1d

III.3.- CONDICIONES NECESARIAS PARA LA CONVERGENCIA DEL ERROR CUADRATICO MEDIO

La convergencia del error cuadrático medio $\sigma_e^2(n)$ estará garantizada cuando la desigualdad $\sigma_e^2(n+1) < \sigma_e^2(n)$ se cumpla para cualquier índice temporal n . Esto fuerza, según la expresión (iii.2.1.8), a que todo el condicionamiento necesario para la convergencia se centre en el polinomio cuadrático $P(\mu, n)$ definido en (iii.2.1.7). Obviamente, para ello, dicho polinomio ha de cumplir el siguiente requisito: ser menor que la unidad para cualquier n , permitiendo así que la secuencia del error cuadrático sea monótona decreciente.

Al forzar el cumplimiento de este requisito, aparecen las condiciones necesarias para la convergencia del error cuadrático medio hacia un valor residual final estable. Antes de forzar la condición anterior es conveniente verificar el carácter no negativo de $P(\cdot)$, ya que la varianza $\sigma_e^2(n)$ no puede ser negativa.

III.3.1.- Estudio del polinomio $P(\mu, n)$

El carácter no negativo del polinomio de segundo grado en μ , $P(\mu, n)$, se determina analizando el carácter no positivo de su discriminante y verificando que tome al menos un valor positivo. Según (iii.2.1.7) esto último se cumple para $\mu=0$, puesto que $P(0, n)=1$. Con respecto a su discriminante se ha de cumplir:

$$D = S_1^2(n) - 4S_2(n) \leq 0 \quad (\text{iii.3.1.1})$$

o equivalentemente:

$$B = S_1^2(n)/4S_2(n) \leq 1 \quad (\text{iii.3.1.2})$$

Sustituyendo en la expresión anterior (iii.2.1.5) y (iii.2.1.6) se obtiene:

$$\beta = \frac{E^2[\epsilon(n) \operatorname{sgn}(\epsilon(n)+u) |\epsilon(n)+u|^{k-1}]}{N E[\epsilon^2(n)] E[|\epsilon(n)+u|^{2k-2}]} \quad (\text{iii.3.1.3})$$

donde se ha usado $E[\epsilon^2(n)]$ por $\sigma_\epsilon(n)$.

Atendiendo a que $|\epsilon(n)| \geq \epsilon(n) \operatorname{sgn}(\epsilon(n)+u)$ para cualquier $\epsilon(n)$ y poniendo $\epsilon^2(n) = |\epsilon(n)|^2$ se consigue:

$$\beta \leq \frac{E^2[|\epsilon(n)| |\epsilon(n)+u|^{k-1}]}{N E[|\epsilon(n)|^2] E[|\epsilon(n)+u|^{2k-2}]} \quad (\text{iii.3.1.4})$$

y utilizando la desigualdad de Cauchy-Schwartz ($E^2[XY] \leq E[X^2]E[Y^2]$) en el numerador de la expresión anterior se obtiene $\beta \leq 1/N$, donde para $N=2$ (la longitud de planta mínima) $\beta < 1$ y por tanto el polinomio $P(\mu, n)$ es definido positivo para todo μ y n .

El requisito que debe de cumplir $P(\cdot)$ para garantizar la convergencia en varianza ($P(\mu, n) < 1$) es, según (iii.2.1.7) equivalente a la siguiente acotación superior de μ :

$$(0 <) \mu < \min(S_1(n)/S_2(n)) \quad (\text{iii.3.1.5})$$

la acotación inferior (< 0) ha sido ya supuesta, ya que el paso de adaptación ha de ser positivo.

En adelante, y por de sencillez, se denominará:

$$C_k(n) = S_1(n)/S_2(n) \quad (\text{iii.3.1.6})$$

a la razón principal de la cota superior de μ para el algoritmo L_k .

Debido a la gran cantidad de situaciones posibles originadas por: el orden del algoritmo, la presencia o no de ruido de planta, su caracterización estadística, etc; las expresiones correspondientes de $S_1(n)$ y $S_2(n)$ son bastante diferentes, con lo que para ordenar el análisis se divide nuevamente el estudio en dos partes, abordando en cada una de ellas la acotación del paso de adaptación y la evaluación de la varianza asintótica conseguida cuando se elige un μ apropiado para el algoritmo.

III.3.2.- Ausencia de ruido de planta

Partiendo de las expresiones (iii.2.2.5) y (iii.2.2.6) es fácil obtener las siguientes relaciones para $C_k(n)$ [FIGUEIRAS-VIDAL, 1986]:

$$C_k(n) = \begin{cases} 2\sqrt{2}/\pi \sigma_\epsilon(n)/N\sigma_x^2 & ; k=1 & \text{(iii.3.2.1a)} \\ (2(k-1)!!/Nk(2k-3)!!) \sigma_x^2 \sigma_\epsilon^{k-2}(n) & ; k \text{ par} & \text{(iii.3.2.1b)} \\ (2\sqrt{2}/\pi(k-1)!!/Nk(2k-3)!!) \sigma_x^2 \sigma_\epsilon^{k-2}(n) & ; k \text{ impar} & \text{(iii.3.2.1c)} \end{cases}$$

de las que se derivan fácilmente las correspondientes cotas superiores (sus valores mínimos) para μ debido a sus características monótonas (decreciente para $k=1$ y creciente para $k>2$).

Para la semirecta $[0, \infty)$ los correspondientes mínimos absolutos se obtienen en:

$$\mu_1 < C_1(\infty) \quad (k=1) \quad \text{(iii.3.2.2a)}$$

$$\mu_2 < C_2(n) = 1/N \sigma_x^2 \quad (k=2) \quad \text{(iii.3.2.2b)}$$

$$\mu_3 < C_k(0) \quad (k>2) \quad \text{(iii.3.2.2c)}$$

donde se observa el hecho ya conocido de la aparición de $1/N\sigma_x^2$ como cota única e independiente del instante n para el algoritmo L2. Se destaca, para $k>2$, la dependencia de la cota de las condiciones medias iniciales del proceso y para $k=1$ se verá que la cota está ligada al valor final de

convergencia y que fuerza la elección del paso de adaptación utilizado.

En general, si se decide utilizar un paso de adaptación μ variable para acelerar la convergencia, este no debe sobrepasar localmente el valor actual de $C_k(n)$, ya que de otro modo la secuencia $\sigma_\epsilon^2(n)$ divergería, esto es, $\sigma_\epsilon^2(n+1) > \sigma_\epsilon^2(n)$.

Si se utilizan, no obstante, pasos de adaptación constantes que verifiquen las anteriores acotaciones absolutas, el valor del polinomio de convergencia $P(\mu, n)$, será variable para los casos donde $k \neq 2$ y constante si $k=2$. Esto conduce a velocidades de convergencia variable cuando $k \neq 2$ y constante para $k=2$ respectivamente.

Puesto que el $\lim_{n \rightarrow \infty} P(\mu, n) = 1$ ($k \neq 2$), tal como se verá al final de este Apartado, la velocidad de convergencia asintótica de los algoritmos no-cuadráticos es cero: esto es, existe un efecto de 'saturación' en la convergencia de estos algoritmos. Esto, sin embargo, no excluye la posibilidad de velocidades locales de convergencia superiores a la ofrecida por el algoritmo cuadrático L2. En efecto, esto ocurre con relativa frecuencia en las primeras iteraciones de los algoritmos con $k > 2$, aunque el efecto de saturación puede ser una desventaja cuando el punto inicial de partida sea también suficientemente bueno para el algoritmo L2.

A continuación se evalúa la varianza (el error cuadrático medio en nuestro estudio) asintótica de los algoritmos bajo las condiciones de convergencia ya presentadas para un paso de adaptación μ constante en el caso de ausencia de ruido de planta [PAEZ-BORRILLO, 1986].

a) Varianza asintótica para $k > 2$.

Como hipótesis de partida supongase $\sigma_{\epsilon}^2(\infty) > 0$. De (iii.2.1.7) se obtiene:

$$\sigma_{\epsilon}^2(\infty) = \sigma_{\epsilon}^2(\infty) (1 - \mu S_1(\infty) + \mu^2 S_2(\infty)) \quad (\text{iii.3.2.3})$$

de donde se deduciría que:

$$\mu = S_1(\infty)/S_2(\infty) = h(k)/\sigma_{\epsilon}^{k-2}(\infty) \quad (\text{iii.3.2.4})$$

una función inversamente proporcional a $\sigma_{\epsilon}^{k-2}(\infty)$ con $h(k)$ una función de k . Como es necesario, para que exista convergencia, que se cumpla $\sigma_{\epsilon}^{k-2}(\infty) < \sigma_{\epsilon}^{k-2}(0)$, la expresión (iii.3.2.4) no es compatible con (iii.3.2.3c) con lo que la hipótesis inicial no es cierta y se deduce que la varianza final en este caso es $\sigma_{\epsilon}^2(0) = 0$.

b) Varianza asintótica para $k=2$.

Puesto que para $k=2$, el polinomio $P(\mu, n)$ es constante y menor que la unidad, utilizando (iii.2.1.8) iterativamente se obtiene:

$$\sigma_{\epsilon}^2(n+1) = P^n(\mu, n) \sigma_{\epsilon}^2(0) \quad (\text{iii.3.2.5})$$

de donde tomando límites en ambas partes:

$$\lim_{n \rightarrow \infty} \sigma_{\epsilon}^2(n+1) = \sigma_{\epsilon}^2(0) \lim_{n \rightarrow \infty} P^n(\mu, n) = 0 \quad (\text{iii.3.2.6})$$

c) Varianza asintótica para $k=1$.

Sea μ definido como:

$$\mu = (2\sqrt{2}/\pi/N \sigma_x^2) T = (S_1(n)/S_2(n)) T / \sigma_{\epsilon}^2(n) \quad (\text{iii.3.2.7})$$

con $0 < T^2 < \sigma_{\epsilon}^2(0)$. Esto es equivalente a una cota para μ

$$\mu < (2\sqrt{2}/\pi/N \sigma_x^2) \sigma_{\epsilon}^2(0) \quad (\text{iii.3.2.8})$$

si se sustituye (iii.3.2.7) en (iii.2.1.8) se obtiene para la primera recurrencia:

$$\begin{aligned} \sigma_{\epsilon}^2(1) &= \sigma_{\epsilon}^2(0)(1-TS_1^2(0)/\sigma_{\epsilon}(0)S_2(0)+T^2S_1^2(0)/\sigma_{\epsilon}^2(0)S_2(0)) = \\ &= \sigma_{\epsilon}^2(0) - (\sigma_{\epsilon}(0)-T)TS_1^2(0)/S_2(0) < \sigma_{\epsilon}^2(0) \quad (\text{iii.3.2.9}) \end{aligned}$$

donde la última desigualdad es debida a que $\sigma_{\epsilon}^2(0)$ está acotada inferiormente por T^2 . Si se sustrae de ambas partes de la expresión anterior T^2 resulta:

$$\begin{aligned} \sigma_{\epsilon}^2(1)-T^2 &= (\sigma_{\epsilon}^2(0)-T^2) - (\sigma_{\epsilon}(0)-T)TS_1^2(0)/S_2(0) = \\ &= (\sigma_{\epsilon}(0)-T)(\sigma_{\epsilon}(0)+T-TS_1^2(0)/S_2(0)) \quad (\text{iii.3.2.10}) \end{aligned}$$

pero en este caso (con $k=1$)

$$S_1^2(0)/S_2(0) = (8\sigma_x^4/\pi)/N\sigma_x^4 = 8/N\pi \quad (\text{iii.3.2.11})$$

por tanto para $N \geq 2$ (la longitud de planta mínima) la anterior expresión es siempre menor que 2 y de acuerdo con esto y (iii.3.2.10) se deduce:

$$\sigma_{\epsilon}^2(1) - T^2 > (\sigma_{\epsilon}(0) - T)^2 > 0 \quad (\text{iii.3.2.12})$$

con lo que $T < \sigma_{\epsilon}(1)$. Repitiendo iterativamente este argumento se llega a:

$$T^2 < \sigma_{\epsilon}^2(n+1) < \sigma_{\epsilon}^2(n) \quad (\text{iii.3.2.13})$$

con lo que se concluye que la secuencia $\sigma_{\epsilon}^2(n)$ es monótona decreciente y esta limitada por $T^2 \neq 0$ ($\sigma_{\epsilon}^2(\infty)=T^2$), siempre y cuando éste se elija según la expresión (iii.3.2.7).

A continuación se verifica la afirmación que $\lim_{n \rightarrow \infty} P(\mu, n) = 1$ hecha anteriormente en los casos de $k \neq 2$.

- para $k \geq 2$.

Puesto que en este caso $\sigma_{\epsilon}^2(\infty)=0$ y $S_1(n)$ y $S_2(n)$ son dadas por (iii.2.2.5a,b) y (iii.2.2.6b), queda claro que $\lim_{n \rightarrow \infty} S_1(n)=0$ y $\lim_{n \rightarrow \infty} S_2(n)=0$ y por consiguiente el límite para $P(\mu, n)$ es la unidad.

- para $k = 1$.

En este caso, según (iii.3.2.13) $\sigma_\epsilon^2(\infty) = \Gamma^2$ y el paso de adaptación $\mu = (S_1(\infty)/S_2(\infty))\Gamma/\sigma_\epsilon^2(\infty)$. Si ambas expresiones son sustituidas en (iii.2.1.8) cuando n tiende a infinito, se obtiene:

$$1 - S_1^2(\infty)\Gamma/S(\infty)\sigma_\epsilon(\infty) + S_1^2(\infty)\Gamma/S_2(\infty)\sigma_\epsilon^2(\infty) = 1 \quad (\text{iii.3.2.14})$$

con lo que queda verificado que el límite de $P(\mu, n)$ es la unidad.

III.3.3.- Presencia de ruido de planta

A diferencia del caso anterior, cuando el ruido aditivo de planta no es nulo, el error residual de identificación no tiende a cero para valores finitos del paso de adaptación μ . Para verificar esta afirmación se procede de la forma siguiente:

Supóngase que $\sigma_\epsilon^2(n) \rightarrow 0$, por tanto se cumple $\epsilon(n) \rightarrow 0$. Tomando límites en (iii.2.1.4):

$$\lim_{\sigma_\epsilon(n) \rightarrow 0} \sigma_\epsilon^2(n+1) = \lim_{\sigma_\epsilon(n) \rightarrow 0} \sigma_\epsilon^2(n)(1 - \mu S_1(n) + \mu^2 S_2(n)) \quad (\text{iii.3.3.1})$$

considerando las expresiones (iii.2.1.5) y (iii.2.1.6) y tomando los límites siguientes:

$$\lim_{\sigma_\epsilon(n) \rightarrow 0} \sigma_\epsilon^2(n) S_1(n) \rightarrow 0 \quad (\text{iii.3.3.2})$$

$$\lim_{\sigma_\epsilon(n) \rightarrow 0} \sigma_\epsilon^2(n) S_1(n) \rightarrow k^2 N \sigma_x^4 E[u^{2k-2}] \quad (\text{iii.3.3.3})$$

se observa que la expresión (iii.3.3.1) tendría un límite no nulo, no siendo compatible con la hipótesis inicial a menos que el paso de adaptación $\mu \rightarrow 0$. Por consiguiente se concluye que en presencia de ruido de planta la varianza final del error residual tiende a un valor final no nulo.

En este Apartado sólo se consideran los 4 primeros algoritmos y los ruidos de planta ya mencionados. Usando los valores de $S_1(n)$ y $S_2(n)$ en T3.1, se calculan las correspondientes expresiones de $C_k(n)$ ($k=1,2,3,4$) que acotan localmente a μ , discutiendo sus características particulares según el orden k del algoritmo y tipo de ruido de planta.

a) Acotación de μ para $k=1$.

En este caso:

$$C_1(n) = \begin{cases} K_1 s^2 / \sqrt{1+s^2} & \text{(iii.3.3.4a)} \\ K_2 s^2 (1 - \text{erf}(s)) \exp(s^2) & \text{(iii.3.3.4b)} \\ K_3 s^2 \text{erf}(\sqrt{3/2} s^2) & \text{(iii.3.3.4c)} \\ k_4 s \exp(-1/2 s^2) & \text{(iii.3.3.4d)} \end{cases}$$

Las constantes K_1 a K_4 no dependen de s , tan solo de las potencias de ruido, del proceso de entrada y la longitud de la planta N . En la Figura F3.1 se observa como estas cuatro curvas crecen monótonamente con s (s está en dB).

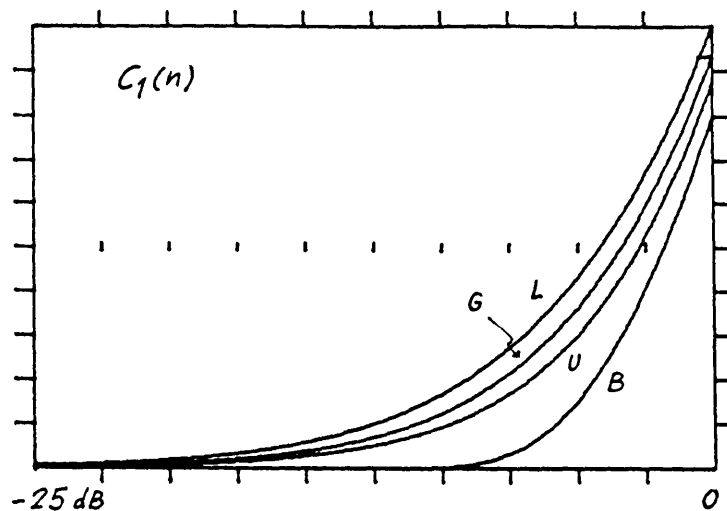


Figura F3.1

Si se elige para $\mu=C_1(s_0)$, y se utiliza (iii.3.1.6), se puede escribir (iii.2.1.4) como:

$$\sigma_e^2(n+1) = \sigma_e^2(n)(1 - C_1(s_0)S_2(n)(C_1(s(n)) - C_1(s_0))) \quad (\text{iii.3.3.5})$$

estando la convergencia garantizada para todo valor de $s(n)$ que permita que $C_1(s(n)) > C_1(s_0) = \mu$ (que μ este acotado localmente). Pero debido al carácter monotono de las curvas anteriores, esto se cumplirá siempre, hasta que $s(n)$ alcance su valor final, que se producirá con la igualdad $C_1(s(\infty))=C_1(s_0)$, siendo su valor final por consiguiente:

$$\sigma_e^2(\infty) = s_0^2 \sigma_u^2 = (C_1^{-1}(\mu))^2 \sigma_u^2 \quad (\text{iii.3.3.6})$$

o sea la elección de μ fija automáticamente la varianza final.

b) Acotación de μ para $k=2$.

En este caso (el algoritmo LMS):

$$C_2(n) = K's / (1+s^2) \quad (\text{iii.3.3.7})$$

siendo independiente de la distribución estadística del ruido de planta u (sólo depende de su potencia). Esta función también crece monótonamente con s , con lo que la discusión es similar a la del caso anterior, esto es, la elección del paso de adaptación fija el valor final de la convergencia y viceversa. En la Figura F3.2 se puede observar como evoluciona dicha curva.

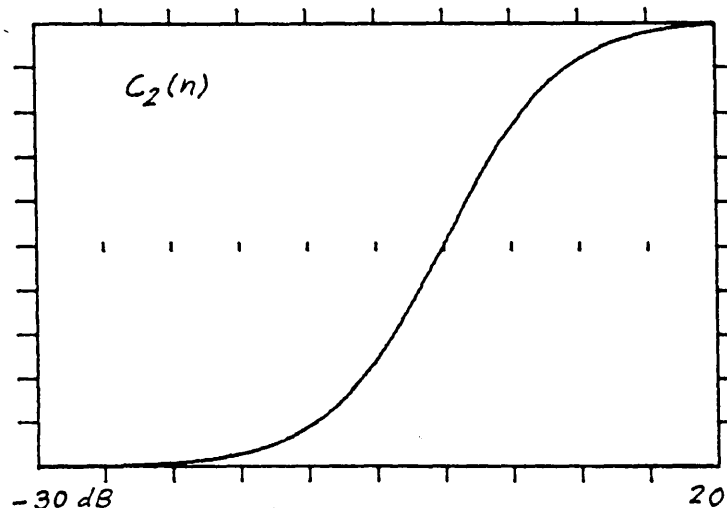


Figura F3.2

c) Acotación de μ para $k=3$ y $k=4$.

En estos casos $C_3(n)$ y $C_4(n)$ no poseen un carácter monótono. Presentan aproximadamente un máximo en:

$k=3$, Gaussiana : $s = 4$ dB

Laplaciana : $s = 4$ dB

Uniforme : $s = 2$ dB

Binaria : $s = 2$ dB

$k=4$, Gaussiana : $s = 0$ dB

Laplaciana : $s = 4$ dB

Uniforme : $s = -2$ dB

Binaria : $s = -7$ dB

decrecen hacia cero cuando $s \rightarrow 0$ ó ∞ . Sus gráficas están representadas en las Figuras F3.3 y F3.4 respectivamente.

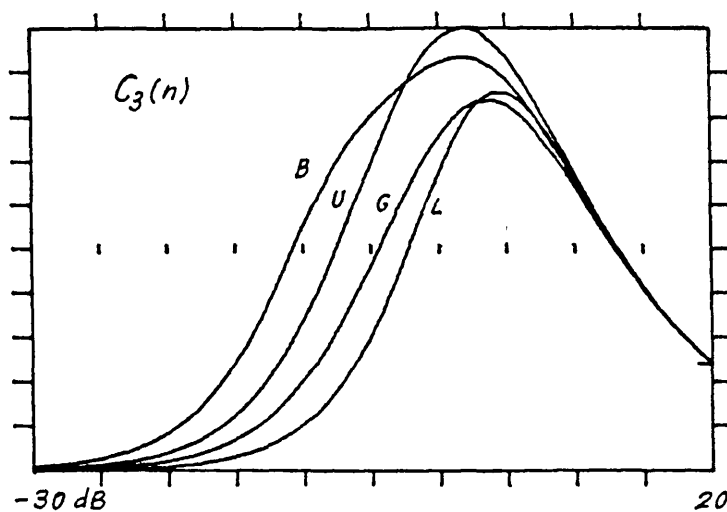


Figura F3.3

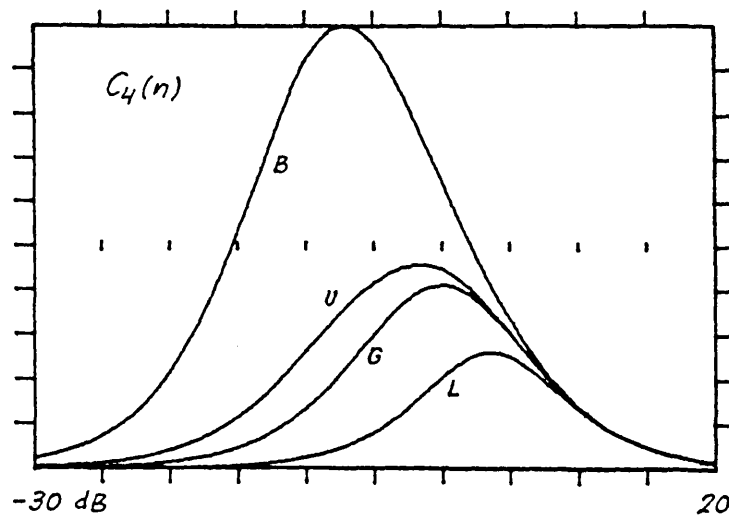


Figura F3.4

Debido a la forma de estas funciones, la discusión seguida en los otros dos casos no es aplicable aquí. En este caso, a diferencia de los anteriores, es obligado examinar los valores iniciales y finales deseados para la varianza del error, ya que $C_k(s(n))$ puede alcanzar su valor mínimo en uno de estos extremos. Walach y Widrow encontraron heurísticamente en [WALACH,1984] que para un mismo paso de adaptación, el proceso convergía al valor deseado o no en función del error inicial. Con errores iniciales mayores el proceso divergía en las primeras iteraciones. Este comportamiento se explica teóricamente a la vista de F3.3 y F3.4 y el siguiente razonamiento:

Supóngase $s_0 < s(0)$, si se elige un $\mu = C_k(s_0)$ con $C_k(s_0) < C_k(s(0))$ la discusión es similar a la de los casos anteriores y la secuencia $\sigma_e^2(n)$ converge hacia s_0 , ya que $C_k(s(n)) > C_k(s_0)$ para cualquier n finito. Sin embargo, cuando el valor inicial de la varianza es tal que $C_k(s(0)) < C_k(s_0)$, la secuencia divergería desde su primera iteración, esto es, $\sigma_e(n+1) > \sigma_e(n)$ (véase (iii.3.3.5) con $C_k(\cdot)$ en vez de $C_1(\cdot)$). Para evitar esto habría que partir de otra condición inicial que disminuyera la varianza inicial o

redefinir $\mu < C_k(s(0))$ y en este caso la secuencia convergería hacia $s(\infty) = C_k^{-1}(\mu)$, siendo este punto deducido de $C_k(s(\infty)) = C_k(s(0))$. Con esto el valor para la varianza final sería:

$$\sigma_\epsilon^2(\infty) = s^2(\infty)\sigma_u^2 = (C_k^{-1}(\mu))^2 \sigma_u^2 < s_0^2 \sigma_u^2 \quad (\text{iii.3.3.8})$$

y la varianza convergería hacia un valor más pequeño que el previamente elegido s_0^2 .

Como nota final, se vuelve a resaltar que el hecho de elegir un paso de adaptación (o equivalentemente elegir una varianza final para la secuencia) en los casos con $k > 2$ requiere examinar el valor correspondiente de $C_k(n)$ en el punto inicial $s(0)$ con el objeto de evitar una posible divergencia debido al carácter no monótono de dichas curvas.

En el Apartado III.6 de conclusiones finales se presentan las curvas teóricas y simulaciones correspondientes derivadas del análisis anterior. Estas curvas pretenden validar el desarrollo anterior además de comparar las prestaciones de convergencia de cada uno de los algoritmos cuando se fija un error residual final conocido.

III.4.-EVALUACION DEL DESAJUSTE FINAL Y RELACION CON LAS CONSTANTES DE TIEMPO DE ADAPTACION

El proceso de adaptación de los coeficientes del sistema a los del filtro transversal sigue una curva de aprendizaje con un desajuste final a la que se le puede ajustar una envolvente exponencial de una determinada constante de tiempo. El desajuste final es debido a que se utiliza un estimador instantáneo del gradiente, el cual lleva implícito un error de estimación que impide que se alcance el mínimo teórico de la superficie objetivo citado en el Apartado III.1.

En el presente Apartado se pretende evaluar las constantes de tiempo correspondientes a los diferentes algoritmos de la familia (excepto el de $k=1$ dado su complejidad de análisis) y su relación con los desajustes medios finales obtenidos. Se sigue el mismo análisis utilizado por Walach y Widrow en [WALACH,1984] extendiéndolo al caso de algoritmos impares, para de esta manera contrastar los resultados actuales con los obtenidos allí por éstos. Este análisis, aunque incompleto, ya que se realiza bajo la hipótesis de un estado de convergencia avanzado (esto es, $u(n) \gg \epsilon(n)$), ofrece resultados satisfactorios en dicho entorno.

El análisis está dividido en dos fases:

- Análisis de la convergencia del valor medio del vector diferencia $y(n)$, definido en (iii.1.7), de donde se derivan las constantes de tiempo de cada coeficiente.

- Determinación de las constantes de tiempo medias de cada uno de los algoritmos y evaluación de los desajustes medios obtenidos

III.4.1.- Análisis de la convergencia en media

Partiendo de las expresiones (iii.1.7) y (iii.1.9) se puede escribir para el vector diferencia $\underline{v}(n)$ la siguiente relación:

$$\underline{v}(n+1) = \underline{v}(n) - \mu k (\epsilon(n) + u(n)) |\epsilon(n) + u(n)|^{k-2} \underline{x}(n) \quad (\text{iii.4.1.1})$$

donde haciendo distinción entre los algoritmos pares e impares:

a) para k par:

$$\underline{v}(n+1) = \underline{v}(n) - \mu k (\epsilon(n) + u(n))^{k-1} \underline{x}(n) \quad (\text{iii.4.1.2})$$

si se desarrolla la potencia en $k-1$ por el binomio de Newton y se considera el error residual $\epsilon(n)$ despreciable frente al ruido de planta $u(n)$, se obtiene aproximadamente:

$$\begin{aligned} \underline{v}(n+1) &= \underline{v}(n) - \mu k (u^{k-1}(n) + (k-1)\epsilon(n)u^{k-2}(n)) \underline{x}(n) = \\ &= \underline{v}(n) - \mu k (k-1) u^{k-2}(n) \underline{x}(n) \underline{x}^t(n) \underline{v}(n) - \mu k u^{k-1}(n) \underline{x}(n) = \\ &= (\underline{I} - \mu k (k-1) u^{k-2}(n) \underline{x}(n) \underline{x}^t(n)) \underline{v}(n) - \mu k u^{k-1}(n) \underline{x}(n) \end{aligned} \quad (\text{iii.4.1.3})$$

siendo \underline{I} la matriz diagonal unidad.

b) para k impar (excepto k=1):

$$\underline{v}(n+1) = \underline{v}(n) - \mu k \text{sgn}(\epsilon(n) + u(n)) (\epsilon(n) + u(n))^{k-1} \underline{x}(n) \quad (\text{iii.4.1.4})$$

donde considerando nuevamente $u(n) \gg \epsilon(n)$ y de forma análoga al caso par se obtiene:

$$\begin{aligned} \underline{v}(n+1) &= (\underline{I} - \mu k (k-1) \text{sgn}(u(n)) u^{k-2}(n) \underline{x}(n) \underline{x}^t(n)) \underline{v}(n) + \\ &+ \mu k (k-1) \text{sgn}(u(n)) u^{k-1}(n) \underline{x}(n) \end{aligned} \quad (\text{iii.4.1.5})$$

Las expresiones (iii.4.1.3) y (iii.4.1.5) se pueden escribir de forma única como:

$$\underline{v}(n+1) = (\underline{I} - f_1(k,u)\underline{x}(n)\underline{x}^t(n))\underline{v}(n) - f_2(k,u)\underline{x}(n) \quad (\text{iii.4.1.6})$$

con

$$f_1(k,u) = \begin{cases} \mu k(k-1)u^{k-2}(n) & ; k \text{ par} \quad (\text{iii.4.1.7a}) \\ \mu k(k-1)\text{sgn}(u(n))u^{k-2}(n) & ; k \text{ impar} \quad (\text{iii.4.1.7b}) \end{cases}$$

$$f_2(k,u) = \begin{cases} \mu k u^{k-1}(n) & ; k \text{ par} \quad (\text{iii.4.1.8a}) \\ \mu k \text{sgn}(u(n))u^{k-1}(n) & ; k \text{ impar} \quad (\text{iii.4.1.8b}) \end{cases}$$

Tomando esperanzas matemáticas en (iii.4.1.6) y recordando las hipótesis H1, H3 y H4 se obtiene en general:

$$E[\underline{v}(n+1)] = \underline{\Omega} E[\underline{v}(n)] \quad (\text{iii.4.1.9})$$

con

$$\underline{\Omega} = \underline{I} - \mu k(k-1)E[|u(n)|^{k-2}] \underline{R} \quad (\text{iii.4.1.10})$$

y

$$\underline{R} = E[\underline{x}(n)\underline{x}^t(n)] \quad (\text{iii.4.1.11})$$

Bajo nuestras hipótesis, la matriz de autocorrelación \underline{R} es diagonal, con lo que la convergencia en media de cada componente $v_i(n)$ es independiente de las demás y la acotación particular de μ para garantizar dicha convergencia idéntica a las del resto de las componentes, ya que el autovalor i -ésimo de \underline{R} coincide con la potencia del proceso σ_x^2 , esto es:

$$\mu < 1/k(k-1)E[|u|^{k-2}] \sigma_x^2 \quad (\text{iii.4.1.12})$$

En la expresión (iii.4.1.9) se observa que el valor medio del vector $\underline{v}(n)$ sigue una expresión geométrica, con lo que es factible el ajustar una exponencial de constante de tiempo conocida a cada componente. Debido a H1 éstas serán idénticas para todas las componentes y valdrán (véase [WIDROW,1976]):

$$\tau_i = 1/\mu k(k-1)E[|u|^{k-2}] \sigma_x^2 \quad (\text{iii.4.1.13})$$

donde se observa que para un valor arbitrario de k , el conjunto de constantes de tiempo es proporcional a las de cualquier otro algoritmo con k diferente y en particular a las del algoritmo LMS, con lo que se puede utilizar este último como patrón de referencia de los diferentes algoritmos. Es de resaltar también la dependencia (para $k > 2$) de los momentos del ruido de planta $u(n)$, con lo que en estos algoritmos la presencia de ruidos de planta de características estadísticas diferentes harán variar las prestaciones de velocidad de los mismos.

III.4.2.-Desajustes medios finales

Con objeto de tener un criterio de comparación válido relativo a la velocidad de convergencia de cada uno de los algoritmos de la familia se introduce el parámetro n_{med} , el cual determinará el número medio de muestras para convergencia y estará relacionado con las constantes de tiempo de los coeficientes calculadas anteriormente. Debido a que todos los coeficientes son adaptados de forma simultánea, el número medio de muestras para convergencia se define como el promedio armónico de

las constantes de tiempo individuales, esto es:

$$1/n_{med} = (1/N) \sum_{i=1}^N 1/\tau_i \quad (\text{iii.4.2.1})$$

de donde, utilizando (iii.4.1.12) se obtiene:

$$n_{med} = 1/\mu k(k-1)E[|u|^{k-2}] \sigma_x^2 \quad (iii.4.2.2)$$

Iguálmente a como se hace en [WALACH,1984] se define el desajuste medio del algoritmo k-ésimo como la relación:

$$M(k) = (\text{pot. error residual})/(\text{pot. error Wiener}) \quad (iii.4.2.3)$$

o más explícitamente según (iii.2.1.2) como:

$$M(k) = E[\sigma_e^2(\infty)]/E[u^2] = \Sigma (\sigma_x^2/\sigma_u^2) E[v_i^2(\infty)] \quad (iii.4.2.4)$$

expresión que mide el error residual del proceso en régimen permanente, normalizado por la potencia del ruido de planta. Esta función es característica de los procesos adaptativos de gradiente estocástico, ya que mide la degradación final del mismo respecto del valor óptimo alcanzable.

Para relacionar este desajuste final con el tiempo medio de convergencia se procede de la siguiente forma: se traspone (iii.4.1.6), se multiplica el resultado por la derecha miembro a miembro nuevamente con (iii.4.1.6) y se toman esperanzas matemáticas utilizando las hipótesis H1, H3 y H4. Con ello se obtiene la siguiente expresión recurrente para la matriz de autocorrelación del vector diferencia $\underline{v}(n)$:

$$\begin{aligned} E[\underline{v}(n+1)\underline{v}^t(n+1)] &= E[\underline{v}(n)\underline{v}^t(n)] + E[f_2^2(k,u)\underline{x}(n)\underline{x}^t(n)] - \\ &- E[f_1(k,u)(\underline{x}(n)\underline{x}^t(n)\underline{v}(n)\underline{v}^t(n) + \underline{v}(n)\underline{v}^t(n)\underline{x}(n)\underline{x}^t(n))] + \\ &+ E[f_1^2(k,u)\underline{x}(n)\underline{x}^t(n)\underline{v}(n)\underline{v}^t(n)\underline{x}(n)\underline{x}^t(n)] \end{aligned} \quad (iii.4.2.5)$$

donde recordando la hipótesis auxiliar de estado avanzado de convergencia ($\underline{v}(n)$ pequeño) y observando que en el último término aparece $f_1^2(k,u)$ (o de forma equivalente μ^2) además del producto $\underline{v}\underline{v}^t$, se puede despreciar este término frente al antepenúltimo. Aún más, en régimen permanente (con $n \rightarrow \infty$) y con la simplificación anterior, (iii.4.2.5) se queda en:

$$\begin{aligned} \underline{0} &= -E[f_1(k,u)(\underline{x}\underline{x}^t\underline{v}(\infty)\underline{v}^t(\infty) + \underline{v}(\infty)\underline{v}^t(\infty)\underline{x}\underline{x}^t)] + \\ &+ E[f_2^2(k,u)\underline{x}\underline{x}^t] \end{aligned} \quad (iii.4.2.6)$$

de donde, expresando la esperanza como producto de esperanzas y despejando se obtiene la matriz de autocorrelación en régimen permanente del vector $\underline{v}(n)$:

$$E[\underline{v}(\infty)\underline{v}^t(\infty)] = (E[f_2^2(k,u)]/E[f_1(k,u)]) \underline{I} \quad (\text{iii.4.2.7})$$

y sustituyendo (iii.4.1.7) y (iii.4.1.8) se obtiene finalmente para cualquier k arbitrario ($k > 1$):

$$E[\underline{v}(\infty)\underline{v}^t(\infty)] = (\mu k E[u^{2k-2}]/(k-1)E[|u|^{k-2}]) \underline{I} \quad (\text{iii.4.2.8})$$

En la ecuación anterior se observa que las componentes están incorreladas (hecho ya conocido para el LMS [WIDROW,1975], [WIDROW,1967]) y que la única diferencia formal con la expresión obtenida por Walach y Widrow en [WALACH,1984] es que en la esperanza del denominador aparece el valor absoluto al considerar también los algoritmos de k impares.

Según la expresión anterior, los valores de la diagonal principal son:

$$E[v_1^2(\infty)] = \mu k E[u^{2k-2}]/(k-1)E[|u|^{k-2}] \quad (\text{iii.4.2.9})$$

con lo que sustituyéndola en (iii.4.2.4) se obtiene el valor del desajuste medio final $M(k)$ para cualquier $k > 1$. Este resulta:

$$M(k) = (\mu k N \sigma_x^2 E[u^{2k-2}]/(k-1) \sigma_u^2 E[|u|^{k-2}]) \quad (\text{iii.4.2.10})$$

Si se expresa en función del número medio de muestras n_{med} haciéndola independiente del paso de adaptación μ :

$$M(k) = \frac{NE[u^{2k-2}]}{(k-1)^2 \sigma_u^2 E^2[|u|^{k-2}] n_{med}} \quad (\text{iii.4.2.11})$$

de lo que se deduce que cuando se desean convergencias rápidas el desajuste $M(k)$ crece y al contrario con convergencias lentas.

Con objeto de comparar los desajustes finales entre los diferentes algoritmos de la familia manteniendo fijo el número de muestras medio para la convergencia, se define una nueva función $A(k)$ la cual utiliza como patrón de referencia el desajuste final del algoritmo LMS. Dicha función se define como:

$$A(k) = M(2)/M(k) \quad (\text{iii.4.2.12})$$

$$k > 1$$

de donde, utilizando (iii.4.2.11) resulta:

$$A(k) = (k-1)^2 \sigma_u^2 E[|u|^{k-2}] / E[u^{2k-2}] \quad (\text{iii.4.2.13})$$

En la Tabla T3.2 se recogen los valores correspondientes a los algoritmos con $k=2$ a $k=6$ en función de la distribución estadística del ruido de planta presente (se supone la misma potencia de ruido). Para su evaluación se han utilizado las expresiones generales de los momentos centrales calculados en el Apendice A.10.

A(k)	(k)				
	2	3	4	5	6
GAUSSIANA	1	.85	.60	.39	.24
LAPLACIANA	1	.33	.10	.029	.0026
UNIFORME	1	1.67	2.33	3.0	3.67
BINARIA	1	4	9	16	25

Tabla T3.2

En esta Tabla se observa que para ruidos de planta con distribuciones estadísticas concentradas (uniforme y binaria), a igualdad de velocidad de convergencia media, los algoritmos con $k > 2$ ofrecen desajustes finales más pequeños que el algoritmo LMS y al contrario con los ruidos de planta con distribuciones estadísticas no concentradas (Gaussiana y Laplaciana). Este efecto se agudiza cuanto más concentrada o abierta sean las distribuciones respectivamente. Por ejemplo, para un ruido Laplaciano los algoritmos de orden bajo ofrecen el menor desajuste final, pero ocurre totalmente lo contrario cuando el ruido es binario. En la Tabla T3.3 se resumen las mejoras y degradaciones en dB ofrecidas por los anteriores algoritmos dependiendo del ruido de planta presente.

(k)

$10 \log(A(k))$ (dB)	2	3	4	5	6
GAUSSIANA	0	-0.71	-2.22	-4.09	-6.20
LAPLACIANA	0	-4.81	-10.0	-15.4	-22.2
UNIFORME	0	2.22	3.67	4.77	5.65
BINARIA	0	6.02	9.54	12.04	13.98

Tabla T3.3

En el Apartado III.6 de conclusiones de este Capítulo se ofrecen las curvas teóricas y las simulaciones derivadas del análisis anterior, las cuales permiten comprobar la veracidad de los resultados de la Tabla T3.3 cuando se fija el número medio de muestras para la convergencia.

III.5.- OPTIMIZACION DETERMINISTA DEL PASO DE ADAPTACION

El paso de adaptación μ , tal y como ha sido estudiado en los Apartados anteriores, es el máximo responsable de la convergencia del algoritmo, ya que ha de estar convenientemente acotado. Sin embargo, aunque la elección del paso a utilizar en cada proceso puede depender de un desajuste final o una velocidad de convergencia deseados, es también posible calcular en cada iteración el paso de adaptación óptimo que maximice la reducción de la varianza del error residual restante, incrementando así la velocidad de convergencia media. El hecho de ser determinista está ligado a que su ley de variación es conocida, dependiendo esta de la varianza $\sigma_{\epsilon}^2(n)$ (no es una v.a.) a través de $S_1(n)$ y $S_2(n)$.

Para calcular este paso de adaptación óptimo se ha de derivar e igualar a cero el polinomio de convergencia $P(\mu, n)$, esto es:

$$dP(\mu, n)/d\mu \Big|_{\mu=\mu_0(n)} = -S_1(n) + 2\mu_0(n)S_2(n) = 0 \quad (\text{iii.5.1})$$

de donde

$$\mu_0(n) = S_1(n)/2S_2(n) = C_k(n)/2 \quad (\text{iii.5.2})$$

Se observa que la convergencia local (y por tanto la absoluta) está garantizada, ya que $C_k(n)$, según (iii.3.1.6) era la cota local superior de μ necesaria para la convergencia del algoritmo L_k .

Sustituyendo este valor óptimo de μ en la expresión de $P(\mu, n)$ se obtiene la óptima reducción local posible del error residual final:

$$P_0(\mu, n) = P(\mu, n) \Big|_{\mu=\mu_0(n)} = 1 - S_1^2(n)/4S_2(n) \quad (\text{iii.5.3})$$

donde, de acuerdo con (iii.3.1.4) $P_0(\mu, n) < 1$ si $N \geq 2$, con lo que se garantiza la convergencia en varianza tal como se dijo anteriormente.

Debido a que $S_1(n)$ tiende a cero cuando lo hace el error residual $e(n)$, $P_0(\mu, n)$ tiende asintóticamente a la unidad.

El valor instantáneo de μ ha de ser calculado en cada iteración utilizando las expresiones oportunas de $S_1(s(n))$ y $S_2(s(n))$. Con ello se obtiene una secuencia de $\mu(n)$ variables que permiten una mayor reducción local del error mejorando la velocidad de convergencia.

A continuación, se presentará una discusión comparativa de prestaciones, considerando la ya usual separación del estudio en los casos de planta no ruidosa y planta ruidosa.

III.5.1.- Ausencia de ruido de planta

De las expresiones (iii.2.2.5) y (iii.2.2.6) se deduce fácilmente para el caso sin ruido:

$$P_0(\mu, n) = \begin{cases} 1 - 2/\pi N & ; k=1 & \text{(iii.5.1.1a)} \\ 1 - ((k-1)!!)^2 / N(2k-3)!! & ; k \text{ par} & \text{(iii.5.1.1b)} \\ 1 - 2^k ((k-1)/2)!!^2 / \pi N(2k-3)!! & ; k \text{ impar} & \text{(iii.5.1.1c)} \end{cases}$$

Es conveniente resaltar que las expresiones anteriores sólo dependen de la longitud de planta N y del orden del algoritmo, manteniéndose constantes con la iteración n . Esto permite que se alcance un valor de error residual final nulo sin que aparezca un efecto de curvatura o saturación en la convergencia de la varianza del error. La Tabla T3.4 resume la reducción de varianza del error en dB por cada 100 iteraciones para los primeros 6 algoritmos de la familia y 3 longitudes de planta distintas.

$P_0(\mu, n)$	(k)					
	1	2	3	4	5	6
dB						
10	28.7	45.8	38.6	26.9	17.3	10.6
(N)						
20	14.1	22.3	18.6	13.2	8.33	5.24
40	7.00	11.0	9.22	6.56	4.36	2.61

Tabla T3.4

Nótese que el algoritmo optimizado L2 presenta el mejor comportamiento y que el incrementar el orden del algoritmo k hace disminuir la velocidad de convergencia, aunque este efecto va desapareciendo paulatinamente conforme la longitud de planta del filtro aumenta.

Es conveniente en este punto resaltar la expresión particular del paso de adaptación óptimo. De acuerdo con (iii.5.2) resulta ser:

$$\mu_0(n) = F(k)/N \sigma_x^2 \sigma_\epsilon^{k-2}(n) \quad (\text{iii.5.1.2})$$

donde F(k) es una función del orden k. Este paso óptimo depende de la iteración n, excepto cuando k=2 que es constante, lo cual es una ventaja adicional de este algoritmo ya que no es necesario un cálculo local del μ a utilizar. Cuando k=1, el paso decrece con n de la misma manera que lo hace $\sigma_\epsilon(n)$ y con k>2 crecería indefinidamente hasta infinito si desea mantener el carácter óptimo hasta alcanzar un error residual final nulo.

De acuerdo con lo anterior, se puede concluir nuevamente que el algoritmo L2 optimizado es la mejor elección que se puede hacer si hay ausencia de ruido de planta en el sistema a identificar, ya que ofrece la

mayor velocidad de convergencia del error cuadrático medio para un paso de adaptación óptimo constante.

III.5.2.- Presencia de ruido de planta

Con objeto de evitar la tediosa labor de reescribir las correspondientes expresiones de los pasos óptimos en función de $S_1(n)$ y $S_2(2)$ (según se dan en las tablas I3.1) se ha optado por presentar las curvas correspondientes de la reducción óptima del error residual $P_0(\mu, n)$ en dB para cada uno de los ruidos de planta bajo estudio con los 4 primeros algoritmos de la familia.

En dichas curvas se explora la banda de $s=\sigma_e/\sigma_u$ desde los 20 dB a -30 dB asumiendo una longitud de planta de $N=20$. El resumen de figuras es el siguiente:

- Figura F3.5: ruido Gaussiano. L1 a L4
- Figura F3.6: ruido Laplaciano. L1 a L4
- Figura F3.7: ruido Uniforme. L1 a L4
- Figura F3.8: ruido Binario. L1 a L4
- Figura F3.9: algoritmo L1. Todos los ruidos
- Figura F3.10: algoritmo L2. Todos los ruidos
- Figura F3.11: algoritmo L3. Todos los ruidos
- F3.12: algoritmo L4. Todos los ruidos

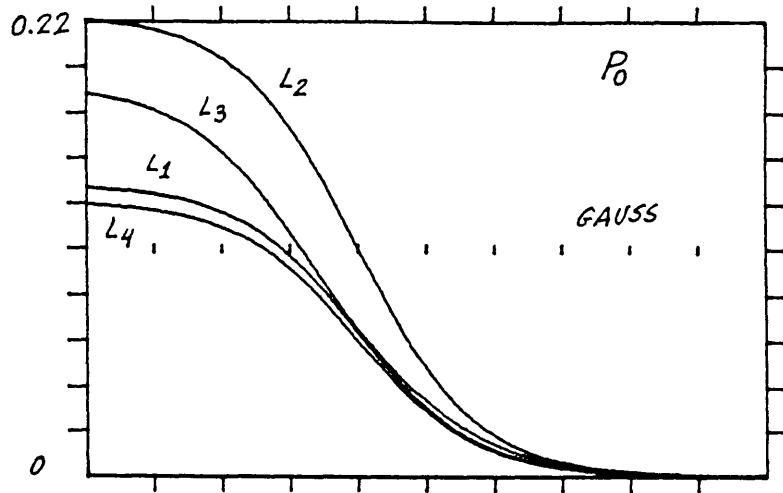


Figura F3.5

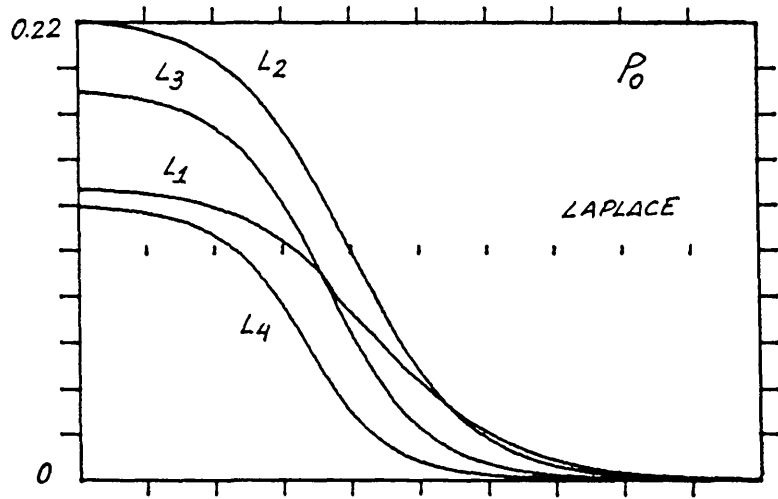


Figura F3.6

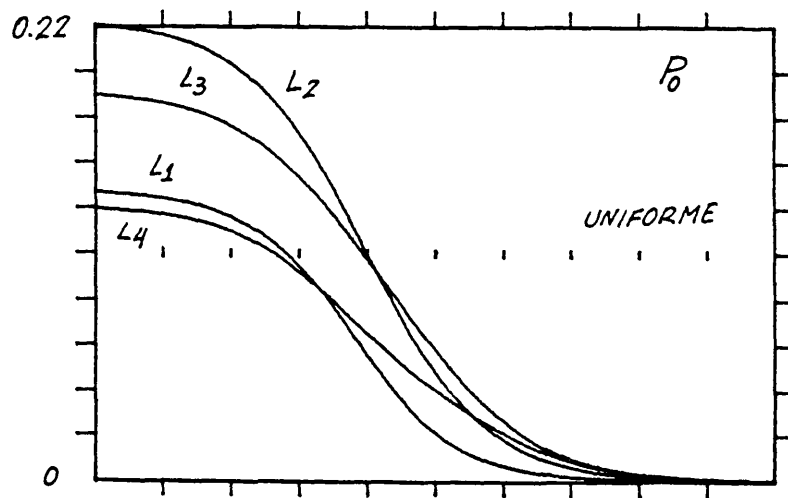


Figura F3.7

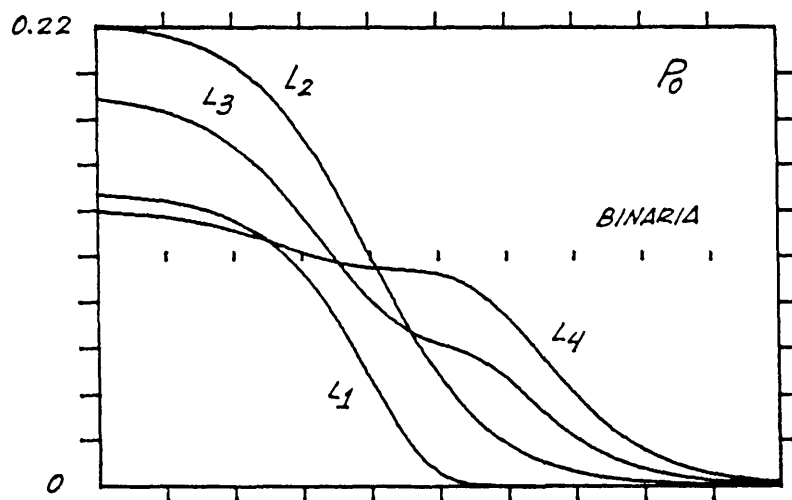


Figura F3.8

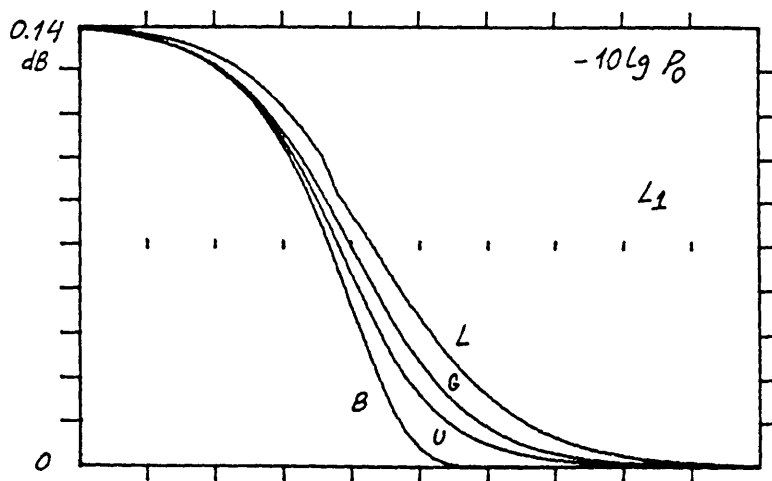


Figura F3.9

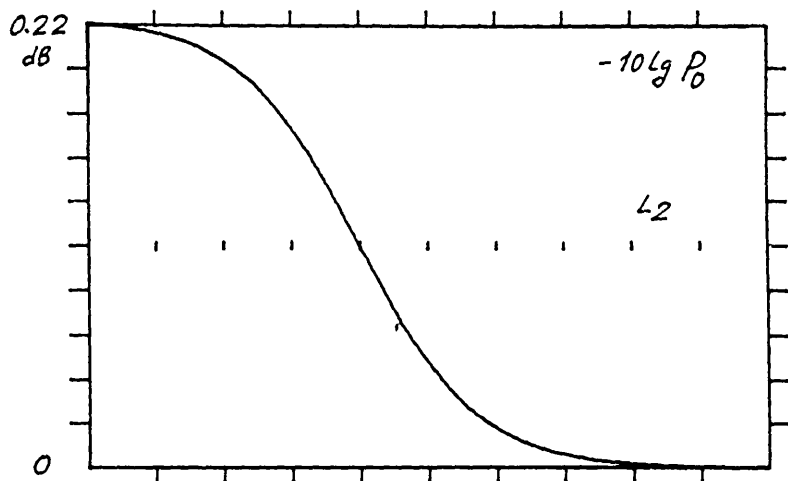


Figura F3.10

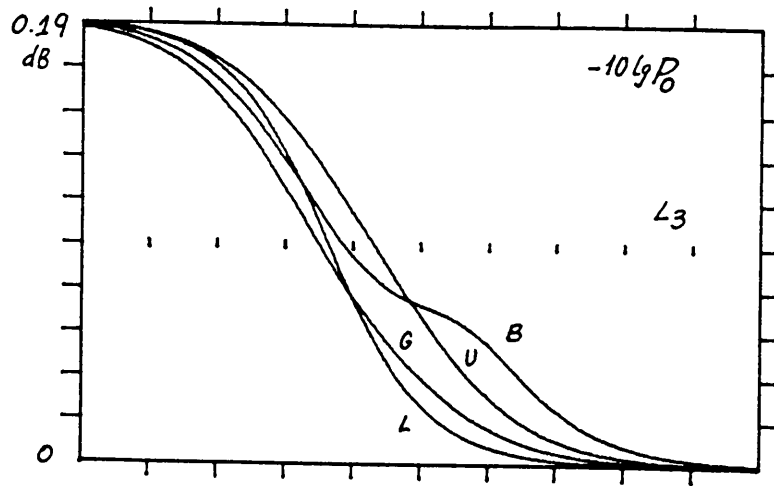


Figura F3.11

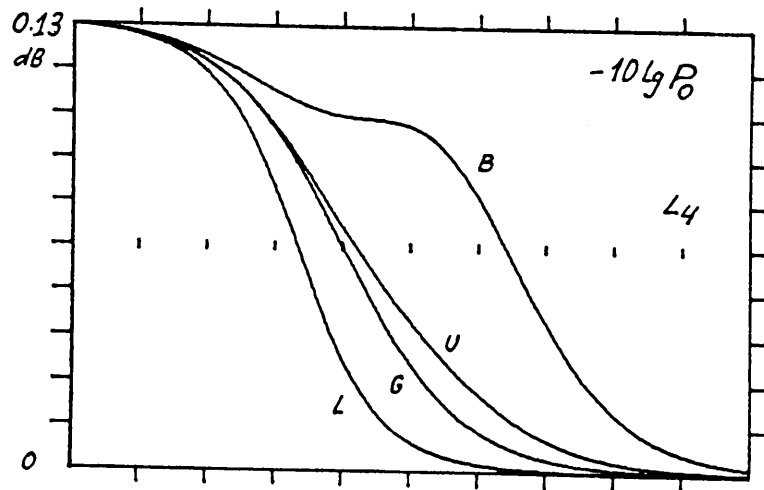


Figura F3.12

En las primera 4 curvas se puede observar que para valores altos de s (comparable a la ausencia de ruido de planta) el algoritmo L2 ofrece mejores prestaciones que el resto, tal como se dedujo anteriormente, sin embargo cuando s decrece existen diferentes comportamientos en función del ruido de planta involucrado:

- * Con ruido Gaussiano, el L2 es el mejor de todos.
- * Con ruido Laplaciano, el L1 tiende a comportarse mejor que el resto cuando s decrece (por debajo de $s = -7$ dB).
- * Con ruido uniforme y binario los algoritmos de orden alto tienden a comportarse mejor cuando el ruido de planta empieza a ser relevante con respecto al error residual. En ambos casos el L4 tiende a ser el mejor de todos, seguido por el resto en orden descendente.

Las 4 curvas siguientes revelan el comportamiento de cada uno de los algoritmos individualmente con respecto a los ruidos bajo estudio. Se observa en la Figura F3.10 la no dependencia del algoritmo L2 con el caracter del ruido. A la vista del resto de las figuras y como comentario general se puede afirmar que los algoritmos de órdenes bajos son más adecuados para ruidos de planta de colas largas y los de orden elevado para los de colas cortas.

Estos resultados son coherentes con los del Apartado anterior y permiten a cualquier diseñador de sistemas adaptativos el elegir en cada caso el algoritmo apropiado.

III.6.- SIMULACION DE LOS ALGORITMOS

En este Apartado se verifican, por medio de simulación, las conclusiones teóricas establecidas en los Apartados anteriores.

Se realizan cuatro grupos de simulaciones diferentes con los cuatro primeros algoritmos de la familia Lk con objeto de:

a) Examinar la evolución del error cuadrático medio con ausencia de ruido planta.

b) Examinar la evolución del error cuadrático medio con presencia de ruido de planta y mismo desajuste final para los 4 algoritmos.

c) Examinar la evolución del error cuadrático medio con presencia de ruido de planta y un paso de adaptación optimizado.

d) Examinar el desajuste final con idéntica velocidad de convergencia.

Los elementos comunes a los cuatro grupos de simulaciones son:

- Potencia media del ruido de planta $\sigma_u^2 = 1$

- Longitud de la planta $N=10$

- Vector de coeficientes para la planta

$$c_p(i) = (-1)^{(i-1)} \{1 - 0.05(i-1)\} \quad (\text{iii.6.1})$$
$$(1 \leq i \leq 10)$$

- Condición inicial para los coeficientes del filtro adaptativo

$$c_i(0) = c_p(i) (1+r(i)) \quad (\text{iii.6.2})$$

con $r(i)$ una v.a. uniforme en $[-.5, .5]$

III.6.1.- Convergencia con ausencia de ruido de planta

Para el primer grupo se usa una secuencia de excitación Gaussiana, blanca de media nula y potencia unidad. Inicialmente se han evaluado las correspondientes cotas para μ según las expresiones (iii.3.2.2a,b,c) donde se ha utilizado $\epsilon^2(0)$ como estima de $\sigma_e^2(0)$.

Las cotas resultantes son:

$$\mu_1 < .046, \mu_2 < .1, \mu_3 < .047, \mu_4 < .031$$

pero una simulación mas exhaustiva ofrece las siguientes cotas prácticas para convergencia:

$$\mu_2 < .095, \mu_3 < .042, \mu_4 < .034$$

(no se referencia la cota para el algoritmo L1 ya que nunca diverge). Se observa que los dos conjuntos de valores anteriores son lo suficiente parecidos como para validar el análisis e hipótesis utilizadas.

Las Figuras F3.13a a F3.13d muestran las curvas de convergencia teórica para la varianza de error residual y un promedio de 10 realizaciones para $\epsilon^2(n)$ del ejemplo considerado. Se han utilizado los siguientes pasos de adaptación:

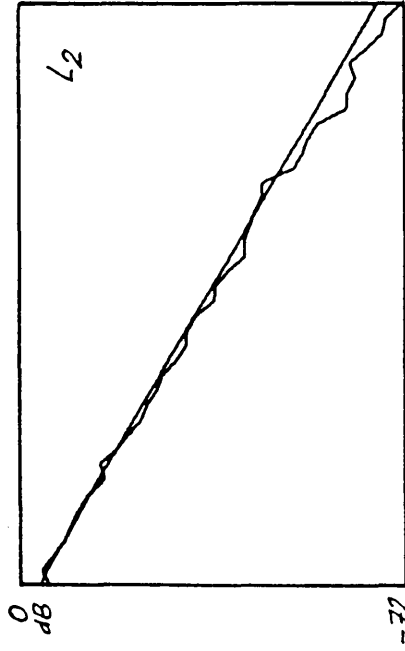
$$\mu_1 = .005, \mu_2 = .01, \mu_3 = .02, \mu_4 = .02$$

se han generado 400 muestras y las curvas presentadas se han suavizado con una ventana móvil uniforme de 10 muestras.

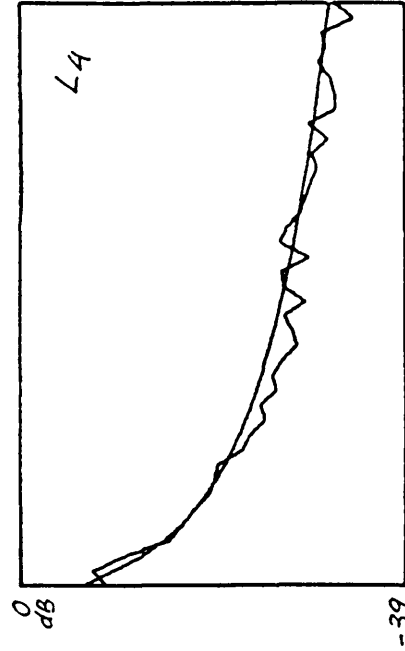
En estas figuras se puede observar:

- a) El comportamiento log-lineal del algoritmo L2 con una μ constante.
- b) La varianza del error residual alcanza un valor estable no nulo en el algoritmo L1.

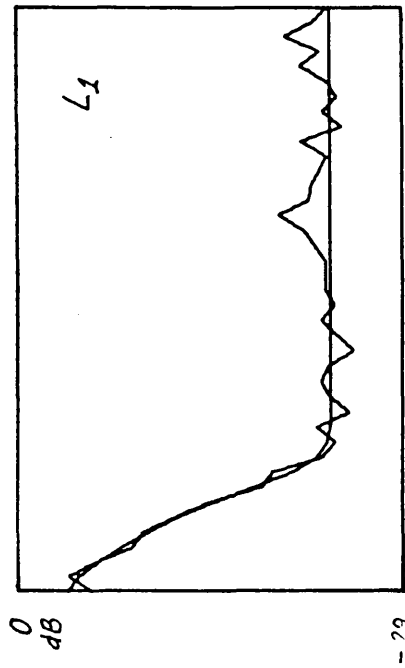
c) El efecto de saturación en los algoritmos L3 y L4 debido a la separación progresiva del μ usado con el óptimo correspondiente.



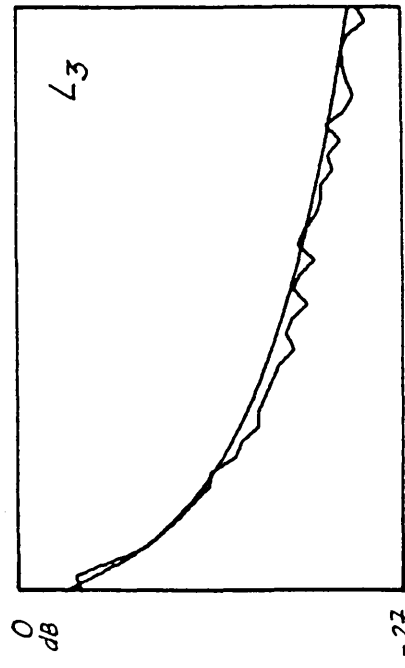
F3.13b



F3.13d



F3.13a



F3.13c

III.6.2.- Convergencia con ruido de planta

Se usa una excitación Gaussiana blanca de media nula y potencia $\sigma_x^2=20$. Se desea igualmente un desajuste final constante igual a -25 dB para todos los ensayos con objeto de poder comparar las curvas de convergencia de los diferentes algoritmos con los 4 ruidos de planta estudiados. Para ello se ha calculado el paso de adaptación como $\mu=C(s(\infty))$ con $s^2(\infty)=-25\text{dB}$. En los casos con $k=3,4$ se cumple además $C(s(\infty))<C(s(0))$, condición necesaria para evitar la divergencia. Todas las curvas simuladas provienen de promediar 4 realizaciones de 10000 muestras cada una y han sido suavizadas con una ventana móvil uniforme de 40 muestras. Las curvas simuladas se superponen a las teóricas correspondientes con objeto de verificar la validez del análisis teórico. Se presentan además con cada conjunto de simulaciones, las velocidades teóricas de convergencia como ayuda en la discusión del comportamiento de las simulaciones.

La Figura F3.14 presenta las velocidades teóricas de convergencia, medida como $-10 \log P(\mu, n)$, para los 4 primeros algoritmos y ruido de planta Gaussiano. Las Figuras F3.15a a F3.15d muestran las convergencias teóricas y simuladas con dicho ruido de planta, donde los valores para los pasos de adaptación, calculados según el Apartado 11.3, para conseguir un desajuste final de -25 dB son:

$$\mu_1 = 2.5 \times 10^{-5} \quad \mu_2 = 1.6 \times 10^{-5} \quad \mu_3 = 5.6 \times 10^{-5} \quad \mu_4 = 1.6 \times 10^{-6}$$

En este conjunto de figuras, así como en las restantes, se destaca la mayor velocidad de convergencia para $k > 2$ en los momentos iniciales de adaptación (errores grandes) así como una saturación progresiva de dicha convergencia conforme el error residual decrece.

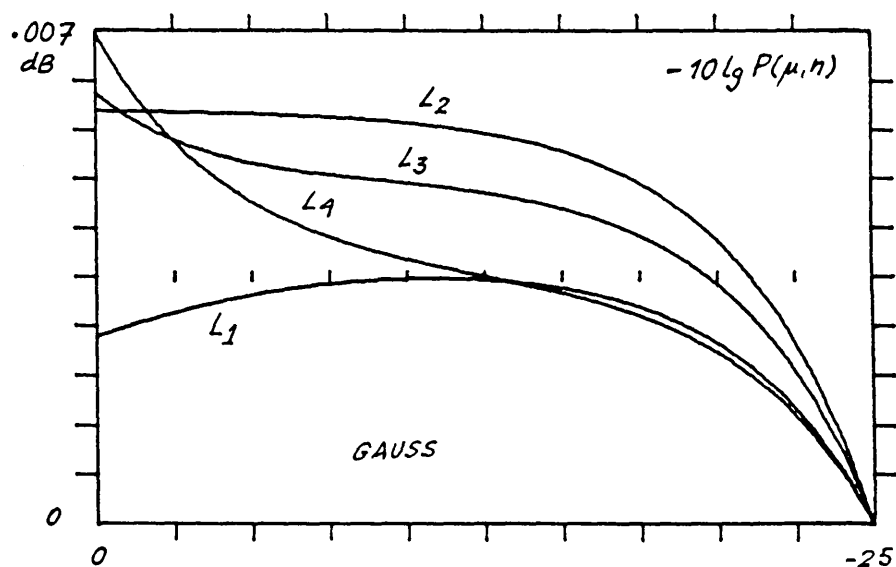
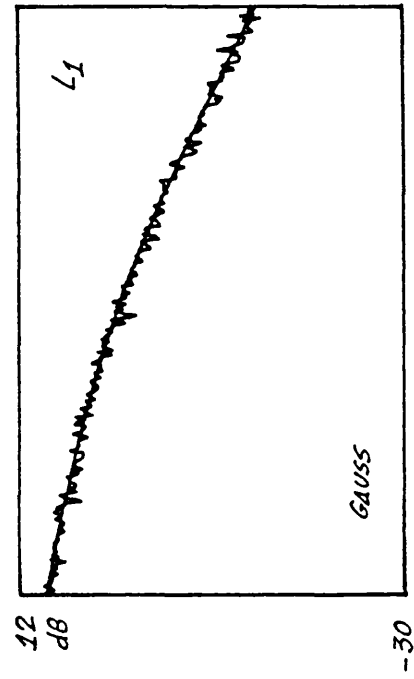
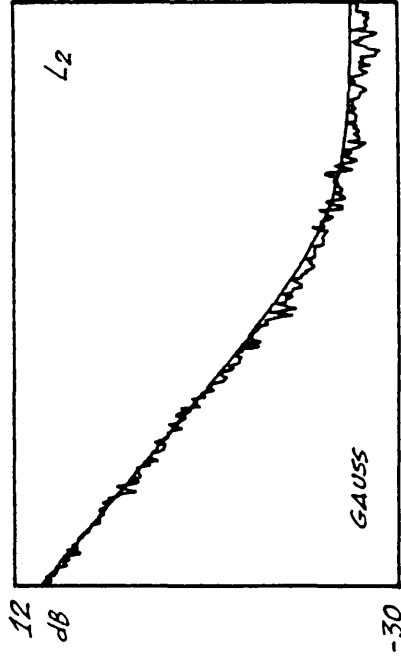


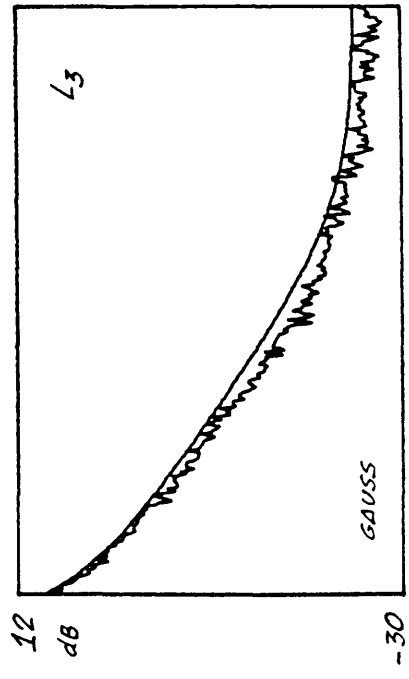
Figura F3.14



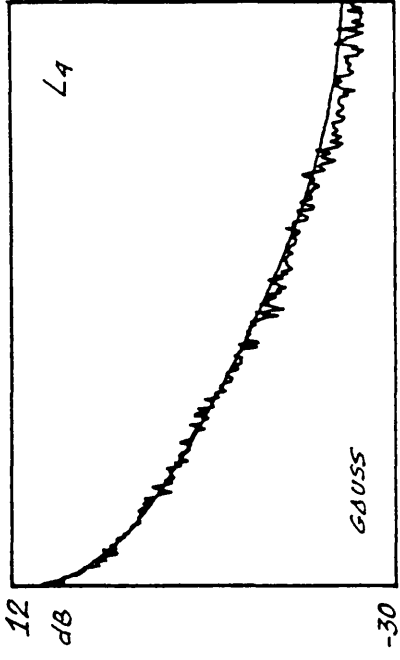
F3.15a



F3.15b



F3.15c



F3.15d

La Figura F3.16 presenta las velocidades teóricas de convergencia para ruido de planta Laplaciano y las Figuras F3.17a a F3.17d las convergencias teóricas y simuladas para dicho ruido. En este caso se observa claramente la inflexión que sufre la convergencia del algoritmo L1 (correspondiente al máximo de velocidad que este algoritmo presenta en algún punto intermedio de su convergencia y que F3.16 resalta). Los pasos de adaptación utilizados son:

$$\mu_1 = 4.2 \times 10^{-5} \quad \mu_2 = 1.6 \times 10^{-5} \quad \mu_3 = 2.5 \times 10^{-6} \quad \mu_4 = 2.6 \times 10^{-7}$$

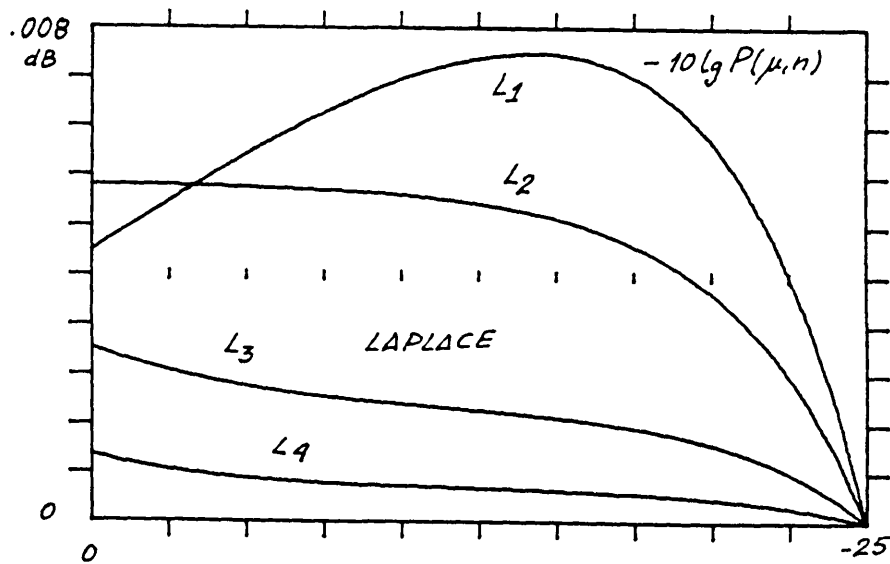
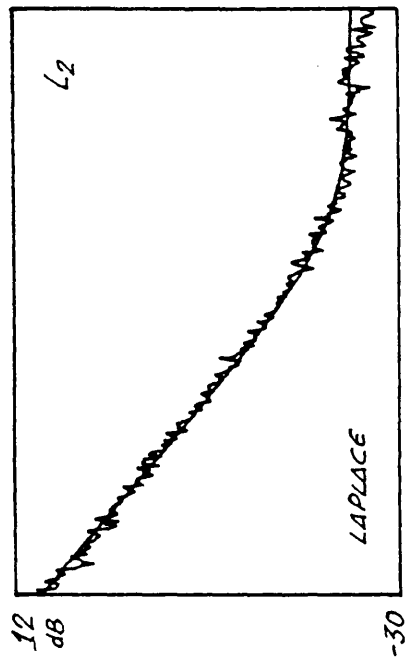
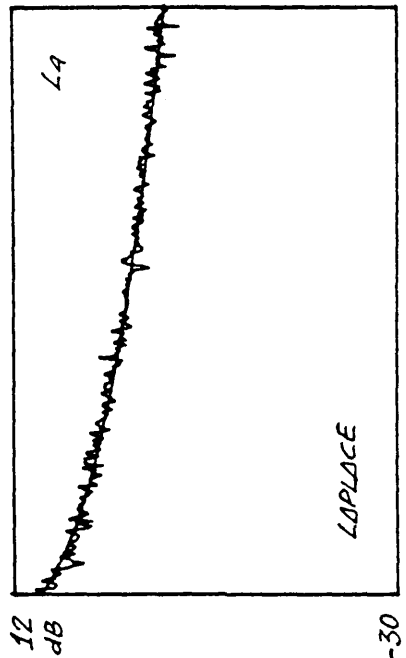


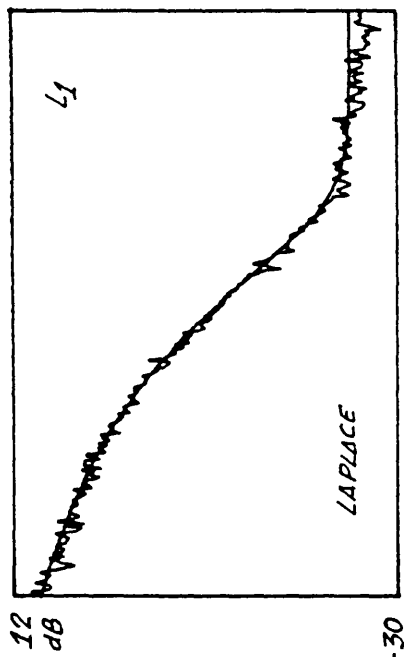
Figura F3.16



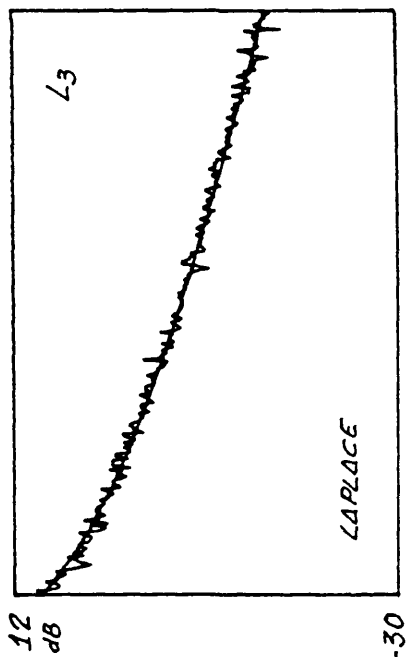
F3.17b



F3.17d



F3.17a



F3.17c

La Figura F3.18 presenta las velocidades teóricas de convergencia para ruido de planta Uniforme y las Figuras F3.19a a F3.19d las convergencias teóricas y simuladas para dicho ruido. Los pasos de adaptación utilizados son:

$$\mu_1 = 1.6 \times 10^{-5} \quad \mu_2 = 1.6 \times 10^{-5} \quad \mu_3 = 1.0 \times 10^{-5} \quad \mu_4 = 4.3 \times 10^{-6}$$

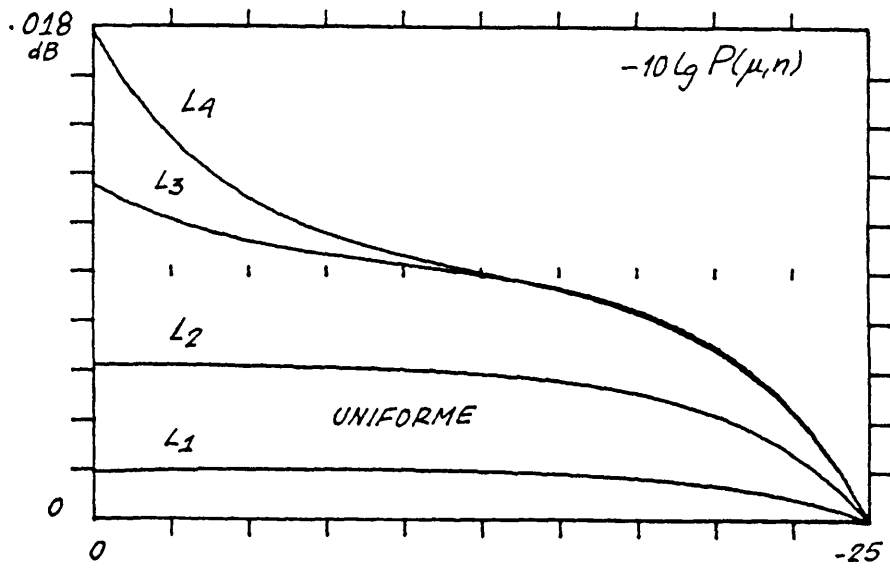
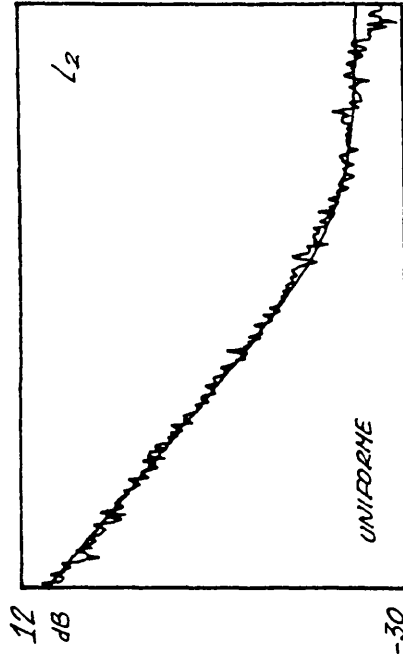
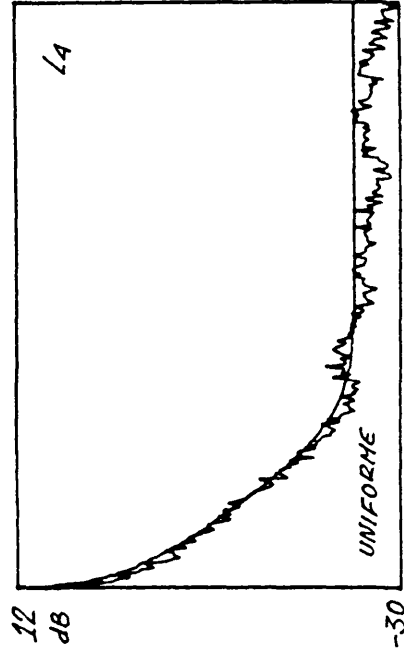


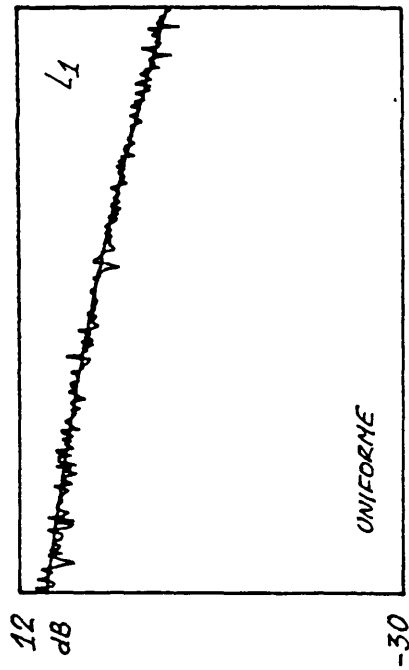
Figura F3.18



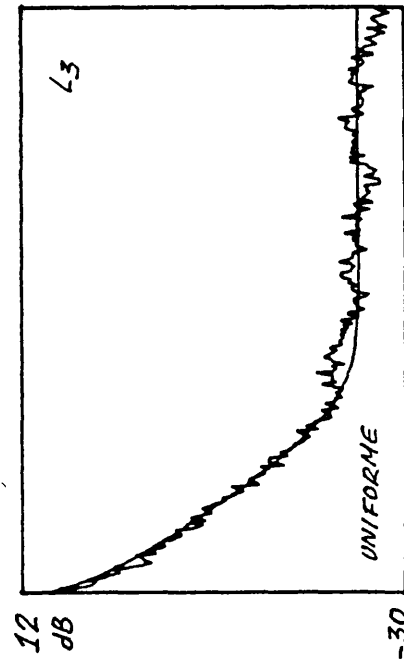
F3.19b



F3.19d



F3.19a



F3.19c

La Figura F3.20 presenta las velocidades teóricas de convergencia para ruido de planta Binario y las figuras F3.21a a F3.21c las convergencias teóricas y simuladas para dicho ruido. En este caso no se considera el algoritmo L1, ya que las 10000 muestras no son suficientes para hacer converger el error residual una cantidad apreciable. Los pasos de adaptación utilizados son:

$$\mu_2 = 1.6 \times 10^{-5} \quad \mu_3 = 2.1 \times 10^{-5} \quad \mu_4 = 2.3 \times 10^{-5}$$

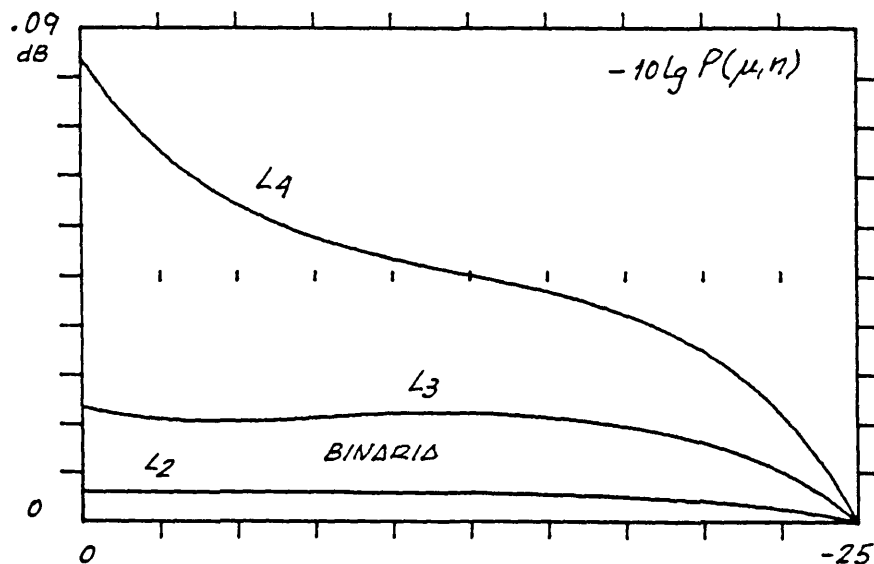
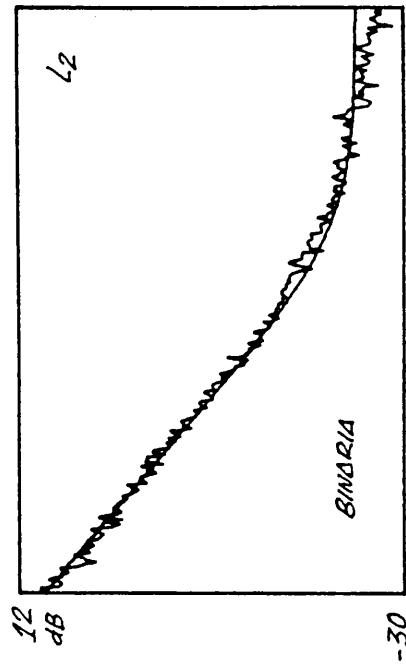
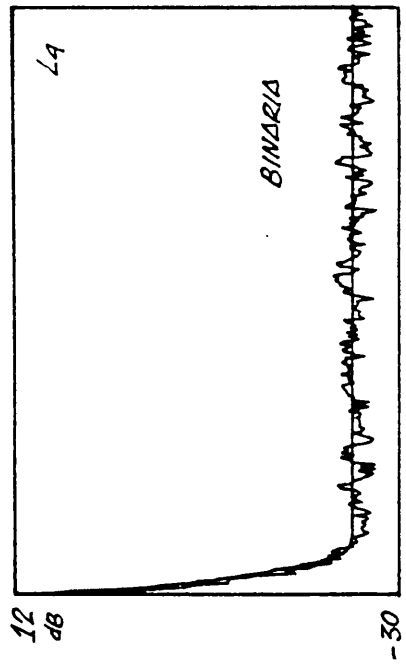


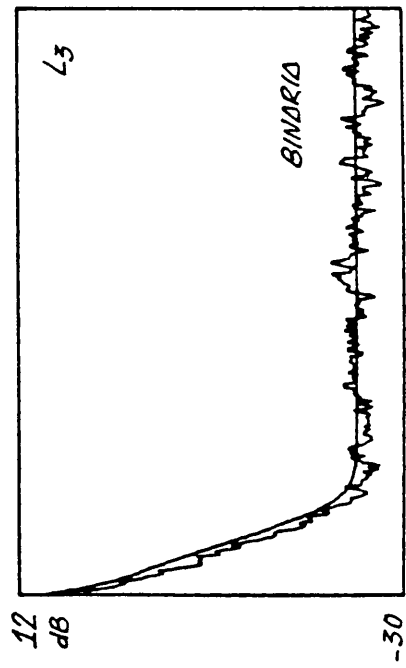
Figura F3.20



F3.21a



F3.21c



F3.21b

Los resultados recogidos en las gráficas anteriores, suscitan los siguientes comentarios generales:

a) Las curvas simuladas se ajustan bastante bien a las calculadas teóricamente.

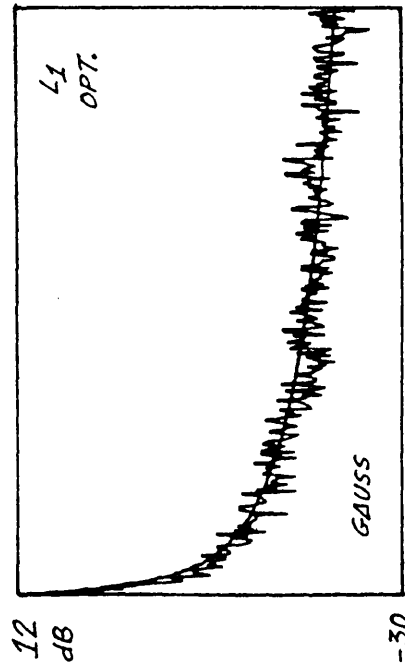
b) Se observa una mayor velocidad de convergencia cuando el orden del algoritmo es alto/bajo y la distribución estadística del ruido aditivo de planta posee colas cortas/largas.

III.6.3.- Convergencia óptima con ruido de planta

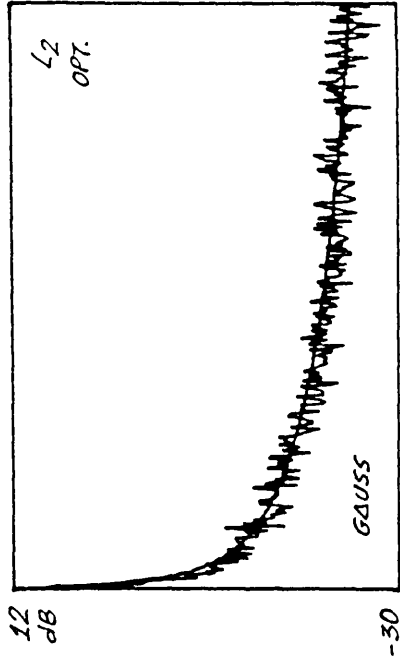
Se usa igualmente excitación Gaussiana blanca de media nula y potencia $\sigma_x^2=20$. En este caso se examina la convergencia de los 4 primeros algoritmos con un paso de adaptación óptimo variable, evaluado como $\mu_0(n)=S_1(s(n))/2S_2(s(n))$ e inicializado a partir de una primera estima de $s^2(0)$ ($\hat{s}^2(0)=\epsilon^2(0)$) obtenida con la ayuda de (iii.6.2). Todas las curvas simuladas provienen de promediar 4 realizaciones de 2000 muestras cada una y han sido suavizadas con una ventana móvil uniforme de 10 muestras. Obsérvese que al igual que en las simulaciones anteriores las curvas teóricas son seguidas por las simuladas, garantizando así la fiabilidad del análisis.

Las figuras de este grupo de simulaciones se resumen:

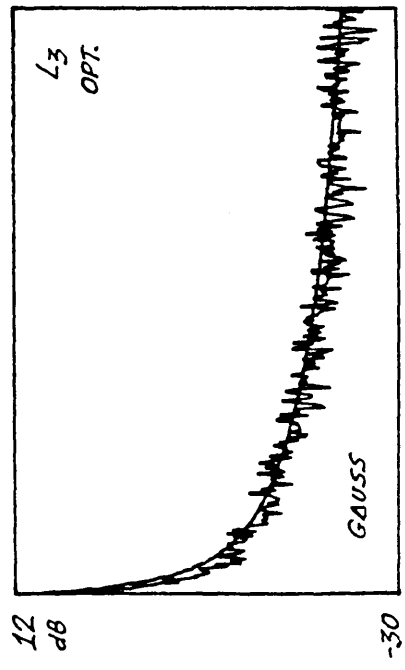
- Figuras F3.22a a F3.22d: Ruido Gaussiano
- Figuras F3.23a a F3.23d: Ruido Laplaciano
- Figuras F3.24a a F3.24d: Ruido Uniforme
- Figuras F3.25a a F3.25d: Ruido Binario



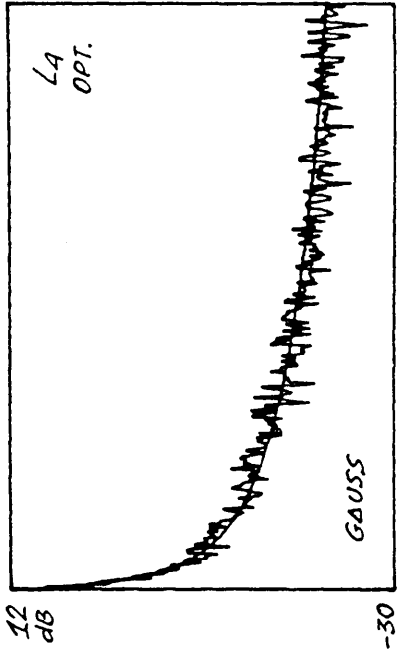
F3.22a



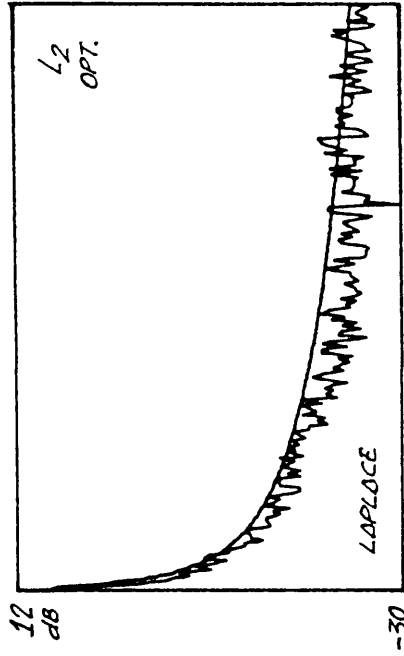
F3.22b



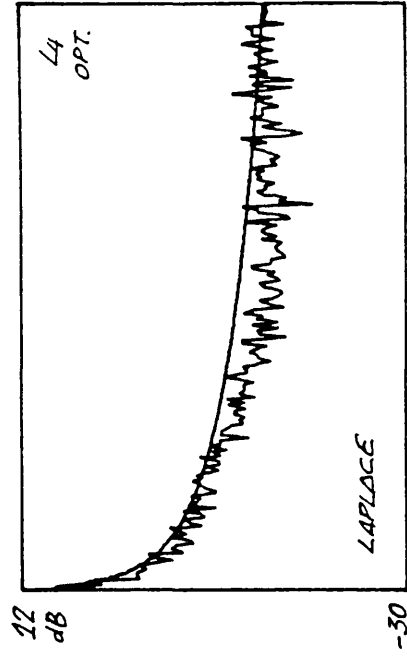
F3.22c



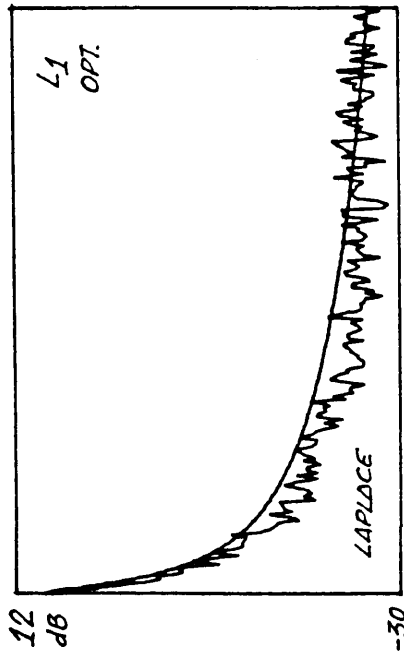
F3.22d



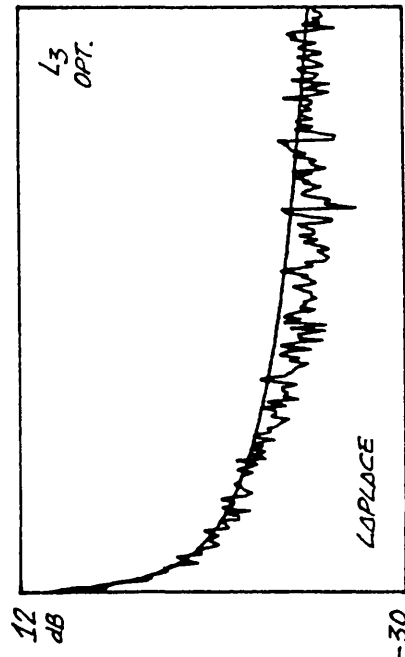
F3.23b



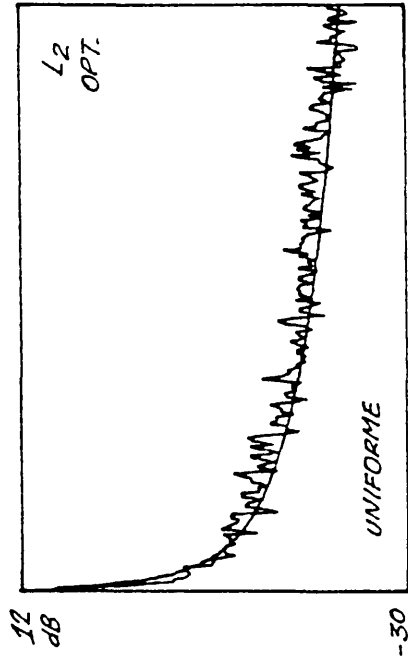
F3.23d



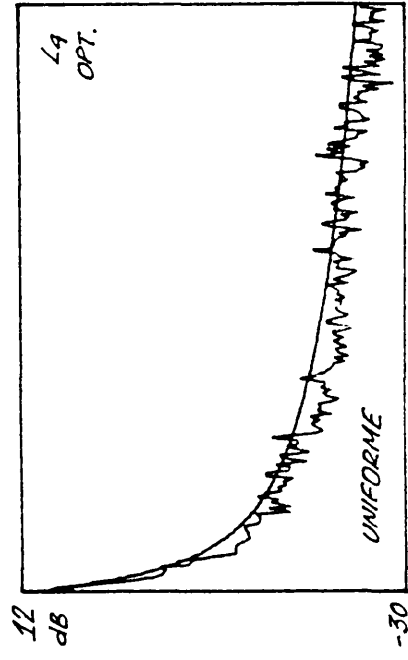
F3.23a



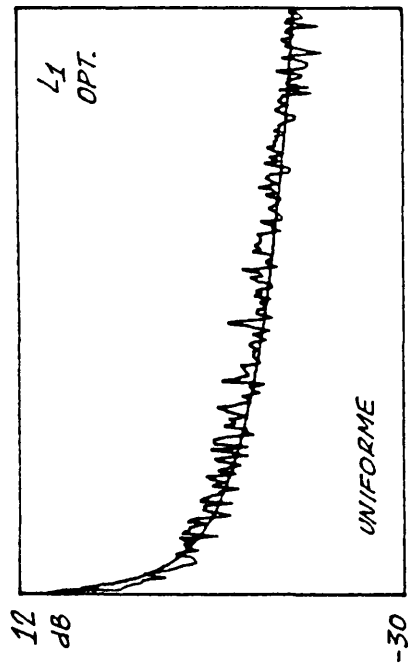
F3.23c



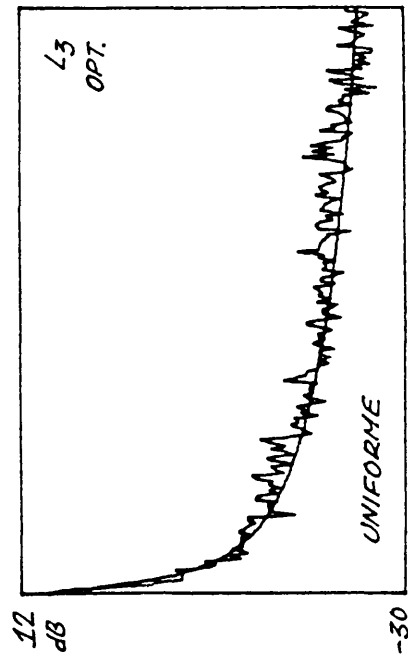
F3.24b



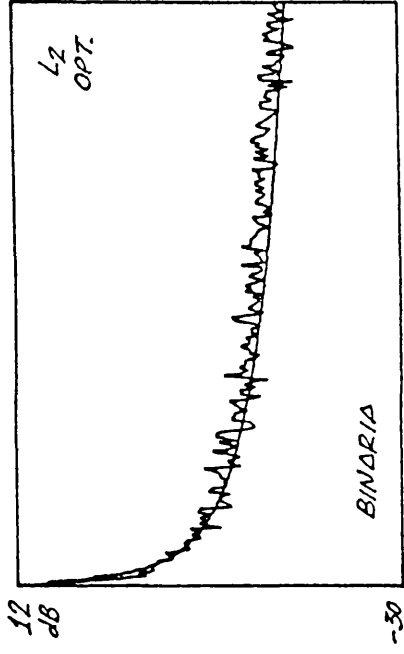
F3.24d



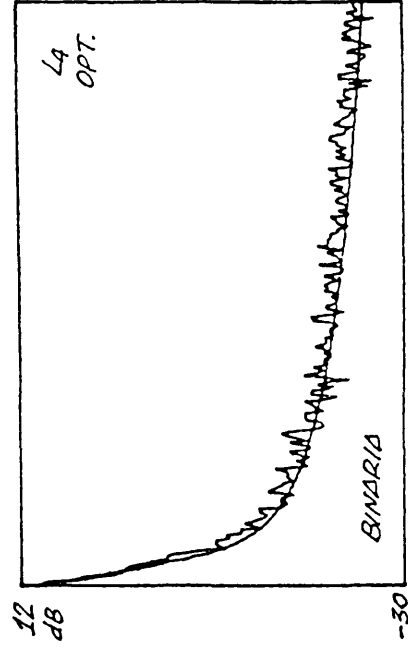
F3.24a



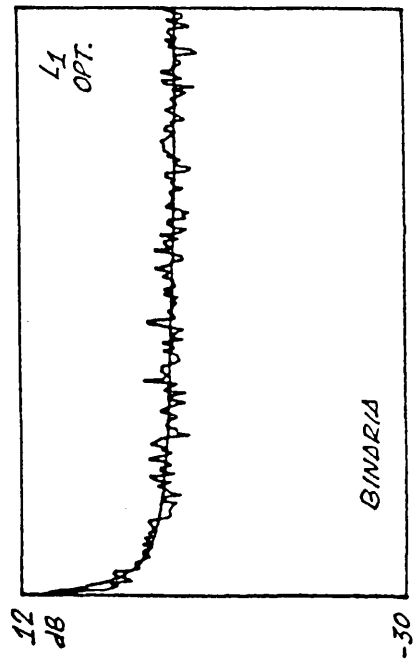
F3.24c



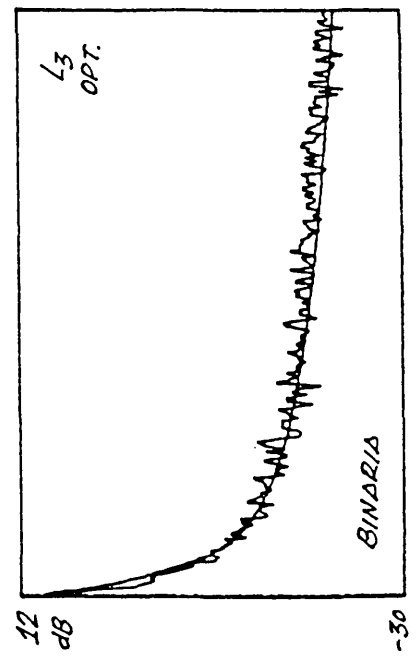
F3.25b



F3.25d



F3.25a



F3.25c

III.6.4.- Idéntica velocidad y diferente desajuste final

Se usa una excitación Gaussiana blanca de media nula y potencia $\sigma_x^2 = 3$. Como base de comparación se utiliza el mismo número medio de muestras para convergencia $n_{med} = 1600$ muestras y se elige un desajuste final para el algoritmo L2 de -25 dB. Se promedian 10 realizaciones de 10000 muestras con un suavizado de 200 con objeto de poder medir gráficamente los diferentes desajustes finales. No se presentan las curvas teóricas ya que en las simulaciones previas se ha observado su clara coincidencia con las simuladas. Y sólo se simulan las plantas de ruido Binario y Laplaciano, ya que para las otras dos (Uniforme y Gaussiano) las diferencias de sus desajustes finales no son excesivas y por tanto difícilmente observables en las gráficas. Contrástense estos resultados con los de la Tabla T3.3.

Las Figuras de este grupo se resumen:

- Figura 3.26 : Ruido Laplaciano
- Figura 3.27 : Ruido Binario

Antes de concluir con los Apartados relativos a simulación es conveniente resaltar que se han simulado otros casos con excitación no Gaussiana, obteniéndose resultados muy similares en comportamiento medio a los teóricos obtenidos. Este hecho implica que el análisis seguido es suficientemente aceptable incluso para predecir funcionamientos medios de los algoritmos bajo otras hipótesis de partida.

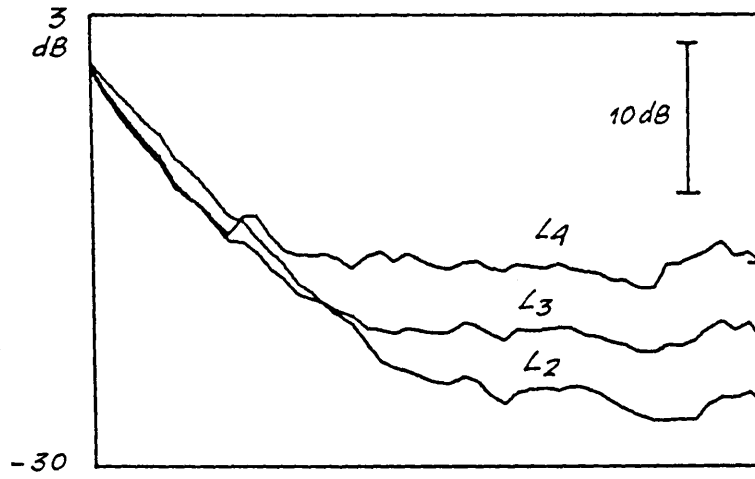


Figura F3.26

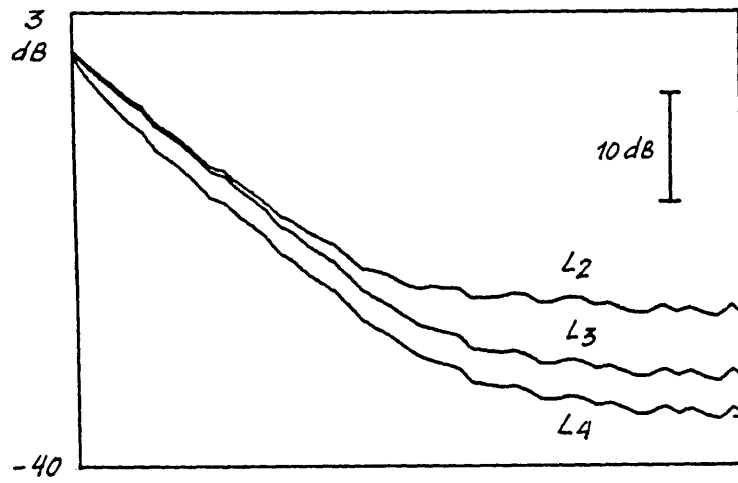


Figura F3.27

III.7.- CONCLUSIONES

Se ha estudiado una familia de algoritmos adaptativos que usan el gradiente instantáneo del momento absoluto k-ésimo del error total (residual más ruido de planta) para actualizar los coeficientes de un filtro transversal en una aplicación de identificación de planta.

Las hipótesis supuestas (H1 a H5) y el análisis de la evolución de la convergencia de la varianza del error residual son validadas por los resultados de simulaciones, ya que las curvas simuladas siguen de forma bastante ajustada a las teóricas.

A continuación se resumen las conclusiones más significativas en función de los resultados obtenidos.

a) En ausencia de ruido de planta:

- Si $k \geq 2$, la condición necesaria para la convergencia a cero de la varianza es el uso de una cota superior para μ , la cual depende de la varianza inicial del error residual, excepto para $k=2$ que es constante.
- Si $k=1$, la varianza del error nunca diverge y su valor final depende del valor elegido para μ . Si se permite el uso de un μ decreciente hacia cero se conseguiría una varianza final nula.
- Un μ constante origina una velocidad de convergencia constante para $k=2$ y variable (llega a saturarse) para $k \neq 2$. La convergencia asintótica del L2 es la más rápida de todos.

- El valor óptimo de μ es constante para $k=2$, una secuencia creciente no acotada para $k>2$ y decreciente para $k=1$. En este caso todos los algoritmos convergen a cero con una velocidad constante, siendo el algoritmo optimizado L2 es el mas rápido de todos.
- Como conclusión final, se recomienda la elección del algoritmo L2 en caso de ausencia de ruido de planta debido a que ofrece una convergencia constante y un μ óptimo constante.

b) Con ruido de planta:

- Con un μ constante, debidamente acotado, la varianza del error residual para todos los algoritmos converge hacia un valor no nulo, el cual depende del paso μ elegido.
- Cuando $k>2$, se ha de considerar además la varianza inicial del error con objeto de evitar posibles divergencias, ya que en estos casos la elegida para conseguir una varianza final deseada ha de satisfacer una condición adicional dependiente de la varianza inicial del error. Como guía práctica, para un μ seleccionado, el riesgo de divergencia aumenta si la varianza inicial aumenta.
- La velocidad de convergencia está relacionada con la distribución estadística del ruido de planta considerado excepto para el caso con $k=2$. Esta, además, varía en función de la razón local potencia error/potencia ruido. Para distribuciones con colas cortas/largas, los esquemas de mayor/menor orden se comportan mejor respectivamente. No obstante, si el ruido de planta es Gaussiano, el algoritmo L2 es el de mejores prestaciones.

- Como corolario del anterior, a igualdad de velocidad de convergencia, los algoritmos de mayor orden permiten un desajuste final más pequeño para ruidos de planta de cola corta y al contrario con los de cola larga.
- Con μ constante, la convergencia en varianza: tiene carácter log-lineal con errores grandes y medios para $k=2$, alcanza un máximo de convergencia para errores medios para $k=1$ y es máxima en los instantes iniciales de adaptación (errores grandes) saturándose conforme el error decrece.
- El uso de un μ optimizado de forma determinista hace crecer la velocidad de convergencia para cualquier algoritmo manteniendo el comportamiento relativo ante colas largas/cortas citado anteriormente cuando el ruido de planta es relevante. En este caso, la convergencia de la varianza tiende asintóticamente a cero.
- El costoso cálculo de un μ optimizado puede ser reducido considerando la posibilidad de aproximar la curva teórica del μ óptimo por tramos según el índice temporal n y un conocimiento estimado de la evolución del error cuadrático. Tales aproximaciones ofrecen mejores resultados que el considerar μ constantes.

Permanecen abiertas líneas de trabajo tales como el introducir un optimizado estocástico, el utilizar uno u otro algoritmo en función de las velocidades y desajustes locales o la posibilidad de utilizar funciones generales de costo (compuestas) en función de unos requisitos de optimización o robustez.

CAPITULO IV

IDENTIFICACION ADAPTATIVA DE PLANTA

EN EL DOMINIO DE LA FRECUENCIA

IV.1. INTRODUCCION DEL ALGORITMO

En este Capítulo se presenta, analiza y simula la misma familia de algoritmos tratada en el Capítulo anterior en el dominio (complejo) de la frecuencia. Se pretende establecer una formulación similar con objeto de predecir el comportamiento cuadrático medio de cualquier miembro de la familia [PAEZ-BORRALLO,1987].

Revisando la literatura de los algoritmos adaptativos definidos en este dominio, la primera versión del algoritmo LMS para señales complejas es, igualmente que en dominio temporal, debida a Widrow [WIDROW,1975a]. Dicha formulación es de gran aplicación en el dominio complejo de la frecuencia, ya que puede ser obtenida como resultado de aplicar tal algoritmo a cualquier componente arbitraria de las DFT's de las señales involucradas en un proceso de filtrado. Considerando además, que la DFT proporciona una serie de componentes aproximadamente ortogonales o incorreladas, los algoritmos correspondientes tienden a implementar procesos de adaptación independientes entre sí [ALVAREZ,1985]. Sin embargo, este algoritmo implica implícitamente el uso de la convolución circular entre la entrada al filtro y su respuesta impulsiva, mientras que la implementación temporal usa una convolución lineal. Este hecho supone que los resultados de ambas implementaciones no sean equivalentes y convergan hacia soluciones diferentes, pero sus prestaciones tanto de convergencia como de régimen permanente están fuertemente relacionados entre sí [REED,1981]. Aún con más detalle en el análisis, en [BERSHAD,1979] y [BERSHAD,1982] se estudia la convergencia de la matriz de autocorrelación de la DFT de los coeficientes de un filtro obteniendo resultados muy similares a los correspondientes en el dominio del tiempo.

En esta Tesis, el análisis en frecuencia está motivado fundamentalmente por dos razones: la separabilidad de la señal en componentes cuasi-ortogonales acompañado de la reducción en carga computacional que experimenta la implementación del clásico LMS debido al procesado en bloque de la DFT, [IDENTINO,1978],[FERRARA,1980] y completar el estudio iniciado en el Capítulo anterior analizando el algoritmo generalizado Lk en el dominio de la frecuencia con la presencia de un ruido de planta.

La configuración conceptual del filtro en el dominio de la frecuencia que se presenta en la Figura F4.1 es deducible de la del dominio del tiempo. La señal de entrada $x(n)$ y el ruido de planta $u(n)$ son acumulados hasta formar bloques de longitud N puntos. Entonces se procesan mediante una FFT de N puntos y, suponiendo $x(n), u(n)$ y la respuesta impulsiva del filtro reales, la mitad de los puntos complejos correspondientes a la FFT de la entrada son ponderados por cada uno de los coeficientes complejos del filtro a identificar. Ambos pares de números complejos se restan entre sí y se suman al correspondiente transformado del ruido de planta y el conjunto total final alimenta a un procesador de FFT inversas, el cual produce la salida corregida del sistema correspondiente al bloque de puntos l -ésimo con un retardo de N puntos.

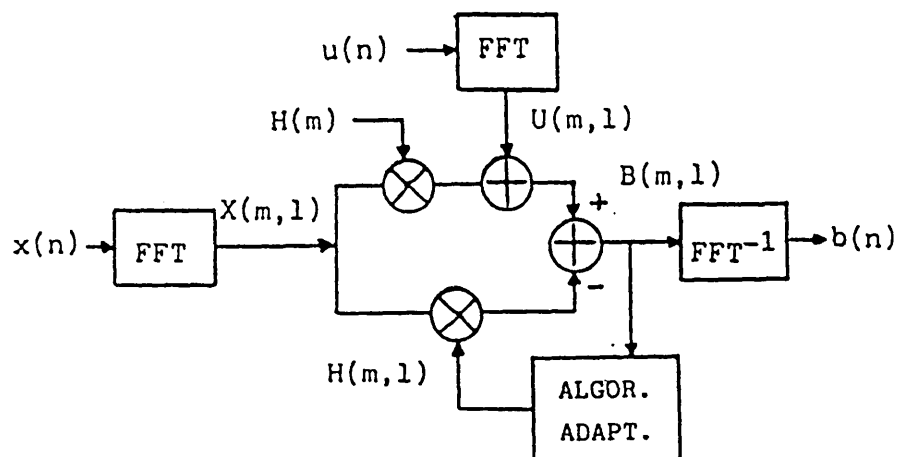


Figura F4.1

En la práctica sólo se tiene acceso a las salidas y entradas de la planta. Esto supone que la implementación efectiva difiera de la conceptual presentada anteriormente, en concreto esta sería tal como se presenta en la figura F4.2.

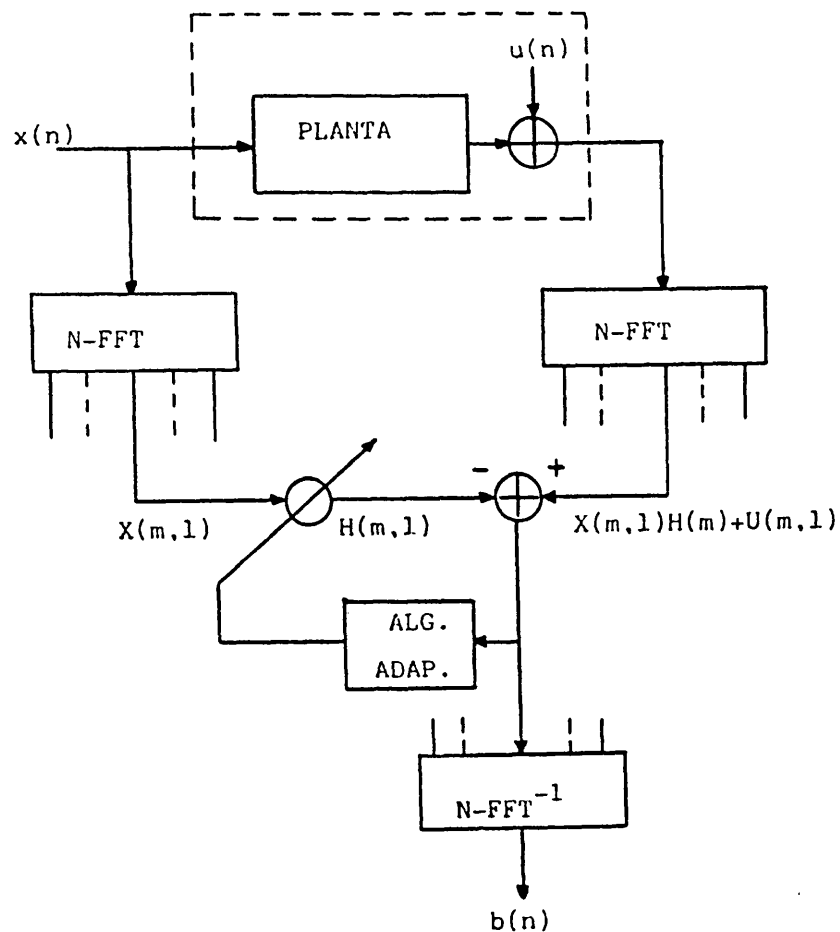


Figura F4.2

Se observa que cada coeficiente complejo es adaptado con independencia de los demás una vez cada N muestras, constituyendo el conjunto completo un banco de $N/2+1$ procesos adaptativos. Esto significa que el número de adaptaciones individuales requerido para obtener una salida se reduce con respecto al algoritmo temporal en un factor de N . Debido a esto, el valor del paso de adaptación debe de ser incrementado por un factor de N con objeto de mantener el mismo orden de magnitud en el error residual. Este mayor valor del paso debido a una cadencia menos frecuente de adaptación permite una reducción en los requisitos de resolución de los coeficientes complejos espectrales [IDENTINO,1978], y por tanto del número de bits para su representación, punto importante en el caso de implementación en aritmética de coma fija.

En este Capítulo se sigue un formato parecido al anterior, ya que se estudia el comportamiento de la convergencia del error cuadrático medio en todo su margen dinámico de variación, así como sus comportamientos asintóticos, estableciendo las condiciones necesarias para la convergencia de los algoritmos de gradiente estocástico derivados de la minimización de la potencia k del valor absoluto del error total. En este estudio, no obstante, se analiza tan sólo ruido de planta de tipo Gaussiano, ya que el efecto de convolución sobre los datos de la FFT hace que distribuciones originales, distintas a la Gaussiana en el dominio del tiempo, se comporten aproximadamente como tal en el dominio complejo de la frecuencia si los bloques son largos. De otra parte, todos aquellos análisis y demostraciones similares a los del Capítulo anterior serán solamente nombrados y referenciados con objeto de evitar una duplicidad en el desarrollo. Según esto, no se aborda el estudio relacionado con el paso de adaptación óptimo ya que su discusión es formalmente idéntica a la del Capítulo anterior.

Según la Figura F4.1 , para la raya espectral m ($0 \leq m \leq N/2$) se denominan:

- $X(m,l)$ la transformada de la señal de entrada en el bloque de datos l -ésimo.
- $U(m,l)$ la transformada del ruido de planta en el bloque l -ésimo.
- $H(m)$ y $H(m,l)$ las transformadas de los coeficientes de la planta y filtro adaptado en el bloque l -ésimo
- $E(m,l)$ el error residual transformado complejo en el bloque l -ésimo.

El algoritmo generalizado L_k trata de minimizar para cada raya espectral m y bloque l el siguiente estadístico objetivo:

$$f(H(m)) = E[|E(m,l)+U(m,l)|^k] \quad (\text{iv.1.1})$$

usando el gradiente del estimador instantáneo, esto es:

$$\hat{\nabla} E[|E(m,l)+U(m,l)|^k] = \nabla (|E(m,l)+U(m,l)|^k) \quad (\text{iv.1.2})$$

donde el error residual se define:

$$E(m,l) = (H(m)-H(m,l))X(m,l) \quad (\text{iv.1.3})$$

Llamando por comodidad en la notación:

$$B(m,l) = E(m,l)+U(m,l) \quad (\text{iv.1.4})$$

al error total en frecuencia y

$$A_r(.) = \text{Re} (A(.)) \quad (\text{iv.1.5a})$$

$$A_i(.) = \text{Im} (A(.)) \quad (\text{iv.1.5b})$$

a las partes real e imaginaria del número complejo $A(.)$, el gradiente estocástico queda como:

$$\nabla |B(m,l)|^k = \nabla (B_r^2(m,l) + B_i^2(m,l))^{k/2} = \frac{k/2 \cdot d[.]}{dH_r(m)} + j \frac{d[.]}{dH_i(m)} \quad (\text{iv.1.6})$$

donde las parciales de las partes real e imaginaria son respectivamente:

$$\begin{aligned} \frac{d[\cdot]}{dH_r(m)} &= -k |B(m,1)|^{k-2} (B_r(m,1)X_r(m,1) + B_i(m,1)X_i(m,1)) = \\ &= -k |B(m,1)|^{k-2} \operatorname{Re}(B(m,1)X^*(m,1)) \end{aligned} \quad (\text{iv.1.7a})$$

$$\begin{aligned} \frac{d[\cdot]}{dH_i(m)} &= k |B(m,1)|^{k-2} (B_r(m,1)X_i(m,1) - B_i(m,1)X_r(m,1)) \\ &= -k |B(m,1)|^{k-2} \operatorname{Im}(B(m,1)X^*(m,1)) \end{aligned} \quad (\text{iv.1.7b})$$

(el símbolo $*$ denota conjugación compleja); por lo que agrupando ambas expresiones según (iv.1.6)

$$\underline{y} |B(m,1)|^k = -k |B(m,1)|^{k-2} B(m,1)X^*(m,1) \quad (\text{iv.1.8})$$

y por tanto el algoritmo de adaptación para la raya m y el bloque l -ésimo será:

$$H(m, l+1) = H(m, l) + k\mu_m |B(m,1)|^{k-2} B(m,1)X^*(m,1) \quad (\text{iv.1.9})$$

en general complejo salvo para la raya espectral $m=0$ y donde μ_m es el paso de adaptación para la raya espectral m .

En la expresión anterior se observa la similitud con el algoritmo definido en el dominio del tiempo. Para $k=2$, el algoritmo anterior coincide con el propuesto por Widrow en [WIDROW,1975a] como versión compleja del LMS.

IV.2.- ANALISIS DEL ERROR RESIDUAL CUADRATICO MEDIO

Previamente al análisis estadístico del algoritmo, y con objeto de facilitar su formulación, se establecen las hipótesis de trabajo para los procesos involucrados en el sistema. Estas son:

- H1) $x(n)$ y $u(n)$ son procesos conjuntamente estacionarios.
- H2) $x(n)$ y $u(n)$ son procesos Gaussianos, blancos de medias nulas y varianzas σ_x^2 y σ_u^2 respectivamente.
- H3) $x(n)$ y $u(n)$ son independientes entre si.
- H4) El error residual en frecuencia $E(m,1)$ es Gaussiano circular complejo de medias nulas.
- H5) Independencia conjunta de las partes reales e imaginarias de $H(m,1), E(m,1), X(m,1)$ y $U(m,1)$.

Al igual que en el Capítulo anterior, la hipótesis H5 es falsa, pero se acepta bajo el supuesto de un estado avanzado de convergencia y un paso de adaptación pequeño. Esto supone que $E(m,1)$ esté prácticamente incorrelada con $X(m,1)$ y este grado de incorrelación se acepta como independencia.

Como consecuencia de la hipótesis H2, y según el Apéndice A.11, las v.a. $X(m,1)$ y $U(m,1)$ resultan ser Gaussianas circulares complejas para $m \neq 0$, es decir:

$$E[X(m,1)] = 0 \quad (\text{iv.2.1a})$$

$$\text{Var}[X_r(m,1)] = \text{Var}[X_i(m,1)] = N \sigma_x^2 / 2 \quad (\text{iv.2.1b})$$

$$E[X_r(m,1)X_i(m,1)] = 0 \quad (\text{iv.2.1c})$$

y equivalentemente para $U(m,1)$. Y para $m=0$ son v. a. Gaussianas blancas de media nula y varianzas $N\sigma_x^2$ y $N\sigma_u^2$ respectivamente.

IV.2.1.- Expresión recurrente generalizada para la varianza del error residual

Definiendo la diferencia de cada raya espectral del filtro con su versión adaptada como:

$$V(m,1) = H(m) - H(m,1) \quad (\text{iv.2.1.1})$$

se obtiene segun (iv.1.9) para cada componente espectral m la siguiente recurrencia:

$$V(m,1+1) = V(m,1) - k \mu_m |B(m,1)|^{k-2} B(m,1) X^*(m,1) \quad (\text{iv.2.1.2})$$

si se conjuga la expresión anterior y se multiplica miembro a miembro por la expresión sin conjugar, se obtiene:

$$\begin{aligned} |V(m,1+1)|^2 &= |V(m,1)|^2 - 2k \mu_m |B(m,1)|^{k-2} \text{Re}(B(m,1)E^*(m,1)) + \\ &+ k^2 \mu_m^2 |B(m,1)|^{2k-2} |X(m,1)|^2 \end{aligned} \quad (\text{iv.2.1.3})$$

utilizando aquí la hipótesis H5 y observando que $H(m,1)$ depende de los bloques anteriores $1-1, 1-2, \dots$, se puede escribir:

$$E[|E(m,1)|^2] = E[|V(m,1)|^2] E[|X(m,1)|^2] \quad (\text{iv.2.1.4})$$

pero segun el Apendice A.11:

$$E[|X(m,1)|^2] = N\sigma_x^2 \quad (\text{iv.2.1.5})$$

y debido a la hipótesis H4, la varianza de $E(m,1)$ coincidirá con su valor cuadrático medio, esto es:

$$\sigma_E^2(m,1) = \text{Var}[E_r(m,1)] + \text{Var}[E_i(m,1)] = E[|E(m,1)|^2] \quad (\text{iv.2.1.6})$$

por tanto, tomando esperanzas en (iv.2.1.3) y multiplicando ambos miembros por $N\sigma_x^2$ se obtiene como expresión recurrente para la varianza del error residual en frecuencia:

$$\sigma_E^2(m,1+1) = \sigma_E^2(m,1)(1 - \mu_m S_1(m,1) + \mu_m^2 S_2(m,1)) \quad (\text{iv.2.1.7})$$

expresión similar a la obtenida en el dominio del tiempo estando en el caso actual $S_1(m,1)$ y $S_2(m,1)$ definidas como:

$$S_1(m,1) = 2kE \left[|B(m,1)|^{k-2} \operatorname{Re}(B(m,1)E^*(m,1)) \right] N \sigma_x^2 / \sigma_E^2(m,1) \quad (\text{iv.2.1.8})$$

$$S_2(m,1) = k^2 E \left[|B(m,1)|^{2k-2} \right] N \sigma_x^4 / \sigma_E^2(m,1) \quad (\text{iv.2.1.9})$$

Igual que en el Capítulo anterior, se define:

$$P(m,1) = 1 - \mu_m S_1(m,1) + \mu_m^2 S_2(m,1) \quad (\text{iv.2.1.10})$$

polinomio cuadrático en μ_m sobre el que se puede estudiar la convergencia en varianza del algoritmo complejo.

Con objeto de simplificar el seguimiento de las expresiones de $S_1(m,1)$ y $S_2(m,1)$ se divide el análisis en dos: los casos de ausencia y presencia de ruido de planta.

iv.2.2.- Caso 1: Planta sin ruido

En este caso $u(n)=0$ y por consiguiente lo será su DFT $U(m,1)$. Con esta simplificación, las expresiones (iv.2.1.8) y (iv.2.1.9) resultan:

$$S_1(m,1) = 2kE \left[|E(m,1)|^k \right] N \sigma_x^2 / \sigma_E^2(m,1) \quad (\text{iv.2.2.1})$$

$$S_2(m,1) = k^2 E \left[|E(m,1)|^{2k-2} \right] N^2 \sigma_x^4 / \sigma_E^2(m,1) \quad (\text{iv.2.2.2})$$

si se observa que, para $m \neq 0$

$$|E(m,1)| = (E_r(m,1) + E_i(m,1))^{\frac{1}{2}} \quad (\text{iv.2.2.3})$$

es una v.a. Rayleigh, y para $m=0$

$$E(m,1) = E_r(m,1) \quad (\text{iv.2.2.4})$$

es una v.a. Gaussiana, según el Apéndice A.11 se obtiene:

$m \neq 0$

$$S_1(m, l) = \begin{cases} T_a(m, l) & ; k \text{ par} & (\text{iv.2.2.5a}) \\ \sqrt{\pi/2} T_a(m, l) & ; k \text{ impar} & (\text{iv.2.2.5b}) \end{cases}$$

con

$$T_a(m, l) = 2kN k!! \sigma_x^2 \sigma_E^{k-2}(m, l) / 2^{k/2} \quad (\text{iv.2.2.6})$$

$$S_2(m, l) = (kN \sigma_x^2)^2 (2k-2)!! \sigma_E^{2k-4}(m, l) / 2^{k-1} \quad (\text{iv.2.2.7})$$

$m = 0$

$$S_1(m, l) = \begin{cases} T_b(m, l) & ; k \text{ par} & (\text{iv.2.2.8a}) \\ \sqrt{2/\pi} T_b(m, l) & ; k \text{ impar} & (\text{iv.2.2.8b}) \end{cases}$$

con

$$T_b(m, l) = 2kN (k-1)!! \sigma_x^2 \sigma_E^{k-2}(m, l) \quad (\text{iv.2.2.9})$$

$$S_2(m, l) = (kN \sigma_x^2)^2 (2k-3)!! \sigma_E^{2k-4}(m, l) \quad (\text{iv.2.2.10})$$

Estas expresiones son también válidas en el caso en que exista ruido de planta y su DFT $U(m, l)$ sea despreciable frente al error residual $E(m, l)$ en la raya espectral m .

iv.2.3.- Caso 2: Planta ruidosa

En este caso las expresiones $S_1(m, l)$ y $S_2(m, l)$ no admiten reducción posible y son directamente las expresadas en (iv.2.1.8) y (iv.2.1.9). Se ha supuesto un ruido de planta Gaussiano independiente, y según el Apéndice A.11, esto supone un ruido en frecuencia Gaussiano circular complejo. Esta limitación, no restringe la validez del análisis, ya que ruidos en el dominio temporal diferentes al Gaussiano se comportarían de forma

aproximadamente Gaussiana en la frecuencia si la longitud de la DFT utilizada fuera suficientemente alta.

En este caso el análisis de las esperanzas matemáticas se realiza en los Apéndices A.12 y A.13 obteniendo las correspondientes expresiones generales. En el primero de ellos se evalúa la expresión:

$$Q_2 = E[|E+U|^{2k-2}] \quad (\text{iv.2.3.1})$$

que, prescindiendo de los índices, es la componente principal de $S_2(m,1)$; y en el segundo de los Apéndices se evalúa:

$$Q_1 = E[|E+U|^{k-2} (|E|^2 + \text{Re}(EU^*))] \quad (\text{iv.2.3.2})$$

que es igualmente la componente principal de $S_1(m,1)$.

Debido a que las expresiones generalizadas obtenidas en los Apéndices anteriores no son cerradas y dada la enorme complejidad que supondría el cálculo de las expresiones Q_1 y Q_2 para cada uno de los algoritmos de la familia, se ha procedido, al igual que en el Capítulo anterior, al análisis de los cuatro primeros algoritmos (L1 a L4). Dicho análisis se lleva a cabo en los Apéndices A.14 y A.15 y sus resultados se resumen en la Tablas T4.1a y T4.1b. Con objeto de simplificar la notación se ha llamado $s^2 = \sigma_E^2(m,1)/N\sigma_U^2$ a la varianza del error residual normalizada por la potencia del ruido de planta en frecuencia.

k Componente espectral $n = 0$

$S_1(\cdot)$	1	$(2\sqrt{2/\pi} N\sigma_x^2/\sigma_u) / \sqrt{s^2+1}$
	2	$4N\sigma_x^2$
	3	$6\sqrt{2/\pi} N\sqrt{N}\sigma_x^2\sigma_u\sqrt{s^2+1}(1+s^2/(s^2+1)+1/(s^2+1)^2)$
	4	$24N^2\sigma_x^2\sigma_u^2(s^2+1)$
$S_2(\cdot)$	1	$N\sigma_x^4/\sigma_u^2 s^2$
	2	$4N\sigma_x^4 (s^2+1) / s^2$
	3	$27(N\sqrt{N}\sigma_x^2\sigma_u^2(s^2+1)/s)^2$
	4	$240N^4\sigma_x^4\sigma_u^2(s^2+1)^3 / s$

Tabla I4.1a

k Componente espectral $\mu > 0$

$S_1(\cdot)$	1	$(\sqrt{\pi}\sqrt{N}\sigma_x^2/\sigma_u)/\sqrt{s^2+1}$
	2	$4N\sigma_x^2$
	3	$(9\sqrt{\pi}/2)N\sqrt{N}\sigma_x^2\sigma_u\sqrt{s^2+1}$
	4	$16N^2\sigma_x^2\sigma_u^2(s^2+1)$
$S_2(\cdot)$	1	$N\sigma_x^4/\sigma_u^2s^2$
	2	$4N\sigma_x^4(s^2+1)/s^2$
	3	$18(N\sqrt{N}\sigma_x^2\sigma_u^2(s^2+1)/s)^2$
	4	$96N^4\sigma_x^4\sigma_u^4(s^2+1)^3/s^2$

Tabla I4.1b

IV.3.- CONDICIONES NECESARIAS PARA LA CONVERGENCIA DEL ERROR CUADRATICO MEDIO

La convergencia en varianza del error estará garantizada siempre y cuando se verifique la desigualdad $\sigma_E^2(m, l+1) < \sigma_E^2(m, l)$ para cualquier pareja (m, l) . Este hecho condiciona al polinomio cuadrático $P(m, l)$, definido en (iv.2.1.10), a cumplir el requisito ya exigido igualmente a su homónimo en el dominio del tiempo, esto es: ser menor que la unidad, permitiendo así el carácter monótono decreciente de la secuencia $\sigma_E^2(m, l)$ con el índice temporal l .

Las consecuencias derivadas del cumplimiento de este requisito son las que en la práctica se utilizan como condiciones necesarias para la convergencia. Antes de forzar la condición anterior es conveniente verificar el carácter no negativo de $P(\cdot)$, ya que la varianza $\sigma_E^2(m, l)$ no puede ser negativa.

IV.3.1.- Estudio del polinomio $P(m, l)$

Al igual que en el Capítulo anterior se analiza el carácter no negativo del polinomio cuadrático estudiando el carácter no positivo de su discriminante y verificando además que en al menos en un punto $P(m, l) > 0$. Esto último se cumple para $\mu_m = 0$ y con respecto al discriminante

$$D = S_1^2(m, l) - 4S_2(m, l) \leq 0 \quad (\text{iv.3.1.1})$$

o de forma equivalente:

$$B = S_1^2(m, l) / 4S_2(m, l) \leq 1 \quad (\text{iv.3.1.2})$$

de donde, sustituyendo las expresiones generales (iv.2.1.8) y (iv.2.1.9) se obtiene (prescindiendo de los índices (m, l)):

$$= \frac{E^2[|B|^{k-2} \operatorname{Re}(BE^*)]}{E[|E|^2]E[|B|^{2k-2}]} \quad (\text{iv.3.1.3})$$

donde se ha usado $E[|E|^2]$ por $\sigma_E^2(m,1)$.

Teniendo presente que

$$\operatorname{Re}(BE^*) = |BE| \cos(\operatorname{tg}^{-1}(\operatorname{Im}(BE^*)/\operatorname{Re}(BE^*))) \leq |BE| \quad (\text{iv.3.1.4})$$

$$m=0,1,\dots,N/2$$

se obtiene, utilizando $|BE| = |B| |E|$ la siguiente acotación de β :

$$\beta \leq \frac{E^2[|B|^{k-1} |E|]}{E[|E|^2]E[|B|^{2k-2}]} \quad (\text{iv.3.1.5})$$

y considerando la desigualdad de Cauchy-Schwartz ($E^2(XY) \leq E(X^2)E(Y^2)$)

en el numerador de la expresión anterior, se obtiene $\beta \leq 1$ cumpliéndose por tanto el requisito de polinomio no negativo independientemente de los índices $(m,1)$.

El requisito de monotonía $P(m,1) < 1$ es según (iv.2.1.10) equivalente a la siguiente acotación superior de μ_m :

$$(0 <) \mu_m < \min_1 S_1(m,1)/S_2(m,1) \quad (\text{iv.3.1.6})$$

válida dentro del marco de actuación del índice temporal l y para cualquier raya espectral m . Igualmente se define en este Capítulo la razón

$$C_k(m,1) = S_1(m,1)/S_2(m,1) \quad (\text{iv.3.1.7})$$

como la cota superior local del paso de adaptación μ_m necesaria para la convergencia en varianza.

Para abordar el estudio del mínimo de esta razón se desdobra el estudio en los casos de ausencia y presencia de ruido de planta, ya que las expresiones correspondientes de $S_1(m,1)$ y $S_2(m,1)$ son bastantes diferentes.

En lo relativo al caso de presencia de ruido de planta tan solo se analizaran los 4 primeros algoritmos (L1 a L4), ya que el estudio generalizado es súmamente complicado debido a que no se poseen expresiones cerradas de $S_1(m,1)$ (Apendice A.13).

IV.3.2.- Ausencia de ruido de planta

La coincidencia formal de las expresiones $S_1(m,1)$ y $S_2(m,1)$ con sus homónimas en el dominio del tiempo permite simplificar este Apartado, ya que el análisis y las conclusiones derivadas de él son las mismas. Como resumen, se citan las cotas absolutas en función del orden k del algoritmo para cualquier raya espectral m :

$$\mu_m < C_1(m, \infty) \quad (k=1) \quad (\text{iv.3.2.1a})$$

$$\mu_m < C_2(m,1) = 1/N\sigma_x^2 \quad (k=2) \quad (\text{iv.3.2.1b})$$

$$\mu_m < C_k(m,0) \quad (k>2) \quad (\text{iv.3.2.1c})$$

donde se observa nuevamente que para $k=1$, la cota depende del valor final de convergencia elegido; para $k=2$ es independiente del instante temporal l ; y para $k>2$ la cota está relacionada con la varianza inicial de error residual.

En caso de ausencia de ruido de planta, si el paso de adaptación elegido está convenientemente acotado la varianza del error converge hacia cero salvo en el caso L1 (ver el Apartado III.3.2 del Capítulo anterior para una discusión más detallada).

IV.3.3.- Presencia de ruido de planta

En el caso de existencia de un ruido aditivo de planta, el error residual de identificación no se anula totalmente para valores finitos del paso de adaptación. Esta afirmación se verifica de la siguiente manera:

Supóngase que $\sigma_E(m,1) \rightarrow 0$, por tanto, se cumple $E(m,1) \rightarrow 0$.

Tomando límites en (iv.2.1.7):

$$\lim \sigma_E^2(m,1+1) = \lim \sigma_E^2(m,1)(1 - \mu_m S_1(m,1) + \mu_m^2 S_2(m,1)) \quad (\text{iv.3.3.1})$$

$$\sigma_E(m,1) \rightarrow 0 \quad \sigma_E(m,1) \rightarrow 0$$

considerando (iv.2.1.8) y (iv.2.1.9) y tomando límites, se tiene:

$$\lim \sigma_E^2(m,1) S_1(m,1) \rightarrow 0 \quad (\text{iv.3.3.2})$$

$$\sigma_E(m,1) \rightarrow 0$$

$$\lim \sigma_E^2(m,1) S_2(m,1) \rightarrow k^2 N^2 \sigma_x^4 E[|U(m,1)|^{2k-2}] \quad (\text{iv.3.2.3})$$

$$\sigma_E(m,1) \rightarrow 0$$

por lo que (iv.3.3.1) tendría un límite no nulo, no siendo compatible con la hipótesis de partida a menos que el paso de adaptación fuera nulo. Con esto, por tanto, se concluye que la varianza final no puede ser nula cuando existe ruido aditivo de planta.

Iguálmente que en el dominio del tiempo tan solo se analizan los 4 primeros algoritmos de la familia. Para el cálculo de sus cotas locales $C_K(m,1)$ se utilizan las expresiones de $S_1(m,1)$ y $S_2(m,1)$ de la Tabla I4.1

En las Figuras F4.3, F4.4, F4.5 y F4.6 se presentan las curvas correspondientes a $C_K(m,1)$ (para $m=0$ y $m \neq 0$) observándose lo siguiente:

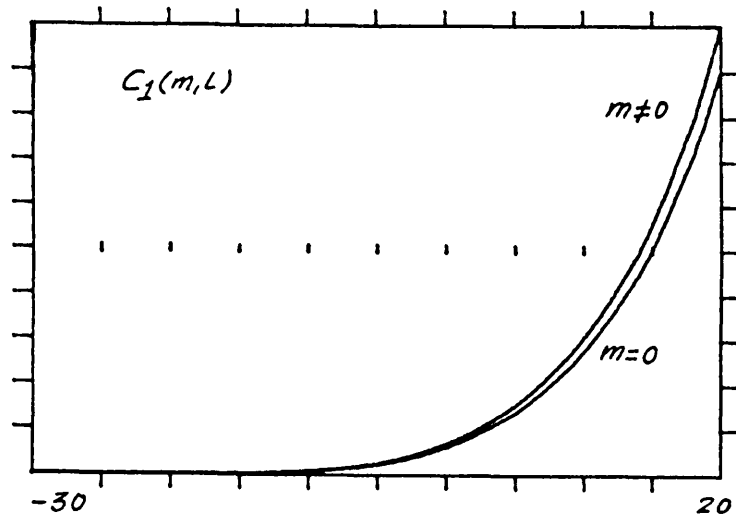


Figura F4.3

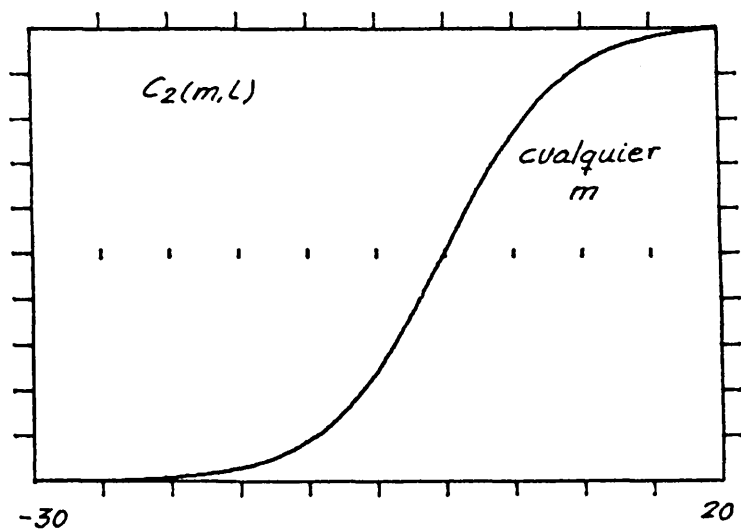


Figura F4.4

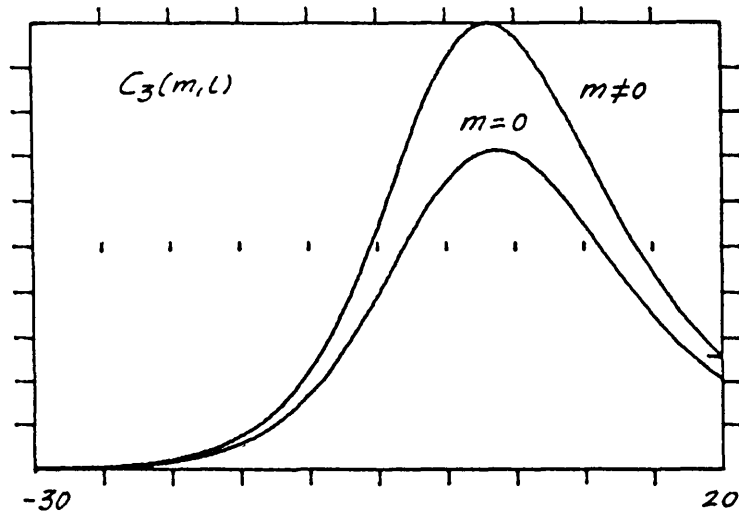


Figura F4.5

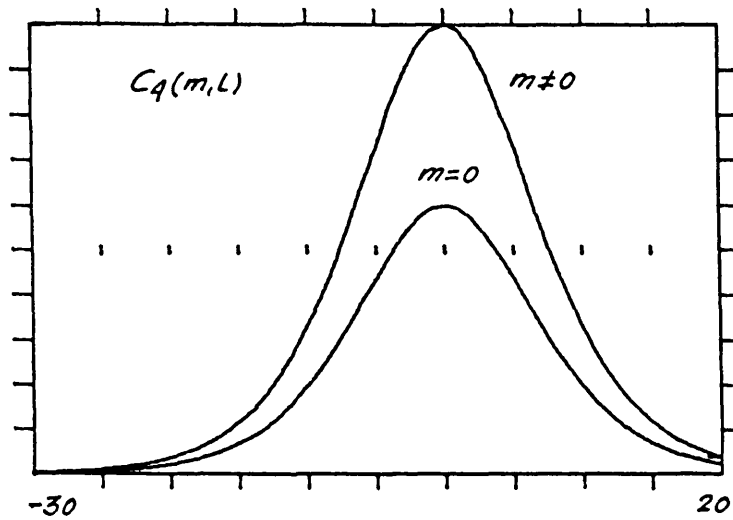


Figura F4.6

- Para $k=1$ y $k=2$ existe monotonía con s al igual que en el dominio temporal. Esto significa que los mínimos absolutos se encontrarán en las ordenadas correspondientes al valor final de convergencia, esto es, para $s(\infty)$.

- Para $k>2$ existe un máximo, por lo que al igual que en el dominio temporal, hay que considerar las ordenadas correspondientes a los valores extremos $s(0)$ y $s(\infty)$ de tal forma que se cumpla la condición $\min C_k(m,1)=C_k(m,\infty)<C_k(m,0)$.

IV.4.- SIMULACION DE LOS ALGORITMOS

En este Apartado se verifica, por medio de simulación, la evolución de la convergencia media en varianza descrita analíticamente en los Apartados anteriores.

En este caso, a diferencia con el dominio temporal, se presenta un único paquete de simulaciones. Este examina el comportamiento dinámico del error cuadrático medio en presencia de ruido de planta, habiendo elegido los pasos de adaptación adecuados para lograr el mismo desajuste final para cada raya espectral m en cada uno de los 4 primeros algoritmos.

Las condiciones utilizadas para la simulación son:

- Excitación Gaussiana, blanca, de media cero y potencia $\sigma_x^2=1$
- Ruido de planta Gaussiano, blanco, de media nula y potencia $\sigma_u^2=1$
- Longitud de la planta $N=32$

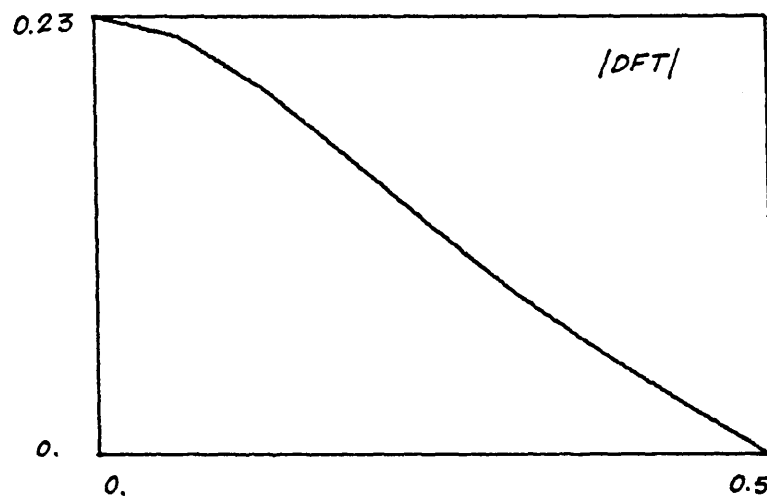


Figura F4.7

- Vector de coeficientes para la planta en el dominio temporal:

$$h(i) = \begin{cases} h(33-i) & ; i=1,16 \\ \exp(-2(i-17)) & ; i=17,32 \end{cases} \quad (\text{iv.4.1})$$

cuya DFT se presenta en la figura F4.7.

- Condición inicial para los coeficientes en frecuencia del filtro adaptativo.

$$H_r(m,0) = 0 \quad ; m=0, N/2 \quad (\text{iv.4.2})$$

$$H_i(m,0) = 0$$

IV.4.1.- Convergencia con ruido de planta

Se desea un desajuste final constante para cada raya espectral igual a -16 dB para todos los ensayos, con objeto de poder comparar las curvas de convergencia de los diferentes algoritmos entre sí.

Para ello se ha calculado el paso de adaptación como $\mu = C(m, \infty)$ con $s^2(\infty) = -16$ dB. En los casos con $k=3,4$ se cumple además $C(m, \infty) < C(m, 0)$, condición necesaria para evitar la divergencia. Todas las curvas simuladas provienen de promediar 10 realizaciones de 4800 muestras (150 bloques de 32 muestras cada uno) y el resultado ha sido suavizado con una ventana móvil uniforme de 6 bloques. Los pasos de adaptación utilizados, calculados según las recomendaciones del Apartado IV.3, para conseguir un desajuste final de -16 dB en cada raya espectral son:

k=1

$$\mu_m = \begin{cases} 6.7 \times 10^{-3} & ; m=0 \\ 7.8 \times 10^{-3} & ; m>0 \end{cases}$$

k=2

$$\mu_m = 7.7 \times 10^{-4} \quad ; \text{cualquier } m$$

k=3

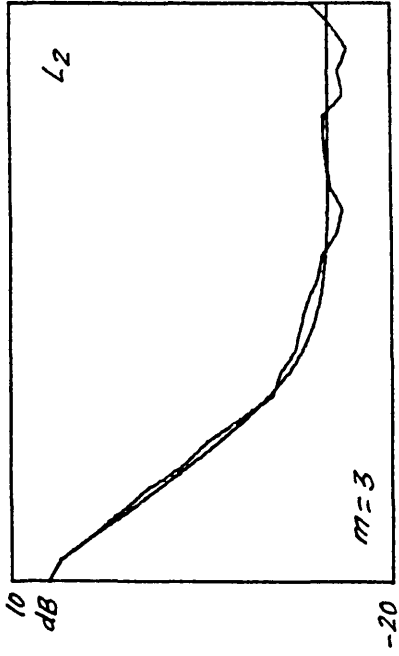
$$\mu_m = \begin{cases} 4.7 \times 10^{-5} & ; m=0 \\ 5.9 \times 10^{-5} & ; m>0 \end{cases}$$

k=4

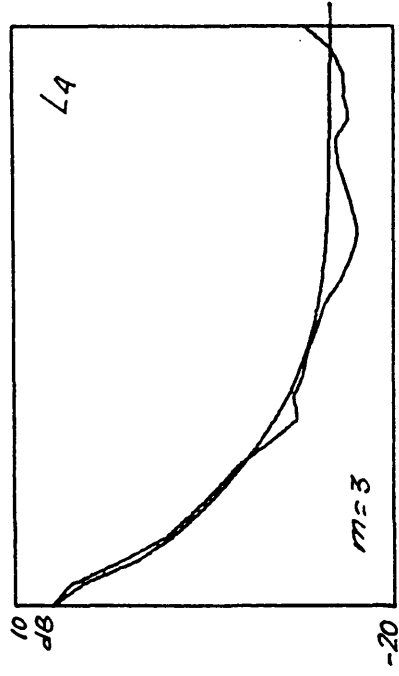
$$\mu_m = \begin{cases} 2.3 \times 10^{-6} & ; m=0 \\ 3.9 \times 10^{-6} & ; m>0 \end{cases}$$

Las Figuras F4.8a a F4.8d muestran las convergencias teóricas y simuladas del error cuadrático en frecuencia para la raya espectral $m=3$ y F4.9a a F4.9d las correspondientes a $m=0$. Puede observarse, al igual que en el dominio temporal, que las curvas simuladas siguen fielmente, en

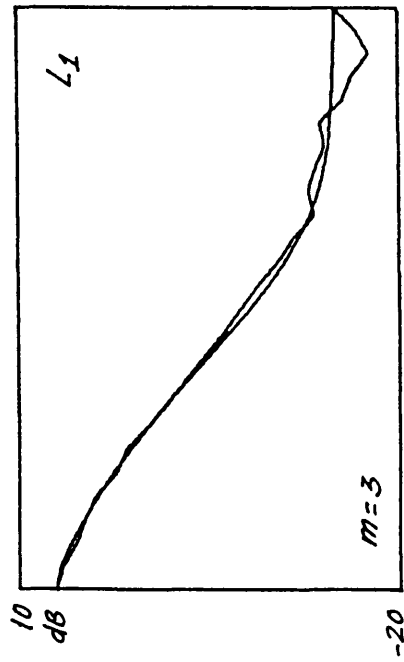
media, a sus respectivas teóricas. Este resultado, por tanto, valida el análisis realizado en todo el margen dinámico de aplicación y además proporciona el medio de diseño del paso de adaptación en función del desajuste final deseado.



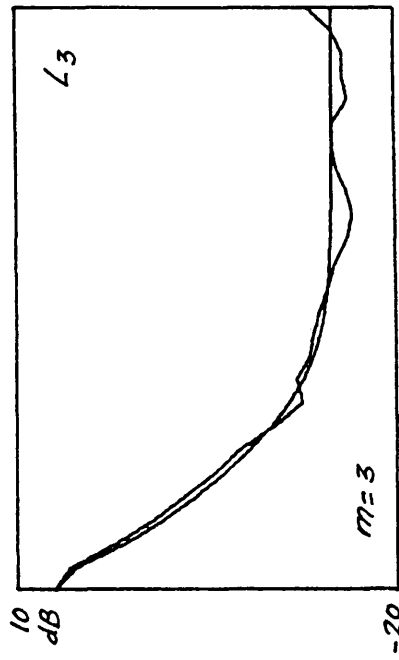
F4.8b



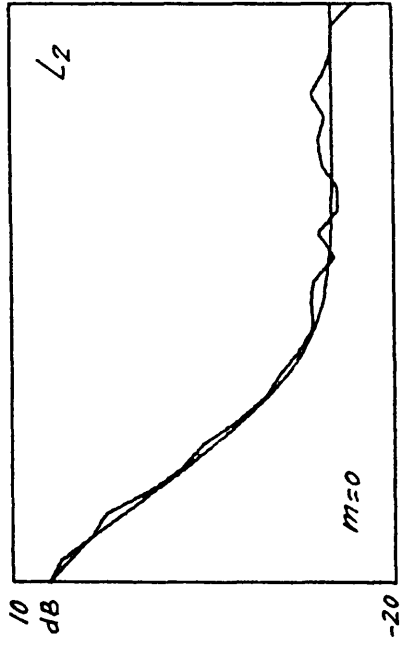
F4.8d



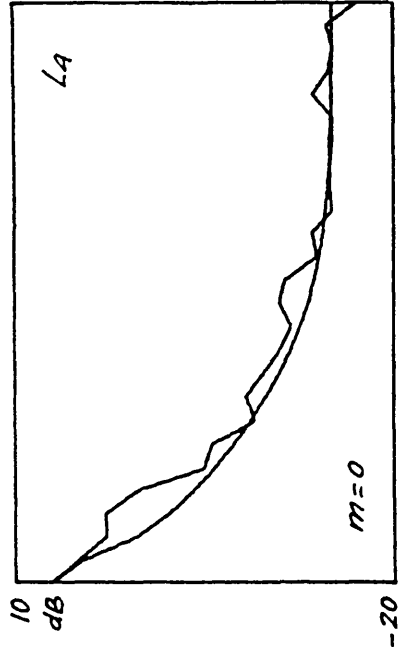
F4.8a



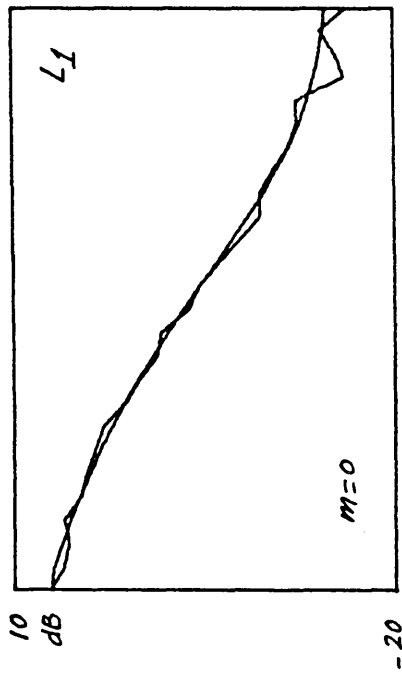
F4.8c



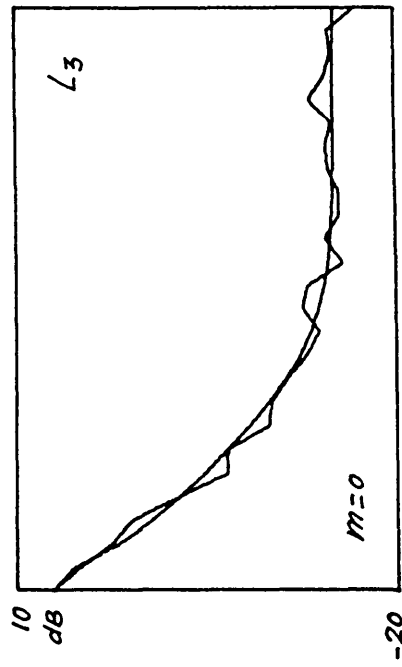
F4.9b



F4.9d

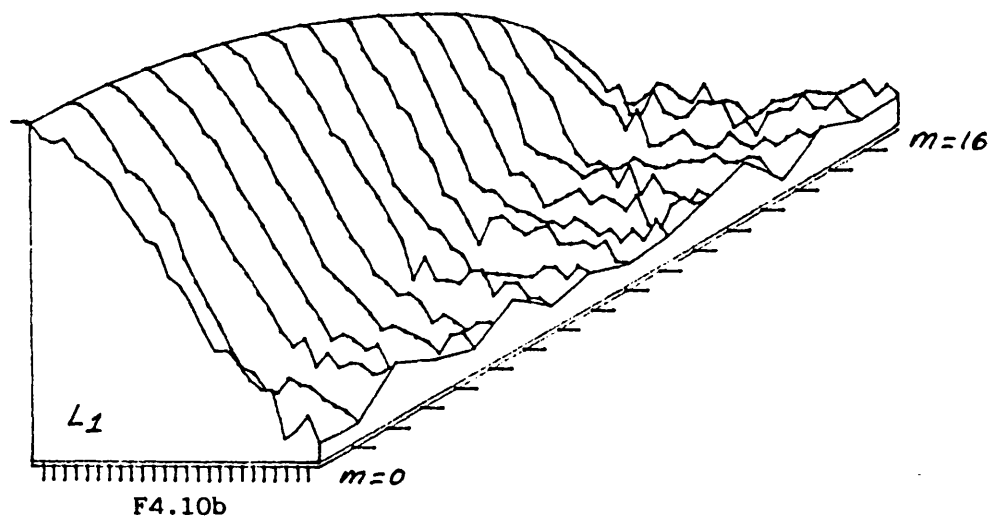
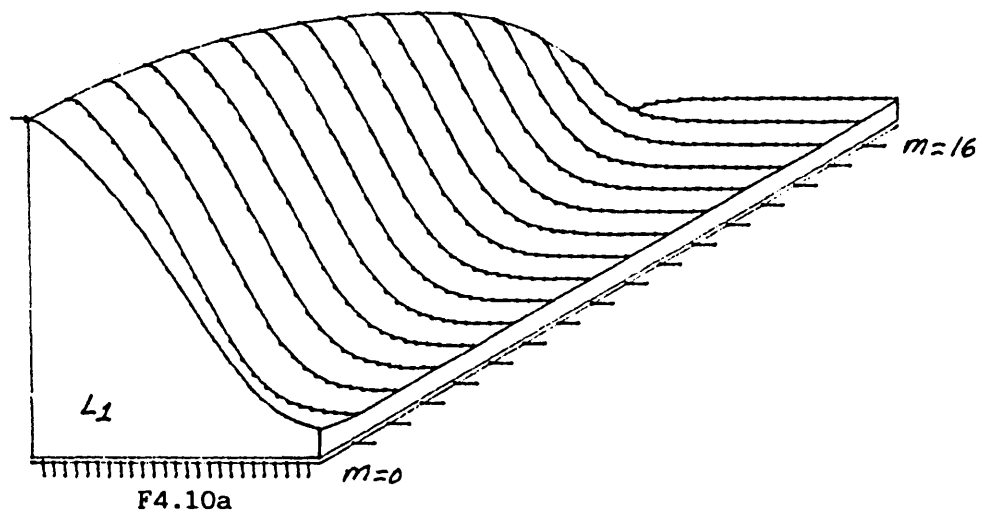


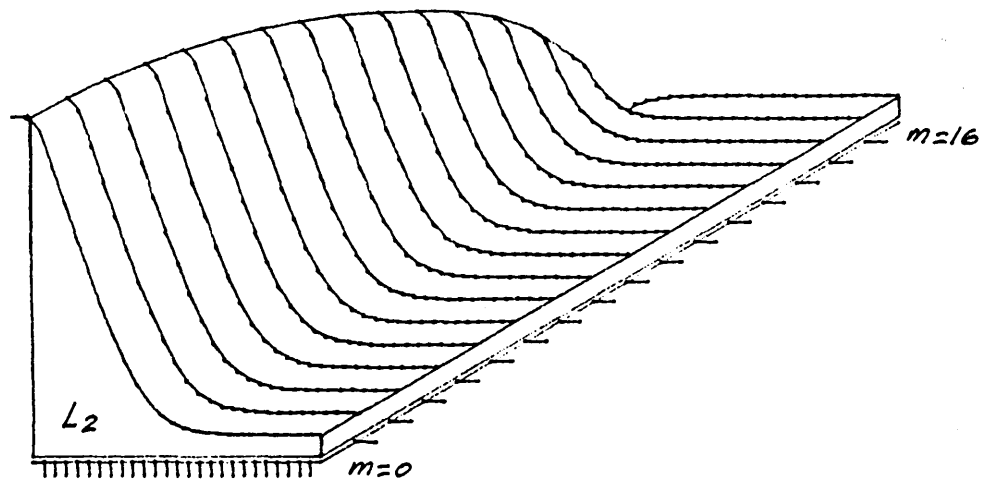
F4.9a



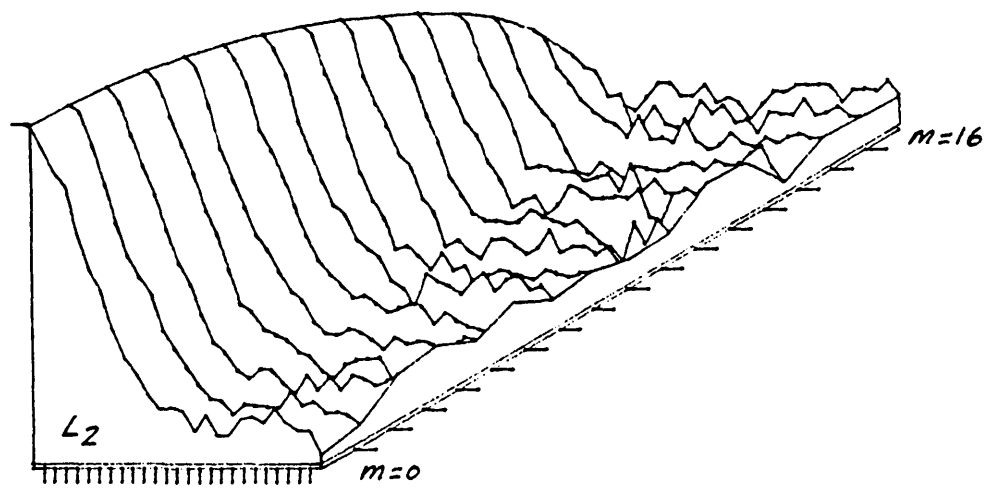
F4.9c

Con objeto de evaluar el comportamiento simultáneo del banco de procesos adaptativos en paralelo se presenta el conjunto de Figuras F4.10a, F4.10b, F4.11a, F4.11b, F4.12a, F4.12b, F4.13a y F4.13b en las que se recoge la evolución de las convergencias teóricas y simuladas correspondientes a todas las rayas espectrales ($m=0, N/2$) y los 4 primeros algoritmos de la familia. En ellas se destaca la diferencia de velocidad de convergencia entre los diferentes algoritmos. Así, para $k > 2$, la convergencia es más rápida en los instantes iniciales (errores grandes) y tiende a saturarse conforme el error decrece. Es el mismo efecto aparecido en el dominio temporal. Igualmente se observa la inflexión que sufre la convergencia del algoritmo L1 (el máximo de velocidad de convergencia lo obtiene en un punto intermedio) y el carácter log-lineal del algoritmo L2 durante la mayor parte de la convergencia.

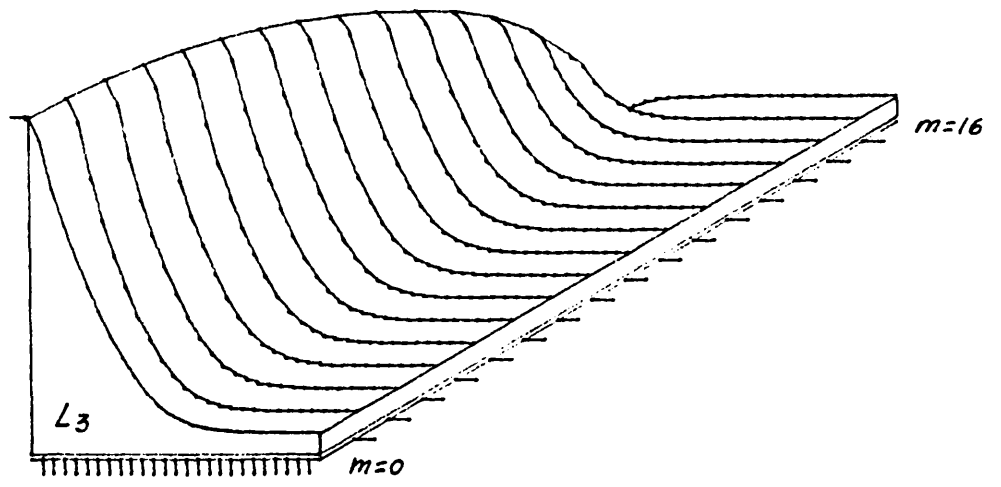




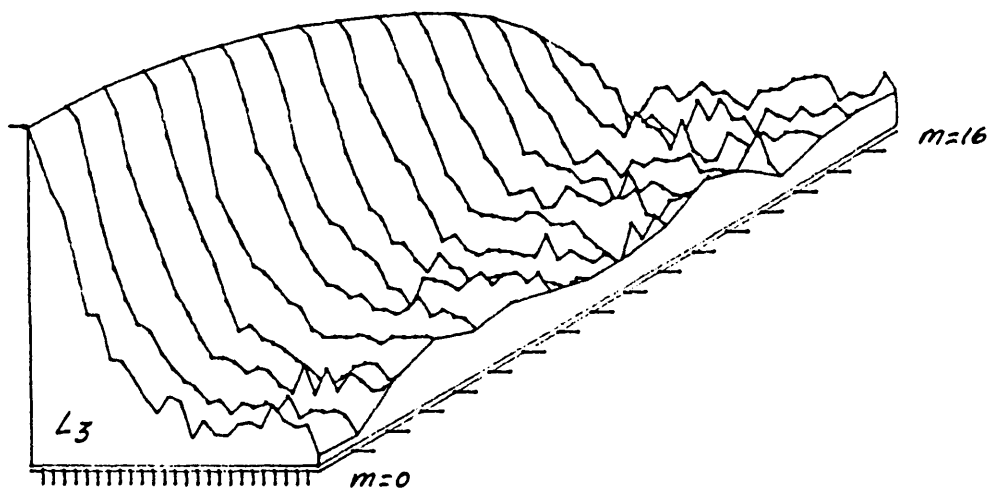
F4.11a



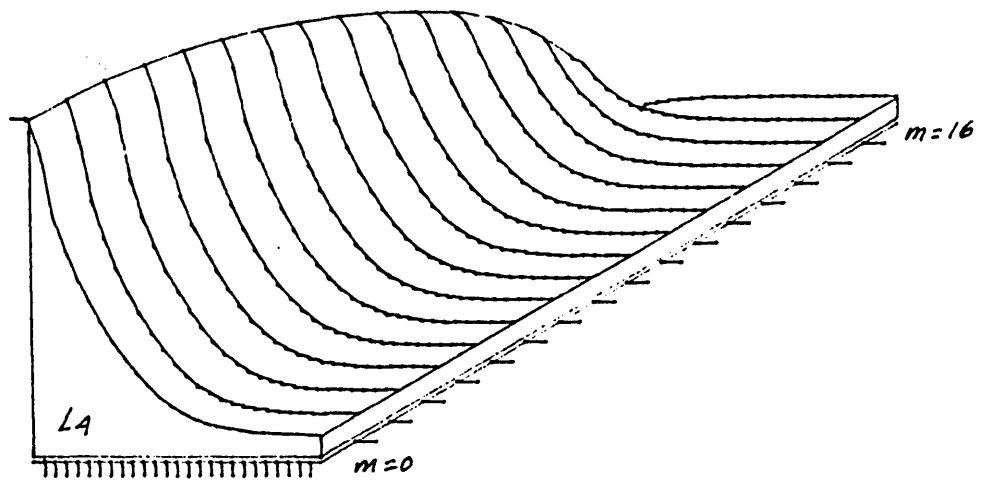
F4.11b



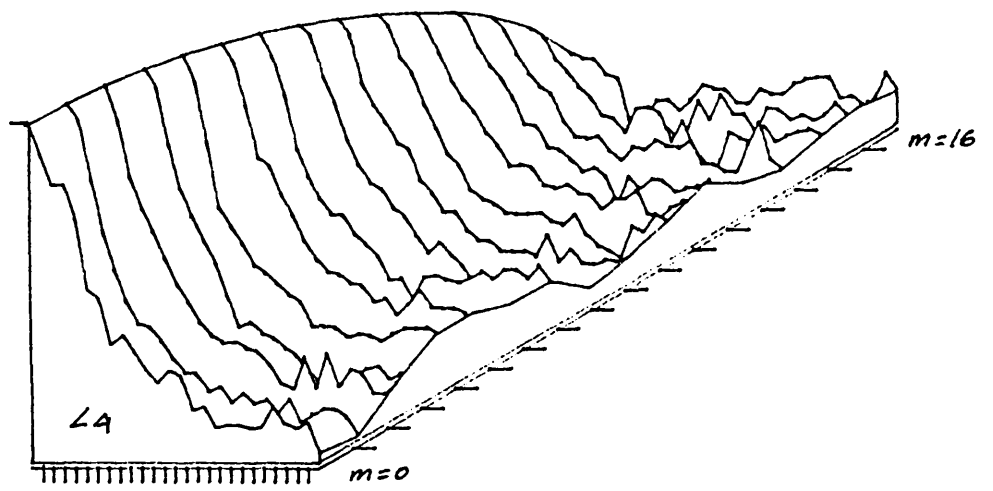
F4.12a



F4.12b



F4.13a



F4.13b

IV.5.- CONCLUSIONES

Se ha presentado y analizado una familia de algoritmos adaptativos en el dominio complejo de la frecuencia que usa el gradiente instantáneo del momento absoluto k-ésimo del error total en cada componente espectral (error residual de adaptación más ruido aditivo de planta). Este valor es utilizado localmente en la actualización de la correspondiente raya espectral adaptativa de un filtro en una aplicación de identificación de planta.

Las hipótesis de partida (H1 a H5) y el análisis de la evolución de la convergencia de la varianza del error residual son validadas por los resultados de la simulación, ya que las curvas simuladas encajan de forma bastante ajustadas en las teóricas.

A continuación se resumen las conclusiones más significativas, unas derivadas de la propia implementación en el dominio de la frecuencia y otras en función de los resultados obtenidos en el análisis teórico y simulación.

a) Derivadas de la implementación:

- A diferencia con el dominio temporal, los algoritmos son aplicables a cualquier tipo de planta.
- Existe separabilidad (ortogonalización) de las señales en frecuencia, permitiendo el adaptar independientemente cada raya espectral. Esto implica que se pueden realizar seguimientos espectralmente localizados de alguna raya aislada con objeto de algún funcionamiento particular. Igualmente existe más libertad en la elección del paso de adaptación,

ya que se dispone de un vector completo.

- Debido al procesado bloque de la DFT sobre los registros de entrada, las rayas espectrales se adaptan cada N muestras siendo necesarios $N/2+1$ procesos adaptativos simultáneos (suponiendo datos reales).
- Si N aumenta, el número de multiplicaciones necesarias aumenta más lentamente que en la implementación temporal [DENTINO, 1978]

b) Derivadas del análisis (estas son bastante similares a las obtenidas en el dominio temporal):

Con ausencia de ruido de planta:

- Si $k \geq 2$ la condición necesaria para que la varianza del error tienda a un valor nulo es el uso de una cota superior de μ , la cual depende de la varianza de error residual en frecuencia inicial, excepto para $k=2$ que es constante.
- Si $k = 1$, la varianza del error nunca diverge y su valor final depende del valor elegido para μ . Si se permite el uso de un μ decreciente hacia cero se conseguiría una varianza final nula.
- Un μ constante origina una velocidad de convergencia constante para $k=2$ y variable (llega a saturarse) para $k \neq 2$, por lo que la convergencia asintótica del L_2 es la más rápida de todos.

Con ruido de planta:

- Con un μ constante, debidamente acotado, la varianza del error residual para cualquier raya espectral y cualquier algoritmo converge hacia un valor no nulo, el cual depende del μ elegido.

- Cuando $k > 2$, la varianza inicial del error residual debe de ser considerada con objeto de evitar posibles divergencias, ya que la elección del paso μ ha de cumplir una condición adicional relacionada con el estado inicial del sistema adaptativo.
- Los algoritmos de orden superior ofrecen convergencias más rápidas cuando el error residual es grande pero se saturan a medida que el error decrece hacia el valor de régimen permanente.
- El algoritmo L1 presenta una inflexión en la convergencia para valores medios del error residual. Esta inflexión se manifiesta como un máximo local de velocidad de convergencia.

Apartados que recojan el análisis de un μ optimizado variable o medidas del desajuste final en frecuencia han sido deshechados por su similitud con el análisis en el dominio temporal.

Queda abierto el análisis teórico en frecuencia del comportamiento medio de la misma familia de algoritmos con otros ruidos de planta, definidos no Gaussianos, en el dominio complejo de la frecuencia. No obstante, en calidad de conjetura, parece razonable esperar un comportamiento similar al obtenido en el dominio del tiempo, esto es, mayor velocidad de convergencia para los algoritmos de orden bajo/alto con distribuciones de ruido en frecuencia de cola larga/corta.

CAPITULO V

APLICACIONES DE LA NUEVA FAMILIA
DE ALGORITMOS

V.1.- INTRODUCCION

El marco de aplicación de los algoritmos introducidos en esta Tesis es potencialmente el mismo donde suele actuar el LMS. A modo de ejemplo cabe citar las siguientes aplicaciones típicas:

- Cancelación de ecos
- Reducción de interferencias
- Igualación de canales
- Codificación lineal predictiva, etc.

Todos estos sistemas suelen utilizar técnicas adaptativas en la búsqueda de la solución óptima que minimice la función objetivo. No obstante, la elección del algoritmo adaptativo a implementar depende de factores tales como complejidad, carga computacional, velocidad de convergencia, mínimo error alcanzable, etc., y en cada caso se han de primar los factores favorables con objeto de facilitar dicha elección. Olvidándonos de los factores de complejidad y carga computacional, las características de velocidad y desajuste de los algoritmos tratados en esta Tesis dependen fundamentalmente de la estadística del error total observado (el de adaptación más ruido aditivo). Por tanto la elección del algoritmo adecuado para una determinada aplicación dependerá de su situación actual (señales y sistemas involucrados) y podrá ser diferente bajo situaciones distintas.

Además de las simulaciones efectuadas en los Capítulos correspondientes, que pueden ser consideradas como aplicaciones reales de identificación de sistemas, se desarrollan en este Capítulo dos aplicaciones típicas dentro del campo de las comunicaciones (modelables como identificadores de planta), las cuales hacen uso de diferentes algoritmos de la familia. Tales aplicaciones son las de cancelación de ecos de voz y datos en línea telefónica ([SONDHI,1980], [MESSERSCHMITT,1984]) Ambas están definidas en el dominio temporal y serán simuladas en ordenador utilizando registros reales de señal (voz) y sistemas (híbrida) semejantes a los reales e implementadas en formato de coma flotante y precisión simple. Se implementan, para cada aplicación, varios algoritmos de la familia con idénticas condiciones de partida, evaluando cualitativamente por comparación las prestaciones ofrecidas por cada una de las implementaciones. El criterio de evaluación está basado en la evolución del error cuadrático instantáneo de cancelación, de forma similar a la presentada en Capítulos anteriores.

V.2.- CANCELACION DE ECOS CON TRANSMISION DE VOZ

Como es conocido, debido a las transiciones 4 hilos-2 hilos existentes en la red telefónica pública, aparece el fenómeno perturbador del eco originado por la falta de desacoplo (en el circuito de 4 hilos) que proporcionan las híbridas sitas en dichas transiciones. En este caso, sólo nos referiremos al primer eco, o eco del hablante, el cual se origina en la híbrida del extremo oyente y se define como la porción de señal del hablante, que atravesando la híbrida del extremo oyente retorna al extremo hablante perturbando la audición de este (tolerable si el retardo es inferior a 50 ms) [SONDHI,1980]. Para evitar esta perturbación de utilizan clásicamente dos procedimientos conocidos como supresión y cancelación de ecos. Estos dan origen a sistemas que se introducen en el extremo oyente con objeto de: suprimir el eco, los cuales desconectan el circuito oyente-hablante cuando no se detecta señal oyente, o cancelar el eco, el cual trata de reproducir una réplica instantánea del eco, con objeto de contrarrestarlo in situ.

El supresor (se suele utilizar para ecos con retardos comprendidos entre 50 y 200 ms) esta basado en un detector de señal, el cual compara las señales en los circuitos hablante-oyente y oyente-hablante decidiendo si se ha de abrir o cerrar este último circuito en función de la presencia de solo eco o señal oyente más eco. Este sistema tiene el inconveniente de las posibles interrupciones en la conversación por parte del oyente, las cuales aparecen de forma imprevista y son difíciles de detectar, motivando perdidas en la comunicación al estar el circuito aún abierto.

El cancelador se implementa como un sistema paralelo a la híbrida, de respuesta parecida a ésta, el cual se excita con la señal hablante y ofrece como salida una réplica instantánea del eco que se sustrae de la señal del circuito oyente-hablante (se suelen utilizar para ecos con retardos superiores a 200 ms). Este procedimiento no impide el paso de la señal oyente aunque puede introducir una perturbación adicional si la respuesta del sistema es muy diferente a la de la híbrida. Debido al desconocimiento y/o degradación de la caracterización de la híbrida, el cancelador suele ser implementado como un sistema adaptativo, el cual trata de seguir, lo más fielmente posible, la salida de la híbrida. Por sencillez, su implementación suele ser transversal. En nuestra aplicación, utilizaremos la estructura de la Figura F5.1, donde H denota la híbrida y H' el cancelador adaptativo.

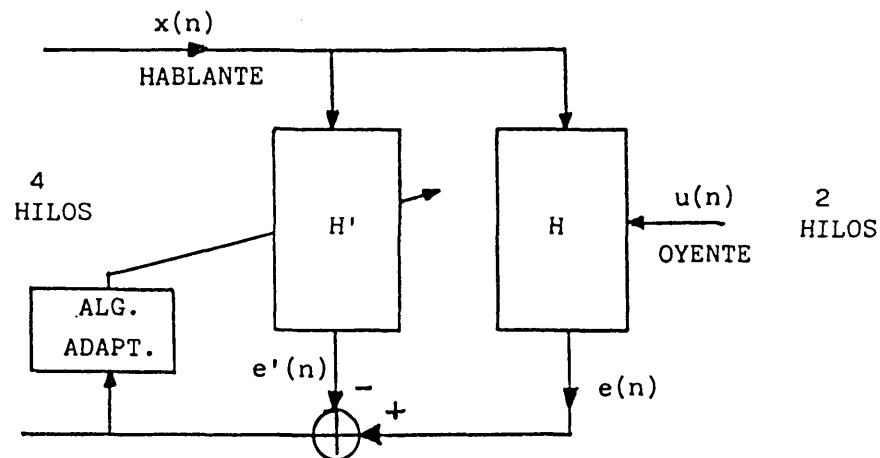


Figura F5.1

Habitualmente, los canceladores adaptativos de eco sólo actúan cuando no existe señal del oyente en la línea (disponen de sensores de nivel que *inhibe su funcionamiento*), ya que esta señal, al estar incorrelada con el eco y ser de nivel similar o superior, altera el proceso de adaptación de

los coeficientes. En nuestro caso, se pretende todo lo contrario, esto es, permitir la adaptación con la presencia de la señal del oyente (comunicación full-duplex), con objeto de evaluar las prestaciones de los primeros 4 algoritmos de la familia ante un eco más 'ruido' de forma análoga a la estudiada anteriormente.

En esta aplicación hay que resaltar, en general, que las señales de voz involucradas en el proceso no satisfacen las hipótesis de Gaussianness, estacionariedad e incorrelación establecidas en el estudio generalizado, con lo que se admite, a priori, una posible desviación de los resultados obtenidos con los simulados en el Capítulo III. Con objeto de poder controlar los procesos de adaptación para una posterior comparación objetiva, se utilizan, tanto para la señal del hablante como del oyente, tramos de sonidos vocalizados de potencia estable (vocales sostenidas). De igual manera y dado que la señal de voz se puede caracterizar estadísticamente (de forma aproximada) por una distribución de cola larga, se diseña un conjunto de pasos de adaptación, a partir de la distribución Laplaciana, que garanticen el mismo desajuste final para el error de cancelación.

A continuación se presentan los datos utilizados para la simulación:

- Híbrida de 32 coeficientes cuyas respuesta impulsiva y DFT se presentan en las Figuras F5.2 y F5.3 respectivamente.
- Algoritmo utilizado:

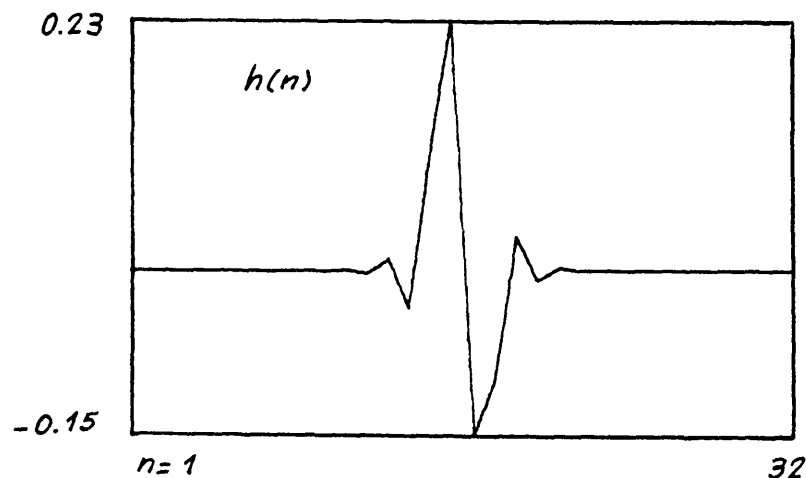
$$\underline{h}'(n+1) = \underline{h}'(n) + \mu |r(n)|^{k-2} r(n) \underline{x}(n) \quad (v.1.1)$$

$$\text{con } r(n) = e(n) - e'(n) + u(n) \quad (v.1.2)$$

- 24000 muestras de señal, correspondientes a tramos de duración de 2.66 seg. (frec. muestreo 9 KHz).
- Señal del hablante: /e/ sostenida de potencia media 5 dB. En la Figura F5.4 se presenta un tramo de 1200 muestras.
- Señal del oyente: /a/ sostenida de potencia media 0 dB. En la Figura F5.5 se presenta un tramo de 1200 muestras.
- Pasos de adaptación , diseñados según distribución Laplaciana, para un desajuste final: Pot. eco/Pot. oyente = -25 dB.

$$\begin{aligned} \mu_1 &= 1.48 \times 10^{-4} & \mu_2 &= 9.85 \times 10^{-5} \\ \mu_3 &= 2.76 \times 10^{-5} & \mu_4 &= 5.21 \times 10^{-6} \end{aligned}$$

- Cancelador transversal de 32 coeficientes inicialmente a cero.



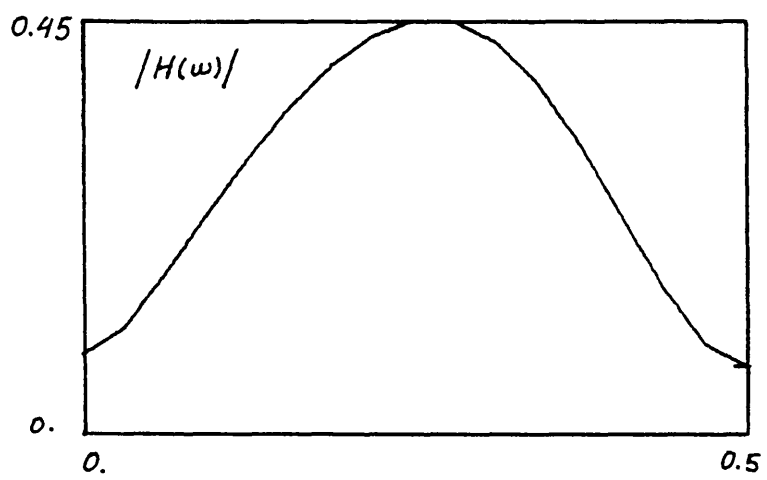


Figura F5.3

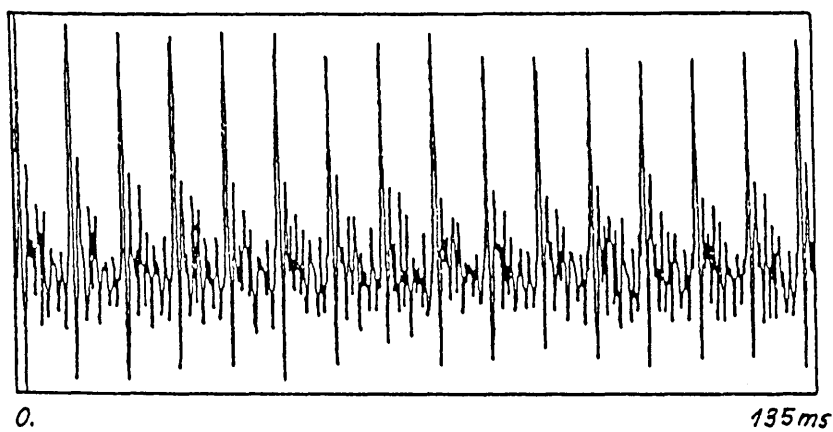


Figura F5.4

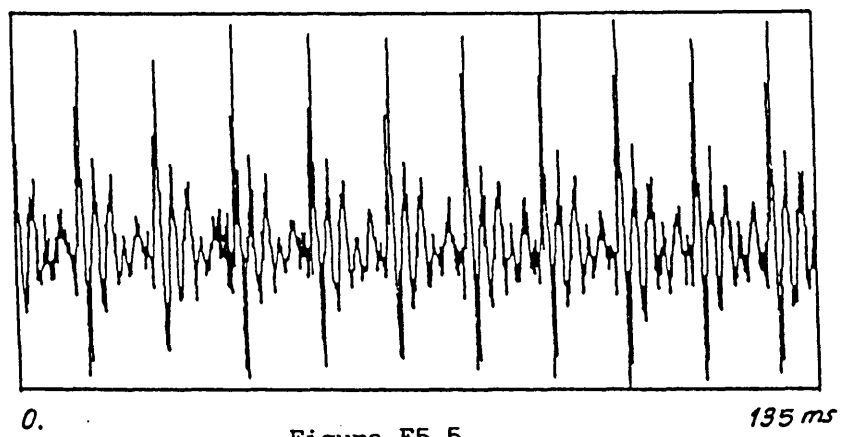
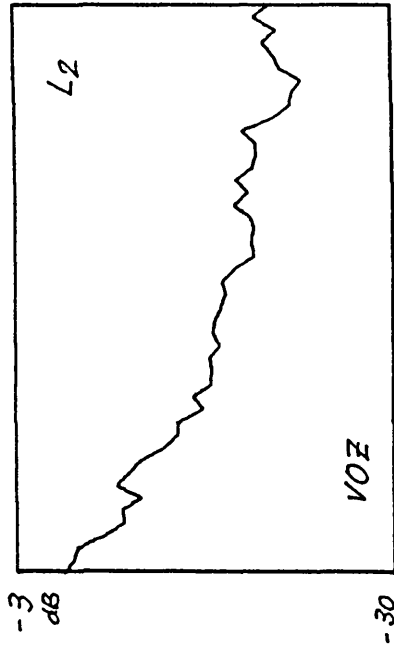


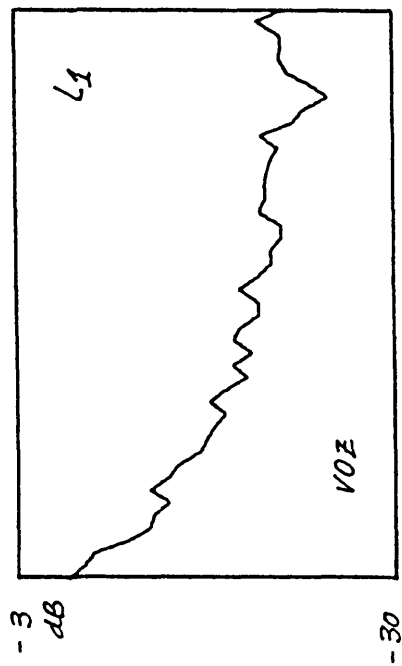
Figura F5.5

Los resultados del proceso de adaptación se presentan en la Figura F5.6. Representan el error cuadrático instantáneo normalizado por la potencia media de la señal del oyente. Las curvas han sido suavizadas con una ventana móvil uniforme de 500 muestras. En ellas se observa la ligera ventaja de utilizar el algoritmo L1 frente al L2 y de ambos frente a los de orden superior. Se señala también que el desajuste final correspondiente al algoritmo L1 es muy próximo a -25 dB, confirmando la hipótesis de distribución de cola larga para la señal de voz.

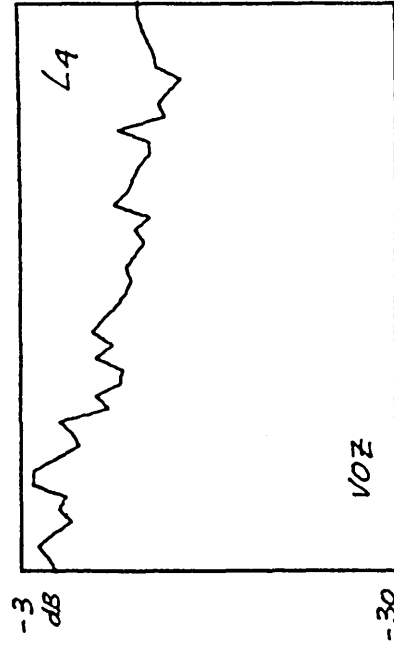
En la práctica, ante un tramo más largo de voz, sin control de las potencias de las señales y admitiendo no estacionariedades en las mismas (tramos vocalizados/no vocalizados, etc), el diseño de los pasos de adaptación adecuados es difícil. No obstante, ante una situación controlada (estacionaria), el algoritmo L1 ha demostrado ser superior al resto en una comunicación full-duplex (señal oyente presente). Considerando este hecho, además de su facilidad de implementación y estabilidad, cabría pensar en la posibilidad de utilizar algoritmos L1 en lugar de inhibir el funcionamiento de los canceladores adaptativos L2 habituales, cuando los sensores de nivel detectan la presencia de señal del oyente en la línea.



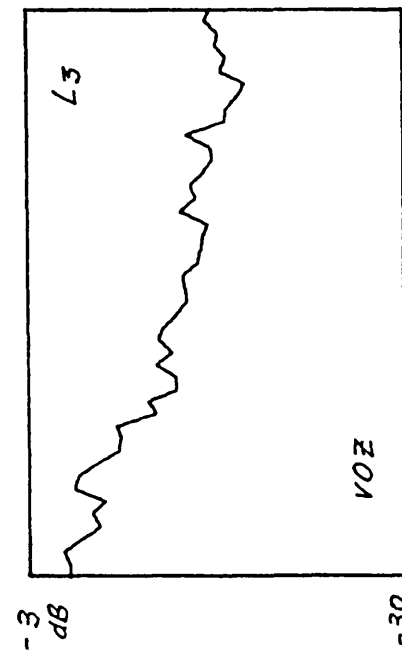
F5.6a



F5.6b



F5.6c



F5.6d

V.3.- CANCELACION DE ECOS CON TRANSMISION DE DATOS

Un fenómeno similar al anterior aparece en la transmisión full-duplex de datos por medio de un canal telefónico. El terminal transmisor/receptor de datos se haya conectado al lazo de abonado de la red pública (2 hilos) por medio de una híbrida, ya que dado el caracter full-duplex de la comunicación se hace necesaria la separación física de los canales de transmisión y recepción. Evidentemente el desacoplo entre ambos canales no es perfecto, entre 10 y 20 dB en la práctica, y por tanto aparece un eco (sin retardo) de la señal transmitida en la entrada del propio receptor. Este eco no tiene efecto alguno en una comunicación semi-dúplex, ya que el receptor se encuentra inhibido durante la transmisión, pero es altamente perturbador en una comunicación simultánea full-duplex. Con objeto de reducir al máximo este eco se introduce igualmente un cancelador en paralelo con la híbrida. La estructura final, similar a la aplicación anterior, se presenta en la Figura F5.7.

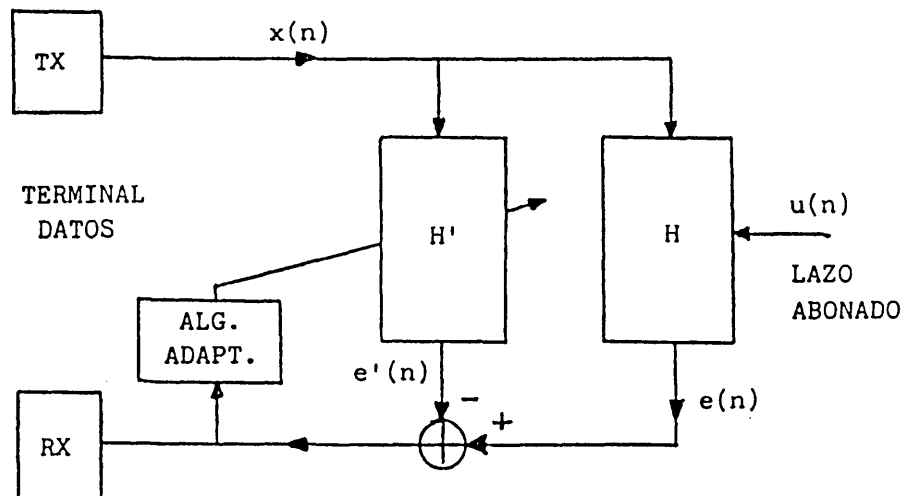


Figura F5.7

Existen en la actualidad dos modalidades claramente distintas de transmisión de datos por línea telefónica. La primera es una transmisión digital banda-base de pulsos, convenientemente conformados, y de alta velocidad que utiliza todo el ancho de banda disponible del par. Esta modalidad es hoy de gran interés, ya que es el soporte de la Red Digital de Servicios Integrados (RDSI). La segunda es una transmisión paso-banda de datos (modulando una portadora), de baja o media velocidad, dentro de la banda vocal (300,3400 Hz), y que es utilizada ampliamente por una gran variedad de terminales de datos.

La aplicación que se presenta utiliza la primera modalidad, ya que las señales involucradas son prácticamente de carácter binario (salvando la conformación y adecuación del pulso al canal) similares a las estudiadas en el Capítulo III. Se prescinde del algoritmo L1, ya que se conoce a priori su mal comportamiento ante este tipo de señales. Igual que en la aplicación anterior, no se satisface la hipótesis de Gaussianidad para la señal de entrada, con lo que existe la posibilidad de desviaciones cuantitativas de los resultados finales.

Los datos utilizados para la simulación son los siguientes:

- Idéntica caracterización de la híbrida a la utilizada en el Apartado anterior.
- Algoritmo utilizado: idéntico al del Apartado anterior.
- 10000 muestras aleatorizadas de señal binaria (1 muestra por símbolo) para las señales de transmisión y recepción.

- Potencia media señal transmisión: 15 dB.
- Potencia media señal recepción: 1 dB.
- Pasos de adaptación , diseñados según distribución binaria, para un desajuste final normalizado : Pot. eco/Pot. recepción = -25 dB

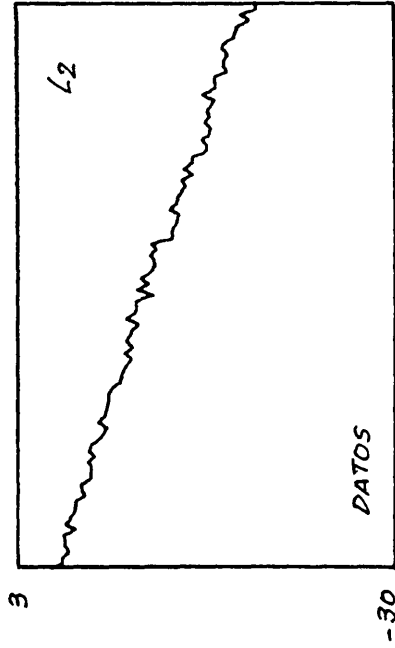
$$\mu_2 = 9.85 \times 10^{-5}$$

$$\mu_3 = 7.27 \times 10^{-4} \quad \mu_4 = 4.49 \times 10^{-3}$$

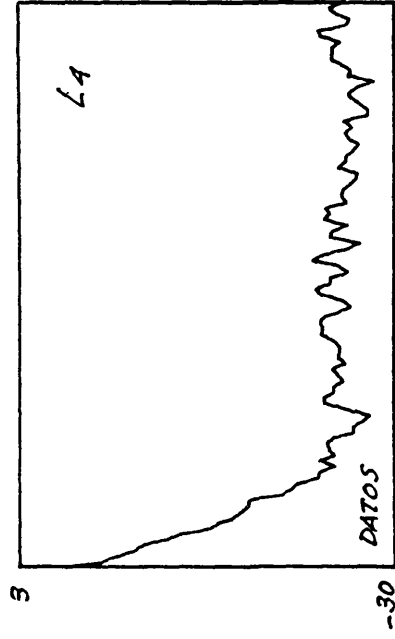
- Cancelador transversal de 32 coeficientes inicialmente a cero.

Los resultados del proceso de adaptación se presentan en la Figura F5.8. Estos evidencian la ventaja de utilizar algoritmos de orden alto cuando las señales involucradas en la transmisión son de carácter concentrado (distribuciones de cola corta).

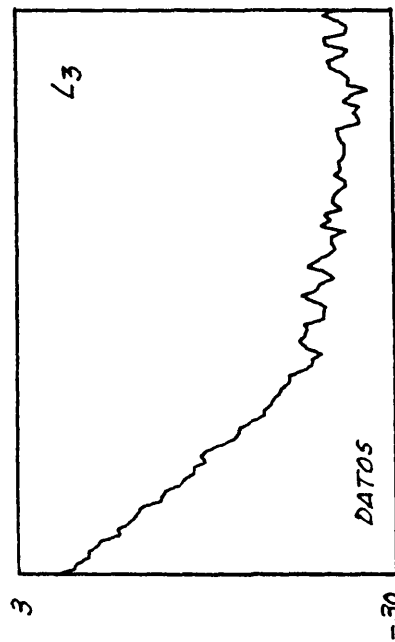
Se observa además que los pasos de adaptación seleccionados para un mismo desajuste final ofrecen aproximadamente tal valor a pesar de no contar con entradas Gaussianas.



F5.8a



F5.8c



F5.8b

CAPITULO VI

**CONCLUSIONES FINALES Y LINEAS
DE TRABAJO**

VI.1.- CONCLUSIONES

En esta Tesis se ha pretendido contribuir a la ampliación de los criterios clásicos de optimización dentro del marco de aplicaciones relacionado con el Procesado Digital de la señal (PDS). Se han planteado soluciones bloques y adaptativas alternativas a de la de error cuadrático mínimo, derivadas de otros criterios de optimización. Estas se han aplicado a procedimientos y algoritmos ya conocidos dentro del PDS, los cuales fueron originariamente desarrollados y analizados bajo el criterio cuadrático debido a su sencillez y optimalidad en determinadas condiciones.

Dentro del ámbito de los procedimientos bloque se han contrastado las prestaciones de las soluciones de mínima norma L1 y norma L2 en métodos particularmente eficaces de Estimación Espectral Paramétrica para sinusoides en ruido. Los procedimientos presentados están basados en la resolución de un sistema sobredimensionado de ecuaciones lineales en los que el vector de términos independientes (o vector de error) debe ser minimizado. Debido a que dicho vector posee un carácter poco distribuido y cuenta con la presencia de valores claramente desplazados del resto, la solución L1 ofrece claras ventajas (no computacionales), ya que el coste de valor absoluto no grava tanto los errores altos como el cuadrático, el cual tiende a distribuir el error total sobre toda la solución con objeto de minimizar el cuadrado del vector final. Los procedimientos mencionados han sido elegidos con el objeto de demostrar que en aplicaciones concretas y bajo una serie de condicionamientos, ya discutidos en el Capítulo correspondiente, la solución norma L1 mínima es una clara alternativa a la ampliamente extendida solución norma L2. Tales condicionamientos reducen la robustez de la solución L1, ya que limitan el margen de aplicación, pero

hay que considerar que dicho condicionamiento puede ser conocido en el momento de decidir el método a emplear y obviamente sugeriría una solución norma L1. No obstante, la solución norma L1 exige procedimientos de Programación Lineal (en la Tesis, el Simplex), los cuales son computacionalmente más costosos que el conocido de error cuadrático mínimo o solución L2. Este hecho puede limitar sus uso en aquellos casos en que las dimensiones del sistema sean excesivamente grandes, aunque en algunas situaciones, en las que se dispone de algún conocimiento adicional, se puedan eliminar ecuaciones prácticamente inactivas y que aportan poco a la solución final. Así ocurre, por ejemplo, en casos (analizados en la Tesis) en los que se dispone de conocimiento sobre la localización aproximada de las rayas espectrales buscadas.

Los métodos particulares implicados, tales como el método de Prony-Hildebrand (definido en el dominio del tiempo) y el segundo método de Cadzow (definido en el dominio del desplazamiento de la autocorrelación) son apropiados para la búsqueda de una solución norma L1, ya que en sus vectores de error aparecen picos claramente desplazados del resto de los valores, los cuales permanecen casi intactos en la solución L1, no distorsionando así el vector de coeficientes buscado, mientras que en la solución L2 se distribuyen con el resto del error, afectando, por tanto, al valor deseado del vector de coeficientes.

De igual manera se ha desarrollado una formulación similar a la anterior en el dominio de la frecuencia, la cual permite la aplicación de soluciones norma L1 y L2. Los comentarios sobre la elección de la solución a aplicar son idénticos a los citados anteriormente, aunque cabe destacar que el desarrollo de ambos métodos en este dominio permite introducir directamente el conocimiento previo espectral del que se disponga, tan solo

eliminando las ecuaciones definidas en las bandas que no interesen. Esta reducción del número de ecuaciones favorece al resultado final, ya que se prescinde de ecuaciones de ruido y además disminuye la carga computacional requerida para la obtención de la solución final, especialmente importante en el planteamiento de la solución norma L1.

En el marco de las soluciones adaptativas, el esfuerzo se ha centrado dentro del área de los algoritmos de gradiente estocástico. Se ha introducido un criterio de optimización generalizado como alternativa al cuadrático que conducía al clásico algoritmo LMS. Tal criterio, está basado en la minimización de una potencia entera k del valor absoluto del error total observado por medio de su gradiente instantáneo. De este criterio se deriva una familia de algoritmos adaptativos, de la que el algoritmo LMS es un miembro más ($k=2$) y donde la complejidad computacional del algoritmo crece directamente con el orden de la potencia entera elegida, concretamente $(k-1)N$ multiplicaciones por iteración (contrástese con los métodos más actuales LS que necesitan $7N$).

Con el objeto de poder elaborar un estudio ordenado sobre las posibilidades de estos algoritmos, se ha elegido el problema de identificación de planta, tanto trabajando en el dominio del tiempo como en el de la frecuencia, como soporte de la implementación del algoritmo generalizado L_k . Esta estructura admite, de manera sencilla, la inclusión de un ruido aditivo de planta, el cual, según su estadística, juega un papel muy importante en las prestaciones de la familia de algoritmos bajo análisis. Para el estudio en el dominio del tiempo se adopta además la hipótesis de planta transversal. En este caso y con objeto de facilitar el desarrollo teórico de las expresiones que predicen el comportamiento medio del algoritmo L_k , se han utilizado hipótesis simplificadoras (ampliamente

usadas en la literatura especializada) que establecen la estacionariedad conjunta e independencia de los procesos estocásticos implicados en el algoritmo. Tales hipótesis han permitido conseguir unas expresiones finales cerradas y fiables para la evolución de la varianza del algoritmo L_k en un amplio margen dinámico del error de identificación.

Debido al carácter adaptativo de los algoritmos existe el riesgo de divergencia, en lugar de convergencia al mínimo de la función de error elegida. Este riesgo se corre fundamentalmente al elegir el paso de adaptación, el cual regula el grado de adaptación local obtenido. En la Tesis se ha estudiado la máxima variabilidad permitida a este parámetro con objeto de garantizar la convergencia en varianza, así como su mejor elección localmente posible para lograr la mayor reducción de la varianza del error residual. El paso de adaptación es, igualmente, el responsable de la velocidad de adaptación del algoritmo y desajuste final del error. Este hecho plantea al diseñador un compromiso en su elección, ya que mayor velocidad y mínimo desajuste final son incompatibles. En este punto y como resultado importante del estudio realizado, hay que señalar con respecto a la garantía de convergencia, que en la elección del paso de adaptación para los algoritmos con $k=1,2$, tan solo es necesario considerar el desajuste final deseado, mientras que en los algoritmos con $k>2$ hay que considerar además el error inicial, ya que el paso de adaptación a utilizar ha de cumplir un requisito adicional que depende conjuntamente de las condiciones inicial y final de la convergencia. Este hecho, aunque un inconveniente, no es grave en la práctica y no reduce por ello el uso de algoritmos de k elevado.

Como resumen de las propiedades más importantes de los algoritmos de esta familia cabe resaltar: la mayor velocidad de adaptación de los algoritmos de k alto con errores de adaptación grandes, así como su gradual saturación cuando el error decrece; el mejor comportamiento en velocidad y/o desajuste final de los algoritmos de k alto/bajo cuando el error residual está contaminado con ruido aditivo de planta de distribución estadística de cola corta/larga; y la independencia del clásico algoritmo LMS frente a la estadística del ruido aditivo de planta así como su carácter óptimo en ausencia de éste. Estas conclusiones y el conjunto de expresiones que regulan el comportamiento medio de los algoritmos, ayudan a un diseñador en la elección del criterio de optimización y el diseño del paso de adaptación ante un problema semejante o modelable como el de identificación de planta transversal en el dominio del tiempo. No se descarta, no obstante, la posibilidad de implementar algoritmos diferentes (distintos k) para diferentes tramos de la convergencia utilizando las mejores prestaciones locales de cada uno de ellos.

Idéntico estudio ha sido llevado a cabo en el dominio complejo de la frecuencia, del cual se derivan algoritmos muy parecidos a los obtenidos en el dominio del tiempo. En este caso las hipótesis de partida son algo diferentes, aunque fundamentalmente están basadas en el carácter de independencia estadística conjunta de las variables aleatorias complejas involucradas. El estudio se fundamenta en el análisis, en el dominio de la frecuencia, del comportamiento medio del algoritmo adaptativo generalizado L_k para cada uno de los coeficientes espectrales del filtro a identificar. Esta formulación se deriva del uso de la DFT sobre los datos, la cual permite además elegir el tamaño del identificador adaptativo espectral en función de la aplicación. Además, debido al procesamiento bloque (N muestras) de la DFT, las adaptaciones se realizan cada N instantes de tiempo, de tal

manera que el número de operaciones necesario para ellas aumenta más lentamente, con N , que el necesario en la implementación temporal. Las conclusiones generales derivadas del análisis del algoritmo generalizado complejo L_k son, en relación con la convergencia y el paso de adaptación, cualitativamente similares a las obtenidas en el dominio del tiempo, por lo que no serán repetidas. Esta similitud no implica una equivalencia, ya que el análisis en frecuencia, debido al procesado de la DFT, introduce implícitamente una convolución circular entre la entrada al sistema y su respuesta impulsiva, claramente distinta a la convolución lineal utilizada en el dominio temporal.

La implementación en paralelo, en el dominio de la frecuencia, permite utilizar simultáneamente algoritmos adaptativos diferentes para cada una de las rayas espectrales consideradas. Esto supone una mejor adecuación de las prestaciones de los algoritmos a las situaciones particulares de cada componente espectral, como por ejemplo mayor o menor velocidad de convergencia y/o mayor desajuste final en el error. Todas estas ventajas, derivadas de la implementación paralelo en frecuencia, unidas a las propias del criterio de optimización considerado, hacen de esta familia de algoritmos una clara alternativa a la implementación temporal donde se desee una identificación espectral localizada.

Como punto final, se reitera que el objetivo fundamental de esta Tesis ha sido el de contribuir con criterios de optimización alternativos al de error cuadrático mínimo, desarrollando apropiadamente su formulación y expresiones correspondientes, los cuales pueden ser aplicados, con ventaja, en situaciones propicias. Como ejemplo de situaciones se han presentado procedimientos clásicos de estimación espectral e identificación adaptativa, los cuales estaban mejor condicionados para soluciones o

algoritmos diferentes al cuadrático. Este condicionamiento ha sido claramente resaltado en la Tesis con objeto de establecer principios que faciliten la elección del criterio de optimización adecuado ante otros procedimientos de PDS, diferentes a los tratados, que presenten características similares. No obstante, debido a que el criterio de optimización más adecuado para un problema de estimación determinado depende del propio problema, las conclusiones presentadas en este trabajo contribuyen, solo parcialmente, a engrosar el conjunto de recomendaciones de soluciones y algoritmos más propicios que el cuadrático, para problemas cuyas caracterizaciones sean reconocidas como similares a las aquí presentadas.

VI.2.- SUGERENCIAS DE FUTURAS LINEAS DE TRABAJO

En este trabajo se realizan algunas aportaciones concretas en la utilización de criterios de minimización diferentes al cuadrático. No obstante, existe un amplio abanico de posibles nuevas líneas de trabajo. A continuación se detallan algunas de ellas:

- Como complemento al trabajo realizado:

- * Análisis de la familia de algoritmos adaptativos con entrada y ruido de planta coloreados.
- * Formulación y análisis de la misma familia de algoritmos adaptativos en la frecuencia, con otros ruidos de planta (no Gaussianos).
- * Análisis de prestaciones, en el dominio de la frecuencia, con enventanado de datos distintos del uniforme.
- * Uso de algoritmos mixtos (diferentes criterios de optimización en función del tramo de convergencia) con objeto de acelerar la velocidad de convergencia y reducir el desajuste final.
- * Análisis de prestaciones con pasos de adaptación normalizados por una potencia entera del proceso de entrada.
- * Aproximación por tramos de los pasos de adaptación calculados como óptimos.
- * Formulación y análisis del uso de pasos de adaptación estocásticos (dependientes del error residual) que aceleren la convergencia.

- Como nuevas líneas de trabajo:

- * Formulación y análisis del algoritmo de máxima pendiente (determinista) con superficies no cuadráticas con objeto de promover versiones estocásticas alternativas a la presentada.
- * Formulación y análisis de la misma familia de algoritmos en filtros con estructura en celosía con el objeto de evitar la degradación del algoritmo con la dispersión de los autovalores del proceso de entrada.
- * Formulación y análisis de la misma familia de algoritmos en el problema de deconvolución adaptativa, con aplicación inmediata en igualación de canales y predicción lineal.
- * Análisis de prestaciones de la familia en aplicaciones tales como: extrapolación de secuencias, seguimiento de tonos, reconocimiento de señal a partir de su módulo o fase (métodos de relajación), intensificación de tonos (ALE), eliminación de interferencias, etc.

VII

REFERENCIAS

REFERENCIAS

- [ABRAMOWITZ,1970] Abramowitz,M., Stegun,A.I.; Handbook of Mathematical Functions, 9th edition, Dover, 1970
- [ALEXANDER,1986] Alexander,S.T.; Adaptive Signal Processing: Theory and Applications, Ed. Springer-Verlag, 1986
- [ALVAREZ,1985] Alvarez, A.M.,Grant,P.M.; Comparison of Frequency Domain Adaptive Equaliser Structures. Proc. IASTED Symp. on Applied Digital Filtering, 91-94; Paris, Junio 1985
- [BERSHAD,1979] Bershad,N.J., Feintuch,B.; Analysis of the Frequency Domain Adaptive Filter, Proc. IEEE, vol. 67, pp. 1658-1659, Dec. 1979
- [BERSHAD,1982] Bershad,N.J., Feintuch,P.L.; Correlation Function for the Weigth Sequence of the Frequency Domain Adaptive Filter, IEEE Trans. on ASSP, vol. 30, no. 5, pp. 801-804,Oct. 1982
- [BERSHAD,1986] Bershad,N.J.; Analysis of the Normalized LMS Algorithm with Gaussian Inputs, IEEE Trans. on ASSP, vol. 34, no. 4, pp. 793-806, Aug., 1986
- [BITMEAD,1980] Bitmead,R.R., Anderson,B.,D.,O.; Performance of Adaptive Estimation Algorithms in Dependent Random Enviroments, IEEE Trans. on Aut. Control, vol. 25, pp. 788-794, Aug. 1980
- [BLOOMFIELD,1976] Bloomfield,P.; Fourier Analysis of Time Series: An Introduction, J. Wiley, 1976
- [BRY,1982] Bry,K.J.,Le Roux,J.; Comparison of Some Algorithms for Identifying Autoregressive Signals in the Presence of Observation

Noise, Proc. ICASSP-82, pp. 224-227, Paris, Mayo 1982

[CADZOW,1979] Cadzow,J.A.; ARMA Spectral Estimation: An Efficient Closed Form Procedure, Proc. RADC Spectrum Estimation Workshop, pp. 81-97, Oct. 1979

[CADZOW,1982a] Cadzow,J.A.; Spectral Estimation: An Overdetermined Approach, Proc. IEEE, vol. 70, pp. 907-938, Sep. 1982

[CADZOW,1982B] Cadzow, J.A.; ARMA Modeling of Time Series, IEEE Trans. on PAMI, vol. PAMI-4, no. 2, pp. 256-261, Mayo 1982

[CLAASEN,1981] Claassen,T.A.C.M., Mecklenbraeuker,W.F.G.; Comparison of the Convergence of Two Algorithms for Adaptive FIR Digital Filters, IEEE Trans. on ASSP, vol. 29, no. 3, pp. 670-678, Jun. 1981

[CLAASEN,1983] Claassen,T.A.C.M., Mecklenbraeuker,W.F.G.; Overview of Adaptive Techniques in Signal Processing, Proc. EUSIPCO-83, pp. 747-754, Erlangen, West-Germany, Sep. 1983

[CHILDERS,1978] Childers,D.G.; Modern Spectrum Analysis, IEEE Press, New York, 1978

[DANIELL,1970] Daniell,T.; Adaptive Estimation with Mutually Correlated Training Samples, IEEE Trans. on SSC, vol. SSC-6, pp. 12-19, Enero 1970

[DENTINO,1978] Dentino,M.,McCool,J.,Widrow,B.; Adaptive Filtering in the Frequency Domain, Proc. IEEE, vol. 66, no. 12, pp. 1658- 1659, Dec. 1978

[ERDELYI,1954] Erdélyi,A.; Tables of Integral Transforms, McGraw Hill Book Company Inc., 1954

- [EYKHOFF,1974] Eykhoff,P.; System Identification, John Wiley and Sons,1974
- [FALCONER,1978] Falconer,D.D., Ljung,L.; Application of Fast Kalman Estimation to Adaptive Estimation, IEEE Trans. on Communications, vol. COM-26, pp. 1439-1445, Oct. 1978
- [FARDEN,1981] Farden,D.C.; Stochastic Approximation with Correlated Data, IEEE Trans. on Inf. Theory,vol. 27,pp. 105-113, Jan. 1981
- [FERRARA,1980] Ferrara,E.R.; Fast Implementation of LMS Adaptive Filters, IEEE Trans. on ASSP, vol. 28, no. 4, pp. 474-475, Aug. 1980
- [FEUER,1985] Feuer,A., Weinstein,E.; Convergence Analysis of LMS Filters with Uncorrelated Gaussian Data, IEEE Trans. on ASSP, vol. 33,no. 1, pp. 222-229, Feb. 1985
- [FIGUEIRAS-VIDAL,1985a] Figueiras-Vidal,A.R.,Casar-Corredera,J.R., Garcia-Gomez,R.,Páez-Borrallo,J.M.; L1-Norm vs. L2-Norm Minimization in Parametric Spectral Analysis: A General Discussion, Int. Conf. ASSP, pp. 9.1.1-9.1.4,Tampa,Mar. 1985.
- [FIGUEIRAS-VIDAL,1985b] Figueiras-Vidal,A.R., Páez-Borrallo,J.M., Ruiz-Mezcua,B.; Using A Priori Information and L1-Norm Minimization in Cadzow's and Prony's Line Analysis Methods, Int. Conf. on Mathematics in Signal Processing, Bath (UK), Sep.,1985
- [FIGUEIRAS-VIDAL,1986] Figueiras-Vidal,A.R.,Páez-Borrallo,J.M., Garcia-Gómez,R.,Hernández-Gómez,L.A.; Lk-Norm Adaptive Transversal Filters, remitido para número especial sobre Modelling-Identification Techniques in Signal Processing a: Automatic, Control and Production Systems (AFCT) y Signal Processing (GRETSI)

- [GARDNER,1984] Gardner,W.A.; Learning Characteristics of Stochastic Gradient Descent Algorithms: A General Study, Analysis and Critique, Signal Processing no. 6, pp. 113-133, 1984
- [GERSHO,1969] Gersho,A.; Some aspects of linear estimation with non-mean-square error criteria, Conf. Rec., Third Asilomar Conf. Circuits and Systems, pp. 202-206, Dec. 1969
- [GERSHO,1984] Gersho,A.; Adaptive Filtering with Binary Reinforcement, IEEE Trans. on Inf. Theory, vol. IT-30, no. 2, pp. 191-199, Marzo, 1984
- [GIBSON,1980] Gibson,C., Haykin,S.; Learning Characteristics of Adaptive Lattice Filtering Algorithms, IEEE Trans. on ASSP, vol. ASSP-28, pp. 681-691, Dic. 1980
- [GITLIN,1973] Gitlin,R.D., Mazo,J.E., Taylor,M.G.; On the Design of Gradient Algorithms for Digitally Implemented Adaptive Filters, IEEE Trans. on Cir. Theory, vo. 20, no. 2, pp. 125-136, Mar. 1973
- [GOLDSTEIN,1977] Goldstein,E.G.,Yudin,D.B.; Programacion Lineal. Problemas y Aplicaciones, Editorial Paraninfo, 1977
- [GRIFFITHS,1969] Griffiths,L.J.; A Simple Adaptive Algorithm for Real Time Processing in Antenna Arrays, Proc. IEEE, pp. 2143-2159, 1969
- [GRIFFITHS,1977] Griffiths,L.J.; A Continuously Adaptive Filter Implemented as a Lattice Structure, Proc. ICASSP-77, Hartford CT, Abril 1977
- [HASTINGS,1969] Hastings-James,R.; Recursive Generalized Least-Squares Procedure for Online Identification of Process Parameters, Proc. IEE, vol. 116, pp. 2057-2062, Dic. 1969

- [HAYKIN,1979] Haykin,S.; Nonlinear Methods of Spectral Analysis, Topics in Applied Physics, vol. 34, Ed. Springer-Verlag, 1979
- [HAYKIN,1984] Haykin,S.; Introduction to Adaptive Filters, McMillan Publishing Company, Londres, 1984
- [HOLTZ,1973] Holtz,H.; Prony's Methods and Related Approaches to Exponential Approximation, Aerospace Corp. Rep. ATR-73(9990)-5, Junio 1973
- [HONIG,1981] Honig,M.L.,Messerschmitt,D.G.; Convergence Properties of an Adaptive Lattice Filter, IEEE Trans. on ASSP, vol. ASSP-29, pp. 642-653, Junio 1981
- [HONIG,1984] Honig,M.L. and Messerschmitt,D.G.; Adaptive Filters. Structures, Algorithms and Applications, Kluwer Academic Publishers, 1984
- [HILDEBRAND,1956] Hildebrand,F.B., Introduction to Numerical Analysis, ch. 9, McGraw Hill,1956
- [HSIA,1983] Hsia,T.C., Convergence Analysis of LMS and NMLS Adaptive Algorithms, Proc. IEEE ICASSP Boston,pp. 667-670, Apr. 1983
- [KEYBL,1983] Keybl,J.,Eichman,G.; Estimation of Two Closely Spaced Frequencies Buried in White Noise Using Linear Programming, Proc. Topical Meeting on Signal Recovery and Synthesis with Incomplete Information and Partial Constrains, Incline Village Nevada, pp. WA18-1/-/WA18-4, Jan. 1983
- [KIM,1975] Kim,J.K.,Dacisson,L.D.; Adaptive Linear Estimation for Stationary M-dependent Processes, IEEE Trans. on Inf. Theory, vol. 21,PP. 23-31, Jan. 1975

- [KUNZI,1971] Kunzi H.P. y otros; Numerical Methods of Mathematical Optimization, Academic Press, 1971
- [LACOSS,1971] Lacoss,R.T.; Data Adaptive Spectral Analysis Methods, Geophysics, Abril 1971
- [LARIMORE,1983] Larimore,M.G., Treichler, J.R.; Convergence Behaviour of the Constant Modulus Algorithm, Proc. IEEE ICASSP Boston, vol. 1, pp. 13-16, Mar. 1983
- [LEVY,1982] Levy,S.,Walker,C.,Ulrych,T.J.,Fullagar,P.K.; A Linear Programming Approach to the Estimation of the Power Spectra of Harmonic Processes, IEEE Trans. ASSP,vol. ASSP-30, pp. 675-679, Aug. 1982
- [LJUNG,1983] Ljung,L., Soederstroem,T.; Theory and Practice of Recursive Identification, Ed. MIT Press, 1983
- [LUENBERGER,1973] Luenberger,D.G.: Introduction to Linear and Nonlinear Programming, Addison-Wesley Publishing Company, 1973
- [MAKHOUL,1975] Makhoul,J.; Linear Prediction: A Tutorial Review, Proc. IEEE, vol. 63, pp. 561-580, Abril 1975
- [MARPLE,1979] Marple Jr.,S.L.; Spectral Analysis by Pisarenko and Prony Methods, Rec. ICASSP-1979, pp. 159-161
- [MARTINELLI,1986] Martinelli,G., Orlandi,G., Burrascano,P.; Spectral Estimation by Repetitive L1-Norm Minimization, Proc. IEEE, vol. 74, no. 3, pp. 523-524, marzo 1984
- [MEDAUGH,1981] Medaugh,R.S., Griffiths,L.J.; A Comparison of Two Fast Linear Predictors, Proc. ICASSP-81, Atlanta GA, Mayo 1981

- [MESSERSCHMITT,1984] Messerschmitt,D.G.; Echo Cancellation in Speech and Data Transmission, IEEE Journal on SAC, vol. SAC-2, no. 2, pp. 283-293, Marzo 1984
- [MONTAGNA,1982] Montagna, R., Nebia,L.; Comparison of Some Algorithms for Tap Weigth Evaluation in Adaptive Echo Cancellers, Proc. of ICASSP-82 Paris, vol. 3 ,pp. 1404-1407, Mar. 1982
- [PAEZ-BORRALLO,1985a] Páez-Borrallo,J.M., Figueiras-Vidal,A.R.; Minimum L1-Norm Spectral Estimation in the Frequency Domain, MELECON-85, Madrid, Oct. 1985
- [PAEZ-BORRALLO,1985b] Paez-Borrallo,J.M.,Figueiras-Vidal,A.R., Hernandez-Gomez,L.A.; L-odd versus L-even Adaptive Algorithms: An Introduction, Proc. IASTED Symp. on Applied Digital Filtering, 32-35; Paris, Junio 1985
- [PAEZ-BORRALLO,1986] Paez-Borrallo,J.M.,Figueiras-Vidal,A.R.,Docampo-Amoedo,D.,Casar-Corredera,J.R.; Convergence of Minimum Lk-Norm Gradient-Type Algorithms in Noise-Free Plant Identification, EUSIPCO-86, vol. 1, pp. 105-108, La Haya, Sep. 1986.
- [PAEZ-BORRALLO,1987] Páez-Borrallo,J.M.; Figueiras-Vidal,A.R., Vergara-Domínguez,L.; Analysis of Lk-Frequency-Adaptive Transversal Filters in Plant Identification, a presentar en ICASSP-87, Dallas, Abril 1987
- [PROAKIS,1983] Proakis,J.G.; Digital Communications, McGraw-Hill, New York, 1983
- [RAO,1984] Rao,S.S.; Optimization. Theory and Applications, 2nd edition, Wiley Eastern Limited, 1984

- [CREED,1981] Reed,F.A, Feintuch,P.L.; A Comparison of LMS Adaptive Cancellers Implemented in the Frequency Domain an the Time Domain, IEEE Trans. on Circuits and Systems, vol. CAS-28, no. 6, pp. 610-615, Jun. 1981
- [RICE,1964] Rice,J.R.; Norms for Smoothing and Estimation, SIAM Rev., vol. 6, pp. 243-256, Julio 1964
- [SARI,1982] Sari,H.; Performance Evaluation of Three Adaptive Equalization Algorithms, Proc. IEEE ICASSP Paris,vol. 3 pp. 1385-1389, 1982
- [SONDHI,1980] Sonfhi,M.M., Berkley,D.A.; Silencing Echoes on The Telephone Network, Proc. IEEE, vol. 68, no. 8, pp. 948-963, Aug. 1980
- [SRINATH,1979] Srinath,M.D. and Rajasekaran,P.K.; An Introduction to Statistical Signal Processing with Applications, J. Wiley, 1979
- [SULLIVAN,1978] Sullivan,T.M., Frost,O.L., Treichler,J.R.; High Resolution Signal Estimation, ARGO Systems, Internal Rep., Junio 1978
- [VAN BLARICUM,1978] Van Blaricum,M.L., A Review of Prony's Method Techniques for Parameter Estimation, Proc. Rome Air Development Center Spectrum Estimation Workshop, Griffiths, PP. 126-139, Mayo 1978
- [VAN TREES,1968] Van Trees,H.L.; Detection, Estimation and Modulation Theory, J. Wiley, 1968
- [WEISS,1963] Weiss,L., McDonough,R.N.; Prony's Method, Z-Transforms and Pade Approximations, SIAM Rev., vol. 5, pp. 145-149, Abril 1963
- [WIDROW,1960] Widrow,B.,Hoff,M.E.; Adaptive Switching Circuits, WESCON Conv. Rec.,pt. 4, pp. 96-140,1960

- [WIDROW,1967] Widrow,B.,Mantey,P.E.,Griffiths,L.J.,Goode,B.B.; Adaptive Antenna Systems, Proc. IEEE vol. 55, no. 12, pp. 2143-2159, Dec. 1967
- [WIDROW,1975a] Widrow,B., Mc Coll,J., Ball,M.; The Complex LMS Algorithm, Proc. of IEEE, vol. 63, pp. 719-720, April 1975
- [WIDROW,1975b] Widrow,B. et al.; Adaptive Noise Cancelling: Principles and Applications, Proc. IEEE, vol. 63, no. 12, pp. 1692-1716, Dec. 1975
- [WIDROW,1976] Widrow,B.,McCool,J.M.,Larimore,M.G. and Johnson Jr.,C. R.; Stationary and Nonstationary Characteristics of the LMS Adaptive Filter, Proc. IEEE, vol. 64, no. 8, pp. 1151-1162, Aug. 1976.
- [WIDROW,1984] Widrow,B.,Walach,E; On the Statistical Efficiency of the LMS Algorithm with Nonstationary Inputs, IEEE Trans. Inf. Theory, vol. 30, no. 2, part 1, PP. 211-221, Mar. 1984.
- [WIDROW,1985] Widrow,B. and Stearns S.D.; Adaptive Signal Processing, Prentice-Hall Inc.,1984

VIII

APPENDICES

A.1- CALCULO DEL MINIMO DE LA FUNCION OBJETIVO

Sea

$$f(\underline{c}) = E(|(\underline{c}_p - \underline{c})^t \underline{x} + u|^k) \quad (\text{a.1.1})$$

igualando a cero su gradiente se determinan sus maximos y minimos:

$$\underline{0} = \nabla(E(|(\underline{c}_p - \underline{c})^t \underline{x} + u|^k)) \quad (\text{a.1.2})$$

cambiando el orden de la diferenciación por el del cálculo del valor medio y utilizando (iii.1.6), se obtiene:

$$\underline{0} = E(-\underline{x}(\epsilon + u) |\epsilon + u|^{k-2}) \quad (\text{a.1.3})$$

Si se calcula primeramente el valor medio con respecto a u, condicionado a \underline{x} :

$$\underline{0} = E(-\underline{x} E((\epsilon + u) |\epsilon + u|^{k-2}) / \underline{x}) \quad (\text{a.1.4})$$

para k = par:

$$(\epsilon + u) |\epsilon + u|^{k-2} = (\epsilon + u)^{k-1} \quad (\text{a.1.5})$$

desarrollando (a.1.5) según el binomio de Newton:

$$(\epsilon + u)^{k-1} = \sum_{i=0}^{k-1} \epsilon^{k-1-i} u^i \quad (\text{a.1.6})$$

y si la d.f.p. condicional $f(u/x)$ es par respecto al origen (hipótesis H2 y H3 del Capítulo III), los momentos impares de u son nulos, con lo que:

$$E((\epsilon + u)^{k-1} / \underline{x}) = \epsilon^{k-1} + \epsilon^{k-3} E(u^2 / \underline{x}) + \dots + \epsilon E(u^{k-2} / \underline{x}) \quad (\text{a.1.7})$$

Observando (a.1.7) y (a.1.4) se comprueba que $\epsilon = 0$ es una solución que anula el gradiente, para todo k.

para k impar:

$$(\epsilon + u) |\epsilon + u|^{k-2} = \text{sgn}(\epsilon + u) |\epsilon + u|^{k-1} \quad (\text{a.1.8})$$

en este caso el valor medio condicionado es:

$$E(./\underline{x}) = - \int_{-\infty}^{-\epsilon} (\epsilon+u)^{k-1} f(u/\underline{x}) du + \int_{-\epsilon}^{\infty} (\epsilon+u)^{k-1} f(u/\underline{x}) du \quad (\text{a.1.9})$$

que igualmente se anula, para todo k , cuando $\epsilon=0$.

En ambos planteamientos se ha encontrado al menos una solución a (a.1.2) (siendo la misma $\epsilon=0$). Como en el Apéndice A.2 se demuestra la convexidad estricta de la superficie bajo estudio, la solución anterior ha de ser un mínimo y además la única solución que anule su gradiente.

Como según (iii.1.6) y (iii.1.7), ϵ se puede expresar:

$$\epsilon = \left(\frac{c}{p} - c \right) \underline{x} \quad (\text{a.1.10})$$

la solución $\epsilon=0$, válida para cualquier \underline{x} , fuerza:

$$\frac{c}{\text{min}} = \frac{c}{p} \quad (\text{a.1.11})$$

y sustituyendo esta última expresión en (a.1.1) se obtiene:

$$f_{\text{min}}(\underline{c}) = E(|u|^k) \quad (\text{a.1.12})$$

A.2- ESTUDIO DE LA CONVEXIDAD DE LA FUNCION OBJETIVO

Para el estudio de la convexidad de $f(\underline{c})$ se analiza el carácter positivo de su Hessiano [RAO, 1984] definido como:

$$\underline{H}(\underline{c}) = \underline{v} \underline{v}^t (f(\underline{c})) \quad (\text{a.2.1})$$

sustituyendo $f(\underline{c})$ por (a.1.1) se obtiene para el elemento i, j de \underline{H} :

$$h_{ij} = \frac{\partial}{\partial c_i \partial c_j} E(|u + (\underline{c}_p - \underline{c})^t \underline{x}|^k) \quad (1 \leq i, j \leq N) \quad (\text{a.2.2})$$

cambiando el orden de la diferenciación por el del valor medio y expresando este en forma condicionada a x , se obtiene:

$$h_{ij} = E\left(E\left(\frac{\partial}{\partial c_i \partial c_j} |u + (\underline{c}_p - \underline{c})^t \underline{x}|^k / x\right)\right) \quad (\text{a.2.3})$$

diferenciando una vez:

$$h_{ij} = -E\left(x_i \frac{\partial}{\partial c_j} E\left((u + (\underline{c}_p - \underline{c})^t |u + (\underline{c}_p - \underline{c})^t \underline{x}|^{k-2} / x\right)\right) \quad (\text{a.2.4})$$

En este punto, se divide el análisis en dos partes: casos de k par e impar. Así, utilizando (iii.1.6):

para $k = \text{par}$:

utilizando (a.1.5), la parcial anterior se puede escribir:

$$\begin{aligned} \frac{\partial}{\partial c_j} (.) &= \frac{d}{d\varepsilon} (.) \frac{d\varepsilon}{dc_j} = -x_j \frac{d\varepsilon}{dc_j} \int_{-\infty}^{\infty} (\varepsilon+u)^{k-1} f(u/x) du = \\ &= -(k-1) x_j \int_{-\infty}^{\infty} (\varepsilon+u)^{k-2} f(u/x) du \end{aligned} \quad (\text{a.2.5})$$

donde dado el carácter par de k , la integral anterior es siempre positiva. Por tanto el elemento i, j del Hessiano se puede expresar:

$$h_{ij} = E(x_i x_j A(\underline{x})) \quad (\text{a.2.6})$$

siendo $A(\underline{x})$ una función positiva del vector \underline{x} .

para k = impar (k 1):

utilizando (a.1.9) la parcial anterior se escribe:

$$\frac{\partial}{\partial c_j}(\cdot) = -x_j \frac{\partial}{d\epsilon} \left(- \int_{-\infty}^{-\epsilon} (\epsilon+u)^{k-1} f(u/x) du + \int_{-\epsilon}^{\infty} (\epsilon+u)^{k-1} f(u/x) du \right) \quad (\text{a.2.7})$$

derivando y reagrupando términos se obtiene:

$$\frac{\partial}{\partial c_j}(\cdot) = - (k-1)x_j \left(\int_{-\epsilon}^{\infty} (\epsilon+u)^{k-2} f(u/x) du - \int_{-\infty}^{-\epsilon} (\epsilon+u)^{k-2} f(u/x) du \right) \quad (\text{a.2.8})$$

donde debido al dominio de integración y las potencias de los respectivos integrandos la primera integral es positiva y la segunda negativa. Por tanto, igualmente que en el caso de k=par, el elemento (i,j) del Hessiano se puede expresar por (a.2.6).

para k = impar (k=1):

en este caso la parcial se expresa como:

$$\frac{\partial}{\partial c_j}(\cdot) = -x_j \frac{d}{d\epsilon} \left(- \int_{-\infty}^{-\epsilon} f(u/x) du + \int_{-\epsilon}^{\infty} f(u/x) du \right) \quad (\text{a.2.9})$$

derivando y reagrupando términos se obtiene.

$$\frac{\partial}{\partial c_j}(\cdot) = -2x_j f(-\epsilon/x) \quad (\text{a.2.10})$$

donde obviamente $f(-\epsilon/x)$ es positiva dado su carácter de f.d.p., y por consiguiente el elemento (i,j) del Hessiano se puede expresar según (a.2.6).

Según [GERSHO, 1969] , [GERSHO, 1984] si la matriz de covarianza de x es definida positiva lo es también la matriz compuesta por los elementos h_{ij} . Esta condición se cumple en nuestro caso, ya que según la hipótesis H1 del Capítulo III, $x(n)$ es un proceso estacionario, blanco de media nula. Por consiguiente, se concluye que el Hessiano de la superficie objetivo tiene un carácter positivo y por tanto la curva es convexa.

A.3- CALCULO DE $S_1(n)$ PARA RUIDO DE PLANTA GAUSSIANO

Definida la f.d.p. de u como:

$$f(u) = (1/\sqrt{2\pi\sigma_u^2}) \exp(-u^2/2\sigma_u^2) \quad (\text{a.3.1})$$

Primeramente, se analiza la expresión de $g_j(n)$ (iii.2.2.13).

Para ello llámese:

$$I_1(m, \epsilon) = \int_0^{\epsilon} u^m \exp(-u^2/2\sigma_u^2) du \quad (\text{a.3.2})$$

$$I_2(m, \epsilon) = \int_{\epsilon}^{\infty} u^m \exp(-u^2/2\sigma_u^2) du \quad , (m=0,1,\dots) \quad (\text{a.3.3})$$

haciendo uso de las siguientes relaciones [ABRAMOWITZ, 1970] :

$$P(y^2/v) + Q(y^2/v) = 1 \quad (\text{a.3.4})$$

$$Q(y^2/v) = F(v) \int_{y^2}^{\infty} t^{v/2-1} \exp(-t/2) dt \quad (\text{a.3.5})$$

$$F(v) = 1/(2^{v/2} \Gamma(v/2)) \quad (\text{a.3.6})$$

con $\Gamma(\cdot)$ la conocida función Gamma

Realizando en (a.3.5) el cambio $t=u^2/\sigma_u^2$ se obtiene:

$$\begin{aligned} Q(y^2/v) &= (2F(v)/\sigma_u^v) \int_{|y|\sigma_u}^{\infty} u^{v-1} \exp(-u^2/2\sigma_u^2) du = \\ &= 1 - (2F(v)/\sigma_u^v) \int_0^{|y|\sigma_u} u^{v-1} \exp(-u^2/2\sigma_u^2) du \end{aligned} \quad (\text{a.3.7})$$

la cual combinándola con (a.3.2), (a.3.3) y (a.3.4) permite escribir:

$$I_1(m, \epsilon) = 2^{(m-1)/2} \sigma_u^{m+1} \Gamma(\frac{m+1}{2}) P((\epsilon/\sigma_u)^2/m+1) \quad (\text{a.3.8})$$

$$I_2(m, \epsilon) = 2^{(m-1)/2} \sigma_u^{m+1} \Gamma(\frac{m+1}{2}) Q((\epsilon/\sigma_u)^2/m+1) \quad (\text{a.3.9})$$

donde la función Gamma $\Gamma(m+1/2)$ se puede poner como:

$$\Gamma(\frac{m+1}{2}) = \begin{cases} [(m-1)!! \sqrt{\pi}/2^{m/2}] ; m \text{ par} & (\text{a.3.10.a}) \\ (m-1)!!/2^{(m-1)/2} ; m \text{ impar} & (\text{a.3.10.b}) \end{cases}$$

Atendiendo a (a.3.10) y sustituyendo (a.3.8) y (a.3.9) en (iii.2.2.13) con $m=k-1-j$ se obtiene para $g_j[\epsilon(n)]$:

$$g_j(\epsilon) = \begin{matrix} (m-1)!! \sigma_u^m \operatorname{sgn}(\epsilon) P((\epsilon/\sigma_u)^2/m+1); & m \text{ par} & \text{(a.3.11.a)} \\ (2(m-1)!! \sigma_u^m / \sqrt{2\pi}) Q((\epsilon/\sigma_u)^2/m+1); & m \text{ impar} & \text{(a.3.11.b)} \end{matrix}$$

Según [ABRAMOWITZ, 1970] la expresión $Q(\cdot)$ toma los siguientes valores dependiendo del carácter par de m :

para m par:

$$Q[(\epsilon/\sigma_u)^2/m+1] = \frac{2}{\sqrt{\pi}} \sum_{r=1}^{m/2} (|\epsilon|/\sigma_u)^{2r-1} \exp(-\epsilon^2/2\sigma_u^2) / (2r-1)!! + 1 - \operatorname{erf}[|\epsilon|/\sqrt{2}\sigma_u] \quad \text{(a.3.12a)}$$

para m impar:

$$Q[(\epsilon/\sigma_u)^2/m+1] = \exp(-\epsilon^2/2\sigma_u^2) \left[1 + \sum_{r=1}^{(m-1)/2} (\epsilon/\sigma_u)^{2r} / (2r)!! \right] \quad \text{(a.3.12b)}$$

donde $\operatorname{erf}(\cdot)$ es la función de error definida como:

$$\operatorname{erf}(t) = (2/\sqrt{\pi}) \int_0^t \exp(-s^2) ds \quad \text{(a.3.13)}$$

con la expresión generalizada (a.3.11) se procede seguidamente al análisis de $S_1(n)$. Solo se llevarán a cabo los cálculos correspondientes a los algoritmos L_1 y L_3 , debido a que la forma de proceder con los algoritmos impares de orden superior es similar y su cómputo tedioso.

A.3.1.- Particularización para L_1

En este caso:

$$S_1(n) = 2\sigma_x^2 \phi_0(n) / \sigma_\epsilon^2(n) \quad \text{(a.3.14)}$$

con

$$\phi_0(n) = \int_{-\infty}^{\infty} \epsilon(n) g_0[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) \quad (\text{a.3.15})$$

si se sustituye en la expresión anterior (2.3.11a) con $m=0$, recordando que $f[\epsilon(n)]$ es Gaussiana y obviando de ahora en adelante en las fórmulas la dependencia de ϵ con n se obtiene:

$$\phi_0(n) = (2/\sqrt{2\pi}\sigma_\epsilon) I_A \quad (\text{a.3.16})$$

$$\text{con } I_A = \int_0^{\infty} \epsilon \text{erf}(|\epsilon|/\sqrt{2}\sigma_u) \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon \quad (\text{a.3.17})$$

desarrollado en el Apéndice A.7.

Sustituyendo, finalmente, las expresiones adecuadas se obtiene:

$$S_1(n) = 2\sqrt{2/\pi}\sigma_x^2/\sigma_u \sqrt{1+s^2} \quad (\text{a.3.18})$$

donde se ha sustituido $s^2 = \sigma_\epsilon^2(n)/\sigma_u^2$.

A.3.2.- Particularización para L_3

En este caso:

$$S_1(n) = (6\sigma_x^2/\sigma_\epsilon^2(n))(\phi_0(n)+2\phi_1(n)+\phi_2(n)) \quad (\text{a.3.19})$$

donde utilizando sucesivamente $m=0,1,2$ en (a.3.11) se obtiene:

con $m=2$

$$\begin{aligned} \phi_0(n) &= \int_{-\infty}^{\infty} \epsilon g_0(\epsilon) f(\epsilon) d\epsilon = \\ &= (2\sigma_u^2/\sqrt{2\pi}\sigma_\epsilon) \int_0^{\infty} \epsilon P[(\epsilon/\sigma_u)^2/3] \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon \end{aligned} \quad (\text{a.3.20})$$

donde según (a.13.12a) y (a.3.4) se puede reescribir como:

$$\phi_0(n) = (2\sigma_u^2/\sqrt{2\pi}\sigma_\epsilon) I_A + (2\sigma_u/\pi\sigma_\epsilon) I_B \quad (\text{a.3.21})$$

con I_A expresado en (a.3.17) e

$$I_B = \int_0^{\infty} \epsilon^2 \exp(-\epsilon^2/2(1/\sigma_\epsilon^2+1/\sigma_u^2)) d\epsilon \quad (\text{a.3.22})$$

desarrollados en el Apéndice A.7.

Igualmente (con m=1):

$$\begin{aligned}\phi_1(n) &= \int_{-\infty}^{\infty} \epsilon^2 g_1(\epsilon) f(\epsilon) d\epsilon = \\ &= (\sigma_u / \pi \sigma_\epsilon) \int_0^{\infty} \epsilon^2 Q[(\epsilon/\sigma_u)^2/2] \exp(-\epsilon^2/2 \sigma_\epsilon^2) d\epsilon\end{aligned}\quad (\text{a.3.23})$$

donde sustituyendo (a.3.12b) para m=1 se obtiene:

$$\phi_1(n) = (\sigma_u / \pi \sigma_\epsilon) I_B \quad (\text{a.3.24})$$

con I_B expresado en (a.3.22).

Finalmente, con m=2:

$$\begin{aligned}\phi_2(n) &= \int_{-\infty}^{\infty} \epsilon^3 g_2(\epsilon) f(\epsilon) d\epsilon = \\ &= (2/\pi/\sigma_\epsilon) \int_0^{\infty} \epsilon^3 P[(\epsilon/\sigma_u)^2/1] \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon\end{aligned}\quad (\text{a.3.25})$$

y sustituyendo (a.3.12a) para m=0

$$\phi_2(n) = (\sqrt{2/\pi}/\sigma_\epsilon) I_C \quad (\text{a.3.26})$$

donde I_C es desarrollado en el Apéndice A.7. y se expresa:

$$I_C = \int_0^{\infty} \epsilon^3 \operatorname{erf}(\epsilon/\sqrt{2} \sigma_u) \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon \quad (\text{a.3.27})$$

Reagrupando términos según (a.3.19) y sustituyendo las expresiones para I_A , I_B e I_C calculadas en el Apéndice A.7., se obtiene para $S_1(n)$:

$$S_1(n) = 6\sqrt{2/\pi} \sigma_u^2 [1+2s^2+s^2/(1+s^2)] / \sqrt{1+s^2} \quad (\text{a.3.28})$$

donde nuevamente $s^2 = \sigma_\epsilon^2(n) / \sigma_u^2$

A.4.- CALCULO DE $S_1(n)$ PARA RUIDO DE PLANTA LAPLACIANO

Definida la f.d.p. de u como:

$$f(u) = (1/\sqrt{2} \sigma_u) \exp(-\sqrt{2}|u|/\sigma_u) \quad (\text{a.4.1})$$

se evalúa la expresión de $g_j(n)$ (iii.2.2.13) por medio de sus integrales:

$$I_1(m, \epsilon) = \int_0^{|\epsilon|} u^m \exp(-u/\lambda) du \quad (\text{a.4.2})$$

$$I_2(m, \epsilon) = \int_{|\epsilon|}^{\infty} u^m \exp(-u/\lambda) du, \quad (m=0,1,\dots) \quad (\text{a.4.3})$$

llamando

$$r = u^m \quad (\text{a.4.4a})$$

$$ds = \exp(-u/\lambda) du \quad (\text{a.4.4b})$$

y derivando e integrando respectivamente se obtiene

$$dv = m u^{m-1} \quad (\text{a.4.5a})$$

$$s = -\lambda \exp(-u/\lambda) \quad (\text{a.4.5b})$$

con lo que $I_1(m, \epsilon)$ e $I_2(m, \epsilon)$ resultan al integrar por partes:

$$I_1(m, \epsilon) = -\lambda u^m \exp(-\epsilon/\lambda) + \lambda m I_1(m-1, \epsilon) \quad (\text{a.4.6})$$

$$I_2(m, \epsilon) = \lambda u^m \exp(-\epsilon/\lambda) + \lambda m I_2(m-1, \epsilon) \quad (\text{a.4.7})$$

dos expresiones recurrentes en m.

Las expresiones iniciales son respectivamente

$$I_1(0, \epsilon) = \int_0^{|\epsilon|} \exp(-u/\lambda) du = \lambda - \exp(-|\epsilon|/\lambda) \quad (\text{a.4.8})$$

$$I_2(0, \epsilon) = \int_{|\epsilon|}^{\infty} \exp(-u/\lambda) du = \lambda \exp(-|\epsilon|/\lambda) \quad (\text{a.4.9})$$

En el Apéndice A.8 se obtienen expresiones cerradas, no recurrentes, para $I_1(m, \epsilon)$ e $I_2(m, \epsilon)$. Estas son:

$$I_1(m, \epsilon) = \lambda^{m+1} m!^{-m!w} \sum_{r=0}^m \lambda^{m+1-r} |\epsilon|^r / (r!) \quad (\text{a.4.10})$$

$$I_2(m, \epsilon) = m!w \sum_{r=0}^m \lambda^{m+1-r} |\epsilon| / (r!) \quad (\text{a.4.11})$$

$$\text{donde } w = \exp(-|\epsilon|/\lambda) \quad (\text{a.4.12})$$

De estas expresiones se obtiene la final para $g_j[\epsilon(n)]$ tan solo sustituyendo $\lambda = \sigma_u / \sqrt{2}$ y teniendo presente que $m = k-1-j$. Esta es:

para m par:

$$g_j(\epsilon) = m! \text{sgn}(\epsilon) \left[\sigma_u^m / 2^{m/2} - \exp(-|\epsilon| \sqrt{2} / \sigma_u) \sum_{r=0}^m \sigma_u^{m-r} |\epsilon|^r / (r! 2^{(m-r)/2}) \right] \quad (\text{a.4.13a})$$

para m impar:

$$g_j(\epsilon) = m! \exp(-|\epsilon| \sqrt{2} / \sigma_u) \sum_{r=0}^m \sigma_u^{m-r} |\epsilon|^r / (r! 2^{(m-r)/2}) \quad (\text{a.4.13b})$$

con estas expresiones se procede al cálculo de $S_1(n)$, particularizado para los algoritmos L_1 y L_3 .

A.4.1.- Particularización para L_1

En este caso: $S_1(n) = 2\sigma_x^2 \phi_0(n) / \sigma_\epsilon^2(n)$ con:

$$\phi_0(n) = \int_{-\infty}^{\infty} \epsilon(n) g_0(n) f[\epsilon(n)] d\epsilon(n) \quad (\text{a.4.14})$$

si se sustituye (a.4.13a) con $m=0$ se obtiene:

$$\phi_0(n) = 2 / (\sqrt{2\pi} \sigma_\epsilon) \int_0^{\infty} \epsilon (1 - \exp(-\epsilon \sqrt{2} / \sigma_u)) \exp(-\epsilon^2 / 2\sigma_\epsilon^2) d\epsilon \quad (\text{a.4.15a})$$

donde desarrollando e integrando

$$\phi_0(n) = \sqrt{2/\pi} \sigma_\epsilon - (\sqrt{2/\pi}/\sigma_\epsilon) I_D \quad (\text{a.4.15b})$$

$$\text{con } I_D = \int_0^\infty \epsilon \exp(-\epsilon\sqrt{2}/\sigma_u - \epsilon^2/2\sigma_\epsilon^2) d\epsilon \quad (\text{a.4.16})$$

desarrollado en el Apéndice A.8.

Sustituyendo el valor de I_D en (a.4.15b) se obtiene finalmente para $S_1(n)$:

$$S_1(n) = (2\sqrt{2}\sigma_x^2/\sigma_u) [1-\text{erf}(s)] \exp(s^2) \quad (\text{a.4.17})$$

$$\text{con } s^2 = \sigma_\epsilon^2(n)/\sigma_u^2.$$

A.4.2.- Particularización para L_3

Para este caso:

$$S_1(n) = (6\sigma_x^2/\sigma_\epsilon^2(n)) (\phi_0(n)+2\phi_1(n)+\phi_2(n)) \quad (\text{a.4.18})$$

sustituyendo sucesivamente $m=0,1,2$ en (a.4.13) se obtiene:

para $m=2$

$$\begin{aligned} \phi_0(n) &= \int_{-\infty}^{\infty} \epsilon(n) g_0[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) = \\ &= (2\sigma^2/\sqrt{2\pi}\sigma_\epsilon) \int_{-0}^{\infty} \epsilon(1-\exp(-\epsilon\sqrt{2}/\sigma_u)) \exp(-\epsilon^2/2\sigma_\epsilon^2) - \\ &\quad - (2\sigma_u^2/\sqrt{\pi}\sigma_\epsilon) \int_{-0}^{\infty} \epsilon^2 \exp(-M(\epsilon)) d\epsilon - (\sqrt{2/\pi}/\sigma_\epsilon) \int_0^\infty \exp(-M(\epsilon)) d\epsilon \end{aligned} \quad (\text{a.4.19})$$

donde por simplicidad en la notación se ha llamado

$$M(\epsilon) = \epsilon\sqrt{2}/\sigma_u + \epsilon^2/2\sigma_\epsilon^2 \quad (\text{a.4.20})$$

Desarrollando (a.4.19):

$$\phi_0(n) = \sqrt{2/\pi}\sigma_u^2\sigma_\epsilon - (\sqrt{2/\pi}/\sigma_\epsilon) (\sigma_u^2 I_D + \sqrt{2}\sigma_u I_E + I_F) \quad (\text{a.4.21})$$

con

$$I_E = \int_0^{\infty} \epsilon^2 \exp(-M(\epsilon)) d\epsilon \quad (\text{a.4.22})$$

$$I_F = \int_0^{\infty} \epsilon^3 \exp(-M(\epsilon)) d\epsilon \quad (\text{a.4.23})$$

desarrolladas en el Apéndice A.8

para $m=1$

$$\begin{aligned} \phi_1(n) &= \int_{-\infty}^{\infty} \epsilon^2(n) g_1(\epsilon(n)) f[\epsilon(n)] d\epsilon(n) = \\ &= (\sigma_u / \sqrt{\pi} \sigma_\epsilon) \int_0^{\infty} \epsilon^2 \exp(-M(\epsilon)) d\epsilon + (\sqrt{2/\pi} / \sigma_\epsilon) \int_0^{\infty} \epsilon^3 \exp(-M(\epsilon)) d\epsilon \end{aligned} \quad (\text{a.4.24})$$

que se puede escribir según el Apéndice A.8:

$$\phi_1(n) = (2/\sqrt{2\pi} \sigma_\epsilon) [(\sigma_u/\sqrt{2}) I_E + I_F] \quad (\text{a.4.25})$$

para $m=2$

$$\begin{aligned} \phi_2(n) &= \int_{-\infty}^{\infty} \epsilon^3(n) g_2[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) = \\ &= (\sqrt{2/\pi} / \sigma_\epsilon) \int_0^{\infty} \epsilon^3 \exp[-\epsilon^2/2\sigma_\epsilon^2] d\epsilon - (\sqrt{2/\pi} / \sigma_\epsilon) \int_0^{\infty} \epsilon^3 \exp(-M(\epsilon)) d\epsilon \end{aligned} \quad (\text{a.4.26})$$

donde igualmente se puede escribir:

$$\phi_2(n) = \frac{4\sigma_\epsilon^3}{\sqrt{2\pi}} - (\sqrt{2/\pi} / \sigma_\epsilon) I_F \quad (\text{a.4.27})$$

Reagrupando los términos según (a.4.18) y utilizando los resultados para I_D , I_E e I_F según el Apéndice A.8, se obtiene finalmente para S_1 , la siguiente expresión:

$$S_1 = 6\sqrt{2}\sigma_x^2 \sigma_u [2s/\sqrt{\pi} + (1-\text{erf}(s))\exp s^2] \quad (\text{a.4.28})$$

donde nuevamente $s^2 = \sigma_\epsilon^2(n) / \sigma_u^2$.

A.5.- CALCULO DE $S_1(n)$ PARA RUIDO DE PLANTA UNIFORME

Definida la f.d.p. de u como

$$f(u) = \begin{cases} 1/2 \sqrt{3} \sigma_u & ; |u| \leq \sqrt{3} \sigma_u \\ 0 & ; |u| > \sqrt{3} \sigma_u \end{cases} \quad \begin{matrix} \text{(a.5.1a)} \\ \text{(a.5.1b)} \end{matrix}$$

se evalúa la expresión de $g_j[\epsilon(n)]$ (iii.2.2.13) por medio de sus dos integrales en los siguientes rangos de ϵ .

1. con $|\epsilon| < \sqrt{3} \sigma_u$

$$I_1(m, \epsilon) = \int_0^{|\epsilon|} u^m / 2 \sqrt{3} \sigma_u du = u^{m+1} / 2(m+1) \sqrt{3} \sigma_u \quad \text{(a.5.2a)}$$

$$I_2(m, \epsilon) = \int_{|\epsilon|}^{\sqrt{3} \sigma_u} u^m / 2 \sqrt{3} \sigma_u du = [(\sqrt{3} \sigma_u)^{m+1} - |\epsilon|^{m+1}] / 2(m+1) \sqrt{3} \sigma_u \quad \text{(a.5.2b)}$$

con estos resultados:

$$g_j(\epsilon) = \begin{cases} \text{sgn}(\epsilon) |\epsilon|^{m+1} / (m+1) \sqrt{3} \sigma_u & ; m=\text{par} \\ [(\sqrt{3} \sigma_u)^{m+1} - |\epsilon|^{m+1}] / (m+1) \sqrt{3} \sigma_u & ; m=\text{impar} \end{cases} \quad \begin{matrix} \text{(a.5.3a)} \\ \text{(a.5.3b)} \end{matrix}$$

2. con $|\epsilon| > \sqrt{3} \sigma_u$

$$I_1(m, \epsilon) = \int_0^{\sqrt{3} \sigma_u} u^m / 2 \sqrt{3} \sigma_u du = (\sqrt{3} \sigma_u)^{m+1} / 2(m+1) \quad \text{(a.5.4a)}$$

$$I_2(m, \epsilon) = \int_{|\epsilon|}^{\infty} u^m / 2 \sqrt{3} \sigma_u du = 0 \quad \text{(a.5.4b)}$$

De donde se obtiene

$$g_j(\epsilon) = \begin{cases} \text{sgn}(\epsilon) (\sqrt{3} \sigma_u)^m / (m+1) & ; m=\text{par} \\ 0 & ; m=\text{impar} \end{cases} \quad \begin{matrix} \text{(a.5.5a)} \\ \text{(a.5.1b)} \end{matrix}$$

con las expresiones (a.5.3) y (a.5.5) se procede al cálculo de $S_1(n)$, particularizando para los algoritmos L_1 y L_3 .

A.5.1.- Particularización para L_1

En este caso $S_1(n) = 2\sigma_x^2\phi_0(n)/\sigma_\epsilon^2(n)$ con:

$$\phi_0(n) = \int_{-\infty}^{\infty} \epsilon(n) g_0[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) \quad (\text{a.5.6})$$

donde sustituyendo las expresiones adecuadas de $g_0[\epsilon(n)]$ con $m=0$ se obtiene:

$$\phi_0(n) = (\sqrt{2/\pi}/\sigma_\epsilon) [(1/\sqrt{3}\sigma_u) I_H + I_G] \quad (\text{a.4.7})$$

donde según el Apéndice A.9.

$$I_G = \int_{-\infty}^{\infty} \frac{\epsilon \exp(-\epsilon^2/2\sigma_\epsilon^2)}{\sqrt{3}\sigma_u} d\epsilon \quad (\text{a.5.8})$$

$$I_H = \int_0^{\sqrt{3}\sigma_u} \epsilon^2 \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon \quad (\text{a.5.9})$$

sustituyendo los resultados de estas integrales se obtiene para $S_1(n)$:

$$S_1(n) = 2\sigma_x^2 \text{erf}(\sqrt{3}/2s^2)/\sqrt{3}\sigma_u \quad (\text{a.5.10})$$

A.5.2.- Particularización para L_3

Para este algoritmo:

$$S_1(n) = (6\sigma_x^2/\sigma_\epsilon^2(n)) [\phi_0(n) + 2\phi_1(n) + \phi_2(n)] \quad (\text{a.5.11})$$

sustituyendo sucesivamente $m=0,1$, y 2 en (a.5.3) y (a.5.5) se obtiene:

para $m=2$

$$\begin{aligned} \phi_0(n) &= \int_{-\infty}^{\infty} \epsilon(n) g_0[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) = \\ &= (\sqrt{2/\pi}/\sigma_\epsilon) [(1/3\sqrt{3}\sigma_u) I_j + \sigma_u^2 I_G] \end{aligned} \quad (\text{a.5.12})$$

donde según el Apéndice A.9:

$$I_J = \int_0^{\sqrt{3}\sigma_u} \epsilon^4 \exp(-\epsilon^2/2 \sigma_\epsilon^2) d\epsilon \quad (\text{a.5.13})$$

e I_G esta definida en (a.5.8)

para $m=1$

$$\begin{aligned} \phi_1(n) &= \int_{-\infty}^{\infty} \epsilon^2(n) g_1[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) = \\ &= (1/\sqrt{2\pi}\sigma_\epsilon) [\sqrt{3}\sigma_u I_G - (1/\sqrt{3}\sigma_u) I_J] \end{aligned} \quad (\text{a.5.14})$$

donde I_G e I_J han sido ya definidas

para $m=0$

$$\begin{aligned} \phi_2(n) &= \int_{-\infty}^{\infty} \epsilon^3(n) g_2[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) = \\ &= (\sqrt{2/\pi}/\sigma_\epsilon) [(1/\sqrt{3}\sigma_u) I_J + I_I] \end{aligned} \quad (\text{a.5.15})$$

con I_I definida en el Apéndice A.9. como:

$$I_I = \int_{\sqrt{3}\sigma_u}^{\infty} \epsilon^3 \exp(-\epsilon^2/2 \sigma_\epsilon^2) d\epsilon \quad (\text{a.5.16})$$

Reagrupando los términos $\phi_0(n)$, $\phi_1(n)$, y $\phi_2(n)$ según expresa (a.5.11) y sustituyendo los resultados para I_G , I_H , I_I e I_J según el Apéndice A.9 se obtiene finalmente:

$$S_1(n) = 6\sigma_x^2 \sigma_u [(\sqrt{3}+s^2/\sqrt{3})\text{erf}(\sqrt{3}/2s^2) + \sqrt{2/\pi}\exp(-3s^2/2)] \quad (\text{a.5.17})$$

donde $s^2 = \sigma_\epsilon^2(n) / \sigma_u^2$.

A.6.- CALCULO DE $S_1(n)$ PARA RUIDO DE PLANTA BINARIO

Definida la f.d.p. de u como:

$$f(u) = [\delta(u+\sigma_u) + \delta(u-\sigma_u)]/2 \quad (\text{a.6.1})$$

se evalúa la expresión de $g_j[\epsilon(n)]$ (iii.2.2.13) por medio de sus dos integrales para los siguientes rangos de ϵ :

1. con $|\epsilon| < \sigma_u$

$$I_1(m, \epsilon) = \int_0^{|\epsilon|} u^m [\delta(u+\sigma_u) + \delta(u-\sigma_u)] / 2 du = 0 \quad (\text{a.6.2a})$$

$$I_2(m, \epsilon) = \int_{|\epsilon|}^{\infty} u^m [\delta(u+\sigma_u) + \delta(u-\sigma_u)] / 2 du = \sigma_u^m / 2 \quad (\text{a.6.2b})$$

con estos resultados:

$$g_j(\epsilon) = \begin{cases} 0 & ; m=\text{par} \\ \sigma_u^m & ; m=\text{impar} \end{cases} \quad (\text{a.6.3a})$$

$$(\text{a.6.3b})$$

2. con $|\epsilon| > \sigma_u$

$$I_1(m, \epsilon) = \int_0^{|\epsilon|} u^m [\delta(u+\sigma_u) + \delta(u-\sigma_u)] / 2 du = \sigma_u^m / 2 \quad (\text{a.6.4a})$$

$$I_2(m, \epsilon) = \int_{|\epsilon|}^{\infty} u^m [\delta(u+\sigma_u) + \delta(u-\sigma_u)] / 2 du = 0 \quad (\text{a.6.4b})$$

con estos resultados:

$$g_j(\epsilon) = \begin{cases} \text{sgn}(\epsilon) \sigma_u^m & ; m=\text{par} \\ 0 & ; m=\text{impar} \end{cases} \quad (\text{a.6.5a})$$

$$(\text{a.6.5b})$$

con estas expresiones para $g_j[\epsilon(n)]$ se procede al cálculo de $S_1(n)$, particularizando para los algoritmos L_1 y L_3 .

A.6.1. Particularización para L_1

En este caso $S_1(n) = 2\sigma_x^2 \phi_0(n) / \sigma_\epsilon^2(n)$ con:

$$\phi_0(n) = \int_{-\infty}^{\infty} \epsilon(n) g_0[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) \quad (\text{a.6.6})$$

donde sustituyendo las expresiones adecuadas de $g_0[\epsilon(n)]$ con $m=0$ se obtiene:

$$\phi_0(n) = (\sqrt{2/\pi}/\sigma_\epsilon) I_G \quad (\text{a.6.7})$$

donde según el Apéndice A.9

$$I_G = \int_{\sigma_u}^{\infty} \epsilon \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon \quad (\text{a.6.8})$$

sustituyendo el resultado de esta integral en la expresión de $S_1(n)$ se obtiene:

$$S_1(n) = (2\sqrt{2/\pi} \sigma_x^2 / \sigma_u) \exp(-s^2/2) / s \quad (\text{a.6.9})$$

$$\text{con } s^2 = \sigma_\epsilon^2(n) / \sigma_u^2$$

A.6.2- Particularización para L_3

Para este algoritmo

$$S_1(n) = (6\sigma_x^2 / \sigma_\epsilon^2(n)) [\phi_0(n) + 2\phi_1(n) + \phi_2(n)] \quad (\text{a.6.10})$$

Sustituyendo sucesivamente $m=0,1$ y 2 en (a.6.3) y (a.6.5) se obtiene:

para $m=2$

$$\phi_0(n) = \int_{-\infty}^{\infty} \epsilon(n) g_0[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) = (\sqrt{2/\pi} \sigma_u^2 / \sigma_\epsilon) I_G \quad (\text{a.6.11})$$

donde I_G fue definida en (a.6.8)

para $m=1$

$$\phi_1(n) = \int_{-\infty}^{\infty} \epsilon^2(n) g_1[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) = (\sqrt{2/\pi} \sigma_u / \sigma_\epsilon) I_H \quad (\text{a.6.12})$$

donde según el Apéndice A.9:

$$I_H = \int_{\sigma_u}^{\infty} \epsilon^2 \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon \quad (\text{a.6.13})$$

para $m=0$

$$\phi_2(n) = \int_{-\infty}^{\infty} \epsilon^3(n) g_2[\epsilon(n)] f[\epsilon(n)] d\epsilon(n) = (2/\pi/\sigma_\epsilon) I_I \quad (\text{a.6.14})$$

con I_I definida nuevamente en el Apéndice A.9 como:

$$I_I = \int_{\sigma_u}^{\infty} \epsilon^3 \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon \quad (\text{a.6.15})$$

si se reagrupan los términos $\phi_0(n)$, $\phi_1(n)$ y $\phi_2(n)$ según expresa (a.6.10) y se sustituyen los resultados de integrar I_G , I_H e I_I se obtiene para $S_1(n)$:

$$S_1(n) = 12\sigma_x^2 \sigma_u \left[\text{erf} \left(\frac{1}{2s} + \sqrt{2/\pi} s \right) \exp(-1/2s^2) \right] \quad (\text{a.6.16})$$

A.7.- CALCULO DE LAS INTEGRALES I_A , I_B E I_C UTILIZADAS EN EL APENDICE A.3

Se define

$$I_n(a,b) = \int_0^\infty x^{n+1} \operatorname{erf}(ax) \exp(-b^2 x^2) dx \quad (\text{a.7.1})$$

$$(n=0,2,\dots)$$

como $\operatorname{erf}(.) = 2/\sqrt{\pi} \int e^{-t^2} dt$, la expresión anterior queda:

$$I_n(a,b) = 2/\sqrt{\pi} \int_0^\infty x^{n+1} \left[\int_0^{ax} \exp(-t^2) dt \right] \exp(-b^2 x^2) dx \quad (\text{a.7.2})$$

Llamando

$$u = x^n \int_0^{ax} \exp(-t^2) dt \quad (\text{a.7.3a})$$

$$dv = x \exp(-b^2 x^2) dx \quad (\text{a.7.3b})$$

resulta al derivar e integrar respectivamente:

$$du = \left[(\sqrt{\pi}/2) n x^{n-1} \operatorname{erf}(ax) + a x^n \exp(-a^2 x^2) \right] dx \quad (\text{a.7.4a})$$

$$v = - \frac{1}{2b^2} \exp(-b^2 x^2) \quad (\text{a.7.4b})$$

con lo que integrando (a.7.2) por partes:

$$I_n(a,b) = (1/\sqrt{\pi b^2}) \left[n\sqrt{\pi}/2 \int_0^\infty x^{n-1} \operatorname{erf}(ax) \exp(-b^2 x^2) dx + \right. \\ \left. + a \int_0^\infty x^n \exp(-(a^2+b^2)x^2) dx \right] \quad (\text{a.7.5})$$

Donde se observa que la primera integral es $I_{n-2}(a,b)$ y la segunda se deduce fácilmente a partir del momento central n-simo de una f.p.d. Gaussiana. Esto es, haciendo para la segunda integral el cambio $x^2(a^2+b^2) = y^2$ se obtiene:

$$\int_0^\infty x^n \exp(-(a^2+b^2)x^2) dx = (1/(a^2+b^2)^{(n+1)/2}) \int_0^\infty y^n \exp(-y^2/2) dy \quad (\text{a.7.6})$$

Pero según [PAPOULIS, 1984], para v.a. gaussiana, con n par $E\{y^n\} = (n-1)!! \sigma^n$ con lo que, considerando $\sigma=1$ se puede escribir

$$I_n(a,b) = \frac{n}{2b^2} I_{n-2}(a,b) + a(n-1)!! / (2b^2)^{n/2} (a^2+b^2)^{(n+1)/2} \quad (a.7.7)$$

siendo el valor inicial

$$I_0(a,b) = (a/\sqrt{\pi}b^2) \int_0^\infty \exp(-(a^2+b^2)x^2) dx = a/(2b^2 \sqrt{a^2+b^2}) \quad (a.7.8)$$

con estas ecuaciones recurrentes, el cálculo de las integrales I_A e I_C es inmediato.

$$I_A = \int_0^\infty \epsilon \operatorname{erf}(\epsilon/\sqrt{2}\sigma_u) \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon = I_0(1/\sqrt{2}\sigma_u, 1/\sqrt{2}\sigma_\epsilon) \quad (a.7.9)$$

$$I_C = \int_0^\infty \epsilon^3 \operatorname{erf}(\epsilon/\sqrt{2}\sigma_u) \exp(-\epsilon^2/2\sigma_\epsilon^2) d\epsilon = I_2(1/\sqrt{2}\sigma_u, 1/\sqrt{2}\sigma_\epsilon) \quad (a.7.10)$$

Sustituyendo adecuadamente, se obtiene:

$$I_A = (\sigma_\epsilon^3/\sigma_u) / \sqrt{1+\sigma_\epsilon^2/\sigma_u^2} \quad (a.7.11)$$

$$\begin{aligned} I_C &= 2\sigma_\epsilon^2 I_0(1/\sqrt{2}\sigma_u, 1/\sqrt{2}\sigma_\epsilon) + (\sigma_\epsilon^2/2\sigma_u) / (1/2\sigma_\epsilon^2 + 1/2\sigma_u^2)^{3/2} = \\ &= (\sigma_\epsilon^5/\sigma_u) [2 + 1/(1+\sigma_\epsilon^2/\sigma_u^2)] / \sqrt{1+\sigma_\epsilon^2/\sigma_u^2} \end{aligned} \quad (a.7.12)$$

Definiendo igualmente:

$$I'_n(a) = \int_0^\infty x^n \exp(-a^2 x^2) dx \quad (n=2,4,\dots) \quad (a.7.13)$$

y llamando $t = 2ax$ se puede expresar (a.7.13) como:

$$I'_n(a) = (1/2)^{(n+1)/2} a^{n+1} \int_0^\infty t^n \exp(-t^2/2) dt \quad (a.7.14)$$

donde deduciendo nuevamente la integral anterior del momento central n -simo de una f.p.d. gaussiana centrada en el origen se obtiene para $I'_n(a)$:

$$I'_n(a) = \sqrt{\pi} (n-1)!! / (2)^{(n/2+1)} a^{n+1} \quad (a.16.15)$$

Expresión no recurrente que depende exclusivamente de a y n .

La integral buscada I_B puede expresarse en términos de $I_n(a)$ como:

$$I_B = \int_0^\infty \epsilon^2 \exp(-\epsilon^2/2(\sigma_\epsilon^2 + \sigma_u^2)) d\epsilon = I_2'(\sqrt{1/2\sigma_\epsilon^2 + 1/2\sigma_u^2}) \quad (\text{a.17.16})$$

con lo que utilizando (a.17.15)

$$I_B = \sqrt{\pi}/2\sigma_\epsilon^3 / (1 + \sigma_\epsilon^2/\sigma_u^2)^{3/2} \quad (\text{a.17.17})$$

A.8- CALCULO DE LAS INTEGRALES I_D , I_E E I_F UTILIZADAS EN EL APENDICE A.4

Se define:

$$I_n(a,b) = \int_0^{\infty} x^n \exp[-b^2(ax+x^2)] dx \quad (a.8.1)$$

completando el cuadrado del exponente queda:

$$I_n(a,b) = \exp(a^2b^2/4) \int_0^{\infty} x^n \exp[-(bx+ab/2)^2] dx \quad (a.8.2)$$

haciendo el cambio $bx=u$, la anterior igualdad resulta:

$$I_n(a,b) = (1/b^{n+1}) \exp(a^2b^2/4) \int_0^{\infty} u^n \exp[-(u+ab/2)^2] du \quad (a.8.3)$$

Pero según [ABRAMOWITZ, 1970] (pag. 299), se cumple la siguiente relación integral para la función de error complementaria [erfc(.)=1-erf(.)]:

$$(2/\sqrt{\pi}n!) \int_z^{\infty} (t-z)^n \exp(-t^2) dt = i^n \operatorname{erfc}(z) \quad (a.8.4)$$

donde además, se puede expresar de forma recursiva: (véase nueva [ABRAMOWITZ, 1970] pag. 299):

$$i^n \operatorname{erfc}(ab/2) = -(ab/2n) i^{n-1} \operatorname{erfc}(ab/2) + (1/2n) i^{n-2} \operatorname{erfc}(ab/2) \quad (a.8.5)$$

($n \geq 1$)

con las condiciones iniciales:

$$i^0 \operatorname{erfc}(ab/2) = \operatorname{erfc}(ab/2) = 1 - \operatorname{erf}(ab/2) \quad (a.8.6a)$$

$$i^{-1} \operatorname{erfc}(ab/2) = (2/\sqrt{\pi}) \exp(-a^2b^2/4) \quad (a.8.6b)$$

Si en (a.8.4) se hace el cambio $t-z=u$, se obtiene:

$$i^n \operatorname{erfc}(z) = (2/\sqrt{\pi}n!) \int_0^{\infty} u^n \exp[-(u+z)^2] du \quad (a.8.7)$$

y si se combinan (a.8.3) y (a.8.7) resulta finalmente para $I_n(a,b)$:

$$I_n(a,b) = (n! \sqrt{\pi} / 2b^{n+1}) \exp(a^2b^2/4) i^n \operatorname{erfc}(ab/2) \quad (a.8.8)$$

con esta expresión las integrales referenciadas en el Apéndice A.4. quedan:

$$I_D = \int_0^{\infty} \epsilon \exp[-(\sqrt{2}\epsilon/\sigma_u + \epsilon^2/2\sigma_\epsilon^2)] d\epsilon = I_1(2\sqrt{2}\sigma_\epsilon^2/\sigma_u, 1/\sqrt{2}\sigma_\epsilon) \quad (\text{a.8.9})$$

$$I_E = \int_0^{\infty} \epsilon^2 \exp[-(\sqrt{2}\epsilon/\sigma_u + \epsilon^2/2\sigma_\epsilon^2)] d\epsilon = I_2(2\sqrt{2}\sigma_\epsilon^2/\sigma_u, 1/\sqrt{2}\sigma_\epsilon) \quad (\text{a.8.10})$$

$$I_F = \int_0^{\infty} \epsilon^3 \exp[-(\sqrt{2}\epsilon/\sigma_u + \epsilon^2/2\sigma_\epsilon^2)] d\epsilon = I_3(2\sqrt{2}\sigma_\epsilon^2/\sigma_u, 1/\sqrt{2}\sigma_\epsilon) \quad (\text{a.8.11})$$

Aplicando la expresión recursiva (a.8.5) con las condiciones iniciales (a.8.6) se obtiene:

$$I_D = \sigma_\epsilon^2 - \sqrt{\pi} \sigma_\epsilon^3 \exp(\sigma_\epsilon^2/\sigma_u^2) \operatorname{erfc}(\sigma_\epsilon/\sigma_u) / \sigma_u \quad (\text{a.8.12})$$

$$I_E = \sqrt{2\pi} \sigma_\epsilon^3 (1/2 + \sigma_\epsilon^2/\sigma_u^2) \exp(\sigma_\epsilon^2/\sigma_u^2) \operatorname{erfc}(\sigma_\epsilon/\sigma_u) - \sqrt{2} \sigma_\epsilon^4 / \sigma_u \quad (\text{a.8.13})$$

$$I_F = -2\sqrt{\pi} \sigma_\epsilon^4 (3\sigma_\epsilon/2\sigma_u + \sigma_u^3/\sigma_\epsilon^3) \exp(\sigma_\epsilon^2/\sigma_u^2) \operatorname{erfc}(\sigma_\epsilon/\sigma_u) + 2\sigma_\epsilon^4 (1 + \sigma_\epsilon^2/\sigma_u^2) \quad (\text{a.8.14})$$

A.9- CALCULO DE LAS INTEGRALES I_G , I_H , I_I , E I_J UTILIZADAS EN LOS APENDICES A.5 Y A.6

Se define

$$I_n(a,b) = \int_a^\infty x^{n+1} \exp(-b^2 x^2) dx \quad (n=0,2,\dots) \quad (a.9.1)$$

haciendo los cambios:

$$u=x^n \quad (a.9.2a)$$

$$dv=x \exp(-b^2 x^2) dx \quad (a.9.2b)$$

se obtiene integrando por partes:

$$I_n(a,b) = (a^n/2b^2) \exp(-a^2 b^2) + (n/2b^2) \int_a^\infty x^{n-1} \exp(-b^2 x^2) dx \quad (a.9.3)$$

o igualmente:

$$I_n(a,b) = (a^n/2b^2) \exp(-a^2 b^2) + (n/2b^2) I_{n-2}(a,b) \quad (a.9.4)$$

con

$$I_0(a,b) = \int_a^\infty x \exp(-b^2 x^2) dx = (1/2b^2) \exp(-a^2 b^2) \quad (a.9.5)$$

Con esta expresión recursiva, las integrales I_G e I_I quedan:

$$I_G = \int_a^\infty \epsilon \exp(-\epsilon^2/2 \sigma_\epsilon^2) = I_0(a, 1/\sqrt{2} \sigma_\epsilon) \quad (a.9.6)$$

$$I_I = \int_a^\infty \epsilon^3 \exp(-\epsilon^2/2 \sigma_\epsilon^2) = I_2(a, 1/\sqrt{2} \sigma_\epsilon) \quad (a.9.7)$$

donde para el Apéndice A.5 $a = \sqrt{3} \sigma_u$, y para el Apéndice A.6. $a = \sigma_u$.

Sustituyendo adecuadamente resulta:

Para el Apéndice A.5:

$$I_G = \sigma_\epsilon^2 \exp(-3 \sigma_u^2/2 \sigma_\epsilon^2) \quad (a.9.8)$$

$$I_I = (2 + 3\sigma_u^2 / \sigma_\epsilon^2) \sigma_\epsilon^4 \exp(-3\sigma_u^2 / 2\sigma_\epsilon^2) \quad (\text{a.9.9})$$

Para el Apéndice A.6:

$$I_G = \sigma_\epsilon^2 \exp(-\sigma_u^2 / 2\sigma_\epsilon^2) \quad (\text{a.9.10})$$

$$I_I = (2 + \sigma_u^2 / \sigma_\epsilon^2)^4 \exp(-\sigma_u^2 / 2\sigma_\epsilon^2) \quad (\text{a.9.11})$$

Definiendo igualmente:

$$I'_n(a, b) = \int_0^a x^n e^{-bx^2} dx \quad (n=2, 4, \dots) \quad (\text{a.9.12})$$

haciendo el cambio $(t/\sqrt{2})=bx$ se obtiene:

$$I'_n(a, b) = (1/2)^{(n+1)/2} b^{(n+1)} \int_0^{ab} t^n \exp(-t^2/2) dt \quad (\text{a.9.13})$$

y según [ABRAMOWITZ, 1970] (pag. 942)

$$\int_0^{ab} t^n \exp(-t^2/2) dt = \sqrt{\pi/2} (n-1)!! P(2a^2 b^2 / n+1) \quad (\text{a.9.14})$$

donde utilizando las expresiones (a.3.4) y (a.3.12a) del Apéndice A.3 se obtiene:

$$P(2a^2 b^2 / n+1) = 1 - Q(2a^2 b^2 / n+1) \quad (\text{a.9.15})$$

$$I'_n(a, b) = \sqrt{\pi/2} (n-1)!! / 2^{(n+1)/2} b^{n+1} [\text{erf}(ab) - \sqrt{2/\pi} \exp(-a^2 b^2) \sum_{r=1}^{n/2} (\sqrt{2} ab)^{2r-1} / (2r-1)!!] \quad (\text{a.9.16})$$

con lo que las integrales I_H e I_J resultan:

$$I_H = \int_0^a \epsilon^2 \exp(-\epsilon^2 / 2\sigma_\epsilon^2) d\epsilon = I'_2(a_1 \ 1/\sqrt{2} \ \sigma_\epsilon) \quad (\text{a.9.17})$$

$$I_J = \int_0^a \epsilon^4 \exp(-\epsilon^2 / 2\sigma_\epsilon^2) d\epsilon = I'_4(a_1 \ 1/\sqrt{2} \ \sigma_\epsilon) \quad (\text{a.9.18})$$

donde de nuevo, $a = \sqrt{3} \sigma_\epsilon$ para el Apéndice A.5. y $a = \sigma_\epsilon$ para el A.6. Sustituyendo adecuadamente resulta:

Para el Apéndice A.5.:

$$I_H = \sqrt{\pi/2} \sigma_\epsilon^3 \operatorname{erf}(\sqrt{3}\sigma_u / \sqrt{2}\sigma_\epsilon) - \sqrt{3}\sigma_u \sigma_\epsilon^2 \exp(-3\sigma_u^2 / 2\sigma_\epsilon^2) \quad (\text{a.9.19})$$

$$I_I = -3\sqrt{3} \sigma_\epsilon^5 (\sigma_u / \sigma_\epsilon + \sigma_u^3 / \sigma_\epsilon^3) \exp[-3\sigma_u^2 / 2\sigma_\epsilon^2] + \\ + 3\sqrt{\pi/2} \sigma_\epsilon^5 \operatorname{erf}(3\sigma_u / 2\sigma_\epsilon) \quad (\text{a.9.20})$$

Para el Apéndice A.6.:

$$I_H = \sqrt{\pi/2} \sigma_\epsilon^3 \operatorname{erf}(\sigma_u / \sqrt{2}\sigma_\epsilon) - \sigma_u \sigma_\epsilon^2 \exp(-\sigma_u^2 / 2\sigma_\epsilon^2) \quad (\text{a.9.21})$$

A.10- CALCULO DE $S_2(N)$ CON RUIDO DE PLANTA

Según la expresión (iii.2.2.29), $S_2(n)$ queda definida como:

$$S_2(n) = [k^2 N \sigma_x^4 / \sigma_\epsilon^2(n)] \sum_{j=0}^{2k-2} \binom{2k-2}{j} E\{\epsilon^j(n)\} E\{u^{2k-2-j}\} \quad (\text{a.10.1})$$

donde se observa la dependencia con respecto a los momentos centrales de las v.a. $\epsilon(n)$ y u . Debido a que se admiten f.d.p. pares para $\epsilon(n)$ (Gaussiana) y u , los momentos impares son nulos. Los valores correspondientes a la v.a. $\epsilon(n)$ son, dado su carácter Gaussiano, los mencionados en (iii.2.2.3), esto es:

$$E\{\epsilon^j(n)\} = (j-1)!! \sigma_\epsilon^j(n) \quad (j=\text{par}) \quad (\text{a.10.2})$$

para la v.a. u , según las diferentes f.d.p. adoptadas resultan:

a) u es Gaussiana

Igualmente, según (iii.2.2.3)

$$E\{u^{2k-2-j}\} = (2k-3-j)!! \sigma_u^{2k-2-j}, \quad (2k-2-j=\text{par}) \quad (\text{a.10.3})$$

b) u es Laplaciana

Se define el momento central m -simo como:

$$E\{u^m\} = (1/\sqrt{2} \sigma_u) \int_{-\infty}^{\infty} u^m \exp(-\sqrt{2}|u|/\sigma_u) du \quad (\text{a.10.4})$$

$$\text{como } I(a) = \int_0^{\infty} \exp(-ax) dx = \frac{1}{a} \quad (\text{a.10.5})$$

se cumple derivando n veces:

$$d^n I(a) / da^n = (-1)^n \int_0^{\infty} x^n \exp(-ax) dx = (-1)^n n! / a^{n+1} \quad (\text{a.10.6})$$

y por tanto:

$$E\{u^m\} = \sqrt{2}/\sigma_u \int_0^{\infty} u^m \exp(-\sqrt{2}u/\sigma_u) du = (\sqrt{2}/\sigma_u) d^m I(a) / da^m \Big|_{a=\sqrt{2}/\sigma_u} \quad (\text{a.10.7})$$

sustituyendo:

$$E\{u^m\} = (-1)^m m! \sigma_u^m / 2^{m/2} \quad (\text{a.10.8})$$

pero para m par:

$$E\{u^m\} = (m! / 2^{m/2}) \sigma_u^m \quad (\text{a.10.9})$$

c) u es Uniforme

Se define el momento central m-simo como:

$$E\{u^m\} = (1/2 \sqrt{3} \sigma_u) \int_{-\sqrt{3}\sigma_u}^{\sqrt{3}\sigma_u} u^m du \quad (\text{a.10.10})$$

con lo que para m=par

$$E\{u^m\} = (3^{m/2} / m+1) \sigma_u^m \quad (\text{a.10.11})$$

d) u es Binaria

Se define el momento central m-simo como:

$$E\{u^m\} = \frac{1}{2} \int_{-\infty}^{\infty} u^m [\delta(u+\sigma_u) + \delta(u-\sigma_u)] du \quad (\text{a.10.12})$$

con lo que para m=par:

$$E\{u^m\} = \sigma_u^m \quad (\text{a.10.13})$$

Particularizando las expresiones anteriores para los 4 primeros algoritmos de la familia, esto es L_1 a L_4 , se obtiene:

1. Para L_1

$$S_2(n) = N_x^4 / \sigma_u^2 s^2 \quad (\text{a.10.14})$$

2. Para L_2

$$S_2(n) = 4N_x^4 (1+s^2) / s^2 \quad (\text{a.10.15})$$

3. Para L₃

$$S_2(n) = \begin{cases} 27N\sigma_x^4\sigma_u^2(1+s^2)^2/s^2; & \text{Gaussiana} & \text{(a.10.16a)} \\ 27N\sigma_x^4\sigma_u^2(2+2s^2+s^4)/s^2; & \text{Laplaciana} & \text{(a.10.16b)} \\ 27N\sigma_x^4\sigma_u^2(3/s+2s^2+s^4)/s^2; & \text{Uniforme} & \text{(a.10.16c)} \\ 27N\sigma_x^4\sigma_u^2(1/3+2s^2+s^4)/s^2; & \text{Binaria} & \text{(a.10.16d)} \end{cases}$$

4. Para L₄

$$S_2(n) = \begin{cases} 240N\sigma_x^4\sigma_u^4(1+s^2)^3/s^2; & \text{Gaussiana} & \text{(a.10.17a)} \\ 240N\sigma_x^4\sigma_u^4(6+6s^2+3s^4+s^6)/s^2; & \text{Laplaciana} & \text{(a.10.17b)} \\ 48N\sigma_x^4\sigma_u^4(9/5+15s^2+15s^4+5s^6)/s^2; & \text{Uniforme} & \text{(a.10.17c)} \\ 240N\sigma_x^4\sigma_u^4(1/15+s^2+3s^2+3s^4+s^6)/s^2; & \text{Binaria} & \text{(a.10.17d)} \end{cases}$$

A.11- CARACTERIZACION DE LA V.A. COMPLEJA $X(m,1)$ Y CALCULO DE SU MOMENTO ABSOLUTO DE ORDEN r .

Sea la DFT de un registro de L puntos de $x(n)$:

$$X(m) = \sum_{n=0}^{L-1} x(n) \exp(-j2\pi mn/L) = X_r(m) - jX_i(m) \quad (\text{a.11.1})$$

donde:

$$X_r(m) = \sum_{n=0}^{L-1} x(n) \cos(2\pi mn/L) \quad (\text{a.11.2a})$$

$$X_i(m) = \sum_{n=0}^{L-1} x(n) \text{sen}(2\pi mn/L) \quad (\text{a.11.2b})$$

Para caracterizar conjuntamente a $X_r(m)$ y $X_i(m)$ se parte de las v.a. $x(n)$ definidas independientes y Gaussianas $N(0, \sigma_x^2)$ según la hipótesis H1 del Capítulo IV.

Al estar $X_r(m)$ y $X_i(m)$ formadas por la suma de v.a. Gaussianas ellas mismas son Gaussianas con:

$$E\{X_r(m)\} = 0 \quad (\text{a.11.3a})$$

$$E\{X_i(m)\} = 0 \quad (\text{a.11.3b})$$

$$\text{Var}\{X_r(m)\} = \sum_{n=0}^{L-1} \sigma_x^2 \cos^2(2\pi nm/L) \quad (\text{a.11.3c})$$

$$\text{Var}\{X_i(m)\} = \sum_{n=0}^{L-1} \sigma_x^2 \text{sen}^2(2\pi nm/L) \quad (\text{a.11.3d})$$

donde:

$$\text{Var}\{X_r(m)\} + \text{Var}\{X_i(m)\} = L\sigma_x^2 \quad (\text{a.11.4})$$

y por tanto, por simetría circular:

$$\text{Var}\{X_r(m)\} = \text{Var}\{X_i(m)\} = \sigma_x^2 L/2 \quad (\text{a.11.5})$$

La correlación cruzada es:

$$E\{X_r(m)X_i(m)\} = \sum_{n=0}^{L-1} \sigma_x^2 \cos(2\pi nm/L) \text{sen}(2\pi nm/L) = E\{X_i(m)X_r(m)\} \quad (\text{a.11.6})$$

Si se llama $W_L = \exp(j2\pi/L)$ se cumple:

$$\sum_{n=0}^{L-1} (W_L^{-nm})^2 = (1 - W_L^{-2Lm}) / (1 - W_L^{-2m}) \quad (\text{a.11.7})$$

pero como

$$W_L^{-2Lm} = \exp(-j4\pi m) = 1 \quad (\text{a.11.8})$$

se obtiene para (a.11.7):

$$0 = \sum_{n=0}^{L-1} (\cos^2(2\pi nm/L) - \text{sen}^2(2\pi nm/L) - 2j \text{sen}(2\pi nm/L) \cos(2\pi nm/L)) \quad (\text{a.11.9})$$

de donde se deduce:

$$\sum_{n=0}^{L-1} \text{sen}(2\pi nm/L) \cos(2\pi nm/L) = 0 \quad (\text{a.11.10})$$

con lo que:

$$E\{X_r(m)X_i(m)\} = E\{X_i(m)X_r(m)\} = 0 \quad (\text{a.11.11})$$

y por lo tanto se concluye que la v.a. $X(m)$ es Gaussiana circular compleja, esto es:

$$E\{X_r(m)\} = E\{X_i(m)\} = 0 \quad (\text{a.11.12a})$$

$$\text{Var}\{X_r(m)\} = \text{Var}\{X_i(m)\} = L\sigma_x^2 \quad (\text{a.11.12b})$$

$$E\{X_r(m)X_i(m)\} = E\{X_i(m)X_r(m)\} = 0 \quad (\text{a.11.12c})$$

En el caso en que $m=0$, $X_i(m)=0$, con lo que

$X(m)=X_r(m)$ es una v.a. normal de $N(0, L\sigma_x^2)$.

Para el caso del momento absoluto r -ésimo se procede:

$$E\{|X(m)|^r\} = E\{(X_r^2(m) + X_i^2(m))^{r/2}\} \quad (m \neq 0) \quad (\text{a.11.13a})$$

$$E\{|X(m)|^r\} = E\{|X_r(m)|^r\} \quad (m=0) \quad (\text{a.11.13b})$$

En el primer caso ($m \neq 0$), con el cambio $z = X_r^2(m) + X_i^2(m)$ se obtiene una v.a. Rayleigh y en el segundo caso es v.a. Gaussiana. Según [Papoulis, 1984], los momentos absolutos de ambas variables aleatorias son:

Para $m \neq 0$

$$E\{|X(m)|^r\} = \begin{cases} (L/2)^{r/2} \sqrt{\pi/2} r!! \sigma_x^r; & r=\text{impar} \\ (L/2)^{r/2} r!! \sigma_x^r; & r=\text{par} \end{cases} \quad (\text{a.11.14a})$$

$$(\text{a.11.14b})$$

Para $m=0$

$$E\{|X(m)|^r\} = \begin{cases} L^{r/2} \sqrt{2/\pi} (r-1)!! \sigma_x^r; & r=\text{impar} \\ L^{r/2} (r-1)!! \sigma_x^r; & r=\text{par} \end{cases} \quad (\text{a.11.15a})$$

$$(\text{a.11.15b})$$

A.12- EVALUACION DE LA ESPERANZA $E\{|E+U|^{2k-2}\}$

Sea

$$Q_2 = E\{|E+U|^{2k-2}\} \quad (\text{a.12.1})$$

donde E y U son

caso 1: 2 v.a. Gaussianas circulares complejas e independientes con $N(0, \sigma_E^2)$ y $N(0, \sigma_U^2)$

caso 2: 2 v.a. Gaussianas reales e independientes

A.12.1 Análisis del caso 1

Q_2 se puede escribir como:

$$Q_2 = E\{[(E_r + U_r)^2 + (E_i + U_i)^2]^{2k-2}\} \quad (\text{a.12.2})$$

y donde realizando los cambios:

$$A = E_r + U_r \quad (\text{a.12.3a})$$

$$B = E_i + U_i \quad (\text{a.12.3b})$$

A y B son 2 v.a. Gaussianas independientes con $N(0, \sigma_E^2 + \sigma_U^2)$ y realizando nuevamente el cambio:

$$z = \sqrt{A^2 + B^2} \quad (\text{a.12.4})$$

Q_2 queda como:

$$Q_2 = E\{z^{2k-2}\} \quad (\text{a.12.5})$$

es el momento par de una Rayleigh, que según [PAPOULIS, 1984]

$$Q_2 = (2k-2)!! (\sigma_E^2 + \sigma_U^2)^{k-1} = (2k-2)!! (1 + \sigma_U^2 / \sigma_E^2)^{k-1} \sigma_E^{2k-2} \quad (\text{a.12.6})$$

A.12.2 Análisis del caso 2

Q_2 se puede escribir como:

$$Q_2 = E\{(E_r + U_r)^{2k-2}\} \quad (\text{a.12.7})$$

que nuevamente haciendo el cambio

$$z = E_r + U_r \quad (\text{a.12.8})$$

resulta una v.a. Gaussiana con $N(0, \sigma_E^2 + \sigma_U^2)$ y según [PAPOULIS, 1984], el momento par de una v.a. Gaussiana.

$$Q_2 = (2k-3)!! (\sigma_E^2 + \sigma_U^2)^{k-1} = (2k-3)!! (1 + \sigma_U^2 / \sigma_E^2)^{k-1} \sigma_E^{2k-2} \quad (\text{a.12.9})$$

A.13- EVALUACION DE LA ESPERANZA $E\{|E+U|^{k-2}(|E|^2 + \text{Re}(EU^*))\}$

$$\text{Sea } Q_1 = E\{|E+U|^{k-2}(|E|^2 + \text{Re}(EU^*))\} \quad (\text{a.13.1})$$

donde E y U son:

caso 1: 2 v.a. Gaussianas complejas circularmente e independientes de media nula y varianzas σ_E^2 y σ_U^2

caso 2: 2 v.a. Gaussianas reales e independientes de media nula y varianzas σ_E^2 σ_U^2

Ambos casos se analizan secuencialmente.

a.13.1. Análisis del caso 1

Se puede escribir en este caso:

$$Q_1 = E\{[(E_r + U_r)^2 + (E_i + U_i)^2]^{(k-2)/2} (E_r^2 + E_i^2 + E_r U_r + E_i U_i)\} \quad (\text{a.13.2})$$

o igualmente:

$$Q_1 = \int_{E_r} \int_{E_i} I(E) f(E_r) f(E_i) dE_r dE_i \quad (\text{a.13.3})$$

donde I(E) es la esperanza condicionada a E_r e E_i :

$$I(E) = E_{U_r U_i} \{[(E_r + U_r)^2 + (E_i + U_i)^2]^{(k-2)/2} (E_r^2 + E_i^2 + E_r U_r + E_i U_i)\} \quad (\text{a.13.4})$$

Si, admitiendo independencia y Gaussianness para U_r e U_i , se realizan los cambios:

$$E_r + U_r = A \quad (\text{a.13.5a})$$

$$E_i + U_i = B \quad (\text{a.13.5b})$$

$$z = \sqrt{A^2 + B^2} \quad (\text{a.13.5c})$$

$$\vartheta = \text{tg}^{-1}(B/A) \quad (\text{a.13.5d})$$

El Jacobiano resulta $J=1/Z$ y la nueva función de densidad conjunta:

$$f_{z\theta} = (z/2\pi\sigma_U^2) \exp[-(z^2 + \alpha^2 - 2zE_r \cos\theta - 2zE_i \sin\theta) / 2\sigma_U^2] \quad (\text{a.13.6})$$

donde se ha usado $\alpha^2 = E_r^2 + E_i^2$.

Con esto la expresión para $I(E)$ queda:

$$I(E) = (1/2\pi\sigma_U^2) \int_0^{2\pi} \int_0^\infty z^k (E_r \cos\theta + E_i \sin\theta) \cdot \exp[-(z^2 + \alpha^2 - 2z(E_r \cos\theta + E_i \sin\theta)) / 2\sigma_U^2] dz d\theta \quad (\text{a.13.7})$$

como según [ABRAMOWITZ, 1970] la función modificada de Bessel $I_0(x)$ se puede expresar:

$$I_0(x) = (1/2\pi) \int_0^{2\pi} \exp(x \cos\theta) d\theta \quad (\text{a.13.8})$$

haciendo en (a.13.7) el cambio

$$(E_r \cos\theta + E_i \sin\theta) / \sqrt{E_r^2 + E_i^2} = \cos\phi \quad (\text{a.13.9})$$

o equivalentemente

$$\phi = \theta - \text{tg}^{-1}(E_i/E_r) \quad (\text{a.13.10})$$

se obtiene:

$$I(E) = \exp(-\alpha^2/2\sigma_U^2) \int_0^\infty z^{k-1} \exp(-z^2/2\sigma_U^2) \cdot ((1/2\pi) \int_0^{2\pi} (z\alpha/\sigma_U^2) \cos\phi \exp(z\alpha \cos\phi/\sigma_U^2) d\phi) dz \quad (\text{a.13.11})$$

pero de (a.13.8) se deduce derivando con respecto a x :

$$x I_0'(x) = (1/2\pi) \int_0^{2\pi} x \cos\theta \exp(x \cos\theta) d\theta \quad (\text{a.13.12})$$

con lo que $I(E)$ se puede escribir:

$$I(E) = (\alpha \exp(-\alpha^2/2\sigma_U^2)) \int_0^\infty z^k I_0'(z\alpha I_0'(z\alpha/\sigma_U^2) \exp(-z^2/2\sigma_U^2)) dz \quad (\text{a.13.13})$$

Según [ABRAMOWITZ, 1970] se cumplen las siguientes soluciones:

$$I_0'(x) = I_1(x) \quad (\text{a.13.14a})$$

$$jI_1 = J_1(jx) \quad (\text{a.13.14b})$$

$$\int_0^\infty \exp(-a^2 t^2) t^{\mu-1} J_1(bt) dt = (b/2a) \Gamma(1/2 + \mu/2) M(\mu/2 + 1/2, 2, -b^2/4a^2) / 2a^\mu \quad (\text{a.13.14c})$$

con lo que combinándolas convenientemente y simplificando para $I(E)$ la siguiente relación:

$$I(E) = 2^{(k/2-1)} \alpha^2 \sigma_U^{k-2} \exp(-\alpha^2/2\sigma_U^2) \Gamma(k/2+1) M(k/2+1, 2, \alpha^2/2\sigma_U^2) \quad (\text{a.13.15})$$

donde $\Gamma(\cdot)$ es la función Gamma y $M(\cdot, \cdot, \cdot)$ es la función hipergeométrica confluyente definida como:

$$M(a, b, z) = \sum_{n=0}^{\infty} (a)_n z^n / (b)_n n! \quad (\text{a.13.16})$$

con $(s)_n = \Gamma(n+s) / \Gamma(s)$

como este caso particular: $a=k/2+1$, $b=2$, se obtiene

$$(a)_n = \Gamma(n+1+k/2) / \Gamma(k/2+1) \quad (\text{a.13.17a})$$

$$(b)_n = (n+1)! \quad (\text{a.13.17b})$$

con lo que:

$$M(k/2+1, 2, \alpha^2/2\sigma_U^2) = \sum_{n=0}^{\infty} (\alpha/\sigma_U)^{2n} \Gamma(n+1+k/2) / 2^n n! (n+1)! \quad (\text{a.13.18})$$

y por tanto:

$$I(E) = 2^{k/2-1} \alpha^2 \sigma_U^{k-2} \exp(-\alpha^2/2\sigma_U^2) \sum_{n=0}^{\infty} (\alpha/\sigma_U)^{2n} \Gamma(n+1+k/2) / 2^n n! (n+1)! \quad (\text{a.13.19})$$

sustituyendo esta expresión en (a.13.13), teniendo en cuenta que E_r y E_i son Gaussianas independientes y haciendo el cambio $\alpha^2 = E_r^2 + E_i^2$ se obtiene:

$$Q_1 = (2^{k/2-1} \sigma_U^{k-2} \sigma_E^2 / \sigma^2) \sum_{n=0}^{\infty} \Gamma(n+1+k/2) / 2^n n! (n+1)! \cdot \int_0^{\infty} (\alpha^{2n+3} / \sigma_U^{2n} \sigma^2) \exp(-\alpha^2 / 2\sigma^2) d\alpha \quad (\text{a.13.20})$$

donde $1/\sigma^2 = 1/\sigma_U^2 + 1/\sigma_E^2$ (a.13.21)

y la función integral es el momento $2n+2$ -simo de una v.a. Rayleigh.

Como según [PAPOULIS, 1984], el momento par de una Rayleigh es:

$$\int_0^{\infty} (\alpha^{2n+3} / \sigma^2) \exp(-\alpha^2 / 2\sigma^2) d\alpha = (2n+2)!! \sigma^{2n+2} = 2^{n+1} \Gamma(n+2) \sigma^{2n+2} \quad (\text{a.13.22})$$

queda finalmente para Q_1 la siguiente expresión no cerrada:

$$Q_1 = [2^{k/2} \sigma_U^k / (1 + \sigma_E^2 / \sigma_U^2) (1 + \sigma_U^2 / \sigma_E^2)] \sum_{n=0}^{\infty} \Gamma(n+1+k/2) / n! (1 + \sigma_U^2 / \sigma_E^2)^n \quad (\text{a.13.23})$$

a.13.2. Análisis del caso 2

En este caso:

$$Q_1 = E\{|E_r + U_r|^{k-2} E_r(E_r + U_r)\} \quad (\text{a.13.24})$$

donde igualmente que en el caso anterior:

$$Q_1 = \int_{E_r} I(E) f(E_r) dE_r \quad (\text{a.13.25})$$

con

$$I(E) = E_{U_r} \{|E_r + U_r|^{k-2} E_r(E_r + U_r)\} \quad (\text{a.13.26})$$

Desarrollando la potencia en $k-2$ por el binomio de Newton y recordando que U_r es $N(0, \sigma_u^2)$ se obtiene:

para $k=\text{par}$

$$I(E) = \sum_{j=0}^{k-1} \binom{k-j}{j} E_r^{j+1} \int_{-\infty}^{\infty} (1/\sqrt{2\pi\sigma_u^2}) U_r^{k-1-j} \exp(-U_r^2/2\sigma_u^2) dU_r \quad (\text{a.13.27})$$

pero la integral es el momento $k-i-j$ -ésimo μ_{k-i-j} de una Gaussiana con lo que:

$$I(E) = \sum_{j=0}^{k-1} \binom{k-1}{j} \mu_{k-1-j} E_r^{j+1} \quad (\text{a.13.28})$$

para $k=\text{impar}$

$$I(E) = \sum_{j=0}^{k-1} \binom{k-1}{j} E_r^{j+1} \int_{-\infty}^{\infty} (1/\sqrt{2\pi\sigma_U^2}) \text{sgn}(E_r + U_r) U_r^{k-1-j} \exp(-U_r^2/2\sigma_U^2) dU_r \quad (\text{a.13.29})$$

la integral anterior es conocida como $g_j(E_r)$, véase (iii.2.2.11) y en el caso de U_r gaussiana resulta según (a.3.11):

$$g_j(E_r) = \begin{cases} (q-1)!! \sigma_U^q \text{sgn}(E_r) P[(E_r/\sigma_U)^2/q+1]; & q \text{ par} & (\text{a.13.30a}) \\ (2(q-1)!! \sigma_U^q / \sqrt{2\pi}) Q[(E_r/\sigma_U)^2/q+1]; & q \text{ impar} & (\text{a.13.30b}) \end{cases}$$

con lo que

$$I(E) = \sum_{j=0}^{k-1} \binom{k-1}{j} E_r^{j+1} g_j(E_r) \quad (\text{a.13.31})$$

con estas expresiones para $I(E)$ queda finalmente:

para $k=\text{par}$

$$Q_1 = \sum_{j=0}^{k-1} \binom{k-1}{j} \mu_{k-1-j} \int E_r^{j+1} f(E_r) dE_r = \sum_{j=0}^{k-1} \binom{k-1}{j} \mu_{k-1-j} \mu_{j+1} \quad (\text{a.13.32})$$

donde μ_i es el momento i -ésimo Gaussiano. Se hace notar que los momentos impares son nulos, al ser E_r y U_r de medias nulas.

para $k=\text{impar}$

$$Q_1 = \sum_{j=0}^{k-1} \binom{k-1}{j} \phi_j \quad (\text{a.13.33})$$

con $\phi_j(n)$ definida en (iii.2.2.10) como:

$$\phi_j = \int E_r^{j+1} g_j(E_r) f(E_r) dE_r \quad (\text{a.13.34})$$

integral resuelta para algunos casos particulares ($j=0,1,2$) en el Apéndice A.3.

Las expresiones calculadas para Q_1 son las generales en el caso de elección arbitraria del orden del algoritmo k . No obstante, en los siguientes dos Apéndices se desarrollan las expresiones cerradas de Q_1 particularizadas para los primeros cuatro algoritmos L_1 a L_4 .

A.14- EVALUACION DE Q_1 PARA LOS ALGORITMOS L_2 Y L_4

Se analizan los casos con:

- a) $m=0$: v.a. reales
- b) $m \neq 0$: v.a. complejas

-En el primer caso con $m=0$ se obtienen respectivamente según (a.13.32):

para $k=2$

$$Q_1 = \sigma_E^2 \quad (\text{a.14.1})$$

para $k=4$

$$Q_1 = 3\sigma_U^2 \sigma_E^2 + 3\sigma_E^4 = 3\sigma_E^4 (1 + \sigma_U^2 / \sigma_E^2) \quad (\text{a.14.2})$$

-En el segundo caso con $m \neq 0$ partiendo de $I(E)$ (a.3.15) se evalúa primeramente la función hipergeométrica confluyente $M(k/2+1, 2, \alpha^2/2\sigma_U^2)$. Para ello según [ABRAMOWITZ, 1970] se parte de las siguientes relaciones:

$$M(1, 1, z) = \exp z \quad (\text{a.14.3a})$$

$$M'(a, b, z) = (a/b)M(a+1, b+1, z) \quad (\text{a.14.3b})$$

$$M(a+1, b, z) = M(a, b, z) + (z/a)M'(a, b, z) \quad (\text{a.14.3c})$$

donde para $k=2$:

$$M(2, 2, z) = \exp z \quad (\text{a.14.4})$$

y para $k=4$:

$$M(3, 2, z) = M(2, 2, z) + (z/2)M'(2, 2, z) = (1+z/2) \exp z \quad (\text{a.14.5})$$

se ha utilizado $z = \alpha^2 / 2\sigma_U^2$

considerando además que $\Gamma(n+1) = n!$, se obtienen las siguientes expresiones

para k=2

$$I(E) = \alpha^2 \quad (\text{a.14.6})$$

para k=4

$$I(E) = 4 \alpha^2 \sigma_U^2 (1 + \alpha^2 / 4\sigma_U^2) \quad (\text{a.14.7})$$

sustituyendo estas expresiones en (a.13.3), teniendo presente que E_r y E_i son v.a. Gaussianas independientes y haciendo el cambio $\alpha = \sqrt{E_r^2 + E_i^2}$ (v.a. Rayleigh) se obtiene según [PAPOULIS, 1984] :

para k=2

$$Q_1 = \int_0^\infty (\alpha^3 / \sigma_E^2) \exp(-\alpha^2 / 2\sigma_E^2) d\alpha = 2\sigma_E^2 \quad (\text{a.14.8})$$

para k=4

$$Q_1 = \int_0^\infty (1/\sigma_E^2) (4\alpha^3 \sigma_U^2 + \alpha^5) \exp(-\alpha^2 / 2\sigma_E^2) d\alpha = 8\sigma_E^2 \sigma_U^2 + 8\sigma_E^4 = 8\sigma_E^4 (1 + \sigma_U^2 / \sigma_E^2) \quad (\text{a.14.9})$$

A.15- EVALUACION DE Q_1 PARA LOS ALGORITMOS L_1 Y L_3

Se analizan los casos con

- a) $m=0$: v.a. reales
- b) $m \neq 0$: v.a. complejas

-En el primer caso con $m=0$ se obtiene

para $k=1$

$$Q_1 = \phi_0 \quad (\text{a.15.1})$$

donde según el Apéndice A.3, expresión (a.3.16)

$$Q_1 = (\sqrt{2/\pi}/\sigma_U) \sigma_E^2 / \sqrt{1 + \sigma_E^2/\sigma_U^2} \quad (\text{a.15.2})$$

para $k=3$

$$Q_1 = \phi_0 + 2\phi_1 + \phi_3 \quad (\text{a.15.3})$$

donde según el Apéndice A.3 utilizando las expresiones (a.3.19) y (a.3.27) se obtiene:

$$Q_1 = 6 \sqrt{2/\pi} \sigma_x^2 \sigma_U [1 + 2\sigma_E^2/\sigma_U^2 + (\sigma_E^2/\sigma_U^2)/(1 + \sigma_E^2/\sigma_U^2)] / \sqrt{1 + \sigma_E^2/\sigma_U^2} \quad (\text{a.15.4})$$

-En el segundo caso con $m \neq 0$ partiendo nuevamente de $I(E)$ (a.13.15) y evaluando igualmente $M(k/2+1, 2, \alpha^2/2 \sigma_U^2)$ usando (a.14.3c) con [ABRAMOWITZ, 1970]):

$$M(1/2, 1, z) = \exp(z/2) I_0(z/2) \quad (\text{a.15.5})$$

se obtiene para $k=1$ y $k=3$:

$k=1$

$$M(3/2, 2, z) = 2M'(1/2, 1, z) = [I_0(z/2) + I_1(z/2)] \exp(z/2) \quad (\text{a.15.6})$$

k=3

$$M(5/2, 2, z) = M(3/2, 2, z) + (2z/3)M'(3/2, 2, z) \quad (\text{a.15.7})$$

derivando (a.15.5) y considerando la relación (según [ABRAMOWITZ, 1970]):

$$(1/zd/dz)^r [z^{-q} I_q(z)] = z^{-(q+r)} I_{q+r}(z) \quad (\text{a.15.8})$$

se obtiene:

$$M'(3/2, 2, z) = [3I_0(z/2)/4 + I_1(z/2) + I_2(z/2)/4] \exp(z/2) \quad (\text{a.15.9})$$

con lo que sustituyéndola en (a.15.7) además de (a.15.6) se obtiene:

$$M(5/2, 2, z) = \exp(z/2) [(1+z/2)I_0(z/2) + (1+2z/3)I_1(z/2) + (z/6)I_2(z/2)] \quad (\text{a.15.10})$$

con este último resultado y considerando que $\Gamma(n+1/2) = (2n-1)!! \sqrt{\pi}/2^n$, se obtiene según (a.3.15) las siguientes expresiones particulares:

para k=1

$$I(E) = z \sqrt{\pi/2} [I_0(z/2) + I_1(z/2)] \exp(-z/2) \quad (\text{a.15.11})$$

para k=3

$$I(E) = 3z \sqrt{\pi/2} [(1+z/2)I_0(z/2) + (1+2z/3)I_1(z/2) + (z/6)I_2(z/2)] \quad (\text{a.15.12})$$

con $z = \alpha^2/2 \sigma_u^2$

con estas expresiones de $I(E)$ se obtienen las correspondientes de Q_1 en los casos de $k=1$ y $k=3$. Para ambos casos se sustituye en (a.13.3) con $\alpha = \sqrt{E_r^2 + E_i^2}$ (v.a. Rayleigh) y se obtiene:

para k=1

$$Q_1(\sqrt{\pi/2}/2 \sigma_u \sigma_E) \int_0^\infty \alpha^3 \exp(-\alpha^2/2\sigma^2) [I_0(\alpha^2/4\sigma_u^2) + I_1(\alpha^2/4\sigma_u^2)] dx \quad (\text{a.15.13})$$

donde se ha utilizado

$$1/\sigma^2 = 1/\sigma_E^2 + 1/\sigma_U^2 \quad (\text{a.15.14})$$

haciendo el cambio $\alpha^2/4\sigma_U = x$ se obtiene para Q_1 :

$$Q_1 (4\sqrt{\pi/2}\sigma_U^3/\sigma_E^2) \int_0^\infty x \exp(-2\sigma_U^2 x/\sigma^2) [I_0(x) + I_1(x)] dx \quad (\text{a.15.15})$$

si se llama a las integrales

$$I_P = \int_0^\infty x \exp(-bx) I_0(x) dx \quad (\text{a.15.16a})$$

$$I_Q = \int_0^\infty x \exp(-bx) I_1(x) dx \quad (\text{a.15.16b})$$

como según [ERDELYI, 1954] se cumple:

$$\int_0^\infty x^n \exp(-ax) J_0(xy) dx = \frac{n!}{(a^2 + y^2)^{(n+1)/2}} P_n(a/\sqrt{a^2 + y^2}) \quad (\text{a.15.17})$$

con $J_0(jx) = I_0(x)$ y $P_n(\cdot)$ el polinomio de Legendre de orden n , se obtiene para I_P :

$$I_P = b/(b^2 - 1)^{3/2} \quad (\text{a.15.18})$$

como además según [ERDELYI, 1954]

$$\int_0^\infty x \exp(-ax) J_1(xy) dx = y/(a^2 + y^2)^{3/2} \quad (\text{a.15.19})$$

con $J_1(jx) = jI_1(x)$ se obtiene para I_Q :

$$I_Q = 1/(b^2 - 1)^{3/2} \quad (\text{a.15.20})$$

con estos resultados para I_P e I_Q , haciendo $b = 2\sigma_U^2/\sigma^2$ y sustituyéndolos en (a.15.15) se obtiene para Q_1 :

$$Q_1 = \sqrt{\pi/2} \sigma_E / \sqrt{1 + \sigma_U^2/\sigma_E^2} \quad (\text{a.15.21})$$

para $k=3$

$$Q_1 = (3 \pi / 2 \sigma_U / 2 \sigma_E^2) \int_0^\infty \alpha^3 \exp(-\alpha^2 / 2 \sigma^2) [(1 + \alpha^2 / 4 \sigma_U) I_0(\alpha^2 / \sigma_U) + (1 + \alpha^2 / 3 \sigma_U^2) I_1(\alpha^2 / 4 \sigma_U) + (\alpha^2 / 12 \sigma_U I_2)(\alpha^2 / 4 \sigma_U)] dx \quad (\text{a.15.22})$$

donde se ha utilizado nuevamente (a.15.14). Haciendo los cambios $\alpha^2 / 4 \sigma_U = x$ y $b = 2 \sigma_U^2 / \sigma^2$ se obtiene para Q_1 :

$$Q_1 = (12 \sqrt{\pi} / 2 \sigma_U^5 / \sigma_E^2) \int_0^\infty \exp(-bx) [(x + x^2) I_0(x) + (x + 4x^2 / 3) I_1(x) + (x^2 / 3) I_2(x)] dx \quad (\text{a.15.23})$$

llámese a las integrales

$$I_R = \int_0^\infty x^2 \exp(-bx) I_0(x) dx \quad (\text{a.15.24a})$$

$$I_S = \int_0^\infty x^2 \exp(-bx) I_1(x) dx \quad (\text{a.15.24b})$$

$$I_T = \int_0^\infty x^2 \exp(-bx) I_2(x) dx \quad (\text{a.15.24c})$$

donde utilizando la relación (a.15.17) para resolver I_R se obtiene:

$$I_R = (2b^2 + 1) / (b^2 - 1)^{5/2} \quad (\text{a.15.25})$$

nuevamente según [ERDELYI, 1954]

$$\int_0^\infty x^2 \exp(-ax) J_1(xy) dx = 3ay / (a^2 + y^2)^{5/2} \quad (\text{a.15.26})$$

$$\int_0^\infty x^2 \exp(-ax) J_2(xy) dx = 3y^2 / (a^2 + y^2)^{5/2}$$

con lo que se obtiene para I_S e I_T :

$$I_S = 3b / (b^2 - 1)^{5/2} \quad (\text{a.15.27})$$

$$I_T = 3 / (b^2 - 1)^{5/2} \quad (\text{a.15.28})$$

La expresión (a.15.23) de Q_1 se puede escribir como

$$Q_1 = (12\sqrt{\pi/2}\sigma_U^5/\sigma_E^2) [I_P + I_Q + I_R + (4/3)I_S + (1/3)I_T] \quad (\text{a.15.29})$$

con lo que sustituyendo las expresiones oportunas y simplificando se llega a la expresión final para Q_1 :

$$Q_1 = 3 \sqrt{\pi/2} \sigma_E^3 \sqrt{1 + \sigma_U^2/\sigma_E^2}$$